Chapter 7: Feature Engineering

Accompanying code for chapter 7 of the book Applied Data Science in Tourism: Interdisciplinary Approches, Methodologies and Applications.

This notebook is available online at https://github.com/DrDub/artfeateng/tree/master/tourism.

MIT License

Copyright 2021 Pablo Duboue

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Open in Colab

OS Setup

PySpark uses an embedded Java virtual machine (JVM). If you need to set a specific JVM, uncomment here changing to the path of the JVM

Details: The error message says something in the line of *java.lang.reflect.InaccessibleObjectException: Unable to make private java.nio.DirectByteBuffer(long,int) accessible* see this stack overflow post Spark does't run in Windows anymore for background.

In [1]:
```
import os

# os.environ['JAVA_HOME'] = '/usr/lib/jvm/java-8-openjdk-amd64'
```

This is the set-up for Google Collab. Skip this cell if running locally.

In [ ]:
```
!wget http://artoffeatureengineering.com/tourism/listings-cleaned2.csv
!wget http://artoffeatureengineering.com/tourism/stops.tsv
!pip install pyspark shap
```

Imports

In [2]:
```
import math

import pyspark

from pyspark.sql import SparkSession
from pyspark.sql.functions import udf, countDistinct
from pyspark.sql.types import IntegerType, ArrayType, StringType, DoubleType, BooleanType

from pyspark.ml import Pipeline, UnaryTransformer
from pyspark.ml.feature import VectorAssembler, CountVectorizer, StringIndexer, Interaction
from pyspark.ml.regression import RandomForestRegressor
from pyspark.ml.linalg import VectorUDT

from pyspark.mllib.evaluation import RegressionMetrics

from IPython.display import HTML, display
```

Setup

In [3]:
```
spark = SparkSession.builder.appName('FeatureEngineeringAirBnB').getOrCreate()
```

Loading the data

This data was downloaded from InsideAirBnB.com, an excelent website created by Murray Cox.

The original data was loaded on a spreadsheet program and the text columns were filtered out as NLP is not the focus of this chapter.

By using "infer schema" Spark will do a two-pass inference of basic data fields.

In [4]:
```
alldata = spark.read.csv("listings-cleaned2.csv", inferSchema=True, header=True)
```

We can see how well the inference worked.

In [5]:
```
alldata.printSchema()
```

```
root
 |-- id: integer (nullable = true)
 |-- host_since: string (nullable = true)
 |-- host_location: string (nullable = true)
 |-- host_response_time: string (nullable = true)
 |-- host_response_rate: string (nullable = true)
 |-- host_acceptance_rate: string (nullable = true)
 |-- host_is_superhost: string (nullable = true)
 |-- host_neighbourhood: string (nullable = true)
 |-- host_listings_count: integer (nullable = true)
 |-- host_total_listings_count: integer (nullable = true)
 |-- host_verifications: string (nullable = true)
 |-- host_has_profile_pic: string (nullable = true)
 |-- host_identity_verified: string (nullable = true)
 |-- neighbourhood: string (nullable = true)
 |-- neighbourhood_cleansed: string (nullable = true)
 |-- neighbourhood_group_cleansed: string (nullable = true)
 |-- latitude: double (nullable = true)
 |-- longitude: double (nullable = true)
 |-- property_type: string (nullable = true)
 |-- room_type: string (nullable = true)
 |-- accommodates: integer (nullable = true)
 |-- bathrooms: string (nullable = true)
 |-- bathrooms_text: string (nullable = true)
 |-- bedrooms: integer (nullable = true)
 |-- beds: integer (nullable = true)
 |-- amenities: string (nullable = true)
 |-- price: integer (nullable = true)
 |-- minimum_nights: integer (nullable = true)
 |-- maximum_nights: integer (nullable = true)
 |-- minimum_minimum_nights: integer (nullable = true)
 |-- maximum_minimum_nights: integer (nullable = true)
 |-- minimum_maximum_nights: integer (nullable = true)
 |-- maximum_maximum_nights: integer (nullable = true)
```

```
 |-- minimum_nights_avg_ntm: double (nullable = true)
 |-- maximum_nights_avg_ntm: double (nullable = true)
 |-- has_availability: string (nullable = true)
 |-- availability_30: integer (nullable = true)
 |-- availability_60: integer (nullable = true)
 |-- availability_90: integer (nullable = true)
 |-- availability_365: integer (nullable = true)
 |-- instant_bookable: string (nullable = true)
 |-- calculated_host_listings_count: integer (nullable = true)
 |-- calculated_host_listings_count_entire_homes: integer (nullable = true)
 |-- calculated_host_listings_count_private_rooms: integer (nullable = true)
 |-- calculated_host_listings_count_shared_rooms: integer (nullable = true)
```

In [6]:

```
len(alldata.columns)
```

Out[6]: 45

There are many columns with usable types. Others seem to be assigned `string` when a more suitable type might work (e.g., `date` for `host_since`). But we need to do some EDA on them and for that we need to keep some held-out data to avoid basing our intuitions in the whole dataset (which will not leave additional data to test whether the intuitions are correct).

Data split

We are using 40% of the data for testing and 30% of the training data as development test:

In [7]:

```
(fulldataset, testData) = alldata.randomSplit([0.8, 0.2], seed=42)
(dataset, devset)   = fulldataset.randomSplit([0.6, 0.4], seed=42)
```

Let us now peek at the training data to gain some insights.

In [8]:

```
dataset.show(vertical=True, truncate=False)
```

```
-RECORD 0-----------------------
```

```
 host_since                                     | 2009/11/07
 host_location                                  | Vancouver, British Columbia, Canada
 host_response_time                             | within a day
 host_response_rate                             | 0.71
 host_acceptance_rate                           | 0.71
 host_is_superhost                              | t
 host_neighbourhood                             | Downtown Vancouver
 host_listings_count                            | 1
 host_total_listings_count                      | 1
 host_verifications                             | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'offline_
government_id', 'selfie', 'government_id']
 host_has_profile_pic                           | f
 host_identity_verified                         | t
 neighbourhood                                  | null
 neighbourhood_cleansed                         | West End
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.28209
 longitude                                      | -123.12515
 property_type                                  | Entire condominium
 room_type                                      | Entire home/apt
 accommodates                                   | 2
 bathrooms                                      | null
 bathrooms_text                                 | 1 bath
 bedrooms                                       | 1
 beds                                           | 1
 amenities                                      | ['Kitchen', 'Oven', 'Gym', 'Paid parking off premises', 'Mic
rowave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Extra pillows and blankets', 'Dish
washer', 'Cooking basics', 'Hot water kettle', 'Smoke alarm', 'Heating', 'Single level home', 'Hair dryer',
 'Dedicated workspace', 'Iron', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Hangers',
'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning', 'El
evator']
 price                                          | 90
 minimum_nights                                 | 30
 maximum_nights                                 | 60
 minimum_minimum_nights                         | 30
 maximum_minimum_nights                         | 30
 minimum_maximum_nights                         | 60
 maximum_maximum_nights                         | 60
 minimum_nights_avg_ntm                         | 30.0
 maximum_nights_avg_ntm                         | 60.0
 has_availability                               | t
 availability_30                                | 11
 availability_60                                | 41
 availability_90                                | 70
 availability_365                               | 70
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
```

```
 calculated_host_listings_count_shared_rooms  | 0
-RECORD 1-----------------------
```

```
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 0
 availability_365                             | 228
 instant_bookable                             | t
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
-RECORD 2-------------------
```

```
 price                                          | 94
 minimum_nights                                 | 3
 maximum_nights                                 | 60
 minimum_minimum_nights                         | 3
 maximum_minimum_nights                         | 3
 minimum_maximum_nights                         | 60
 maximum_maximum_nights                         | 60
 minimum_nights_avg_ntm                         | 3.0
 maximum_nights_avg_ntm                         | 60.0
 has_availability                               | t
 availability_30                                | 0
 availability_60                                | 0
 availability_90                                | 0
 availability_365                               | 0
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
-RECORD 3-----------------------
```

```
 bathrooms_text                                 | 1 private bath
 bedrooms                                       | 1
 beds                                           | 1
 amenities                                      | ['Kitchen', 'Cable TV', 'Wifi', 'Essentials', 'Heating', 'Wa
sher', 'Dryer', 'Long term stays allowed']
 price                                          | 50
 minimum_nights                                 | 30
 maximum_nights                                 | 360
 minimum_minimum_nights                         | 30
 maximum_minimum_nights                         | 30
 minimum_maximum_nights                         | 360
 maximum_maximum_nights                         | 360
 minimum_nights_avg_ntm                         | 30.0
 maximum_nights_avg_ntm                         | 360.0
 has_availability                               | t
 availability_30                                | 9
 availability_60                                | 20
 availability_90                                | 50
 availability_365                               | 325
 instant_bookable                               | f
 calculated_host_listings_count                 | 3
 calculated_host_listings_count_entire_homes    | 0
 calculated_host_listings_count_private_rooms   | 3
 calculated_host_listings_count_shared_rooms    | 0
-RECORD 4-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                             | 21913
 host_since                                     | 2010/02/22
 host_location                                  | Vancouver, British Columbia, Canada
 host_response_time                             | within an hour
 host_response_rate                             | 1
 host_acceptance_rate                           | 1
 host_is_superhost                              | t
 host_neighbourhood                             | Renfrew-Collingwood
 host_listings_count                            | 3
 host_total_listings_count                      | 3
 host_verifications                             | ['email', 'phone', 'reviews', 'offline_government_id', 'self
ie', 'government_id']
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                         | Renfrew-Collingwood
 neighbourhood_group_cleansed                   | null
```

```
 latitude                                      | 49.23427
 longitude                                     | -123.05309
 property_type                                 | Entire house
 room_type                                     | Entire home/apt
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Oven', 'Microwave', 'Wifi', 'Dryer', 'Carbon mo
noxide alarm', 'Stove', 'Hot water', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heatin
g', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Free parking o
n premises', 'Coffee maker', 'Bed linens', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV']
 price                                         | 80
 minimum_nights                                | 28
 maximum_nights                                | 400
 minimum_minimum_nights                        | 28
 maximum_minimum_nights                        | 28
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 28.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 2
 availability_60                               | 13
 availability_90                               | 43
 availability_365                              | 318
 instant_bookable                              | f
 calculated_host_listings_count                | 3
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 2
 calculated_host_listings_count_shared_rooms   | 0
-RECORD 5-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------
 id                                            | 24273
 host_since                                    | 2011/10/09
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | within an hour
 host_response_rate                            | 1
 host_acceptance_rate                          | 0.82
 host_is_superhost                             | f
 host_neighbourhood                            | Kitsilano
```

```
 host_listings_count                          | 2
 host_total_listings_count                    | 2
 host_verifications                           | ['email', 'phone', 'reviews', 'jumio', 'offline_government_i
d', 'government_id']
 host_has_profile_pic                         | f
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Kitsilano
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.2626
 longitude                                    | -123.17308
 property_type                                | Entire apartment
 room_type                                    | Entire home/apt
 accommodates                                 | 6
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 2
 beds                                         | 3
 amenities                                    | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi',
'Dryer', 'Fireplace guards', 'Carbon monoxide alarm', 'Stove', 'Crib', 'Hot water', 'Babysitter recommendati
ons', 'Garden or backyard', 'Extra pillows and blankets', 'Dishwasher', 'Luggage dropoff allowed', 'Cooking
 basics', 'Room-darkening shades', 'Private entrance', 'Smoke alarm', 'Baking sheet', 'Heating', 'Single lev
el home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Coffee maker', 'Be
d linens', 'Keypad', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Shower gel', 'Patio or balcon
y', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Wa
sher', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'TV', 'Ethernet connection']
 price                                        | 200
 minimum_nights                               | 30
 maximum_nights                               | 365
 minimum_minimum_nights                       | 30
 maximum_minimum_nights                       | 30
 minimum_maximum_nights                       | 365
 maximum_maximum_nights                       | 365
 minimum_nights_avg_ntm                       | 30.0
 maximum_nights_avg_ntm                       | 365.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 10
 availability_365                             | 223
 instant_bookable                             | f
 calculated_host_listings_count               | 2
 calculated_host_listings_count_entire_homes  | 2
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
-RECORD 6-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 availability_365                              | 333
 instant_bookable                              | t
 calculated_host_listings_count                | 4
 calculated_host_listings_count_entire_homes   | 4
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
-RECORD 7-----------------------
```

```
 maximum_minimum_nights                        | 90
 minimum_maximum_nights                        | 365
 maximum_maximum_nights                        | 365
 minimum_nights_avg_ntm                        | 90.0
 maximum_nights_avg_ntm                        | 365.0
 has_availability                              | t
 availability_30                               | 30
 availability_60                               | 60
 availability_90                               | 90
 availability_365                              | 365
 instant_bookable                              | f
 calculated_host_listings_count                | 2
 calculated_host_listings_count_entire_homes   | 2
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
-RECORD 8-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                            | 72659
 host_since                                    | 2011/02/03
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | N/A
 host_response_rate                            | N/A
 host_acceptance_rate                          | 0
 host_is_superhost                             | f
 host_neighbourhood                            | Commercial Drive
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | null
 neighbourhood_cleansed                        | Grandview-Woodland
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.27357
 longitude                                     | -123.07024
 property_type                                 | Entire condominium
 room_type                                     | Entire home/apt
 accommodates                                  | 4
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 2
 beds                                          | 2
 amenities                                     | ['Kitchen', 'Lake access', 'Oven', 'Paid parking off premise
```

```
s', 'Gym', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water',
 'Dishwasher', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm', 'He
ating', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Coffee mak
er', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Hangers',
 'Refrigerator', 'Essentials', 'Washer', 'BBQ grill', 'TV', 'Elevator']
 price                                       | 210
 minimum_nights                              | 7
 maximum_nights                              | 21
 minimum_minimum_nights                      | 7
 maximum_minimum_nights                      | 7
 minimum_maximum_nights                      | 21
 maximum_maximum_nights                      | 21
 minimum_nights_avg_ntm                      | 7.0
 maximum_nights_avg_ntm                      | 21.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 0
 availability_90                             | 0
 availability_365                            | 40
 instant_bookable                            | f
 calculated_host_listings_count              | 1
 calculated_host_listings_count_entire_homes | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms | 0
-RECORD 9-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                          | 77157
 host_since                                  | 2011/02/27
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | within a few hours
 host_response_rate                          | 1
 host_acceptance_rate                        | 1
 host_is_superhost                           | f
 host_neighbourhood                          | Arbutus Ridge
 host_listings_count                         | 2
 host_total_listings_count                   | 2
 host_verifications                          | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'offline_
government_id', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                      | Arbutus Ridge
 neighbourhood_group_cleansed                | null
```

```
 latitude                                     | 49.24675
 longitude                                    | -123.16711
 property_type                                | Private room in guest suite
 room_type                                    | Private room
 accommodates                                 | 3
 bathrooms                                    | null
 bathrooms_text                               | 1 private bath
 bedrooms                                     | 1
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Microwave', 'Wifi', 'Laundromat nearby', 'Drye
r', 'Hot water', 'Garden or backyard', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking bas
ics', 'Private entrance', 'Smoke alarm', 'Heating', 'Mini fridge', 'Single level home', 'Hair dryer', 'Lockb
ox', 'Dedicated workspace', 'First aid kit', 'Iron', 'Body soap', 'Free street parking', 'Coffee maker', 'Be
d linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Barbecue utensils', 'Shower gel', 'Hang
ers', 'Cable TV', 'Essentials', 'Washer', 'Conditioner', 'BBQ grill', 'Beach essentials', 'TV', 'Long term s
tays allowed']
 price                                        | 75
 minimum_nights                               | 3
 maximum_nights                               | 60
 minimum_minimum_nights                       | 3
 maximum_minimum_nights                       | 3
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
 minimum_nights_avg_ntm                       | 3.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 3
 availability_60                              | 3
 availability_90                              | 3
 availability_365                             | 267
 instant_bookable                             | t
 calculated_host_listings_count               | 2
 calculated_host_listings_count_entire_homes  | 0
 calculated_host_listings_count_private_rooms | 2
 calculated_host_listings_count_shared_rooms  | 0
-RECORD 10-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------
 id                                           | 111077
 host_since                                   | 2011/05/09
 host_location                                | North Vancouver, British Columbia, Canada
 host_response_time                           | within a few hours
 host_response_rate                           | 1
 host_acceptance_rate                         | 0.33
```

```
 host_is_superhost                              | f
 host_neighbourhood                             | Grandview-Woodland
 host_listings_count                            | 1
 host_total_listings_count                      | 1
 host_verifications                             | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'governme
nt_id']
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                         | Grandview-Woodland
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.26726
 longitude                                      | -123.05969
 property_type                                  | Entire guest suite
 room_type                                      | Entire home/apt
 accommodates                                   | 3
 bathrooms                                      | null
 bathrooms_text                                 | 1 bath
 bedrooms                                       | 2
 beds                                           | 2
 amenities                                      | ['Kitchen', 'Indoor fireplace', 'Microwave', 'Wifi', 'Drye
r', 'Stove', 'Hot water', 'Cooking basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Iro
n', 'Free street parking', 'Coffee maker', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed']
 price                                          | 125
 minimum_nights                                 | 3
 maximum_nights                                 | 365
 minimum_minimum_nights                         | 3
 maximum_minimum_nights                         | 3
 minimum_maximum_nights                         | 365
 maximum_maximum_nights                         | 365
 minimum_nights_avg_ntm                         | 3.0
 maximum_nights_avg_ntm                         | 365.0
 has_availability                               | t
 availability_30                                | 4
 availability_60                                | 34
 availability_90                                | 64
 availability_365                               | 339
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
-RECORD 11-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 calculated_host_listings_count_entire_homes  | 0
 calculated_host_listings_count_private_rooms | 1
 calculated_host_listings_count_shared_rooms  | 0
-RECORD 12-----------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                           | 129163
 host_since                                   | 2011/05/17
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | within a few hours
 host_response_rate                           | 1
 host_acceptance_rate                         | 1
 host_is_superhost                            | f
 host_neighbourhood                           | Renfrew-Collingwood
 host_listings_count                          | 2
 host_total_listings_count                    | 2
 host_verifications                           | ['email', 'phone', 'reviews', 'jumio', 'offline_government_i
d', 'government_id']
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | null
 neighbourhood_cleansed                       | Renfrew-Collingwood
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.24507
 longitude                                    | -123.04941
 property_type                                | Private room in house
 room_type                                    | Private room
 accommodates                                 | 1
 bathrooms                                    | null
 bathrooms_text                               | 1 shared bath
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Hot water', 'Paid parking on premises', 'Lock o
n bedroom door', 'Free street parking', 'Cable TV', 'Smoke alarm', 'Wifi', 'Essentials', 'Heating', 'Washe
r', 'Dryer', 'Carbon monoxide alarm', 'Dishes and silverware', 'TV', 'Long term stays allowed', 'Host greets
you', 'Hair dryer', 'Shampoo', 'Fire extinguisher']
 price                                        | 45
 minimum_nights                               | 30
 maximum_nights                               | 280
 minimum_minimum_nights                       | 30
 maximum_minimum_nights                       | 30
 minimum_maximum_nights                       | 280
 maximum_maximum_nights                       | 280
 minimum_nights_avg_ntm                       | 30.0
```

```
 maximum_nights_avg_ntm                       | 280.0
 has_availability                             | t
 availability_30                              | 30
 availability_60                              | 60
 availability_90                              | 90
 availability_365                             | 365
 instant_bookable                             | t
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 0
 calculated_host_listings_count_private_rooms | 1
 calculated_host_listings_count_shared_rooms  | 0
-RECORD 13-------------------
```

```
 price                                          | 100
 minimum_nights                                 | 30
 maximum_nights                                 | 365
 minimum_minimum_nights                         | 30
 maximum_minimum_nights                         | 30
 minimum_maximum_nights                         | 365
 maximum_maximum_nights                         | 365
 minimum_nights_avg_ntm                         | 30.0
 maximum_nights_avg_ntm                         | 365.0
 has_availability                               | t
 availability_30                                | 0
 availability_60                                | 0
 availability_90                                | 0
 availability_365                               | 224
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
-RECORD 14-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                             | 200643
 host_since                                     | 2011/08/16
 host_location                                  | Vancouver, British Columbia, Canada
 host_response_time                             | within an hour
 host_response_rate                             | 1
 host_acceptance_rate                           | 1
 host_is_superhost                              | f
 host_neighbourhood                             | Strathcona
 host_listings_count                            | 4
 host_total_listings_count                      | 4
 host_verifications                             | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                         | Downtown Eastside
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.27384
 longitude                                      | -123.08289
 property_type                                  | Private room in house
 room_type                                      | Private room
 accommodates                                   | 4
 bathrooms                                      | null
```

```
 bathrooms_text                                | 2 shared baths
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi',
'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Garden or backyard', 'Dishwasher', 'Extra pillows a
nd blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Room-darkening shades', 'Smoke alarm', 'Heatin
g', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Free street parking', 'Coffee maker', 'Be
d linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Essentials',
'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
 price                                         | 87
 minimum_nights                                | 1
 maximum_nights                                | 1125
 minimum_minimum_nights                        | 1
 maximum_minimum_nights                        | 1
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 1.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 30
 availability_60                               | 60
 availability_90                               | 90
 availability_365                              | 365
 instant_bookable                              | t
 calculated_host_listings_count                | 4
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms  | 4
 calculated_host_listings_count_shared_rooms   | 0
-RECORD 15-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                            | 248014
 host_since                                    | 2011/10/16
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | within a day
 host_response_rate                            | 0.5
 host_acceptance_rate                          | N/A
 host_is_superhost                             | f
 host_neighbourhood                            | Grandview-Woodland
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
```

```
 neighbourhood                                  | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                         | Grandview-Woodland
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.29103
 longitude                                      | -123.05782
 property_type                                  | Entire house
 room_type                                      | Entire home/apt
 accommodates                                   | 3
 bathrooms                                      | null
 bathrooms_text                                 | 1 bath
 bedrooms                                       | 2
 beds                                           | 0
 amenities                                      | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi',
'Host greets you', 'Dryer', 'Stove', 'Carbon monoxide alarm', 'Hot water', 'Garden or backyard', 'Dishwashe
r', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iro
n', 'Free street parking', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Cable TV', 'Re
frigerator', 'Hangers', 'Essentials', 'Washer', 'Long term stays allowed', 'BBQ grill', 'TV', 'Cleaning befo
re checkout']
 price                                          | 100
 minimum_nights                                 | 31
 maximum_nights                                 | 180
 minimum_minimum_nights                         | 31
 maximum_minimum_nights                         | 31
 minimum_maximum_nights                         | 180
 maximum_maximum_nights                         | 180
 minimum_nights_avg_ntm                         | 31.0
 maximum_nights_avg_ntm                         | 180.0
 has_availability                               | t
 availability_30                                | 4
 availability_60                                | 4
 availability_90                                | 4
 availability_365                               | 113
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
-RECORD 16-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------
 id                                             | 258636
 host_since                                     | 2011/08/09
 host_location                                  | Vancouver, British Columbia, Canada
 host_response_time                             | N/A
```

```
 host_response_rate                            | N/A
 host_acceptance_rate                          | 1
 host_is_superhost                             | f
 host_neighbourhood                            | Downtown Vancouver
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | null
 neighbourhood_cleansed                        | Downtown
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.27796
 longitude                                     | -123.12671
 property_type                                 | Entire apartment
 room_type                                     | Entire home/apt
 accommodates                                  | 4
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 2
 amenities                                     | ['Kitchen', 'Iron', 'Elevator', 'Gym', 'Hangers', 'Smoke ala
rm', 'Wifi', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Free parking on premises', 'Long term stays allowe
d', 'Hair dryer', 'Shampoo', 'Dedicated workspace']
 price                                         | 99
 minimum_nights                                | 120
 maximum_nights                                | 365
 minimum_minimum_nights                        | 120
 maximum_minimum_nights                        | 120
 minimum_maximum_nights                        | 365
 maximum_maximum_nights                        | 365
 minimum_nights_avg_ntm                        | 120.0
 maximum_nights_avg_ntm                        | 365.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 11
 availability_90                               | 41
 availability_365                              | 316
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
-RECORD 17-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
-RECORD 18-------------------
```

```
 availability_30                              | 30
 availability_60                              | 60
 availability_90                              | 90
 availability_365                             | 141
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
-RECORD 19-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                           | 359942
 host_since                                   | 2011/09/23
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | within an hour
 host_response_rate                           | 1
 host_acceptance_rate                         | N/A
 host_is_superhost                            | f
 host_neighbourhood                           | Yaletown
 host_listings_count                          | 8
 host_total_listings_count                    | 8
 host_verifications                           | ['email', 'phone', 'facebook', 'reviews']
 host_has_profile_pic                         | t
 host_identity_verified                       | f
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Downtown
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27825
 longitude                                    | -123.11865
 property_type                                | Private room in apartment
 room_type                                    | Private room
 accommodates                                 | 2
 bathrooms                                    | null
 bathrooms_text                               | 1 shared bath
 bedrooms                                     | 2
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Lake access', 'Oven', 'Paid parking on premise
s', 'Paid parking off premises', 'Gym', 'Microwave', 'Wifi', 'Dryer', 'Stove', 'Hot water', 'Dishwasher', 'E
xtra pillows and blankets', 'Cooking basics', 'Heating', 'Lock on bedroom door', 'Hair dryer', 'Iron', 'Coff
ee maker', 'Bed linens', 'Pocket wifi', 'Shampoo', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Esse
ntials', 'Washer', 'Long term stays allowed', 'Elevator']
 price                                        | 240
 minimum_nights                               | 30
```

```
 maximum_nights                                | 1125
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 29
 availability_60                               | 59
 availability_90                               | 89
 availability_365                              | 364
 instant_bookable                              | f
 calculated_host_listings_count                | 7
 calculated_host_listings_count_entire_homes   | 3
 calculated_host_listings_count_private_rooms  | 4
 calculated_host_listings_count_shared_rooms   | 0
only showing top 20 rows
```

The string fields will need special processing before we can make sense of them.

Some of them, such as `property_type` seem to have categorical data (specific strings that belong to a small set), we will have to transform them to indices or vector of boolean features (one-hot-encoded).

Others are duplicates of the same information (e.g., "neighbourhood", "neighbourhood_cleansed" and "neighbourhood_group_cleansed"). For such, we will select the one with the highest quality to add to the feature vector.

Finally, `amenities` and `host_verifications` are a set-based features, represented as a strings. That will require quite a bit of processing.

Transformers

Before analysing how these columns behave with respect to our target value (`price`), let us handle the categorical data encoding it to indexes. This is is accomplished in pyspark by using aptly named `Transformer` s named *indexers* (the name is unrelated to deep learning transformers).

In pyspark, we can define different transformers and apply them all together as part of a `Pipeline` . We will be recording the different transformers we will be using in the `stages` list defined below.

In [9]:
```
stages = list()
```

Indexing Categorical Features

For this case study, these three key features will be considered: which neighbourhood the property is in, the type of property and the type of room.

Spark has `StringIndexer` es that are very strict. We will relax them with an "other" category and leave unseen categories and rare categories as other.

In [10]:

```
dataset.groupBy("neighbourhood_cleansed").\
  count().\
  sort("count").\
  show(999, truncate=False)
```

```
+------------------------+-----+
|neighbourhood_cleansed  |count|
+------------------------+-----+
|Strathcona              |16   |
|Kerrisdale              |19   |
|Victoria-Fraserview     |21   |
|Killarney               |27   |
|Arbutus Ridge           |28   |
|Shaughnessy             |31   |
|South Cambie            |31   |
|West Point Grey         |35   |
|Oakridge                |45   |
|Sunset                  |46   |
|Marpole                 |58   |
|Dunbar Southlands       |64   |
|Fairview                |68   |
|Renfrew-Collingwood     |68   |
|Hastings-Sunrise        |84   |
|Grandview-Woodland      |97   |
|Riley Park              |105  |
|Downtown Eastside       |112  |
|Mount Pleasant          |142  |
|Kensington-Cedar Cottage|143  |
|West End                |172  |
|Kitsilano               |191  |
|Downtown                |564  |
+------------------------+-----+
```

In [11]:

```
neighbourhood_reduced = set([ p['neighbourhood_cleansed'] \
                              for p in dataset.groupBy("neighbourhood_cleansed").\
```

```
                                  count().sort("count").toLocalIterator() \
                                if p['count'] > 50])
neighbourhood_reduced
```

Out[11]:

```
{'Downtown',
 'Downtown Eastside',
 'Dunbar Southlands',
 'Fairview',
 'Grandview-Woodland',
 'Hastings-Sunrise',
 'Kensington-Cedar Cottage',
 'Kitsilano',
 'Marpole',
 'Mount Pleasant',
 'Renfrew-Collingwood',
 'Riley Park',
 'West End'}
```

In [12]:

```
class SetReducer(UnaryTransformer):
    def __init__(self, valid_set, other_category):
        super(UnaryTransformer, self).__init__()
        self.valid_set = valid_set
        self.other_category = other_category

    def createTransformFunc(self):
        return lambda x: x if x in self.valid_set else self.other_category

    def outputDataType(self):
        return StringType()

    def validateInputType(self, inputType):
        assert inputType == StringType(), f'Expected StringType() and found {inputType}'


neighbourhood_reducer = SetReducer(neighbourhood_reduced, "OTHER"). \
  setInputCol('neighbourhood_cleansed'). \
  setOutputCol('neighbourhood_reduced')
```

In [13]:

```
neighbourhood_reducer.transform(dataset). \
  groupBy("neighbourhood_reduced"). \
    count(). \
    sort("count"). \
    show(999, truncate=False)
stages.append(neighbourhood_reducer)
```

```
+-------------------------+-----+
|neighbourhood_reduced    |count|
+-------------------------+-----+
|Marpole                  |58   |
|Dunbar Southlands        |64   |
|Renfrew-Collingwood      |68   |
|Fairview                 |68   |
|Hastings-Sunrise         |84   |
|Grandview-Woodland       |97   |
|Riley Park               |105  |
|Downtown Eastside        |112  |
|Mount Pleasant           |142  |
|Kensington-Cedar Cottage|143  |
|West End                 |172  |
|Kitsilano                |191  |
|OTHER                    |299  |
|Downtown                 |564  |
+-------------------------+-----+
```

In [14]:

```
dataset.groupBy("property_type").\
  count().\
  sort("count").\
  show(999, truncate=False)
```

```
+---------------------------------+-----+
|property_type                    |count|
+---------------------------------+-----+
|Camper/RV                        |1    |
|Entire timeshare                 |1    |
|Entire cabin                     |1    |
|Shared room in loft              |1    |
|Private room in boat             |1    |
|Boat                             |1    |
|Private room in hostel           |1    |
|Private room in loft             |1    |
|Entire floor                     |1    |
|Entire cottage                   |1    |
|Entire place                     |2    |
|Room in bed and breakfast        |2    |
|Private room in bungalow         |3    |
|Room in hotel                    |3    |
|Shared room in house             |3    |
|Tiny house                       |3    |
|Shared room in apartment         |3    |
|Entire villa                     |5    |
|Room in aparthotel               |5    |
```

```
|Private room in guesthouse       |6    |
|Private room in bed and breakfast|8    |
|Entire bungalow                  |8    |
|Room in boutique hotel           |10   |
|Private room in guest suite      |12   |
|Private room in villa            |12   |
|Private room in townhouse        |17   |
|Entire guesthouse                |27   |
|Entire serviced apartment        |39   |
|Private room in condominium      |40   |
|Entire townhouse                 |46   |
|Entire loft                      |57   |
|Private room in apartment        |94   |
|Entire guest suite               |266  |
|Private room in house            |278  |
|Entire house                     |317  |
|Entire condominium               |389  |
|Entire apartment                 |502  |
+---------------------------------+-----+
```

In [15]:

```
property_reduced = set([ p['property_type'] \
                            for p in dataset.groupBy("property_type").\
                              count().sort("count").toLocalIterator() \
                            if p['count'] > 50])
property_reduced
```

Out[15]:

```
{'Entire apartment',
 'Entire condominium',
 'Entire guest suite',
 'Entire house',
 'Entire loft',
 'Private room in apartment',
 'Private room in house'}
```

In [16]:

```
property_reducer = SetReducer(property_reduced, "OTHER"). \
  setInputCol('property_type'). \
  setOutputCol('property_reduced')

property_reducer.transform(dataset). \
  groupBy("property_reduced"). \
    count(). \
    sort("count"). \
    show(999, truncate=False)
stages.append(property_reducer)
```

```
+-------------------------+-----+
|property_reduced         |count|
+-------------------------+-----+
|Entire loft              |57   |
|Private room in apartment|94   |
|OTHER                    |264  |
|Entire guest suite       |266  |
|Private room in house    |278  |
|Entire house             |317  |
|Entire condominium       |389  |
|Entire apartment         |502  |
+-------------------------+-----+
```

In [17]:

```
dataset.groupBy("room_type").\
  count().\
  sort("count").\
  show(999, truncate=False)
```

```
+---------------+-----+
|room_type      |count|
+---------------+-----+
|Hotel room     |2    |
|Shared room    |7    |
|Private room   |486  |
|Entire home/apt|1672 |
+---------------+-----+
```

In [18]:

```
room_reduced = set([ p['room_type'] \
                            for p in dataset.groupBy("room_type").\
                              count().sort("count").toLocalIterator() \
                            if p['count'] > 50])
room_reduced
```

Out[18]:

```
{'Entire home/apt', 'Private room'}
```

In [19]:

```
room_reducer = SetReducer(room_reduced, "OTHER"). \
  setInputCol('room_type'). \
  setOutputCol('room_reduced')

room_reducer.transform(dataset). \
  groupBy("room_reduced"). \
    count(). \
```

```
    sort("count"). \
    show(999, truncate=False)
stages.append(room_reducer)
```

```
+---------------+-----+
|room_reduced   |count|
+---------------+-----+
|OTHER          |9    |
|Private room   |486  |
|Entire home/apt|1672 |
+---------------+-----+
```

Once the strings have been reduced with the "OTHER" category, we can feed them into the Spark's indexer. The indexers will throw an error if they find an unseen string (which will not happen as anything unseen will be marked as "OTHER").

In [20]:

```
n_indexer = StringIndexer(inputCol="neighbourhood_reduced", outputCol="neighbourhood_indexed")
fitted_n_indexer = n_indexer.fit(neighbourhood_reducer.transform(dataset))
stages.append(fitted_n_indexer)

p_indexer = StringIndexer(inputCol="property_reduced", outputCol="property_indexed")
fitted_p_indexer = p_indexer.fit(property_reducer.transform(dataset))
stages.append(fitted_p_indexer)

r_indexer = StringIndexer(inputCol="room_reduced", outputCol="room_indexed")
fitted_r_indexer = r_indexer.fit(room_reducer.transform(dataset))
stages.append(fitted_r_indexer)
```

In [21]:

```
print(len(fitted_n_indexer.labels))
fitted_n_indexer.labels
```

```
14
```

Out[21]:

```
['Downtown',
 'OTHER',
 'Kitsilano',
 'West End',
 'Kensington-Cedar Cottage',
 'Mount Pleasant',
 'Downtown Eastside',
 'Riley Park',
 'Grandview-Woodland',
 'Hastings-Sunrise',
 'Fairview',
 'Renfrew-Collingwood',
```

```
 'Dunbar Southlands',
 'Marpole']
```

In [22]:

```
print(len(fitted_p_indexer.labels))
fitted_p_indexer.labels
```

```
8
```

Out[22]:

```
['Entire apartment',
 'Entire condominium',
 'Entire house',
 'Private room in house',
 'Entire guest suite',
 'OTHER',
 'Private room in apartment',
 'Entire loft']
```

In [23]:

```
print(len(fitted_r_indexer.labels))
fitted_r_indexer.labels
```

```
3
```

Out[23]:

```
['Entire home/apt', 'Private room', 'OTHER']
```

We can instantiate a pipeline to see the transformed output.

In [24]:

```
pipeline = Pipeline(stages=stages)
pipeline.fit(dataset).transform(dataset).show(vertical=True)
```

```
-RECORD 0-----------------------------------------------------------
 id                                          | 13358
 host_since                                  | 2009/11/07
 host_location                               | Vancouver, Britis...
 host_response_time                          | within a day
 host_response_rate                          | 0.71
 host_acceptance_rate                        | 0.71
 host_is_superhost                           | t
 host_neighbourhood                          | Downtown Vancouver
 host_listings_count                         | 1
 host_total_listings_count                   | 1
 host_verifications                          | ['email', 'phone'...
 host_has_profile_pic                        | f
 host_identity_verified                      | t
 neighbourhood                               | null
 neighbourhood_cleansed                      | West End
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.28209
```

```
 longitude                                     | -123.12515
 property_type                                 | Entire condominium
 room_type                                     | Entire home/apt
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Oven...
 price                                         | 90
 minimum_nights                                | 30
 maximum_nights                                | 60
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 60
 maximum_maximum_nights                        | 60
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 60.0
 has_availability                              | t
 availability_30                               | 11
 availability_60                               | 41
 availability_90                               | 70
 availability_365                              | 70
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | West End
 property_reduced                              | Entire condominium
 room_reduced                                  | Entire home/apt
 neighbourhood_indexed                         | 3.0
 property_indexed                              | 1.0
 room_indexed                                  | 0.0
-RECORD 1-------------------------------------------------------------
 id                                            | 16254
 host_since                                    | 2009/12/15
 host_location                                 | Vancouver, Britis...
 host_response_time                            | N/A
 host_response_rate                            | N/A
 host_acceptance_rate                          | N/A
 host_is_superhost                             | f
 host_neighbourhood                            | Hastings-Sunrise
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['phone', 'facebo...
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, Britis...
```

```
 neighbourhood_cleansed                       | Hastings-Sunrise
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27734
 longitude                                    | -123.04152
 property_type                                | Entire guest suite
 room_type                                    | Entire home/apt
 accommodates                                 | 4
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 2
 beds                                         | 3
 amenities                                    | ['Kitchen', 'Oven...
 price                                        | 195
 minimum_nights                               | 3
 maximum_nights                               | 31
 minimum_minimum_nights                       | 3
 maximum_minimum_nights                       | 3
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
 minimum_nights_avg_ntm                       | 3.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 0
 availability_365                             | 228
 instant_bookable                             | t
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Hastings-Sunrise
 property_reduced                             | Entire guest suite
 room_reduced                                 | Entire home/apt
 neighbourhood_indexed                        | 9.0
 property_indexed                             | 4.0
 room_indexed                                 | 0.0
-RECORD 2-----------------------------------------------------------
 id                                           | 17765
 host_since                                   | 2010/01/07
 host_location                                | Vancouver, Britis...
 host_response_time                           | N/A
 host_response_rate                           | N/A
 host_acceptance_rate                         | 1
 host_is_superhost                            | f
 host_neighbourhood                           | Mount Pleasant
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['phone', 'facebo...
```

```
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, Britis...
 neighbourhood_cleansed                        | Mount Pleasant
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.26181
 longitude                                     | -123.10905
 property_type                                 | Entire apartment
 room_type                                     | Entire home/apt
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Wifi...
 price                                         | 94
 minimum_nights                                | 3
 maximum_nights                                | 60
 minimum_minimum_nights                        | 3
 maximum_minimum_nights                        | 3
 minimum_maximum_nights                        | 60
 maximum_maximum_nights                        | 60
 minimum_nights_avg_ntm                        | 3.0
 maximum_nights_avg_ntm                        | 60.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 0
 availability_365                              | 0
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Mount Pleasant
 property_reduced                              | Entire apartment
 room_reduced                                  | Entire home/apt
 neighbourhood_indexed                         | 5.0
 property_indexed                              | 0.0
 room_indexed                                  | 0.0
-RECORD 3-------------------------------------------------------------
 id                                            | 18795
 host_since                                    | 2010/01/13
 host_location                                 | Vancouver, Britis...
 host_response_time                            | within a day
 host_response_rate                            | 0.7
 host_acceptance_rate                          | 0.4
 host_is_superhost                             | f
 host_neighbourhood                            | West End
```

```
 host_listings_count                           | 3
 host_total_listings_count                     | 3
 host_verifications                            | ['email', 'phone'...
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | null
 neighbourhood_cleansed                        | West End
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.28251
 longitude                                     | -123.12827
 property_type                                 | Private room in h...
 room_type                                     | Private room
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 private bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Cabl...
 price                                         | 50
 minimum_nights                                | 30
 maximum_nights                                | 360
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 360
 maximum_maximum_nights                        | 360
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 360.0
 has_availability                              | t
 availability_30                               | 9
 availability_60                               | 20
 availability_90                               | 50
 availability_365                              | 325
 instant_bookable                              | f
 calculated_host_listings_count                | 3
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms  | 3
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | West End
 property_reduced                              | Private room in h...
 room_reduced                                  | Private room
 neighbourhood_indexed                         | 3.0
 property_indexed                              | 3.0
 room_indexed                                  | 1.0
-RECORD 4-------------------------------------------------------------
 id                                            | 21913
 host_since                                    | 2010/02/22
 host_location                                 | Vancouver, Britis...
 host_response_time                            | within an hour
 host_response_rate                            | 1
```

```
host_acceptance_rate                          | 1
host_is_superhost                             | t
host_neighbourhood                            | Renfrew-Collingwood
host_listings_count                           | 3
host_total_listings_count                     | 3
host_verifications                            | ['email', 'phone'...
host_has_profile_pic                          | t
host_identity_verified                        | t
neighbourhood                                 | Vancouver, Britis...
neighbourhood_cleansed                        | Renfrew-Collingwood
neighbourhood_group_cleansed                  | null
latitude                                      | 49.23427
longitude                                     | -123.05309
property_type                                 | Entire house
room_type                                     | Entire home/apt
accommodates                                  | 2
bathrooms                                     | null
bathrooms_text                                | 1 bath
bedrooms                                      | 1
beds                                          | 1
amenities                                     | ['Kitchen', 'Oven...
price                                         | 80
minimum_nights                                | 28
maximum_nights                                | 400
minimum_minimum_nights                        | 28
maximum_minimum_nights                        | 28
minimum_maximum_nights                        | 1125
maximum_maximum_nights                        | 1125
minimum_nights_avg_ntm                        | 28.0
maximum_nights_avg_ntm                        | 1125.0
has_availability                              | t
availability_30                               | 2
availability_60                               | 13
availability_90                               | 43
availability_365                              | 318
instant_bookable                              | f
calculated_host_listings_count                | 3
calculated_host_listings_count_entire_homes   | 1
calculated_host_listings_count_private_rooms  | 2
calculated_host_listings_count_shared_rooms   | 0
neighbourhood_reduced                         | Renfrew-Collingwood
property_reduced                              | Entire house
room_reduced                                  | Entire home/apt
neighbourhood_indexed                         | 11.0
property_indexed                              | 2.0
room_indexed                                  | 0.0
-RECORD 5-------------------------------------------------------------
id                                            | 24273
host_since                                    | 2011/10/09
```

```
host_location                                   | Vancouver, Britis...
host_response_time                              | within an hour
host_response_rate                              | 1
host_acceptance_rate                            | 0.82
host_is_superhost                               | f
host_neighbourhood                              | Kitsilano
host_listings_count                             | 2
host_total_listings_count                       | 2
host_verifications                              | ['email', 'phone'...
host_has_profile_pic                            | f
host_identity_verified                          | t
neighbourhood                                   | Vancouver, Britis...
neighbourhood_cleansed                          | Kitsilano
neighbourhood_group_cleansed                    | null
latitude                                        | 49.2626
longitude                                       | -123.17308
property_type                                   | Entire apartment
room_type                                       | Entire home/apt
accommodates                                    | 6
bathrooms                                       | null
bathrooms_text                                  | 1 bath
bedrooms                                        | 2
beds                                            | 3
amenities                                       | ['Kitchen', 'Oven...
price                                           | 200
minimum_nights                                  | 30
maximum_nights                                  | 365
minimum_minimum_nights                          | 30
maximum_minimum_nights                          | 30
minimum_maximum_nights                          | 365
maximum_maximum_nights                          | 365
minimum_nights_avg_ntm                          | 30.0
maximum_nights_avg_ntm                          | 365.0
has_availability                                | t
availability_30                                 | 0
availability_60                                 | 0
availability_90                                 | 10
availability_365                                | 223
instant_bookable                                | f
calculated_host_listings_count                  | 2
calculated_host_listings_count_entire_homes     | 2
calculated_host_listings_count_private_rooms    | 0
calculated_host_listings_count_shared_rooms     | 0
neighbourhood_reduced                           | Kitsilano
property_reduced                                | Entire apartment
room_reduced                                    | Entire home/apt
neighbourhood_indexed                           | 2.0
property_indexed                                | 0.0
room_indexed                                    | 0.0
```

```
-RECORD 6-----------------------------------------------------------
 id                                          | 42645
 host_since                                  | 2010/04/17
 host_location                               | Vancouver, Britis...
 host_response_time                          | within an hour
 host_response_rate                          | 1
 host_acceptance_rate                        | 1
 host_is_superhost                           | t
 host_neighbourhood                          | Downtown Vancouver
 host_listings_count                         | 4
 host_total_listings_count                   | 4
 host_verifications                          | ['email', 'phone'...
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, Britis...
 neighbourhood_cleansed                      | Downtown
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.28001
 longitude                                   | -123.12475
 property_type                               | Entire condominium
 room_type                                   | Entire home/apt
 accommodates                                | 3
 bathrooms                                   | null
 bathrooms_text                              | 1 bath
 bedrooms                                    | 1
 beds                                        | 2
 amenities                                   | ['Kitchen', 'Oven...
 price                                       | 159
 minimum_nights                              | 32
 maximum_nights                              | 365
 minimum_minimum_nights                      | 32
 maximum_minimum_nights                      | 32
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 32.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 29
 availability_90                             | 59
 availability_365                            | 333
 instant_bookable                            | t
 calculated_host_listings_count              | 4
 calculated_host_listings_count_entire_homes | 4
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms | 0
 neighbourhood_reduced                       | Downtown
 property_reduced                            | Entire condominium
 room_reduced                                | Entire home/apt
```

```
 neighbourhood_indexed                        | 0.0
 property_indexed                             | 1.0
 room_indexed                                 | 0.0
-RECORD 7-------------------------------------------------------------
 id                                           | 54261
 host_since                                   | 2010/10/05
 host_location                                | Vancouver, Britis...
 host_response_time                           | a few days or more
 host_response_rate                           | 0
 host_acceptance_rate                         | N/A
 host_is_superhost                            | f
 host_neighbourhood                           | Downtown Vancouver
 host_listings_count                          | 2
 host_total_listings_count                    | 2
 host_verifications                           | ['email', 'phone'...
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | null
 neighbourhood_cleansed                       | Downtown
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27655
 longitude                                    | -123.12485
 property_type                                | Entire condominium
 room_type                                    | Entire home/apt
 accommodates                                 | 4
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Hot ...
 price                                        | 126
 minimum_nights                               | 90
 maximum_nights                               | 365
 minimum_minimum_nights                       | 90
 maximum_minimum_nights                       | 90
 minimum_maximum_nights                       | 365
 maximum_maximum_nights                       | 365
 minimum_nights_avg_ntm                       | 90.0
 maximum_nights_avg_ntm                       | 365.0
 has_availability                             | t
 availability_30                              | 30
 availability_60                              | 60
 availability_90                              | 90
 availability_365                             | 365
 instant_bookable                             | f
 calculated_host_listings_count               | 2
 calculated_host_listings_count_entire_homes  | 2
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
```

```
 neighbourhood_reduced                        | Downtown
 property_reduced                             | Entire condominium
 room_reduced                                 | Entire home/apt
 neighbourhood_indexed                        | 0.0
 property_indexed                             | 1.0
 room_indexed                                 | 0.0
-RECORD 8-------------------------------------------------------------
 id                                           | 72659
 host_since                                   | 2011/02/03
 host_location                                | Vancouver, Britis...
 host_response_time                           | N/A
 host_response_rate                           | N/A
 host_acceptance_rate                         | 0
 host_is_superhost                            | f
 host_neighbourhood                           | Commercial Drive
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['email', 'phone'...
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | null
 neighbourhood_cleansed                       | Grandview-Woodland
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27357
 longitude                                    | -123.07024
 property_type                                | Entire condominium
 room_type                                    | Entire home/apt
 accommodates                                 | 4
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 2
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Lake...
 price                                        | 210
 minimum_nights                               | 7
 maximum_nights                               | 21
 minimum_minimum_nights                       | 7
 maximum_minimum_nights                       | 7
 minimum_maximum_nights                       | 21
 maximum_maximum_nights                       | 21
 minimum_nights_avg_ntm                       | 7.0
 maximum_nights_avg_ntm                       | 21.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 0
 availability_365                             | 40
 instant_bookable                             | f
 calculated_host_listings_count               | 1
```

```
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Grandview-Woodland
 property_reduced                             | Entire condominium
 room_reduced                                 | Entire home/apt
 neighbourhood_indexed                        | 8.0
 property_indexed                             | 1.0
 room_indexed                                 | 0.0
-RECORD 9-----------------------------------------------------------
 id                                           | 77157
 host_since                                   | 2011/02/27
 host_location                                | Vancouver, Britis...
 host_response_time                           | within a few hours
 host_response_rate                           | 1
 host_acceptance_rate                         | 1
 host_is_superhost                            | f
 host_neighbourhood                           | Arbutus Ridge
 host_listings_count                          | 2
 host_total_listings_count                    | 2
 host_verifications                           | ['email', 'phone'...
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, Britis...
 neighbourhood_cleansed                       | Arbutus Ridge
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.24675
 longitude                                    | -123.16711
 property_type                                | Private room in g...
 room_type                                    | Private room
 accommodates                                 | 3
 bathrooms                                    | null
 bathrooms_text                               | 1 private bath
 bedrooms                                     | 1
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Micr...
 price                                        | 75
 minimum_nights                               | 3
 maximum_nights                               | 60
 minimum_minimum_nights                       | 3
 maximum_minimum_nights                       | 3
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
 minimum_nights_avg_ntm                       | 3.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 3
 availability_60                              | 3
 availability_90                              | 3
```

```
 availability_365                              | 267
 instant_bookable                              | t
 calculated_host_listings_count                | 2
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms  | 2
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | OTHER
 property_reduced                              | OTHER
 room_reduced                                  | Private room
 neighbourhood_indexed                         | 1.0
 property_indexed                              | 5.0
 room_indexed                                  | 1.0
-RECORD 10-------------------------------------------------------------
 id                                            | 111077
 host_since                                    | 2011/05/09
 host_location                                 | North Vancouver, ...
 host_response_time                            | within a few hours
 host_response_rate                            | 1
 host_acceptance_rate                          | 0.33
 host_is_superhost                             | f
 host_neighbourhood                            | Grandview-Woodland
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone'...
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, Britis...
 neighbourhood_cleansed                        | Grandview-Woodland
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.26726
 longitude                                     | -123.05969
 property_type                                 | Entire guest suite
 room_type                                     | Entire home/apt
 accommodates                                  | 3
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 2
 beds                                          | 2
 amenities                                     | ['Kitchen', 'Indo...
 price                                         | 125
 minimum_nights                                | 3
 maximum_nights                                | 365
 minimum_minimum_nights                        | 3
 maximum_minimum_nights                        | 3
 minimum_maximum_nights                        | 365
 maximum_maximum_nights                        | 365
 minimum_nights_avg_ntm                        | 3.0
 maximum_nights_avg_ntm                        | 365.0
 has_availability                              | t
```

```
 availability_30                              | 4
 availability_60                              | 34
 availability_90                              | 64
 availability_365                             | 339
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Grandview-Woodland
 property_reduced                             | Entire guest suite
 room_reduced                                 | Entire home/apt
 neighbourhood_indexed                        | 8.0
 property_indexed                             | 4.0
 room_indexed                                 | 0.0
-RECORD 11-------------------------------------------------------------
 id                                           | 125176
 host_since                                   | 2011/05/24
 host_location                                | Vancouver, Britis...
 host_response_time                           | within an hour
 host_response_rate                           | 1
 host_acceptance_rate                         | 1
 host_is_superhost                            | f
 host_neighbourhood                           | Riley Park
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['email', 'phone'...
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, Britis...
 neighbourhood_cleansed                       | Riley Park
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.24648
 longitude                                    | -123.09241
 property_type                                | Private room in h...
 room_type                                    | Private room
 accommodates                                 | 3
 bathrooms                                    | null
 bathrooms_text                               | 1 private bath
 bedrooms                                     | 1
 beds                                         | 0
 amenities                                    | ['Kitchen', 'Hot ...
 price                                        | 75
 minimum_nights                               | 2
 maximum_nights                               | 60
 minimum_minimum_nights                       | 2
 maximum_minimum_nights                       | 2
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
```

```
 minimum_nights_avg_ntm                       | 2.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 2
 availability_365                             | 268
 instant_bookable                             | t
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 0
 calculated_host_listings_count_private_rooms | 1
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Riley Park
 property_reduced                             | Private room in h...
 room_reduced                                 | Private room
 neighbourhood_indexed                        | 7.0
 property_indexed                             | 3.0
 room_indexed                                 | 1.0
-RECORD 12-------------------------------------------------------------
 id                                           | 129163
 host_since                                   | 2011/05/17
 host_location                                | Vancouver, Britis...
 host_response_time                           | within a few hours
 host_response_rate                           | 1
 host_acceptance_rate                         | 1
 host_is_superhost                            | f
 host_neighbourhood                           | Renfrew-Collingwood
 host_listings_count                          | 2
 host_total_listings_count                    | 2
 host_verifications                           | ['email', 'phone'...
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | null
 neighbourhood_cleansed                       | Renfrew-Collingwood
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.24507
 longitude                                    | -123.04941
 property_type                                | Private room in h...
 room_type                                    | Private room
 accommodates                                 | 1
 bathrooms                                    | null
 bathrooms_text                               | 1 shared bath
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Hot ...
 price                                        | 45
 minimum_nights                               | 30
 maximum_nights                               | 280
 minimum_minimum_nights                       | 30
```

```
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 280
 maximum_maximum_nights                        | 280
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 280.0
 has_availability                              | t
 availability_30                               | 30
 availability_60                               | 60
 availability_90                               | 90
 availability_365                              | 365
 instant_bookable                              | t
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms  | 1
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Renfrew-Collingwood
 property_reduced                              | Private room in h...
 room_reduced                                  | Private room
 neighbourhood_indexed                         | 11.0
 property_indexed                              | 3.0
 room_indexed                                  | 1.0
-RECORD 13-----------------------------------------------------------
 id                                            | 174135
 host_since                                    | 2011/07/17
 host_location                                 | Vancouver, Britis...
 host_response_time                            | within a few hours
 host_response_rate                            | 1
 host_acceptance_rate                          | 0
 host_is_superhost                             | f
 host_neighbourhood                            | Downtown Eastside
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone'...
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, Britis...
 neighbourhood_cleansed                        | Downtown Eastside
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.27782
 longitude                                     | -123.08985
 property_type                                 | Entire guest suite
 room_type                                     | Entire home/apt
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 2
 amenities                                     | ['Kitchen', 'Wifi...
 price                                         | 100
```

```
 minimum_nights                                | 30
 maximum_nights                                | 365
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 365
 maximum_maximum_nights                        | 365
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 365.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 0
 availability_365                              | 224
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Downtown Eastside
 property_reduced                              | Entire guest suite
 room_reduced                                  | Entire home/apt
 neighbourhood_indexed                         | 6.0
 property_indexed                              | 4.0
 room_indexed                                  | 0.0
-RECORD 14-------------------------------------------------------------
 id                                            | 200643
 host_since                                    | 2011/08/16
 host_location                                 | Vancouver, Britis...
 host_response_time                            | within an hour
 host_response_rate                            | 1
 host_acceptance_rate                          | 1
 host_is_superhost                             | f
 host_neighbourhood                            | Strathcona
 host_listings_count                           | 4
 host_total_listings_count                     | 4
 host_verifications                            | ['email', 'phone'...
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, Britis...
 neighbourhood_cleansed                        | Downtown Eastside
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.27384
 longitude                                     | -123.08289
 property_type                                 | Private room in h...
 room_type                                     | Private room
 accommodates                                  | 4
 bathrooms                                     | null
 bathrooms_text                                | 2 shared baths
 bedrooms                                      | 1
```

```
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Oven...
 price                                         | 87
 minimum_nights                                | 1
 maximum_nights                                | 1125
 minimum_minimum_nights                        | 1
 maximum_minimum_nights                        | 1
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 1.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 30
 availability_60                               | 60
 availability_90                               | 90
 availability_365                              | 365
 instant_bookable                              | t
 calculated_host_listings_count                | 4
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms  | 4
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Downtown Eastside
 property_reduced                              | Private room in h...
 room_reduced                                  | Private room
 neighbourhood_indexed                         | 6.0
 property_indexed                              | 3.0
 room_indexed                                  | 1.0
-RECORD 15-------------------------------------------------------------
 id                                            | 248014
 host_since                                    | 2011/10/16
 host_location                                 | Vancouver, Britis...
 host_response_time                            | within a day
 host_response_rate                            | 0.5
 host_acceptance_rate                          | N/A
 host_is_superhost                             | f
 host_neighbourhood                            | Grandview-Woodland
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone'...
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, Britis...
 neighbourhood_cleansed                        | Grandview-Woodland
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.29103
 longitude                                     | -123.05782
 property_type                                 | Entire house
 room_type                                     | Entire home/apt
 accommodates                                  | 3
```

```
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 2
 beds                                          | 0
 amenities                                     | ['Kitchen', 'Oven...
 price                                         | 100
 minimum_nights                                | 31
 maximum_nights                                | 180
 minimum_minimum_nights                        | 31
 maximum_minimum_nights                        | 31
 minimum_maximum_nights                        | 180
 maximum_maximum_nights                        | 180
 minimum_nights_avg_ntm                        | 31.0
 maximum_nights_avg_ntm                        | 180.0
 has_availability                              | t
 availability_30                               | 4
 availability_60                               | 4
 availability_90                               | 4
 availability_365                              | 113
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Grandview-Woodland
 property_reduced                              | Entire house
 room_reduced                                  | Entire home/apt
 neighbourhood_indexed                         | 8.0
 property_indexed                              | 2.0
 room_indexed                                  | 0.0
-RECORD 16-------------------------------------------------------------
 id                                            | 258636
 host_since                                    | 2011/08/09
 host_location                                 | Vancouver, Britis...
 host_response_time                            | N/A
 host_response_rate                            | N/A
 host_acceptance_rate                          | 1
 host_is_superhost                             | f
 host_neighbourhood                            | Downtown Vancouver
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone'...
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | null
 neighbourhood_cleansed                        | Downtown
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.27796
 longitude                                     | -123.12671
```

```
 property_type                                  | Entire apartment
 room_type                                      | Entire home/apt
 accommodates                                   | 4
 bathrooms                                      | null
 bathrooms_text                                 | 1 bath
 bedrooms                                       | 1
 beds                                           | 2
 amenities                                      | ['Kitchen', 'Iron...
 price                                          | 99
 minimum_nights                                 | 120
 maximum_nights                                 | 365
 minimum_minimum_nights                         | 120
 maximum_minimum_nights                         | 120
 minimum_maximum_nights                         | 365
 maximum_maximum_nights                         | 365
 minimum_nights_avg_ntm                         | 120.0
 maximum_nights_avg_ntm                         | 365.0
 has_availability                               | t
 availability_30                                | 0
 availability_60                                | 11
 availability_90                                | 41
 availability_365                               | 316
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
 neighbourhood_reduced                          | Downtown
 property_reduced                               | Entire apartment
 room_reduced                                   | Entire home/apt
 neighbourhood_indexed                          | 0.0
 property_indexed                               | 0.0
 room_indexed                                   | 0.0
-RECORD 17-------------------------------------------------------------
 id                                             | 283841
 host_since                                     | 2009/08/10
 host_location                                  | Vancouver, Britis...
 host_response_time                             | within a day
 host_response_rate                             | 0.93
 host_acceptance_rate                           | 0.88
 host_is_superhost                              | f
 host_neighbourhood                             | Coal Harbour
 host_listings_count                            | 49
 host_total_listings_count                      | 49
 host_verifications                             | ['email', 'phone'...
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | null
 neighbourhood_cleansed                         | Downtown
```

```
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.2872
 longitude                                    | -123.12352
 property_type                                | Entire apartment
 room_type                                    | Entire home/apt
 accommodates                                 | 6
 bathrooms                                    | null
 bathrooms_text                               | 3 baths
 bedrooms                                     | 3
 beds                                         | 3
 amenities                                    | ['Kitchen', 'Hot ...
 price                                        | 197
 minimum_nights                               | 90
 maximum_nights                               | 1125
 minimum_minimum_nights                       | 90
 maximum_minimum_nights                       | 90
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
 minimum_nights_avg_ntm                       | 90.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 0
 availability_365                             | 0
 instant_bookable                             | f
 calculated_host_listings_count               | 41
 calculated_host_listings_count_entire_homes  | 41
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Downtown
 property_reduced                             | Entire apartment
 room_reduced                                 | Entire home/apt
 neighbourhood_indexed                        | 0.0
 property_indexed                             | 0.0
 room_indexed                                 | 0.0
-RECORD 18-----------------------------------------------------------
 id                                           | 323090
 host_since                                   | 2012/01/01
 host_location                                | Vancouver, Britis...
 host_response_time                           | N/A
 host_response_rate                           | N/A
 host_acceptance_rate                         | 0.85
 host_is_superhost                            | t
 host_neighbourhood                           | Downtown Vancouver
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['email', 'phone'...
 host_has_profile_pic                         | t
```

```
host_identity_verified                       | t
neighbourhood                                | Vancouver, Britis...
neighbourhood_cleansed                       | Downtown
neighbourhood_group_cleansed                 | null
latitude                                     | 49.28206
longitude                                    | -123.12461
property_type                                | Entire apartment
room_type                                    | Entire home/apt
accommodates                                 | 2
bathrooms                                    | null
bathrooms_text                               | 1 bath
bedrooms                                     | 1
beds                                         | 1
amenities                                    | ['Kitchen', 'Cook...
price                                        | 168
minimum_nights                               | 180
maximum_nights                               | 180
minimum_minimum_nights                       | 180
maximum_minimum_nights                       | 180
minimum_maximum_nights                       | 180
maximum_maximum_nights                       | 180
minimum_nights_avg_ntm                       | 180.0
maximum_nights_avg_ntm                       | 180.0
has_availability                             | t
availability_30                              | 30
availability_60                              | 60
availability_90                              | 90
availability_365                             | 141
instant_bookable                             | f
calculated_host_listings_count               | 1
calculated_host_listings_count_entire_homes  | 1
calculated_host_listings_count_private_rooms | 0
calculated_host_listings_count_shared_rooms  | 0
neighbourhood_reduced                        | Downtown
property_reduced                             | Entire apartment
room_reduced                                 | Entire home/apt
neighbourhood_indexed                        | 0.0
property_indexed                             | 0.0
room_indexed                                 | 0.0
-RECORD 19-------------------------------------------------------------
id                                           | 359942
host_since                                   | 2011/09/23
host_location                                | Vancouver, Britis...
host_response_time                           | within an hour
host_response_rate                           | 1
host_acceptance_rate                         | N/A
host_is_superhost                            | f
host_neighbourhood                           | Yaletown
host_listings_count                          | 8
```

```
 host_total_listings_count                    | 8
 host_verifications                           | ['email', 'phone'...
 host_has_profile_pic                         | t
 host_identity_verified                       | f
 neighbourhood                                | Vancouver, Britis...
 neighbourhood_cleansed                       | Downtown
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27825
 longitude                                    | -123.11865
 property_type                                | Private room in a...
 room_type                                    | Private room
 accommodates                                 | 2
 bathrooms                                    | null
 bathrooms_text                               | 1 shared bath
 bedrooms                                     | 2
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Lake...
 price                                        | 240
 minimum_nights                               | 30
 maximum_nights                               | 1125
 minimum_minimum_nights                       | 30
 maximum_minimum_nights                       | 30
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
 minimum_nights_avg_ntm                       | 30.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 29
 availability_60                              | 59
 availability_90                              | 89
 availability_365                             | 364
 instant_bookable                             | f
 calculated_host_listings_count               | 7
 calculated_host_listings_count_entire_homes  | 3
 calculated_host_listings_count_private_rooms | 4
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Downtown
 property_reduced                             | Private room in a...
 room_reduced                                 | Private room
 neighbourhood_indexed                        | 0.0
 property_indexed                             | 6.0
 room_indexed                                 | 1.0
only showing top 20 rows
```

Set features: Amenities and Host Verifications

Once we have dealt with the categorical features, we can address the set-based columns. Let us take a look at them.

In [25]:

```
dataset.select("amenities").show(truncate=False)
```

```
+-----------------------
```

alarm', 'Stove', 'Crib', 'Hot water', 'Babysitter recommendations', 'Garden or backyard', 'Extra pillows and blankets', 'Dishwasher', 'Luggage dropoff allowed', 'Cooking basics', 'Room-darkening shades', 'Private entrance', 'Smoke alarm', 'Baking sheet', 'Heating', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Coffee maker', 'Bed linens', 'Keypad', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Shower gel', 'Patio or balcony', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'TV', 'Ethernet connection']|
|['Kitchen', 'Oven', 'Paid parking on premises', 'Paid parking off premises', 'Gym', 'Microwave', 'Wifi', 'Host greets you', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Dishwasher', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Coffee maker', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Patio or balcony', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'TV', 'Elevator']
|
|['Kitchen', 'Hot tub', 'Gym', 'Indoor fireplace', 'Wifi', 'Host greets you', 'Dryer', 'Carbon monoxide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Shampoo', 'Fire extinguisher', 'Hangers', 'Cable TV', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning', 'Elevator']
|
|['Kitchen', 'Lake access', 'Oven', 'Paid parking off premises', 'Gym', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Dishwasher', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm', 'Heating', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'BBQ grill', 'TV', 'Elevator']
|
|['Kitchen', 'Microwave', 'Wifi', 'Laundromat nearby', 'Dryer', 'Hot water', 'Garden or backyard', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Mini fridge', 'Single level home', 'Hair dryer', 'Lockbox', 'Dedicated workspace', 'First aid kit', 'Iron', 'Body soap', 'Free street parking', 'Coffee maker', 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Barbecue utensils', 'Shower gel', 'Hangers', 'Cable TV', 'Essentials', 'Washer', 'Conditioner', 'BBQ grill', 'Beach essentials', 'TV', 'Long term stays allowed']
|
|['Kitchen', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dryer', 'Stove', 'Hot water', 'Cooking basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Iron', 'Free street parking', 'Coffee maker', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguisher', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed']
|
|['Kitchen', 'Hot tub', 'Oven', 'Wifi', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Garden or backyard', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Free parking on premises', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Patio or balcony', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Long term stays allowed', 'TV']
|
|['Kitchen', 'Hot water', 'Paid parking on premises', 'Lock on bedroom door', 'Free street parking', 'Cable TV', 'Smoke alarm', 'Wifi', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Carbon monoxide alarm', 'Dishes and silverware', 'TV', 'Long term stays allowed', 'Host greets you', 'Hair dryer', 'Shampoo', 'Fire extinguisher']
|
|['Kitchen', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Hot water', 'Luggage dropoff allowed', 'Private entr

```
ance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bed li
nens', 'Shampoo', 'Fire extinguisher', 'Hangers', 'Cable TV', 'Essentials', 'Washer', 'Long term stays allow
ed', 'TV', 'Air conditioning']
|
|['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Ho
t water', 'Garden or backyard', 'Dishwasher', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cook
ing basics', 'Room-darkening shades', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'First
aid kit', 'Iron', 'Free street parking', 'Coffee maker', 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Dish
es and silverware', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'Pack \u20
19n Play/travel crib']
|
|['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi', 'Host greets you', 'Dryer', 'Stove', 'Carbon m
onoxide alarm', 'Hot water', 'Garden or backyard', 'Dishwasher', 'Cooking basics', 'Smoke alarm', 'Heating',
'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Free street parking', 'Dishes and silverwar
e', 'Fire extinguisher', 'Patio or balcony', 'Cable TV', 'Refrigerator', 'Hangers', 'Essentials', 'Washer',
'Long term stays allowed', 'BBQ grill', 'TV', 'Cleaning before checkout']
|
|['Kitchen', 'Iron', 'Elevator', 'Gym', 'Hangers', 'Smoke alarm', 'Wifi', 'Essentials', 'Heating', 'Washer',
'Dryer', 'Free parking on premises', 'Long term stays allowed', 'Hair dryer', 'Shampoo', 'Dedicated workspac
e']
|
|['Kitchen', 'Hot tub', 'Oven', 'Pool', 'Gym', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dryer', 'Carbon mon
oxide alarm', 'Stove', 'Hot water', 'Dishwasher', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspa
ce', 'Iron', 'Free parking on premises', 'Coffee maker', 'Dishes and silverware', 'Hangers', 'Cable TV', 'Re
frigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning']
|
|['Kitchen', 'Cooking basics', 'Iron', 'Elevator', 'Gym', 'Hangers', 'Cable TV', 'Smoke alarm', 'Essential
s', 'Heating', 'Washer', 'Dryer', 'Carbon monoxide alarm', 'Wifi', 'TV', 'Long term stays allowed', 'Hair dr
yer', 'Shampoo', 'Dishes and silverware', 'Dedicated workspace']
|
|['Kitchen', 'Lake access', 'Oven', 'Paid parking on premises', 'Paid parking off premises', 'Gym', 'Microwa
ve', 'Wifi', 'Dryer', 'Stove', 'Hot water', 'Dishwasher', 'Extra pillows and blankets', 'Cooking basics', 'H
eating', 'Lock on bedroom door', 'Hair dryer', 'Iron', 'Coffee maker', 'Bed linens', 'Pocket wifi', 'Shampo
o', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed',
'Elevator']
|
+-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------+
only showing top 20 rows
```

We want to transform them into multiple columns reusing the machinery in MLlib for NLP.

In [26]:

```python
class FieldSplitter(UnaryTransformer):

    def createTransformFunc(self):
        def f(x):
            if x is None:
                return []
            l = x.split("', '")
            if len(l) > 1:
                l[0] = l[0][2:]
                l[-1] = l[-1][:-2]
            return l
        return f

    def outputDataType(self):
        return ArrayType(StringType())

    def validateInputType(self, inputType):
        assert inputType == StringType(), f'Expected StringType() and found {inputType}'

amenities_splitter = FieldSplitter(). \
  setInputCol('amenities'). \
  setOutputCol('amenities_array')

amenities_splitter.transform(dataset).show(vertical=True, truncate=False)
stages.append(amenities_splitter)
```

```
-RECORD 0-----------------------
```

```
 host_identity_verified                       | t
 neighbourhood                                | null
 neighbourhood_cleansed                       | West End
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.28209
 longitude                                    | -123.12515
 property_type                                | Entire condominium
 room_type                                    | Entire home/apt
 accommodates                                 | 2
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Oven', 'Gym', 'Paid parking off premises', 'Mic
rowave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Extra pillows and blankets', 'Dish
washer', 'Cooking basics', 'Hot water kettle', 'Smoke alarm', 'Heating', 'Single level home', 'Hair dryer',
'Dedicated workspace', 'Iron', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Hangers',
'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning', 'El
evator']
 price                                        | 90
 minimum_nights                               | 30
 maximum_nights                               | 60
 minimum_minimum_nights                       | 30
 maximum_minimum_nights                       | 30
 minimum_maximum_nights                       | 60
 maximum_maximum_nights                       | 60
 minimum_nights_avg_ntm                       | 30.0
 maximum_nights_avg_ntm                       | 60.0
 has_availability                             | t
 availability_30                              | 11
 availability_60                              | 41
 availability_90                              | 70
 availability_365                             | 70
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 amenities_array                              | [Kitchen, Oven, Gym, Paid parking off premises, Microwave, W
ifi, Dryer, Carbon monoxide alarm, Stove, Hot water, Extra pillows and blankets, Dishwasher, Cooking basics,
Hot water kettle, Smoke alarm, Heating, Single level home, Hair dryer, Dedicated workspace, Iron, Coffee mak
er, Bed linens, Shampoo, Dishes and silverware, Hangers, Cable TV, Refrigerator, Essentials, Washer, Long te
rm stays allowed, TV, Air conditioning, Elevator]
-RECORD 1-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 amenities_array                              | [Kitchen, Oven, Microwave, Wifi, Dryer, Carbon monoxide alar
m, Stove, Hot water, Extra pillows and blankets, Luggage dropoff allowed, Cooking basics, Private entrance,
 Smoke alarm, Hair dryer, Lockbox, Dedicated workspace, Iron, Free street parking, Free parking on premises,
Coffee maker, Bed linens, Shampoo, Dishes and silverware, Fire extinguisher, Patio or balcony, Children\u201
9s books and toys, Hangers, Refrigerator, Essentials, Washer, Long term stays allowed, Pack \u2019n Play/tra
vel crib]
-RECORD 2-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                           | 17765
 host_since                                   | 2010/01/07
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | N/A
 host_response_rate                           | N/A
 host_acceptance_rate                         | 1
 host_is_superhost                            | f
 host_neighbourhood                           | Mount Pleasant
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['phone', 'facebook', 'reviews', 'jumio', 'offline_governmen
t_id', 'selfie', 'government_id']
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Mount Pleasant
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.26181
 longitude                                    | -123.10905
 property_type                                | Entire apartment
 room_type                                    | Entire home/apt
 accommodates                                 | 2
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Wifi', 'Host greets you', 'Dryer', 'Carbon mono
xide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated work
space', 'Iron', 'Free street parking', 'Free parking on premises', 'Coffee maker', 'Shampoo', 'Dishes and si
lverware', 'Fire extinguisher', 'Hangers', 'Essentials', 'Washer', 'Long term stays allowed', 'Elevator']
 price                                        | 94
```

```
 minimum_nights                                | 3
 maximum_nights                                | 60
 minimum_minimum_nights                        | 3
 maximum_minimum_nights                        | 3
 minimum_maximum_nights                        | 60
 maximum_maximum_nights                        | 60
 minimum_nights_avg_ntm                        | 3.0
 maximum_nights_avg_ntm                        | 60.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 0
 availability_365                              | 0
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 amenities_array                               | [Kitchen, Wifi, Host greets you, Dryer, Carbon monoxide alar
m, Hot water, Luggage dropoff allowed, Smoke alarm, Heating, Hair dryer, Dedicated workspace, Iron, Free str
eet parking, Free parking on premises, Coffee maker, Shampoo, Dishes and silverware, Fire extinguisher, Hang
ers, Essentials, Washer, Long term stays allowed, Elevator]
-RECORD 3-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                            | 18795
 host_since                                    | 2010/01/13
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | within a day
 host_response_rate                            | 0.7
 host_acceptance_rate                          | 0.4
 host_is_superhost                             | f
 host_neighbourhood                            | West End
 host_listings_count                           | 3
 host_total_listings_count                     | 3
 host_verifications                            | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | null
 neighbourhood_cleansed                        | West End
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.28251
 longitude                                     | -123.12827
 property_type                                 | Private room in house
```

```
 room_type                                     | Private room
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 private bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Cable TV', 'Wifi', 'Essentials', 'Heating', 'Wa
sher', 'Dryer', 'Long term stays allowed']
 price                                         | 50
 minimum_nights                                | 30
 maximum_nights                                | 360
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 360
 maximum_maximum_nights                        | 360
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 360.0
 has_availability                              | t
 availability_30                               | 9
 availability_60                               | 20
 availability_90                               | 50
 availability_365                              | 325
 instant_bookable                              | f
 calculated_host_listings_count                | 3
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms | 3
 calculated_host_listings_count_shared_rooms   | 0
 amenities_array                               | [Kitchen, Cable TV, Wifi, Essentials, Heating, Washer, Drye
r, Long term stays allowed]
-RECORD 4-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                            | 21913
 host_since                                    | 2010/02/22
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | within an hour
 host_response_rate                            | 1
 host_acceptance_rate                          | 1
 host_is_superhost                             | t
 host_neighbourhood                            | Renfrew-Collingwood
 host_listings_count                           | 3
 host_total_listings_count                     | 3
 host_verifications                            | ['email', 'phone', 'reviews', 'offline_government_id', 'self
ie', 'government_id']
```

```
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                        | Renfrew-Collingwood
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.23427
 longitude                                     | -123.05309
 property_type                                 | Entire house
 room_type                                     | Entire home/apt
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Oven', 'Microwave', 'Wifi', 'Dryer', 'Carbon mo
noxide alarm', 'Stove', 'Hot water', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heatin
g', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Free parking o
n premises', 'Coffee maker', 'Bed linens', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV']
 price                                         | 80
 minimum_nights                                | 28
 maximum_nights                                | 400
 minimum_minimum_nights                        | 28
 maximum_minimum_nights                        | 28
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 28.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 2
 availability_60                               | 13
 availability_90                               | 43
 availability_365                              | 318
 instant_bookable                              | f
 calculated_host_listings_count                | 3
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 2
 calculated_host_listings_count_shared_rooms   | 0
 amenities_array                               | [Kitchen, Oven, Microwave, Wifi, Dryer, Carbon monoxide alar
m, Stove, Hot water, Luggage dropoff allowed, Private entrance, Smoke alarm, Heating, Single level home, Hai
r dryer, Dedicated workspace, Iron, Free street parking, Free parking on premises, Coffee maker, Bed linens,
Keypad, Shampoo, Dishes and silverware, Fire extinguisher, Hangers, Cable TV, Refrigerator, Essentials, Wash
er, Long term stays allowed, TV]
-RECORD 5-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 availability_365                              | 223
 instant_bookable                              | f
 calculated_host_listings_count                | 2
 calculated_host_listings_count_entire_homes   | 2
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 amenities_array                               | [Kitchen, Oven, Indoor fireplace, Microwave, Wifi, Dryer, Fi
replace guards, Carbon monoxide alarm, Stove, Crib, Hot water, Babysitter recommendations, Garden or backyar
d, Extra pillows and blankets, Dishwasher, Luggage dropoff allowed, Cooking basics, Room-darkening shades, P
rivate entrance, Smoke alarm, Baking sheet, Heating, Single level home, Hair dryer, Dedicated workspace, Iro
n, Free street parking, Bathtub, Coffee maker, Bed linens, Keypad, Shampoo, Fire extinguisher, Dishes and si
lverware, Shower gel, Patio or balcony, Children\u2019s books and toys, High chair, Hangers, Cable TV, Refri
gerator, Essentials, Washer, Long term stays allowed, Pack \u2019n Play/travel crib, TV, Ethernet connectio
n]
-RECORD 6-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
 id                                            | 42645
 host_since                                    | 2010/04/17
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | within an hour
 host_response_rate                            | 1
 host_acceptance_rate                          | 1
 host_is_superhost                             | t
 host_neighbourhood                            | Downtown Vancouver
 host_listings_count                           | 4
 host_total_listings_count                     | 4
 host_verifications                            | ['email', 'phone', 'facebook', 'reviews']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                        | Downtown
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.28001
 longitude                                     | -123.12475
 property_type                                 | Entire condominium
 room_type                                     | Entire home/apt
 accommodates                                  | 3
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 2
 amenities                                     | ['Kitchen', 'Oven', 'Paid parking on premises', 'Paid parkin
g off premises', 'Gym', 'Microwave', 'Wifi', 'Host greets you', 'Dryer', 'Carbon monoxide alarm', 'Stove',
```

```
 'Hot water', 'Dishwasher', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dry
er', 'Dedicated workspace', 'First aid kit', 'Iron', 'Coffee maker', 'Shampoo', 'Fire extinguisher', 'Dishes
and silverware', 'Patio or balcony', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable TV',
 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'TV',
 'Elevator']
 price                                       | 159
 minimum_nights                              | 32
 maximum_nights                              | 365
 minimum_minimum_nights                      | 32
 maximum_minimum_nights                      | 32
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 32.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 29
 availability_90                             | 59
 availability_365                            | 333
 instant_bookable                            | t
 calculated_host_listings_count              | 4
 calculated_host_listings_count_entire_homes | 4
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms | 0
 amenities_array                             | [Kitchen, Oven, Paid parking on premises, Paid parking off p
remises, Gym, Microwave, Wifi, Host greets you, Dryer, Carbon monoxide alarm, Stove, Hot water, Dishwasher,
 Luggage dropoff allowed, Cooking basics, Smoke alarm, Heating, Hair dryer, Dedicated workspace, First aid k
it, Iron, Coffee maker, Shampoo, Fire extinguisher, Dishes and silverware, Patio or balcony, Children\u2019s
books and toys, High chair, Hangers, Cable TV, Refrigerator, Essentials, Washer, Long term stays allowed, Pa
ck \u2019n Play/travel crib, TV, Elevator]
-RECORD 7-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------
 id                                          | 54261
 host_since                                  | 2010/10/05
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | a few days or more
 host_response_rate                          | 0
 host_acceptance_rate                        | N/A
 host_is_superhost                           | f
 host_neighbourhood                          | Downtown Vancouver
 host_listings_count                         | 2
 host_total_listings_count                   | 2
 host_verifications                          | ['email', 'phone', 'reviews', 'offline_government_id', 'self
```

```
ie', 'government_id', 'identity_manual']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | null
 neighbourhood_cleansed                        | Downtown
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.27655
 longitude                                     | -123.12485
 property_type                                 | Entire condominium
 room_type                                     | Entire home/apt
 accommodates                                  | 4
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Hot tub', 'Gym', 'Indoor fireplace', 'Wifi', 'H
ost greets you', 'Dryer', 'Carbon monoxide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alarm', 'H
eating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Shampoo', 'Fire extinguis
her', 'Hangers', 'Cable TV', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning', 'E
levator']
 price                                         | 126
 minimum_nights                                | 90
 maximum_nights                                | 365
 minimum_minimum_nights                        | 90
 maximum_minimum_nights                        | 90
 minimum_maximum_nights                        | 365
 maximum_maximum_nights                        | 365
 minimum_nights_avg_ntm                        | 90.0
 maximum_nights_avg_ntm                        | 365.0
 has_availability                              | t
 availability_30                               | 30
 availability_60                               | 60
 availability_90                               | 90
 availability_365                              | 365
 instant_bookable                              | f
 calculated_host_listings_count                | 2
 calculated_host_listings_count_entire_homes   | 2
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 amenities_array                               | [Kitchen, Hot tub, Gym, Indoor fireplace, Wifi, Host greets
 you, Dryer, Carbon monoxide alarm, Hot water, Luggage dropoff allowed, Smoke alarm, Heating, Hair dryer, De
dicated workspace, Iron, Free parking on premises, Shampoo, Fire extinguisher, Hangers, Cable TV, Essential
s, Washer, Long term stays allowed, TV, Air conditioning, Elevator]
-RECORD 8-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 amenities_array                              | [Kitchen, Lake access, Oven, Paid parking off premises, Gym,
Indoor fireplace, Microwave, Wifi, Dryer, Carbon monoxide alarm, Stove, Hot water, Dishwasher, Extra pillows
and blankets, Luggage dropoff allowed, Cooking basics, Smoke alarm, Heating, Single level home, Hair dryer,
 Dedicated workspace, Iron, Free street parking, Coffee maker, Bed linens, Shampoo, Dishes and silverware, F
ire extinguisher, Patio or balcony, Hangers, Refrigerator, Essentials, Washer, BBQ grill, TV, Elevator]
-RECORD 9-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                           | 77157
 host_since                                   | 2011/02/27
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | within a few hours
 host_response_rate                           | 1
 host_acceptance_rate                         | 1
 host_is_superhost                            | f
 host_neighbourhood                           | Arbutus Ridge
 host_listings_count                          | 2
 host_total_listings_count                    | 2
 host_verifications                           | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'offline_
government_id', government_id']
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Arbutus Ridge
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.24675
 longitude                                    | -123.16711
 property_type                                | Private room in guest suite
 room_type                                    | Private room
 accommodates                                 | 3
 bathrooms                                    | null
 bathrooms_text                               | 1 private bath
 bedrooms                                     | 1
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Microwave', 'Wifi', 'Laundromat nearby', 'Drye
r', 'Hot water', 'Garden or backyard', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking bas
ics', 'Private entrance', 'Smoke alarm', 'Heating', 'Mini fridge', 'Single level home', 'Hair dryer', 'Lockb
ox', 'Dedicated workspace', 'First aid kit', 'Iron', 'Body soap', 'Free street parking', 'Coffee maker', 'Be
d linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Barbecue utensils', 'Shower gel', 'Hang
ers', 'Cable TV', 'Essentials', 'Washer', 'Conditioner', 'BBQ grill', 'Beach essentials', 'TV', 'Long term s
tays allowed']
```

```
 price                                       | 75
 minimum_nights                              | 3
 maximum_nights                              | 60
 minimum_minimum_nights                      | 3
 maximum_minimum_nights                      | 3
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 3.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 3
 availability_60                             | 3
 availability_90                             | 3
 availability_365                            | 267
 instant_bookable                            | t
 calculated_host_listings_count              | 2
 calculated_host_listings_count_entire_homes | 0
 calculated_host_listings_count_private_rooms | 2
 calculated_host_listings_count_shared_rooms | 0
 amenities_array                             | [Kitchen, Microwave, Wifi, Laundromat nearby, Dryer, Hot wat
er, Garden or backyard, Extra pillows and blankets, Luggage dropoff allowed, Cooking basics, Private entranc
e, Smoke alarm, Heating, Mini fridge, Single level home, Hair dryer, Lockbox, Dedicated workspace, First aid
kit, Iron, Body soap, Free street parking, Coffee maker, Bed linens, Shampoo, Fire extinguisher, Dishes and
 silverware, Barbecue utensils, Shower gel, Hangers, Cable TV, Essentials, Washer, Conditioner, BBQ grill, B
each essentials, TV, Long term stays allowed]
-RECORD 10-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
 id                                          | 111077
 host_since                                  | 2011/05/09
 host_location                               | North Vancouver, British Columbia, Canada
 host_response_time                          | within a few hours
 host_response_rate                          | 1
 host_acceptance_rate                        | 0.33
 host_is_superhost                           | f
 host_neighbourhood                          | Grandview-Woodland
 host_listings_count                         | 1
 host_total_listings_count                   | 1
 host_verifications                          | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'governme
nt_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                      | Grandview-Woodland
```

```
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.26726
 longitude                                    | -123.05969
 property_type                                | Entire guest suite
 room_type                                    | Entire home/apt
 accommodates                                 | 3
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 2
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Indoor fireplace', 'Microwave', 'Wifi', 'Drye
r', 'Stove', 'Hot water', 'Cooking basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Iro
n', 'Free street parking', 'Coffee maker', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed']
 price                                        | 125
 minimum_nights                               | 3
 maximum_nights                               | 365
 minimum_minimum_nights                       | 3
 maximum_minimum_nights                       | 3
 minimum_maximum_nights                       | 365
 maximum_maximum_nights                       | 365
 minimum_nights_avg_ntm                       | 3.0
 maximum_nights_avg_ntm                       | 365.0
 has_availability                             | t
 availability_30                              | 4
 availability_60                              | 34
 availability_90                              | 64
 availability_365                             | 339
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 amenities_array                              | [Kitchen, Indoor fireplace, Microwave, Wifi, Dryer, Stove, H
ot water, Cooking basics, Private entrance, Smoke alarm, Heating, Hair dryer, Iron, Free street parking, Cof
fee maker, Keypad, Shampoo, Dishes and silverware, Fire extinguisher, Hangers, Refrigerator, Essentials, Was
her, Long term stays allowed]
-RECORD 11-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                           | 125176
 host_since                                   | 2011/05/24
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | within an hour
```

```
 host_response_rate                            | 1
 host_acceptance_rate                          | 1
 host_is_superhost                             | f
 host_neighbourhood                            | Riley Park
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone', 'facebook', 'reviews', 'offline_governmen
t_id', 'government_id', 'work_email']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                        | Riley Park
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.24648
 longitude                                     | -123.09241
 property_type                                 | Private room in house
 room_type                                     | Private room
 accommodates                                  | 3
 bathrooms                                     | null
 bathrooms_text                                | 1 private bath
 bedrooms                                      | 1
 beds                                          | 0
 amenities                                     | ['Kitchen', 'Hot tub', 'Oven', 'Wifi', 'Carbon monoxide alar
m', 'Stove', 'Hot water', 'Garden or backyard', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dryer', 'D
edicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Free parking on premises', 'Keypad', 'Shampo
o', 'Dishes and silverware', 'Patio or balcony', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Long
 term stays allowed', 'TV']
 price                                         | 75
 minimum_nights                                | 2
 maximum_nights                                | 60
 minimum_minimum_nights                        | 2
 maximum_minimum_nights                        | 2
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 2.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 2
 availability_365                              | 268
 instant_bookable                              | t
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms  | 1
 calculated_host_listings_count_shared_rooms   | 0
 amenities_array                               | [Kitchen, Hot tub, Oven, Wifi, Carbon monoxide alarm, Stove,
Hot water, Garden or backyard, Cooking basics, Smoke alarm, Heating, Hair dryer, Dedicated workspace, Iron,
 Free street parking, Bathtub, Free parking on premises, Keypad, Shampoo, Dishes and silverware, Patio or ba
```

```
lcony, Hangers, Cable TV, Refrigerator, Essentials, Long term stays allowed, TV]
-RECORD 12-----------------------
```

```
 availability_30                              | 30
 availability_60                              | 60
 availability_90                              | 90
 availability_365                             | 365
 instant_bookable                             | t
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 0
 calculated_host_listings_count_private_rooms | 1
 calculated_host_listings_count_shared_rooms  | 0
 amenities_array                              | [Kitchen, Hot water, Paid parking on premises, Lock on bedro
om door, Free street parking, Cable TV, Smoke alarm, Wifi, Essentials, Heating, Washer, Dryer, Carbon monoxi
de alarm, Dishes and silverware, TV, Long term stays allowed, Host greets you, Hair dryer, Shampoo, Fire ext
inguisher]
-RECORD 13-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                           | 174135
 host_since                                   | 2011/07/17
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | within a few hours
 host_response_rate                           | 1
 host_acceptance_rate                         | 0
 host_is_superhost                            | f
 host_neighbourhood                           | Downtown Eastside
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['email', 'phone', 'reviews', 'offline_government_id', 'self
ie', 'government_id', 'identity_manual']
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Downtown Eastside
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27782
 longitude                                    | -123.08985
 property_type                                | Entire guest suite
 room_type                                    | Entire home/apt
 accommodates                                 | 2
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Hot w
ater', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated wor
```

```
kspace', 'Iron', 'Free street parking', 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Hangers', 'Cable TV',
'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning']
 price                                       | 100
 minimum_nights                              | 30
 maximum_nights                              | 365
 minimum_minimum_nights                      | 30
 maximum_minimum_nights                      | 30
 minimum_maximum_nights                      | 365
 maximum_maximum_nights                      | 365
 minimum_nights_avg_ntm                      | 30.0
 maximum_nights_avg_ntm                      | 365.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 0
 availability_90                             | 0
 availability_365                            | 224
 instant_bookable                            | f
 calculated_host_listings_count              | 1
 calculated_host_listings_count_entire_homes | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms | 0
 amenities_array                             | [Kitchen, Wifi, Dryer, Carbon monoxide alarm, Hot water, Lug
gage dropoff allowed, Private entrance, Smoke alarm, Heating, Hair dryer, Dedicated workspace, Iron, Free st
reet parking, Bed linens, Shampoo, Fire extinguisher, Hangers, Cable TV, Essentials, Washer, Long term stays
allowed, TV, Air conditioning]
-RECORD 14-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                          | 200643
 host_since                                  | 2011/08/16
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | within an hour
 host_response_rate                          | 1
 host_acceptance_rate                        | 1
 host_is_superhost                           | f
 host_neighbourhood                          | Strathcona
 host_listings_count                         | 4
 host_total_listings_count                   | 4
 host_verifications                          | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                      | Downtown Eastside
 neighbourhood_group_cleansed                | null
```

```
 latitude                                     | 49.27384
 longitude                                    | -123.08289
 property_type                                | Private room in house
 room_type                                    | Private room
 accommodates                                 | 4
 bathrooms                                    | null
 bathrooms_text                               | 2 shared baths
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi',
'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Garden or backyard', 'Dishwasher', 'Extra pillows a
nd blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Room-darkening shades', 'Smoke alarm', 'Heatin
g', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Free street parking', 'Coffee maker', 'Be
d linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Essentials',
'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
 price                                        | 87
 minimum_nights                               | 1
 maximum_nights                               | 1125
 minimum_minimum_nights                       | 1
 maximum_minimum_nights                       | 1
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
 minimum_nights_avg_ntm                       | 1.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 30
 availability_60                              | 60
 availability_90                              | 90
 availability_365                             | 365
 instant_bookable                             | t
 calculated_host_listings_count               | 4
 calculated_host_listings_count_entire_homes  | 0
 calculated_host_listings_count_private_rooms | 4
 calculated_host_listings_count_shared_rooms  | 0
 amenities_array                              | [Kitchen, Oven, Indoor fireplace, Microwave, Wifi, Dryer, Ca
rbon monoxide alarm, Stove, Hot water, Garden or backyard, Dishwasher, Extra pillows and blankets, Luggage d
ropoff allowed, Cooking basics, Room-darkening shades, Smoke alarm, Heating, Hair dryer, Dedicated workspac
e, First aid kit, Iron, Free street parking, Coffee maker, Bed linens, Shampoo, Fire extinguisher, Dishes an
d silverware, Hangers, Refrigerator, Essentials, Washer, Long term stays allowed, Pack \u2019n Play/travel c
rib]
-RECORD 15-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
 id                                           | 248014
```

```
 host_since                                     | 2011/10/16
 host_location                                  | Vancouver, British Columbia, Canada
 host_response_time                             | within a day
 host_response_rate                             | 0.5
 host_acceptance_rate                           | N/A
 host_is_superhost                              | f
 host_neighbourhood                             | Grandview-Woodland
 host_listings_count                            | 1
 host_total_listings_count                      | 1
 host_verifications                             | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                         | Grandview-Woodland
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.29103
 longitude                                      | -123.05782
 property_type                                  | Entire house
 room_type                                      | Entire home/apt
 accommodates                                   | 3
 bathrooms                                      | null
 bathrooms_text                                 | 1 bath
 bedrooms                                       | 2
 beds                                           | 0
 amenities                                      | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi',
'Host greets you', 'Dryer', 'Stove', 'Carbon monoxide alarm', 'Hot water', 'Garden or backyard', 'Dishwashe
r', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iro
n', 'Free street parking', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Cable TV', 'Re
frigerator', 'Hangers', 'Essentials', 'Washer', 'Long term stays allowed', 'BBQ grill', 'TV', 'Cleaning befo
re checkout']
 price                                          | 100
 minimum_nights                                 | 31
 maximum_nights                                 | 180
 minimum_minimum_nights                         | 31
 maximum_minimum_nights                         | 31
 minimum_maximum_nights                         | 180
 maximum_maximum_nights                         | 180
 minimum_nights_avg_ntm                         | 31.0
 maximum_nights_avg_ntm                         | 180.0
 has_availability                               | t
 availability_30                                | 4
 availability_60                                | 4
 availability_90                                | 4
 availability_365                               | 113
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
```

```
 amenities_array                              | [Kitchen, Oven, Indoor fireplace, Microwave, Wifi, Host gree
ts you, Dryer, Stove, Carbon monoxide alarm, Hot water, Garden or backyard, Dishwasher, Cooking basics, Smok
e alarm, Heating, Hair dryer, Dedicated workspace, First aid kit, Iron, Free street parking, Dishes and silv
erware, Fire extinguisher, Patio or balcony, Cable TV, Refrigerator, Hangers, Essentials, Washer, Long term
 stays allowed, BBQ grill, TV, Cleaning before checkout]
-RECORD 16-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                           | 258636
 host_since                                   | 2011/08/09
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | N/A
 host_response_rate                           | N/A
 host_acceptance_rate                         | 1
 host_is_superhost                            | f
 host_neighbourhood                           | Downtown Vancouver
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | null
 neighbourhood_cleansed                       | Downtown
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27796
 longitude                                    | -123.12671
 property_type                                | Entire apartment
 room_type                                    | Entire home/apt
 accommodates                                 | 4
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Iron', 'Elevator', 'Gym', 'Hangers', 'Smoke ala
rm', 'Wifi', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Free parking on premises', 'Long term stays allowe
d', 'Hair dryer', 'Shampoo', 'Dedicated workspace']
 price                                        | 99
 minimum_nights                               | 120
 maximum_nights                               | 365
 minimum_minimum_nights                       | 120
 maximum_minimum_nights                       | 120
 minimum_maximum_nights                       | 365
 maximum_maximum_nights                       | 365
 minimum_nights_avg_ntm                       | 120.0
```

```
 maximum_nights_avg_ntm                       | 365.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 11
 availability_90                              | 41
 availability_365                             | 316
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 amenities_array                              | [Kitchen, Iron, Elevator, Gym, Hangers, Smoke alarm, Wifi, E
ssentials, Heating, Washer, Dryer, Free parking on premises, Long term stays allowed, Hair dryer, Shampoo, D
edicated workspace]
-RECORD 17-------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                           | 283841
 host_since                                   | 2009/08/10
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | within a day
 host_response_rate                           | 0.93
 host_acceptance_rate                         | 0.88
 host_is_superhost                            | f
 host_neighbourhood                           | Coal Harbour
 host_listings_count                          | 49
 host_total_listings_count                    | 49
 host_verifications                           | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | null
 neighbourhood_cleansed                       | Downtown
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.2872
 longitude                                    | -123.12352
 property_type                                | Entire apartment
 room_type                                    | Entire home/apt
 accommodates                                 | 6
 bathrooms                                    | null
 bathrooms_text                               | 3 baths
 bedrooms                                     | 3
 beds                                         | 3
 amenities                                    | ['Kitchen', 'Hot tub', 'Oven', 'Pool', 'Gym', 'Indoor firepl
ace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Dishwasher', 'Smoke alar
```

```
m', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Coffee maker', 'Dis
hes and silverware', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowe
d', 'TV', 'Air conditioning']
 price                                       | 197
 minimum_nights                              | 90
 maximum_nights                              | 1125
 minimum_minimum_nights                      | 90
 maximum_minimum_nights                      | 90
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 90.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 0
 availability_90                             | 0
 availability_365                            | 0
 instant_bookable                            | f
 calculated_host_listings_count              | 41
 calculated_host_listings_count_entire_homes | 41
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms | 0
 amenities_array                             | [Kitchen, Hot tub, Oven, Pool, Gym, Indoor fireplace, Microw
ave, Wifi, Dryer, Carbon monoxide alarm, Stove, Hot water, Dishwasher, Smoke alarm, Heating, Hair dryer, Ded
icated workspace, Iron, Free parking on premises, Coffee maker, Dishes and silverware, Hangers, Cable TV, Re
frigerator, Essentials, Washer, Long term stays allowed, TV, Air conditioning]
-RECORD 18-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                          | 323090
 host_since                                  | 2012/01/01
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | N/A
 host_response_rate                          | N/A
 host_acceptance_rate                        | 0.85
 host_is_superhost                           | t
 host_neighbourhood                          | Downtown Vancouver
 host_listings_count                         | 1
 host_total_listings_count                   | 1
 host_verifications                          | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                      | Downtown
```

```
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.28206
 longitude                                    | -123.12461
 property_type                                | Entire apartment
 room_type                                    | Entire home/apt
 accommodates                                 | 2
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Cooking basics', 'Iron', 'Elevator', 'Gym', 'Ha
ngers', 'Cable TV', 'Smoke alarm', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Carbon monoxide alarm', 'Wif
i', 'TV', 'Long term stays allowed', 'Hair dryer', 'Shampoo', 'Dishes and silverware', 'Dedicated workspac
e']
 price                                        | 168
 minimum_nights                               | 180
 maximum_nights                               | 180
 minimum_minimum_nights                       | 180
 maximum_minimum_nights                       | 180
 minimum_maximum_nights                       | 180
 maximum_maximum_nights                       | 180
 minimum_nights_avg_ntm                       | 180.0
 maximum_nights_avg_ntm                       | 180.0
 has_availability                             | t
 availability_30                              | 30
 availability_60                              | 60
 availability_90                              | 90
 availability_365                             | 141
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 amenities_array                              | [Kitchen, Cooking basics, Iron, Elevator, Gym, Hangers, Cabl
e TV, Smoke alarm, Essentials, Heating, Washer, Dryer, Carbon monoxide alarm, Wifi, TV, Long term stays allo
wed, Hair dryer, Shampoo, Dishes and silverware, Dedicated workspace]
-RECORD 19-----------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                           | 359942
 host_since                                   | 2011/09/23
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | within an hour
 host_response_rate                           | 1
```

```
 host_acceptance_rate                          | N/A
 host_is_superhost                             | f
 host_neighbourhood                            | Yaletown
 host_listings_count                           | 8
 host_total_listings_count                     | 8
 host_verifications                            | ['email', 'phone', 'facebook', 'reviews']
 host_has_profile_pic                          | t
 host_identity_verified                        | f
 neighbourhood                                 | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                        | Downtown
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.27825
 longitude                                     | -123.11865
 property_type                                 | Private room in apartment
 room_type                                     | Private room
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 shared bath
 bedrooms                                      | 2
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Lake access', 'Oven', 'Paid parking on premise
s', 'Paid parking off premises', 'Gym', 'Microwave', 'Wifi', 'Dryer', 'Stove', 'Hot water', 'Dishwasher', 'E
xtra pillows and blankets', 'Cooking basics', 'Heating', 'Lock on bedroom door', 'Hair dryer', 'Iron', 'Coff
ee maker', 'Bed linens', 'Pocket wifi', 'Shampoo', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Esse
ntials', 'Washer', 'Long term stays allowed', 'Elevator']
 price                                         | 240
 minimum_nights                                | 30
 maximum_nights                                | 1125
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 29
 availability_60                               | 59
 availability_90                               | 89
 availability_365                              | 364
 instant_bookable                              | f
 calculated_host_listings_count                | 7
 calculated_host_listings_count_entire_homes   | 3
 calculated_host_listings_count_private_rooms  | 4
 calculated_host_listings_count_shared_rooms   | 0
 amenities_array                               | [Kitchen, Lake access, Oven, Paid parking on premises, Paid
 parking off premises, Gym, Microwave, Wifi, Dryer, Stove, Hot water, Dishwasher, Extra pillows and blanket
s, Cooking basics, Heating, Lock on bedroom door, Hair dryer, Iron, Coffee maker, Bed linens, Pocket wifi, S
hampoo, Dishes and silverware, Hangers, Refrigerator, Essentials, Washer, Long term stays allowed, Elevator]
```

```
only showing top 20 rows
```

In [27]:

```python
verifications_splitter = FieldSplitter(). \
  setInputCol('host_verifications'). \
  setOutputCol('verifications_array')

verifications_splitter.transform(dataset).show(vertical=True, truncate=False)
stages.append(verifications_splitter)
```

```
-RECORD 0-----------------------
```

```
'Dedicated workspace', 'Iron', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Hangers',
'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning', 'El
evator']
 price                                          | 90
 minimum_nights                                 | 30
 maximum_nights                                 | 60
 minimum_minimum_nights                         | 30
 maximum_minimum_nights                         | 30
 minimum_maximum_nights                         | 60
 maximum_maximum_nights                         | 60
 minimum_nights_avg_ntm                         | 30.0
 maximum_nights_avg_ntm                         | 60.0
 has_availability                               | t
 availability_30                                | 11
 availability_60                                | 41
 availability_90                                | 70
 availability_365                               | 70
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
 verifications_array                            | [email, phone, facebook, reviews, jumio, offline_government_
id, selfie, government_id]
-RECORD 1-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                             | 16254
 host_since                                     | 2009/12/15
 host_location                                  | Vancouver, British Columbia, Canada
 host_response_time                             | N/A
 host_response_rate                             | N/A
 host_acceptance_rate                           | N/A
 host_is_superhost                              | f
 host_neighbourhood                             | Hastings-Sunrise
 host_listings_count                            | 1
 host_total_listings_count                      | 1
 host_verifications                             | ['phone', 'facebook', 'offline_government_id', 'selfie', 'go
vernment_id', 'identity_manual']
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                         | Hastings-Sunrise
 neighbourhood_group_cleansed                   | null
```

```
 latitude                                     | 49.27734
 longitude                                    | -123.04152
 property_type                                | Entire guest suite
 room_type                                    | Entire home/apt
 accommodates                                 | 4
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 2
 beds                                         | 3
 amenities                                    | ['Kitchen', 'Oven', 'Microwave', 'Wifi', 'Dryer', 'Carbon mo
noxide alarm', 'Stove', 'Hot water', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking basic
s', 'Private entrance', 'Smoke alarm', 'Hair dryer', 'Lockbox', 'Dedicated workspace', 'Iron', 'Free street
 parking', 'Free parking on premises', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Fi
re extinguisher', 'Patio or balcony', 'Children\u2019s books and toys', 'Hangers', 'Refrigerator', 'Essentia
ls', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
 price                                        | 195
 minimum_nights                               | 3
 maximum_nights                               | 31
 minimum_minimum_nights                       | 3
 maximum_minimum_nights                       | 3
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
 minimum_nights_avg_ntm                       | 3.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 0
 availability_365                             | 228
 instant_bookable                             | t
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 verifications_array                          | [phone, facebook, offline_government_id, selfie, government_
id, identity_manual]
-RECORD 2-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------
 id                                           | 17765
 host_since                                   | 2010/01/07
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | N/A
 host_response_rate                           | N/A
```

```
 host_acceptance_rate                          | 1
 host_is_superhost                             | f
 host_neighbourhood                            | Mount Pleasant
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['phone', 'facebook', 'reviews', 'jumio', 'offline_governmen
t_id', 'selfie', 'government_id']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                        | Mount Pleasant
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.26181
 longitude                                     | -123.10905
 property_type                                 | Entire apartment
 room_type                                     | Entire home/apt
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Wifi', 'Host greets you', 'Dryer', 'Carbon mono
xide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated work
space', 'Iron', 'Free street parking', 'Free parking on premises', 'Coffee maker', 'Shampoo', 'Dishes and si
lverware', 'Fire extinguisher', 'Hangers', 'Essentials', 'Washer', 'Long term stays allowed', 'Elevator']
 price                                         | 94
 minimum_nights                                | 3
 maximum_nights                                | 60
 minimum_minimum_nights                        | 3
 maximum_minimum_nights                        | 3
 minimum_maximum_nights                        | 60
 maximum_maximum_nights                        | 60
 minimum_nights_avg_ntm                        | 3.0
 maximum_nights_avg_ntm                        | 60.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 0
 availability_365                              | 0
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 verifications_array                           | [phone, facebook, reviews, jumio, offline_government_id, sel
fie, government_id]
-RECORD 3-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
```

```
-----------------------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------------
 id                                          | 18795
 host_since                                  | 2010/01/13
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | within a day
 host_response_rate                          | 0.7
 host_acceptance_rate                        | 0.4
 host_is_superhost                           | f
 host_neighbourhood                          | West End
 host_listings_count                         | 3
 host_total_listings_count                   | 3
 host_verifications                          | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | null
 neighbourhood_cleansed                      | West End
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.28251
 longitude                                   | -123.12827
 property_type                               | Private room in house
 room_type                                   | Private room
 accommodates                                | 2
 bathrooms                                   | null
 bathrooms_text                              | 1 private bath
 bedrooms                                    | 1
 beds                                        | 1
 amenities                                   | ['Kitchen', 'Cable TV', 'Wifi', 'Essentials', 'Heating', 'Wa
sher', 'Dryer', 'Long term stays allowed']
 price                                       | 50
 minimum_nights                              | 30
 maximum_nights                              | 360
 minimum_minimum_nights                      | 30
 maximum_minimum_nights                      | 30
 minimum_maximum_nights                      | 360
 maximum_maximum_nights                      | 360
 minimum_nights_avg_ntm                      | 30.0
 maximum_nights_avg_ntm                      | 360.0
 has_availability                            | t
 availability_30                             | 9
 availability_60                             | 20
 availability_90                             | 50
 availability_365                            | 325
 instant_bookable                            | f
 calculated_host_listings_count              | 3
 calculated_host_listings_count_entire_homes | 0
```

```
 calculated_host_listings_count_private_rooms | 3
 calculated_host_listings_count_shared_rooms  | 0
 verifications_array                          | [email, phone, reviews, jumio, government_id]
-RECORD 4-------------------
```

```
 minimum_nights_avg_ntm                       | 28.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 2
 availability_60                              | 13
 availability_90                              | 43
 availability_365                             | 318
 instant_bookable                             | f
 calculated_host_listings_count               | 3
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 2
 calculated_host_listings_count_shared_rooms  | 0
 verifications_array                          | [email, phone, reviews, offline_government_id, selfie, gover
nment_id]
-RECORD 5-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------
 id                                           | 24273
 host_since                                   | 2011/10/09
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | within an hour
 host_response_rate                           | 1
 host_acceptance_rate                         | 0.82
 host_is_superhost                            | f
 host_neighbourhood                           | Kitsilano
 host_listings_count                          | 2
 host_total_listings_count                    | 2
 host_verifications                           | ['email', 'phone', 'reviews', 'jumio', 'offline_government_i
d', 'government_id']
 host_has_profile_pic                         | f
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Kitsilano
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.2626
 longitude                                    | -123.17308
 property_type                                | Entire apartment
 room_type                                    | Entire home/apt
 accommodates                                 | 6
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 2
 beds                                         | 3
 amenities                                    | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi',
```

```
'Dryer', 'Fireplace guards', 'Carbon monoxide alarm', 'Stove', 'Crib', 'Hot water', 'Babysitter recommendati
ons', 'Garden or backyard', 'Extra pillows and blankets', 'Dishwasher', 'Luggage dropoff allowed', 'Cooking
 basics', 'Room-darkening shades', 'Private entrance', 'Smoke alarm', 'Baking sheet', 'Heating', 'Single lev
el home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Coffee maker', 'Be
d linens', 'Keypad', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Shower gel', 'Patio or balcon
y', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Wa
sher', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'TV', 'Ethernet connection']
 price                                       | 200
 minimum_nights                              | 30
 maximum_nights                              | 365
 minimum_minimum_nights                      | 30
 maximum_minimum_nights                      | 30
 minimum_maximum_nights                      | 365
 maximum_maximum_nights                      | 365
 minimum_nights_avg_ntm                      | 30.0
 maximum_nights_avg_ntm                      | 365.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 0
 availability_90                             | 10
 availability_365                            | 223
 instant_bookable                            | f
 calculated_host_listings_count              | 2
 calculated_host_listings_count_entire_homes | 2
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms | 0
 verifications_array                         | [email, phone, reviews, jumio, offline_government_id, govern
ment_id]
-RECORD 6-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                          | 42645
 host_since                                  | 2010/04/17
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | within an hour
 host_response_rate                          | 1
 host_acceptance_rate                        | 1
 host_is_superhost                           | t
 host_neighbourhood                          | Downtown Vancouver
 host_listings_count                         | 4
 host_total_listings_count                   | 4
 host_verifications                          | ['email', 'phone', 'facebook', 'reviews']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
```

```
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Downtown
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.28001
 longitude                                    | -123.12475
 property_type                                | Entire condominium
 room_type                                    | Entire home/apt
 accommodates                                 | 3
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Oven', 'Paid parking on premises', 'Paid parkin
g off premises', 'Gym', 'Microwave', 'Wifi', 'Host greets you', 'Dryer', 'Carbon monoxide alarm', 'Stove',
 'Hot water', 'Dishwasher', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dry
er', 'Dedicated workspace', 'First aid kit', 'Iron', 'Coffee maker', 'Shampoo', 'Fire extinguisher', 'Dishes
and silverware', 'Patio or balcony', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable TV',
 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'TV',
 'Elevator']
 price                                        | 159
 minimum_nights                               | 32
 maximum_nights                               | 365
 minimum_minimum_nights                       | 32
 maximum_minimum_nights                       | 32
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
 minimum_nights_avg_ntm                       | 32.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 29
 availability_90                              | 59
 availability_365                             | 333
 instant_bookable                             | t
 calculated_host_listings_count               | 4
 calculated_host_listings_count_entire_homes  | 4
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 verifications_array                          | [email, phone, facebook, reviews]
-RECORD 7-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------
 id                                           | 54261
 host_since                                   | 2010/10/05
```

```
 host_location                                  | Vancouver, British Columbia, Canada
 host_response_time                             | a few days or more
 host_response_rate                             | 0
 host_acceptance_rate                           | N/A
 host_is_superhost                              | f
 host_neighbourhood                             | Downtown Vancouver
 host_listings_count                            | 2
 host_total_listings_count                      | 2
 host_verifications                             | ['email', 'phone', 'reviews', 'offline_government_id', 'self
ie', 'government_id', 'identity_manual']
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | null
 neighbourhood_cleansed                         | Downtown
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.27655
 longitude                                      | -123.12485
 property_type                                  | Entire condominium
 room_type                                      | Entire home/apt
 accommodates                                   | 4
 bathrooms                                      | null
 bathrooms_text                                 | 1 bath
 bedrooms                                       | 1
 beds                                           | 1
 amenities                                      | ['Kitchen', 'Hot tub', 'Gym', 'Indoor fireplace', 'Wifi', 'H
ost greets you', 'Dryer', 'Carbon monoxide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alarm', 'H
eating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Shampoo', 'Fire extinguis
her', 'Hangers', 'Cable TV', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning', 'E
levator']
 price                                          | 126
 minimum_nights                                 | 90
 maximum_nights                                 | 365
 minimum_minimum_nights                         | 90
 maximum_minimum_nights                         | 90
 minimum_maximum_nights                         | 365
 maximum_maximum_nights                         | 365
 minimum_nights_avg_ntm                         | 90.0
 maximum_nights_avg_ntm                         | 365.0
 has_availability                               | t
 availability_30                                | 30
 availability_60                                | 60
 availability_90                                | 90
 availability_365                               | 365
 instant_bookable                               | f
 calculated_host_listings_count                 | 2
 calculated_host_listings_count_entire_homes    | 2
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
 verifications_array                            | [email, phone, reviews, offline_government_id, selfie, gover
```

```
nment_id, identity_manual]
-RECORD 8-----------------------
```

```
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 0
 availability_365                              | 40
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 verifications_array                           | [email, phone, reviews, jumio, government_id]
-RECORD 9-----------------------
```

```
d linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Barbecue utensils', 'Shower gel', 'Hang
ers', 'Cable TV', 'Essentials', 'Washer', 'Conditioner', 'BBQ grill', 'Beach essentials', 'TV', 'Long term s
tays allowed']
 price                                          | 75
 minimum_nights                                 | 3
 maximum_nights                                 | 60
 minimum_minimum_nights                         | 3
 maximum_minimum_nights                         | 3
 minimum_maximum_nights                         | 1125
 maximum_maximum_nights                         | 1125
 minimum_nights_avg_ntm                         | 3.0
 maximum_nights_avg_ntm                         | 1125.0
 has_availability                               | t
 availability_30                                | 3
 availability_60                                | 3
 availability_90                                | 3
 availability_365                               | 267
 instant_bookable                               | t
 calculated_host_listings_count                 | 2
 calculated_host_listings_count_entire_homes    | 0
 calculated_host_listings_count_private_rooms   | 2
 calculated_host_listings_count_shared_rooms    | 0
 verifications_array                            | [email, phone, facebook, reviews, jumio, offline_government_
id, government_id]
-RECORD 10-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
 id                                             | 111077
 host_since                                     | 2011/05/09
 host_location                                  | North Vancouver, British Columbia, Canada
 host_response_time                             | within a few hours
 host_response_rate                             | 1
 host_acceptance_rate                           | 0.33
 host_is_superhost                              | f
 host_neighbourhood                             | Grandview-Woodland
 host_listings_count                            | 1
 host_total_listings_count                      | 1
 host_verifications                             | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'governme
nt_id']
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                         | Grandview-Woodland
 neighbourhood_group_cleansed                   | null
```

```
 latitude                                      | 49.26726
 longitude                                     | -123.05969
 property_type                                 | Entire guest suite
 room_type                                     | Entire home/apt
 accommodates                                  | 3
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 2
 beds                                          | 2
 amenities                                     | ['Kitchen', 'Indoor fireplace', 'Microwave', 'Wifi', 'Drye
r', 'Stove', 'Hot water', 'Cooking basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Iro
n', 'Free street parking', 'Coffee maker', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed']
 price                                         | 125
 minimum_nights                                | 3
 maximum_nights                                | 365
 minimum_minimum_nights                        | 3
 maximum_minimum_nights                        | 3
 minimum_maximum_nights                        | 365
 maximum_maximum_nights                        | 365
 minimum_nights_avg_ntm                        | 3.0
 maximum_nights_avg_ntm                        | 365.0
 has_availability                              | t
 availability_30                               | 4
 availability_60                               | 34
 availability_90                               | 64
 availability_365                              | 339
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 verifications_array                           | [email, phone, facebook, reviews, jumio, government_id]
-RECORD 11-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
 id                                            | 125176
 host_since                                    | 2011/05/24
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | within an hour
 host_response_rate                            | 1
 host_acceptance_rate                          | 1
 host_is_superhost                             | f
 host_neighbourhood                            | Riley Park
```

```
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone', 'facebook', 'reviews', 'offline_governmen
t_id', 'government_id', 'work_email']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                        | Riley Park
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.24648
 longitude                                     | -123.09241
 property_type                                 | Private room in house
 room_type                                     | Private room
 accommodates                                  | 3
 bathrooms                                     | null
 bathrooms_text                                | 1 private bath
 bedrooms                                      | 1
 beds                                          | 0
 amenities                                     | ['Kitchen', 'Hot tub', 'Oven', 'Wifi', 'Carbon monoxide alar
m', 'Stove', 'Hot water', 'Garden or backyard', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dryer', 'D
edicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Free parking on premises', 'Keypad', 'Shampo
o', 'Dishes and silverware', 'Patio or balcony', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Long
 term stays allowed', 'TV']
 price                                         | 75
 minimum_nights                                | 2
 maximum_nights                                | 60
 minimum_minimum_nights                        | 2
 maximum_minimum_nights                        | 2
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 2.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 2
 availability_365                              | 268
 instant_bookable                              | t
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms  | 1
 calculated_host_listings_count_shared_rooms   | 0
 verifications_array                           | [email, phone, facebook, reviews, offline_government_id, gov
ernment_id, work_email]
-RECORD 12-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 calculated_host_listings_count_entire_homes  | 0
 calculated_host_listings_count_private_rooms | 1
 calculated_host_listings_count_shared_rooms  | 0
 verifications_array                          | [email, phone, reviews, jumio, offline_government_id, govern
ment_id]
-RECORD 13-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                           | 174135
 host_since                                   | 2011/07/17
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | within a few hours
 host_response_rate                           | 1
 host_acceptance_rate                         | 0
 host_is_superhost                            | f
 host_neighbourhood                           | Downtown Eastside
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['email', 'phone', 'reviews', 'offline_government_id', 'self
ie', 'government_id', 'identity_manual']
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Downtown Eastside
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27782
 longitude                                    | -123.08985
 property_type                                | Entire guest suite
 room_type                                    | Entire home/apt
 accommodates                                 | 2
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Hot w
ater', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated wor
kspace', 'Iron', 'Free street parking', 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Hangers', 'Cable TV',
'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning']
 price                                        | 100
 minimum_nights                               | 30
 maximum_nights                               | 365
 minimum_minimum_nights                       | 30
 maximum_minimum_nights                       | 30
 minimum_maximum_nights                       | 365
```

```
 maximum_maximum_nights                        | 365
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 365.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 0
 availability_365                              | 224
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 verifications_array                           | [email, phone, reviews, offline_government_id, selfie, gover
nment_id, identity_manual]
-RECORD 14-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                            | 200643
 host_since                                    | 2011/08/16
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | within an hour
 host_response_rate                            | 1
 host_acceptance_rate                          | 1
 host_is_superhost                             | f
 host_neighbourhood                            | Strathcona
 host_listings_count                           | 4
 host_total_listings_count                     | 4
 host_verifications                            | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                        | Downtown Eastside
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.27384
 longitude                                     | -123.08289
 property_type                                 | Private room in house
 room_type                                     | Private room
 accommodates                                  | 4
 bathrooms                                     | null
 bathrooms_text                                | 2 shared baths
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi',
```

```
'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Garden or backyard', 'Dishwasher', 'Extra pillows a
nd blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Room-darkening shades', 'Smoke alarm', 'Heatin
g', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Free street parking', 'Coffee maker', 'Be
d linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Essentials',
'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
 price                                       | 87
 minimum_nights                              | 1
 maximum_nights                              | 1125
 minimum_minimum_nights                      | 1
 maximum_minimum_nights                      | 1
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 1.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 30
 availability_60                             | 60
 availability_90                             | 90
 availability_365                            | 365
 instant_bookable                            | t
 calculated_host_listings_count              | 4
 calculated_host_listings_count_entire_homes | 0
 calculated_host_listings_count_private_rooms | 4
 calculated_host_listings_count_shared_rooms | 0
 verifications_array                         | [email, phone, reviews, jumio, government_id]
-RECORD 15-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
 id                                          | 248014
 host_since                                  | 2011/10/16
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | within a day
 host_response_rate                          | 0.5
 host_acceptance_rate                        | N/A
 host_is_superhost                           | f
 host_neighbourhood                          | Grandview-Woodland
 host_listings_count                         | 1
 host_total_listings_count                   | 1
 host_verifications                          | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                      | Grandview-Woodland
 neighbourhood_group_cleansed                | null
```

```
 latitude                                     | 49.29103
 longitude                                    | -123.05782
 property_type                                | Entire house
 room_type                                    | Entire home/apt
 accommodates                                 | 3
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 2
 beds                                         | 0
 amenities                                    | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi',
'Host greets you', 'Dryer', 'Stove', 'Carbon monoxide alarm', 'Hot water', 'Garden or backyard', 'Dishwashe
r', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iro
n', 'Free street parking', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Cable TV', 'Re
frigerator', 'Hangers', 'Essentials', 'Washer', 'Long term stays allowed', 'BBQ grill', 'TV', 'Cleaning befo
re checkout']
 price                                        | 100
 minimum_nights                               | 31
 maximum_nights                               | 180
 minimum_minimum_nights                       | 31
 maximum_minimum_nights                       | 31
 minimum_maximum_nights                       | 180
 maximum_maximum_nights                       | 180
 minimum_nights_avg_ntm                       | 31.0
 maximum_nights_avg_ntm                       | 180.0
 has_availability                             | t
 availability_30                              | 4
 availability_60                              | 4
 availability_90                              | 4
 availability_365                             | 113
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 verifications_array                          | [email, phone, reviews, jumio, government_id]
-RECORD 16-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------
 id                                           | 258636
 host_since                                   | 2011/08/09
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | N/A
 host_response_rate                           | N/A
 host_acceptance_rate                         | 1
```

```
 host_is_superhost                              | f
 host_neighbourhood                             | Downtown Vancouver
 host_listings_count                            | 1
 host_total_listings_count                      | 1
 host_verifications                             | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | null
 neighbourhood_cleansed                         | Downtown
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.27796
 longitude                                      | -123.12671
 property_type                                  | Entire apartment
 room_type                                      | Entire home/apt
 accommodates                                   | 4
 bathrooms                                      | null
 bathrooms_text                                 | 1 bath
 bedrooms                                       | 1
 beds                                           | 2
 amenities                                      | ['Kitchen', 'Iron', 'Elevator', 'Gym', 'Hangers', 'Smoke ala
rm', 'Wifi', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Free parking on premises', 'Long term stays allowe
d', 'Hair dryer', 'Shampoo', 'Dedicated workspace']
 price                                          | 99
 minimum_nights                                 | 120
 maximum_nights                                 | 365
 minimum_minimum_nights                         | 120
 maximum_minimum_nights                         | 120
 minimum_maximum_nights                         | 365
 maximum_maximum_nights                         | 365
 minimum_nights_avg_ntm                         | 120.0
 maximum_nights_avg_ntm                         | 365.0
 has_availability                               | t
 availability_30                                | 0
 availability_60                                | 11
 availability_90                                | 41
 availability_365                               | 316
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
 verifications_array                            | [email, phone, reviews, jumio, government_id]
-RECORD 17-------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
```

```
-----------------------------------------------------------------------------------------------
 id                                          | 283841
 host_since                                  | 2009/08/10
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | within a day
 host_response_rate                          | 0.93
 host_acceptance_rate                        | 0.88
 host_is_superhost                           | f
 host_neighbourhood                          | Coal Harbour
 host_listings_count                         | 49
 host_total_listings_count                   | 49
 host_verifications                          | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | null
 neighbourhood_cleansed                      | Downtown
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.2872
 longitude                                   | -123.12352
 property_type                               | Entire apartment
 room_type                                   | Entire home/apt
 accommodates                                | 6
 bathrooms                                   | null
 bathrooms_text                              | 3 baths
 bedrooms                                    | 3
 beds                                        | 3
 amenities                                   | ['Kitchen', 'Hot tub', 'Oven', 'Pool', 'Gym', 'Indoor firepl
ace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Dishwasher', 'Smoke alar
m', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Coffee maker', 'Dis
hes and silverware', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowe
d', 'TV', 'Air conditioning']
 price                                       | 197
 minimum_nights                              | 90
 maximum_nights                              | 1125
 minimum_minimum_nights                      | 90
 maximum_minimum_nights                      | 90
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 90.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 0
 availability_90                             | 0
 availability_365                            | 0
 instant_bookable                            | f
 calculated_host_listings_count              | 41
 calculated_host_listings_count_entire_homes | 41
 calculated_host_listings_count_private_rooms | 0
```

```
 calculated_host_listings_count_shared_rooms  | 0
 verifications_array                          | [email, phone, reviews, jumio, government_id]
-RECORD 18-------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
 id                                           | 323090
 host_since                                   | 2012/01/01
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | N/A
 host_response_rate                           | N/A
 host_acceptance_rate                         | 0.85
 host_is_superhost                            | t
 host_neighbourhood                           | Downtown Vancouver
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Downtown
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.28206
 longitude                                    | -123.12461
 property_type                                | Entire apartment
 room_type                                    | Entire home/apt
 accommodates                                 | 2
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Cooking basics', 'Iron', 'Elevator', 'Gym', 'Ha
ngers', 'Cable TV', 'Smoke alarm', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Carbon monoxide alarm', 'Wif
i', 'TV', 'Long term stays allowed', 'Hair dryer', 'Shampoo', 'Dishes and silverware', 'Dedicated workspac
e']
 price                                        | 168
 minimum_nights                               | 180
 maximum_nights                               | 180
 minimum_minimum_nights                       | 180
 maximum_minimum_nights                       | 180
 minimum_maximum_nights                       | 180
 maximum_maximum_nights                       | 180
 minimum_nights_avg_ntm                       | 180.0
 maximum_nights_avg_ntm                       | 180.0
 has_availability                             | t
```

```
 availability_30                              | 30
 availability_60                              | 60
 availability_90                              | 90
 availability_365                             | 141
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 verifications_array                          | [email, phone, reviews, jumio, government_id]
-RECORD 19-----------------------
```

```
 minimum_nights                                | 30
 maximum_nights                                | 1125
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 29
 availability_60                               | 59
 availability_90                               | 89
 availability_365                              | 364
 instant_bookable                              | f
 calculated_host_listings_count                | 7
 calculated_host_listings_count_entire_homes   | 3
 calculated_host_listings_count_private_rooms  | 4
 calculated_host_listings_count_shared_rooms   | 0
 verifications_array                           | [email, phone, facebook, reviews]
only showing top 20 rows
```

Now we can use a vectorizer to transform the arrays of strings to arrays of indexes.

In [28]:

```
cv_amenities = CountVectorizer(inputCol="amenities_array", outputCol="amenities_vector",
                               vocabSize=20, minDF=2.0)
fitted_amenities = cv_amenities.fit(amenities_splitter.transform(dataset))
stages.append(fitted_amenities)

cv_verifications = CountVectorizer(inputCol="verifications_array", outputCol="verifications_vector",
                                   vocabSize=5, minDF=2.0)
fitted_verifications = cv_verifications.fit(verifications_splitter.transform(dataset))
stages.append(fitted_verifications)
```

We keep the number of amenities and verifications under control by setting the vocabulary size.

In [29]:

```
print(len(fitted_amenities.vocabulary))
fitted_amenities.vocabulary
```

```
20
```

Out[29]:

```
['Wifi',
 'Heating',
 'Smoke alarm',
 'Essentials',
 'Kitchen',
```

```
 'Hangers',
 'Long term stays allowed',
 'Washer',
 'Dryer',
 'Hair dryer',
 'TV',
 'Dedicated workspace',
 'Hot water',
 'Shampoo',
 'Iron',
 'Carbon monoxide alarm',
 'Dishes and silverware',
 'Fire extinguisher',
 'Refrigerator',
 'Cooking basics']
```

We can now build a pipeline to see the vectorizers in place.

In [30]:

```
pipeline = Pipeline(stages=stages)
pipeline.fit(dataset).transform(dataset).show(vertical=True)
```

```
-RECORD 0-----------------------------------------------------------
 id                                          | 13358
 host_since                                  | 2009/11/07
 host_location                               | Vancouver, Britis...
 host_response_time                          | within a day
 host_response_rate                          | 0.71
 host_acceptance_rate                        | 0.71
 host_is_superhost                           | t
 host_neighbourhood                          | Downtown Vancouver
 host_listings_count                         | 1
 host_total_listings_count                   | 1
 host_verifications                          | ['email', 'phone'...
 host_has_profile_pic                        | f
 host_identity_verified                      | t
 neighbourhood                               | null
 neighbourhood_cleansed                      | West End
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.28209
 longitude                                   | -123.12515
 property_type                               | Entire condominium
 room_type                                   | Entire home/apt
 accommodates                                | 2
 bathrooms                                   | null
 bathrooms_text                              | 1 bath
 bedrooms                                    | 1
 beds                                        | 1
 amenities                                   | ['Kitchen', 'Oven...
```

```
 price                                          | 90
 minimum_nights                                 | 30
 maximum_nights                                 | 60
 minimum_minimum_nights                         | 30
 maximum_minimum_nights                         | 30
 minimum_maximum_nights                         | 60
 maximum_maximum_nights                         | 60
 minimum_nights_avg_ntm                         | 30.0
 maximum_nights_avg_ntm                         | 60.0
 has_availability                               | t
 availability_30                                | 11
 availability_60                                | 41
 availability_90                                | 70
 availability_365                               | 70
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
 neighbourhood_reduced                          | West End
 property_reduced                               | Entire condominium
 room_reduced                                   | Entire home/apt
 neighbourhood_indexed                          | 3.0
 property_indexed                               | 1.0
 room_indexed                                   | 0.0
 amenities_array                                | [Kitchen, Oven, G...
 verifications_array                            | [email, phone, fa...
 amenities_vector                               | (20,[0,1,2,3,4,5,...
 verifications_vector                           | (5,[0,1,2,3,4],[1...
-RECORD 1-----------------------------------------------------------
 id                                             | 16254
 host_since                                     | 2009/12/15
 host_location                                  | Vancouver, Britis...
 host_response_time                             | N/A
 host_response_rate                             | N/A
 host_acceptance_rate                           | N/A
 host_is_superhost                              | f
 host_neighbourhood                             | Hastings-Sunrise
 host_listings_count                            | 1
 host_total_listings_count                      | 1
 host_verifications                             | ['phone', 'facebo...
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | Vancouver, Britis...
 neighbourhood_cleansed                         | Hastings-Sunrise
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.27734
 longitude                                      | -123.04152
 property_type                                  | Entire guest suite
```

```
 room_type                                     | Entire home/apt
 accommodates                                  | 4
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 2
 beds                                          | 3
 amenities                                     | ['Kitchen', 'Oven...
 price                                         | 195
 minimum_nights                                | 3
 maximum_nights                                | 31
 minimum_minimum_nights                        | 3
 maximum_minimum_nights                        | 3
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 3.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 0
 availability_365                              | 228
 instant_bookable                              | t
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Hastings-Sunrise
 property_reduced                              | Entire guest suite
 room_reduced                                  | Entire home/apt
 neighbourhood_indexed                         | 9.0
 property_indexed                              | 4.0
 room_indexed                                  | 0.0
 amenities_array                               | [Kitchen, Oven, M...
 verifications_array                           | [phone, facebook,...
 amenities_vector                              | (20,[0,2,3,4,5,6,...
 verifications_vector                          | (5,[0,2,3],[1.0,1...
-RECORD 2-----------------------------------------------------------
 id                                            | 17765
 host_since                                    | 2010/01/07
 host_location                                 | Vancouver, Britis...
 host_response_time                            | N/A
 host_response_rate                            | N/A
 host_acceptance_rate                          | 1
 host_is_superhost                             | f
 host_neighbourhood                            | Mount Pleasant
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['phone', 'facebo...
 host_has_profile_pic                          | t
```

```
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, Britis...
 neighbourhood_cleansed                       | Mount Pleasant
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.26181
 longitude                                    | -123.10905
 property_type                                | Entire apartment
 room_type                                    | Entire home/apt
 accommodates                                 | 2
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Wifi...
 price                                        | 94
 minimum_nights                               | 3
 maximum_nights                               | 60
 minimum_minimum_nights                       | 3
 maximum_minimum_nights                       | 3
 minimum_maximum_nights                       | 60
 maximum_maximum_nights                       | 60
 minimum_nights_avg_ntm                       | 3.0
 maximum_nights_avg_ntm                       | 60.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 0
 availability_365                             | 0
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Mount Pleasant
 property_reduced                             | Entire apartment
 room_reduced                                 | Entire home/apt
 neighbourhood_indexed                        | 5.0
 property_indexed                             | 0.0
 room_indexed                                 | 0.0
 amenities_array                              | [Kitchen, Wifi, H...
 verifications_array                          | [phone, facebook,...
 amenities_vector                             | (20,[0,1,2,3,4,5,...
 verifications_vector                         | (5,[0,2,3,4],[1.0...
-RECORD 3-------------------------------------------------------------
 id                                           | 18795
 host_since                                   | 2010/01/13
 host_location                                | Vancouver, Britis...
 host_response_time                           | within a day
 host_response_rate                           | 0.7
```

```
host_acceptance_rate                          | 0.4
host_is_superhost                             | f
host_neighbourhood                            | West End
host_listings_count                           | 3
host_total_listings_count                     | 3
host_verifications                            | ['email', 'phone'...
host_has_profile_pic                          | t
host_identity_verified                        | t
neighbourhood                                 | null
neighbourhood_cleansed                        | West End
neighbourhood_group_cleansed                  | null
latitude                                      | 49.28251
longitude                                     | -123.12827
property_type                                 | Private room in h...
room_type                                     | Private room
accommodates                                  | 2
bathrooms                                     | null
bathrooms_text                                | 1 private bath
bedrooms                                      | 1
beds                                          | 1
amenities                                     | ['Kitchen', 'Cabl...
price                                         | 50
minimum_nights                                | 30
maximum_nights                                | 360
minimum_minimum_nights                        | 30
maximum_minimum_nights                        | 30
minimum_maximum_nights                        | 360
maximum_maximum_nights                        | 360
minimum_nights_avg_ntm                        | 30.0
maximum_nights_avg_ntm                        | 360.0
has_availability                              | t
availability_30                               | 9
availability_60                               | 20
availability_90                               | 50
availability_365                              | 325
instant_bookable                              | f
calculated_host_listings_count                | 3
calculated_host_listings_count_entire_homes   | 0
calculated_host_listings_count_private_rooms  | 3
calculated_host_listings_count_shared_rooms   | 0
neighbourhood_reduced                         | West End
property_reduced                              | Private room in h...
room_reduced                                  | Private room
neighbourhood_indexed                         | 3.0
property_indexed                              | 3.0
room_indexed                                  | 1.0
amenities_array                               | [Kitchen, Cable T...
verifications_array                           | [email, phone, re...
amenities_vector                              | (20,[0,1,3,4,6,7,...
```

```
 verifications_vector                        | (5,[0,1,2,4],[1.0...
-RECORD 4-------------------------------------------------------------
 id                                          | 21913
 host_since                                  | 2010/02/22
 host_location                               | Vancouver, Britis...
 host_response_time                          | within an hour
 host_response_rate                          | 1
 host_acceptance_rate                        | 1
 host_is_superhost                           | t
 host_neighbourhood                          | Renfrew-Collingwood
 host_listings_count                         | 3
 host_total_listings_count                   | 3
 host_verifications                          | ['email', 'phone'...
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, Britis...
 neighbourhood_cleansed                      | Renfrew-Collingwood
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.23427
 longitude                                   | -123.05309
 property_type                               | Entire house
 room_type                                   | Entire home/apt
 accommodates                                | 2
 bathrooms                                   | null
 bathrooms_text                              | 1 bath
 bedrooms                                    | 1
 beds                                        | 1
 amenities                                   | ['Kitchen', 'Oven...
 price                                       | 80
 minimum_nights                              | 28
 maximum_nights                              | 400
 minimum_minimum_nights                      | 28
 maximum_minimum_nights                      | 28
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 28.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 2
 availability_60                             | 13
 availability_90                             | 43
 availability_365                            | 318
 instant_bookable                            | f
 calculated_host_listings_count              | 3
 calculated_host_listings_count_entire_homes | 1
 calculated_host_listings_count_private_rooms | 2
 calculated_host_listings_count_shared_rooms | 0
 neighbourhood_reduced                       | Renfrew-Collingwood
 property_reduced                            | Entire house
```

```
 room_reduced                                   | Entire home/apt
 neighbourhood_indexed                          | 11.0
 property_indexed                               | 2.0
 room_indexed                                   | 0.0
 amenities_array                                | [Kitchen, Oven, M...
 verifications_array                            | [email, phone, re...
 amenities_vector                               | (20,[0,1,2,3,4,5,...
 verifications_vector                           | (5,[0,1,2,3,4],[1...
-RECORD 5-----------------------------------------------------------
 id                                             | 24273
 host_since                                     | 2011/10/09
 host_location                                  | Vancouver, Britis...
 host_response_time                             | within an hour
 host_response_rate                             | 1
 host_acceptance_rate                           | 0.82
 host_is_superhost                              | f
 host_neighbourhood                             | Kitsilano
 host_listings_count                            | 2
 host_total_listings_count                      | 2
 host_verifications                             | ['email', 'phone'...
 host_has_profile_pic                           | f
 host_identity_verified                         | t
 neighbourhood                                  | Vancouver, Britis...
 neighbourhood_cleansed                         | Kitsilano
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.2626
 longitude                                      | -123.17308
 property_type                                  | Entire apartment
 room_type                                      | Entire home/apt
 accommodates                                   | 6
 bathrooms                                      | null
 bathrooms_text                                 | 1 bath
 bedrooms                                       | 2
 beds                                           | 3
 amenities                                      | ['Kitchen', 'Oven...
 price                                          | 200
 minimum_nights                                 | 30
 maximum_nights                                 | 365
 minimum_minimum_nights                         | 30
 maximum_minimum_nights                         | 30
 minimum_maximum_nights                         | 365
 maximum_maximum_nights                         | 365
 minimum_nights_avg_ntm                         | 30.0
 maximum_nights_avg_ntm                         | 365.0
 has_availability                               | t
 availability_30                                | 0
 availability_60                                | 0
 availability_90                                | 10
 availability_365                               | 223
```

```
 instant_bookable                              | f
 calculated_host_listings_count                | 2
 calculated_host_listings_count_entire_homes   | 2
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Kitsilano
 property_reduced                              | Entire apartment
 room_reduced                                  | Entire home/apt
 neighbourhood_indexed                         | 2.0
 property_indexed                              | 0.0
 room_indexed                                  | 0.0
 amenities_array                               | [Kitchen, Oven, I...
 verifications_array                           | [email, phone, re...
 amenities_vector                              | (20,[0,1,2,3,4,5,...
 verifications_vector                          | (5,[0,1,2,3,4],[1...
-RECORD 6-----------------------------------------------------------
 id                                            | 42645
 host_since                                    | 2010/04/17
 host_location                                 | Vancouver, Britis...
 host_response_time                            | within an hour
 host_response_rate                            | 1
 host_acceptance_rate                          | 1
 host_is_superhost                             | t
 host_neighbourhood                            | Downtown Vancouver
 host_listings_count                           | 4
 host_total_listings_count                     | 4
 host_verifications                            | ['email', 'phone'...
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, Britis...
 neighbourhood_cleansed                        | Downtown
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.28001
 longitude                                     | -123.12475
 property_type                                 | Entire condominium
 room_type                                     | Entire home/apt
 accommodates                                  | 3
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 2
 amenities                                     | ['Kitchen', 'Oven...
 price                                         | 159
 minimum_nights                                | 32
 maximum_nights                                | 365
 minimum_minimum_nights                        | 32
 maximum_minimum_nights                        | 32
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
```

```
 minimum_nights_avg_ntm                       | 32.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 29
 availability_90                              | 59
 availability_365                             | 333
 instant_bookable                             | t
 calculated_host_listings_count               | 4
 calculated_host_listings_count_entire_homes  | 4
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Downtown
 property_reduced                             | Entire condominium
 room_reduced                                 | Entire home/apt
 neighbourhood_indexed                        | 0.0
 property_indexed                             | 1.0
 room_indexed                                 | 0.0
 amenities_array                              | [Kitchen, Oven, P...
 verifications_array                          | [email, phone, fa...
 amenities_vector                             | (20,[0,1,2,3,4,5,...
 verifications_vector                         | (5,[0,1,4],[1.0,1...
-RECORD 7-----------------------------------------------------------
 id                                           | 54261
 host_since                                   | 2010/10/05
 host_location                                | Vancouver, Britis...
 host_response_time                           | a few days or more
 host_response_rate                           | 0
 host_acceptance_rate                         | N/A
 host_is_superhost                            | f
 host_neighbourhood                           | Downtown Vancouver
 host_listings_count                          | 2
 host_total_listings_count                    | 2
 host_verifications                           | ['email', 'phone'...
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | null
 neighbourhood_cleansed                       | Downtown
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27655
 longitude                                    | -123.12485
 property_type                                | Entire condominium
 room_type                                    | Entire home/apt
 accommodates                                 | 4
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Hot ...
```

```
 price                                          | 126
 minimum_nights                                 | 90
 maximum_nights                                 | 365
 minimum_minimum_nights                         | 90
 maximum_minimum_nights                         | 90
 minimum_maximum_nights                         | 365
 maximum_maximum_nights                         | 365
 minimum_nights_avg_ntm                         | 90.0
 maximum_nights_avg_ntm                         | 365.0
 has_availability                               | t
 availability_30                                | 30
 availability_60                                | 60
 availability_90                                | 90
 availability_365                               | 365
 instant_bookable                               | f
 calculated_host_listings_count                 | 2
 calculated_host_listings_count_entire_homes    | 2
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
 neighbourhood_reduced                          | Downtown
 property_reduced                               | Entire condominium
 room_reduced                                   | Entire home/apt
 neighbourhood_indexed                          | 0.0
 property_indexed                               | 1.0
 room_indexed                                   | 0.0
 amenities_array                                | [Kitchen, Hot tub...
 verifications_array                            | [email, phone, re...
 amenities_vector                               | (20,[0,1,2,3,4,5,...
 verifications_vector                           | (5,[0,1,2,3,4],[1...
-RECORD 8-----------------------------------------------------------
 id                                             | 72659
 host_since                                     | 2011/02/03
 host_location                                  | Vancouver, Britis...
 host_response_time                             | N/A
 host_response_rate                             | N/A
 host_acceptance_rate                           | 0
 host_is_superhost                              | f
 host_neighbourhood                             | Commercial Drive
 host_listings_count                            | 1
 host_total_listings_count                      | 1
 host_verifications                             | ['email', 'phone'...
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | null
 neighbourhood_cleansed                         | Grandview-Woodland
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.27357
 longitude                                      | -123.07024
 property_type                                  | Entire condominium
```

```
 room_type                                     | Entire home/apt
 accommodates                                  | 4
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 2
 beds                                          | 2
 amenities                                     | ['Kitchen', 'Lake...
 price                                         | 210
 minimum_nights                                | 7
 maximum_nights                                | 21
 minimum_minimum_nights                        | 7
 maximum_minimum_nights                        | 7
 minimum_maximum_nights                        | 21
 maximum_maximum_nights                        | 21
 minimum_nights_avg_ntm                        | 7.0
 maximum_nights_avg_ntm                        | 21.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 0
 availability_365                              | 40
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Grandview-Woodland
 property_reduced                              | Entire condominium
 room_reduced                                  | Entire home/apt
 neighbourhood_indexed                         | 8.0
 property_indexed                              | 1.0
 room_indexed                                  | 0.0
 amenities_array                               | [Kitchen, Lake ac...
 verifications_array                           | [email, phone, re...
 amenities_vector                              | (20,[0,1,2,3,4,5,...
 verifications_vector                          | (5,[0,1,2,4],[1.0...
-RECORD 9-------------------------------------------------------------
 id                                            | 77157
 host_since                                    | 2011/02/27
 host_location                                 | Vancouver, Britis...
 host_response_time                            | within a few hours
 host_response_rate                            | 1
 host_acceptance_rate                          | 1
 host_is_superhost                             | f
 host_neighbourhood                            | Arbutus Ridge
 host_listings_count                           | 2
 host_total_listings_count                     | 2
 host_verifications                            | ['email', 'phone'...
 host_has_profile_pic                          | t
```

```
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, Britis...
 neighbourhood_cleansed                       | Arbutus Ridge
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.24675
 longitude                                    | -123.16711
 property_type                                | Private room in g...
 room_type                                    | Private room
 accommodates                                 | 3
 bathrooms                                    | null
 bathrooms_text                               | 1 private bath
 bedrooms                                     | 1
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Micr...
 price                                        | 75
 minimum_nights                               | 3
 maximum_nights                               | 60
 minimum_minimum_nights                       | 3
 maximum_minimum_nights                       | 3
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
 minimum_nights_avg_ntm                       | 3.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 3
 availability_60                              | 3
 availability_90                              | 3
 availability_365                             | 267
 instant_bookable                             | t
 calculated_host_listings_count               | 2
 calculated_host_listings_count_entire_homes  | 0
 calculated_host_listings_count_private_rooms | 2
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | OTHER
 property_reduced                             | OTHER
 room_reduced                                 | Private room
 neighbourhood_indexed                        | 1.0
 property_indexed                             | 5.0
 room_indexed                                 | 1.0
 amenities_array                              | [Kitchen, Microwa...
 verifications_array                          | [email, phone, fa...
 amenities_vector                             | (20,[0,1,2,3,4,5,...
 verifications_vector                         | (5,[0,1,2,3,4],[1...
-RECORD 10-------------------------------------------------------------
 id                                           | 111077
 host_since                                   | 2011/05/09
 host_location                                | North Vancouver, ...
 host_response_time                           | within a few hours
 host_response_rate                           | 1
```

```
host_acceptance_rate                          | 0.33
host_is_superhost                             | f
host_neighbourhood                            | Grandview-Woodland
host_listings_count                           | 1
host_total_listings_count                     | 1
host_verifications                            | ['email', 'phone'...
host_has_profile_pic                          | t
host_identity_verified                        | t
neighbourhood                                 | Vancouver, Britis...
neighbourhood_cleansed                        | Grandview-Woodland
neighbourhood_group_cleansed                  | null
latitude                                      | 49.26726
longitude                                     | -123.05969
property_type                                 | Entire guest suite
room_type                                     | Entire home/apt
accommodates                                  | 3
bathrooms                                     | null
bathrooms_text                                | 1 bath
bedrooms                                      | 2
beds                                          | 2
amenities                                     | ['Kitchen', 'Indo...
price                                         | 125
minimum_nights                                | 3
maximum_nights                                | 365
minimum_minimum_nights                        | 3
maximum_minimum_nights                        | 3
minimum_maximum_nights                        | 365
maximum_maximum_nights                        | 365
minimum_nights_avg_ntm                        | 3.0
maximum_nights_avg_ntm                        | 365.0
has_availability                              | t
availability_30                               | 4
availability_60                               | 34
availability_90                               | 64
availability_365                              | 339
instant_bookable                              | f
calculated_host_listings_count                | 1
calculated_host_listings_count_entire_homes   | 1
calculated_host_listings_count_private_rooms  | 0
calculated_host_listings_count_shared_rooms   | 0
neighbourhood_reduced                         | Grandview-Woodland
property_reduced                              | Entire guest suite
room_reduced                                  | Entire home/apt
neighbourhood_indexed                         | 8.0
property_indexed                              | 4.0
room_indexed                                  | 0.0
amenities_array                               | [Kitchen, Indoor ...
verifications_array                           | [email, phone, fa...
amenities_vector                              | (20,[0,1,2,3,4,5,...
```

```
 verifications_vector                        | (5,[0,1,2,4],[1.0...
-RECORD 11-------------------------------------------------------------
 id                                          | 125176
 host_since                                  | 2011/05/24
 host_location                               | Vancouver, Britis...
 host_response_time                          | within an hour
 host_response_rate                          | 1
 host_acceptance_rate                        | 1
 host_is_superhost                           | f
 host_neighbourhood                          | Riley Park
 host_listings_count                         | 1
 host_total_listings_count                   | 1
 host_verifications                          | ['email', 'phone'...
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, Britis...
 neighbourhood_cleansed                      | Riley Park
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.24648
 longitude                                   | -123.09241
 property_type                               | Private room in h...
 room_type                                   | Private room
 accommodates                                | 3
 bathrooms                                   | null
 bathrooms_text                              | 1 private bath
 bedrooms                                    | 1
 beds                                        | 0
 amenities                                   | ['Kitchen', 'Hot ...
 price                                       | 75
 minimum_nights                              | 2
 maximum_nights                              | 60
 minimum_minimum_nights                      | 2
 maximum_minimum_nights                      | 2
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 2.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 0
 availability_90                             | 2
 availability_365                            | 268
 instant_bookable                            | t
 calculated_host_listings_count              | 1
 calculated_host_listings_count_entire_homes | 0
 calculated_host_listings_count_private_rooms | 1
 calculated_host_listings_count_shared_rooms | 0
 neighbourhood_reduced                       | Riley Park
 property_reduced                            | Private room in h...
```

```
 room_reduced                                   | Private room
 neighbourhood_indexed                          | 7.0
 property_indexed                               | 3.0
 room_indexed                                   | 1.0
 amenities_array                                | [Kitchen, Hot tub...
 verifications_array                            | [email, phone, fa...
 amenities_vector                               | (20,[0,1,2,3,4,5,...
 verifications_vector                           | (5,[0,1,2,3,4],[1...
-RECORD 12-------------------------------------------------------------
 id                                             | 129163
 host_since                                     | 2011/05/17
 host_location                                  | Vancouver, Britis...
 host_response_time                             | within a few hours
 host_response_rate                             | 1
 host_acceptance_rate                           | 1
 host_is_superhost                              | f
 host_neighbourhood                             | Renfrew-Collingwood
 host_listings_count                            | 2
 host_total_listings_count                      | 2
 host_verifications                             | ['email', 'phone'...
 host_has_profile_pic                           | t
 host_identity_verified                         | t
 neighbourhood                                  | null
 neighbourhood_cleansed                         | Renfrew-Collingwood
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.24507
 longitude                                      | -123.04941
 property_type                                  | Private room in h...
 room_type                                      | Private room
 accommodates                                   | 1
 bathrooms                                      | null
 bathrooms_text                                 | 1 shared bath
 bedrooms                                       | 1
 beds                                           | 1
 amenities                                      | ['Kitchen', 'Hot ...
 price                                          | 45
 minimum_nights                                 | 30
 maximum_nights                                 | 280
 minimum_minimum_nights                         | 30
 maximum_minimum_nights                         | 30
 minimum_maximum_nights                         | 280
 maximum_maximum_nights                         | 280
 minimum_nights_avg_ntm                         | 30.0
 maximum_nights_avg_ntm                         | 280.0
 has_availability                               | t
 availability_30                                | 30
 availability_60                                | 60
 availability_90                                | 90
 availability_365                               | 365
```

```
 instant_bookable                              | t
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms  | 1
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Renfrew-Collingwood
 property_reduced                              | Private room in h...
 room_reduced                                  | Private room
 neighbourhood_indexed                         | 11.0
 property_indexed                              | 3.0
 room_indexed                                  | 1.0
 amenities_array                               | [Kitchen, Hot wat...
 verifications_array                           | [email, phone, re...
 amenities_vector                              | (20,[0,1,2,3,4,6,...
 verifications_vector                          | (5,[0,1,2,3,4],[1...
-RECORD 13-------------------------------------------------------------
 id                                            | 174135
 host_since                                    | 2011/07/17
 host_location                                 | Vancouver, Britis...
 host_response_time                            | within a few hours
 host_response_rate                            | 1
 host_acceptance_rate                          | 0
 host_is_superhost                             | f
 host_neighbourhood                            | Downtown Eastside
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone'...
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, Britis...
 neighbourhood_cleansed                        | Downtown Eastside
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.27782
 longitude                                     | -123.08985
 property_type                                 | Entire guest suite
 room_type                                     | Entire home/apt
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 2
 amenities                                     | ['Kitchen', 'Wifi...
 price                                         | 100
 minimum_nights                                | 30
 maximum_nights                                | 365
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 365
 maximum_maximum_nights                        | 365
```

```
 minimum_nights_avg_ntm                       | 30.0
 maximum_nights_avg_ntm                       | 365.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 0
 availability_365                             | 224
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Downtown Eastside
 property_reduced                             | Entire guest suite
 room_reduced                                 | Entire home/apt
 neighbourhood_indexed                        | 6.0
 property_indexed                             | 4.0
 room_indexed                                 | 0.0
 amenities_array                              | [Kitchen, Wifi, D...
 verifications_array                          | [email, phone, re...
 amenities_vector                             | (20,[0,1,2,3,4,5,...
 verifications_vector                         | (5,[0,1,2,3,4],[1...
-RECORD 14-------------------------------------------------------------
 id                                           | 200643
 host_since                                   | 2011/08/16
 host_location                                | Vancouver, Britis...
 host_response_time                           | within an hour
 host_response_rate                           | 1
 host_acceptance_rate                         | 1
 host_is_superhost                            | f
 host_neighbourhood                           | Strathcona
 host_listings_count                          | 4
 host_total_listings_count                    | 4
 host_verifications                           | ['email', 'phone'...
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, Britis...
 neighbourhood_cleansed                       | Downtown Eastside
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27384
 longitude                                    | -123.08289
 property_type                                | Private room in h...
 room_type                                    | Private room
 accommodates                                 | 4
 bathrooms                                    | null
 bathrooms_text                               | 2 shared baths
 bedrooms                                     | 1
 beds                                         | 1
 amenities                                    | ['Kitchen', 'Oven...
```

```
 price                                        | 87
 minimum_nights                               | 1
 maximum_nights                               | 1125
 minimum_minimum_nights                       | 1
 maximum_minimum_nights                       | 1
 minimum_maximum_nights                       | 1125
 maximum_maximum_nights                       | 1125
 minimum_nights_avg_ntm                       | 1.0
 maximum_nights_avg_ntm                       | 1125.0
 has_availability                             | t
 availability_30                              | 30
 availability_60                              | 60
 availability_90                              | 90
 availability_365                             | 365
 instant_bookable                             | t
 calculated_host_listings_count               | 4
 calculated_host_listings_count_entire_homes  | 0
 calculated_host_listings_count_private_rooms | 4
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Downtown Eastside
 property_reduced                             | Private room in h...
 room_reduced                                 | Private room
 neighbourhood_indexed                        | 6.0
 property_indexed                             | 3.0
 room_indexed                                 | 1.0
 amenities_array                              | [Kitchen, Oven, I...
 verifications_array                          | [email, phone, re...
 amenities_vector                             | (20,[0,1,2,3,4,5,...
 verifications_vector                         | (5,[0,1,2,4],[1.0...
-RECORD 15-------------------------------------------------------------
 id                                           | 248014
 host_since                                   | 2011/10/16
 host_location                                | Vancouver, Britis...
 host_response_time                           | within a day
 host_response_rate                           | 0.5
 host_acceptance_rate                         | N/A
 host_is_superhost                            | f
 host_neighbourhood                           | Grandview-Woodland
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['email', 'phone'...
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, Britis...
 neighbourhood_cleansed                       | Grandview-Woodland
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.29103
 longitude                                    | -123.05782
 property_type                                | Entire house
```

```
 room_type                                     | Entire home/apt
 accommodates                                  | 3
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 2
 beds                                          | 0
 amenities                                     | ['Kitchen', 'Oven...
 price                                         | 100
 minimum_nights                                | 31
 maximum_nights                                | 180
 minimum_minimum_nights                        | 31
 maximum_minimum_nights                        | 31
 minimum_maximum_nights                        | 180
 maximum_maximum_nights                        | 180
 minimum_nights_avg_ntm                        | 31.0
 maximum_nights_avg_ntm                        | 180.0
 has_availability                              | t
 availability_30                               | 4
 availability_60                               | 4
 availability_90                               | 4
 availability_365                              | 113
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Grandview-Woodland
 property_reduced                              | Entire house
 room_reduced                                  | Entire home/apt
 neighbourhood_indexed                         | 8.0
 property_indexed                              | 2.0
 room_indexed                                  | 0.0
 amenities_array                               | [Kitchen, Oven, I...
 verifications_array                           | [email, phone, re...
 amenities_vector                              | (20,[0,1,2,3,4,5,...
 verifications_vector                          | (5,[0,1,2,4],[1.0...
-RECORD 16-----------------------------------------------------------
 id                                            | 258636
 host_since                                    | 2011/08/09
 host_location                                 | Vancouver, Britis...
 host_response_time                            | N/A
 host_response_rate                            | N/A
 host_acceptance_rate                          | 1
 host_is_superhost                             | f
 host_neighbourhood                            | Downtown Vancouver
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone'...
 host_has_profile_pic                          | t
```

```
 host_identity_verified                       | t
 neighbourhood                                | null
 neighbourhood_cleansed                       | Downtown
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27796
 longitude                                    | -123.12671
 property_type                                | Entire apartment
 room_type                                    | Entire home/apt
 accommodates                                 | 4
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Iron...
 price                                        | 99
 minimum_nights                               | 120
 maximum_nights                               | 365
 minimum_minimum_nights                       | 120
 maximum_minimum_nights                       | 120
 minimum_maximum_nights                       | 365
 maximum_maximum_nights                       | 365
 minimum_nights_avg_ntm                       | 120.0
 maximum_nights_avg_ntm                       | 365.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 11
 availability_90                              | 41
 availability_365                             | 316
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 neighbourhood_reduced                        | Downtown
 property_reduced                             | Entire apartment
 room_reduced                                 | Entire home/apt
 neighbourhood_indexed                        | 0.0
 property_indexed                             | 0.0
 room_indexed                                 | 0.0
 amenities_array                              | [Kitchen, Iron, E...
 verifications_array                          | [email, phone, re...
 amenities_vector                             | (20,[0,1,2,3,4,5,...
 verifications_vector                         | (5,[0,1,2,4],[1.0...
-RECORD 17-------------------------------------------------------------
 id                                           | 283841
 host_since                                   | 2009/08/10
 host_location                                | Vancouver, Britis...
 host_response_time                           | within a day
 host_response_rate                           | 0.93
```

host_acceptance_rate	0.88
host_is_superhost	f
host_neighbourhood	Coal Harbour
host_listings_count	49
host_total_listings_count	49
host_verifications	['email', 'phone'...
host_has_profile_pic	t
host_identity_verified	t
neighbourhood	null
neighbourhood_cleansed	Downtown
neighbourhood_group_cleansed	null
latitude	49.2872
longitude	-123.12352
property_type	Entire apartment
room_type	Entire home/apt
accommodates	6
bathrooms	null
bathrooms_text	3 baths
bedrooms	3
beds	3
amenities	['Kitchen', 'Hot ...
price	197
minimum_nights	90
maximum_nights	1125
minimum_minimum_nights	90
maximum_minimum_nights	90
minimum_maximum_nights	1125
maximum_maximum_nights	1125
minimum_nights_avg_ntm	90.0
maximum_nights_avg_ntm	1125.0
has_availability	t
availability_30	0
availability_60	0
availability_90	0
availability_365	0
instant_bookable	f
calculated_host_listings_count	41
calculated_host_listings_count_entire_homes	41
calculated_host_listings_count_private_rooms	0
calculated_host_listings_count_shared_rooms	0
neighbourhood_reduced	Downtown
property_reduced	Entire apartment
room_reduced	Entire home/apt
neighbourhood_indexed	0.0
property_indexed	0.0
room_indexed	0.0
amenities_array	[Kitchen, Hot tub...
verifications_array	[email, phone, re...
amenities_vector	(20,[0,1,2,3,4,5,...

```
 verifications_vector                        | (5,[0,1,2,4],[1.0...
-RECORD 18-------------------------------------------------------------
 id                                          | 323090
 host_since                                  | 2012/01/01
 host_location                               | Vancouver, Britis...
 host_response_time                          | N/A
 host_response_rate                          | N/A
 host_acceptance_rate                        | 0.85
 host_is_superhost                           | t
 host_neighbourhood                          | Downtown Vancouver
 host_listings_count                         | 1
 host_total_listings_count                   | 1
 host_verifications                          | ['email', 'phone'...
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, Britis...
 neighbourhood_cleansed                      | Downtown
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.28206
 longitude                                   | -123.12461
 property_type                               | Entire apartment
 room_type                                   | Entire home/apt
 accommodates                                | 2
 bathrooms                                   | null
 bathrooms_text                              | 1 bath
 bedrooms                                    | 1
 beds                                        | 1
 amenities                                   | ['Kitchen', 'Cook...
 price                                       | 168
 minimum_nights                              | 180
 maximum_nights                              | 180
 minimum_minimum_nights                      | 180
 maximum_minimum_nights                      | 180
 minimum_maximum_nights                      | 180
 maximum_maximum_nights                      | 180
 minimum_nights_avg_ntm                      | 180.0
 maximum_nights_avg_ntm                      | 180.0
 has_availability                            | t
 availability_30                             | 30
 availability_60                             | 60
 availability_90                             | 90
 availability_365                            | 141
 instant_bookable                            | f
 calculated_host_listings_count              | 1
 calculated_host_listings_count_entire_homes | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms | 0
 neighbourhood_reduced                       | Downtown
 property_reduced                            | Entire apartment
```

```
 room_reduced                                   | Entire home/apt
 neighbourhood_indexed                          | 0.0
 property_indexed                               | 0.0
 room_indexed                                   | 0.0
 amenities_array                                | [Kitchen, Cooking...
 verifications_array                            | [email, phone, re...
 amenities_vector                               | (20,[0,1,2,3,4,5,...
 verifications_vector                           | (5,[0,1,2,4],[1.0...
-RECORD 19-------------------------------------------------------------
 id                                             | 359942
 host_since                                     | 2011/09/23
 host_location                                  | Vancouver, Britis...
 host_response_time                             | within an hour
 host_response_rate                             | 1
 host_acceptance_rate                           | N/A
 host_is_superhost                              | f
 host_neighbourhood                             | Yaletown
 host_listings_count                            | 8
 host_total_listings_count                      | 8
 host_verifications                             | ['email', 'phone'...
 host_has_profile_pic                           | t
 host_identity_verified                         | f
 neighbourhood                                  | Vancouver, Britis...
 neighbourhood_cleansed                         | Downtown
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.27825
 longitude                                      | -123.11865
 property_type                                  | Private room in a...
 room_type                                      | Private room
 accommodates                                   | 2
 bathrooms                                      | null
 bathrooms_text                                 | 1 shared bath
 bedrooms                                       | 2
 beds                                           | 1
 amenities                                      | ['Kitchen', 'Lake...
 price                                          | 240
 minimum_nights                                 | 30
 maximum_nights                                 | 1125
 minimum_minimum_nights                         | 30
 maximum_minimum_nights                         | 30
 minimum_maximum_nights                         | 1125
 maximum_maximum_nights                         | 1125
 minimum_nights_avg_ntm                         | 30.0
 maximum_nights_avg_ntm                         | 1125.0
 has_availability                               | t
 availability_30                                | 29
 availability_60                                | 59
 availability_90                                | 89
 availability_365                               | 364
```

```
 instant_bookable                              | f
 calculated_host_listings_count                | 7
 calculated_host_listings_count_entire_homes   | 3
 calculated_host_listings_count_private_rooms  | 4
 calculated_host_listings_count_shared_rooms   | 0
 neighbourhood_reduced                         | Downtown
 property_reduced                              | Private room in a...
 room_reduced                                  | Private room
 neighbourhood_indexed                         | 0.0
 property_indexed                              | 6.0
 room_indexed                                  | 1.0
 amenities_array                               | [Kitchen, Lake ac...
 verifications_array                           | [email, phone, fa...
 amenities_vector                              | (20,[0,1,3,4,5,6,...
 verifications_vector                          | (5,[0,1,4],[1.0,1...
only showing top 20 rows
```

Decomposing complex features: host_since

Finally, we will deal with the date in `host_since` by changing the `string` to an `int` by keeping only the years.

In [31]:

```python
class YearKeeper(UnaryTransformer):

    def createTransformFunc(self):
        def keep_year(s):
            if type(s) is str and len(s)>4:
                return int(s.split("/")[0])
            return None
        return keep_year

    def outputDataType(self):
        return IntegerType()

    def validateInputType(self, inputType):
        assert inputType == StringType(), f'Expected StringType() and found {inputType}'

year_keeper = YearKeeper(). \
  setInputCol('host_since'). \
  setOutputCol('since_year')

year_keeper.transform(dataset).show(vertical=True, truncate=False)
stages.append(year_keeper)
```

```
-RECORD 0-----------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 availability_30                              | 11
 availability_60                              | 41
 availability_90                              | 70
 availability_365                             | 70
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 since_year                                   | 2009
-RECORD 1-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
 id                                           | 16254
 host_since                                   | 2009/12/15
 host_location                                | Vancouver, British Columbia, Canada
 host_response_time                           | N/A
 host_response_rate                           | N/A
 host_acceptance_rate                         | N/A
 host_is_superhost                            | f
 host_neighbourhood                           | Hastings-Sunrise
 host_listings_count                          | 1
 host_total_listings_count                    | 1
 host_verifications                           | ['phone', 'facebook', 'offline_government_id', 'selfie', 'go
vernment_id', 'identity_manual']
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Hastings-Sunrise
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27734
 longitude                                    | -123.04152
 property_type                                | Entire guest suite
 room_type                                    | Entire home/apt
 accommodates                                 | 4
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 2
 beds                                         | 3
 amenities                                    | ['Kitchen', 'Oven', 'Microwave', 'Wifi', 'Dryer', 'Carbon mo
noxide alarm', 'Stove', 'Hot water', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking basic
s', 'Private entrance', 'Smoke alarm', 'Hair dryer', 'Lockbox', 'Dedicated workspace', 'Iron', 'Free street
 parking', 'Free parking on premises', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Fi
re extinguisher', 'Patio or balcony', 'Children\u2019s books and toys', 'Hangers', 'Refrigerator', 'Essentia
```

```
ls', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
 price                                       | 195
 minimum_nights                              | 3
 maximum_nights                              | 31
 minimum_minimum_nights                      | 3
 maximum_minimum_nights                      | 3
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 3.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 0
 availability_90                             | 0
 availability_365                            | 228
 instant_bookable                            | t
 calculated_host_listings_count              | 1
 calculated_host_listings_count_entire_homes | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms | 0
 since_year                                  | 2009
-RECORD 2-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                          | 17765
 host_since                                  | 2010/01/07
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | N/A
 host_response_rate                          | N/A
 host_acceptance_rate                        | 1
 host_is_superhost                           | f
 host_neighbourhood                          | Mount Pleasant
 host_listings_count                         | 1
 host_total_listings_count                   | 1
 host_verifications                          | ['phone', 'facebook', 'reviews', 'jumio', 'offline_governmen
t_id', 'selfie', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                      | Mount Pleasant
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.26181
 longitude                                   | -123.10905
 property_type                               | Entire apartment
```

```
 room_type                                      | Entire home/apt
 accommodates                                   | 2
 bathrooms                                      | null
 bathrooms_text                                 | 1 bath
 bedrooms                                       | 1
 beds                                           | 1
 amenities                                      | ['Kitchen', 'Wifi', 'Host greets you', 'Dryer', 'Carbon mono
xide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated work
space', 'Iron', 'Free street parking', 'Free parking on premises', 'Coffee maker', 'Shampoo', 'Dishes and si
lverware', 'Fire extinguisher', 'Hangers', 'Essentials', 'Washer', 'Long term stays allowed', 'Elevator']
 price                                          | 94
 minimum_nights                                 | 3
 maximum_nights                                 | 60
 minimum_minimum_nights                         | 3
 maximum_minimum_nights                         | 3
 minimum_maximum_nights                         | 60
 maximum_maximum_nights                         | 60
 minimum_nights_avg_ntm                         | 3.0
 maximum_nights_avg_ntm                         | 60.0
 has_availability                               | t
 availability_30                                | 0
 availability_60                                | 0
 availability_90                                | 0
 availability_365                               | 0
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
 since_year                                     | 2010
-RECORD 3-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                             | 18795
 host_since                                     | 2010/01/13
 host_location                                  | Vancouver, British Columbia, Canada
 host_response_time                             | within a day
 host_response_rate                             | 0.7
 host_acceptance_rate                           | 0.4
 host_is_superhost                              | f
 host_neighbourhood                             | West End
 host_listings_count                            | 3
 host_total_listings_count                      | 3
 host_verifications                             | ['email', 'phone', 'reviews', 'jumio', 'government_id']
```

```
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | null
 neighbourhood_cleansed                        | West End
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.28251
 longitude                                     | -123.12827
 property_type                                 | Private room in house
 room_type                                     | Private room
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 private bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Cable TV', 'Wifi', 'Essentials', 'Heating', 'Wa
sher', 'Dryer', 'Long term stays allowed']
 price                                         | 50
 minimum_nights                                | 30
 maximum_nights                                | 360
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 360
 maximum_maximum_nights                        | 360
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 360.0
 has_availability                              | t
 availability_30                               | 9
 availability_60                               | 20
 availability_90                               | 50
 availability_365                              | 325
 instant_bookable                              | f
 calculated_host_listings_count                | 3
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms  | 3
 calculated_host_listings_count_shared_rooms   | 0
 since_year                                    | 2010
-RECORD 4-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
 id                                            | 21913
 host_since                                    | 2010/02/22
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | within an hour
 host_response_rate                            | 1
```

```
 host_acceptance_rate                          | 1
 host_is_superhost                             | t
 host_neighbourhood                            | Renfrew-Collingwood
 host_listings_count                           | 3
 host_total_listings_count                     | 3
 host_verifications                            | ['email', 'phone', 'reviews', 'offline_government_id', 'self
ie', 'government_id']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                        | Renfrew-Collingwood
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.23427
 longitude                                     | -123.05309
 property_type                                 | Entire house
 room_type                                     | Entire home/apt
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Oven', 'Microwave', 'Wifi', 'Dryer', 'Carbon mo
noxide alarm', 'Stove', 'Hot water', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heatin
g', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Free parking o
n premises', 'Coffee maker', 'Bed linens', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV']
 price                                         | 80
 minimum_nights                                | 28
 maximum_nights                                | 400
 minimum_minimum_nights                        | 28
 maximum_minimum_nights                        | 28
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 28.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 2
 availability_60                               | 13
 availability_90                               | 43
 availability_365                              | 318
 instant_bookable                              | f
 calculated_host_listings_count                | 3
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 2
 calculated_host_listings_count_shared_rooms   | 0
 since_year                                    | 2010
-RECORD 5-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
```

```
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
 id                                          | 24273
 host_since                                  | 2011/10/09
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | within an hour
 host_response_rate                          | 1
 host_acceptance_rate                        | 0.82
 host_is_superhost                           | f
 host_neighbourhood                          | Kitsilano
 host_listings_count                         | 2
 host_total_listings_count                   | 2
 host_verifications                          | ['email', 'phone', 'reviews', 'jumio', 'offline_government_i
d', 'government_id']
 host_has_profile_pic                        | f
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                      | Kitsilano
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.2626
 longitude                                   | -123.17308
 property_type                               | Entire apartment
 room_type                                   | Entire home/apt
 accommodates                                | 6
 bathrooms                                   | null
 bathrooms_text                              | 1 bath
 bedrooms                                    | 2
 beds                                        | 3
 amenities                                   | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi',
'Dryer', 'Fireplace guards', 'Carbon monoxide alarm', 'Stove', 'Crib', 'Hot water', 'Babysitter recommendati
ons', 'Garden or backyard', 'Extra pillows and blankets', 'Dishwasher', 'Luggage dropoff allowed', 'Cooking
 basics', 'Room-darkening shades', 'Private entrance', 'Smoke alarm', 'Baking sheet', 'Heating', 'Single lev
el home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Coffee maker', 'Be
d linens', 'Keypad', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Shower gel', 'Patio or balcon
y', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Wa
sher', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'TV', 'Ethernet connection']
 price                                       | 200
 minimum_nights                              | 30
 maximum_nights                              | 365
 minimum_minimum_nights                      | 30
 maximum_minimum_nights                      | 30
 minimum_maximum_nights                      | 365
 maximum_maximum_nights                      | 365
 minimum_nights_avg_ntm                      | 30.0
 maximum_nights_avg_ntm                      | 365.0
 has_availability                            | t
```

```
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 10
 availability_365                             | 223
 instant_bookable                             | f
 calculated_host_listings_count               | 2
 calculated_host_listings_count_entire_homes  | 2
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 since_year                                   | 2011
-RECORD 6-------------------
```

```
 'Elevator']
 price                                       | 159
 minimum_nights                              | 32
 maximum_nights                              | 365
 minimum_minimum_nights                      | 32
 maximum_minimum_nights                      | 32
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 32.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 29
 availability_90                             | 59
 availability_365                            | 333
 instant_bookable                            | t
 calculated_host_listings_count              | 4
 calculated_host_listings_count_entire_homes | 4
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms | 0
 since_year                                  | 2010
-RECORD 7-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                          | 54261
 host_since                                  | 2010/10/05
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | a few days or more
 host_response_rate                          | 0
 host_acceptance_rate                        | N/A
 host_is_superhost                           | f
 host_neighbourhood                          | Downtown Vancouver
 host_listings_count                         | 2
 host_total_listings_count                   | 2
 host_verifications                          | ['email', 'phone', 'reviews', 'offline_government_id', 'self
ie', 'government_id', 'identity_manual']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | null
 neighbourhood_cleansed                      | Downtown
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.27655
 longitude                                   | -123.12485
 property_type                               | Entire condominium
```

```
 room_type                                     | Entire home/apt
 accommodates                                  | 4
 bathrooms                                     | null
 bathrooms_text                                | 1 bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Hot tub', 'Gym', 'Indoor fireplace', 'Wifi', 'H
ost greets you', 'Dryer', 'Carbon monoxide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alarm', 'H
eating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Shampoo', 'Fire extinguis
her', 'Hangers', 'Cable TV', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning', 'E
levator']
 price                                         | 126
 minimum_nights                                | 90
 maximum_nights                                | 365
 minimum_minimum_nights                        | 90
 maximum_minimum_nights                        | 90
 minimum_maximum_nights                        | 365
 maximum_maximum_nights                        | 365
 minimum_nights_avg_ntm                        | 90.0
 maximum_nights_avg_ntm                        | 365.0
 has_availability                              | t
 availability_30                               | 30
 availability_60                               | 60
 availability_90                               | 90
 availability_365                              | 365
 instant_bookable                              | f
 calculated_host_listings_count                | 2
 calculated_host_listings_count_entire_homes   | 2
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 since_year                                    | 2010
-RECORD 8-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
 id                                            | 72659
 host_since                                    | 2011/02/03
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | N/A
 host_response_rate                            | N/A
 host_acceptance_rate                          | 0
 host_is_superhost                             | f
 host_neighbourhood                            | Commercial Drive
 host_listings_count                           | 1
 host_total_listings_count                     | 1
```

```
 host_verifications                          | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | null
 neighbourhood_cleansed                      | Grandview-Woodland
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.27357
 longitude                                   | -123.07024
 property_type                               | Entire condominium
 room_type                                   | Entire home/apt
 accommodates                                | 4
 bathrooms                                   | null
 bathrooms_text                              | 1 bath
 bedrooms                                    | 2
 beds                                        | 2
 amenities                                   | ['Kitchen', 'Lake access', 'Oven', 'Paid parking off premise
s', 'Gym', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water',
 'Dishwasher', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm', 'He
ating', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Coffee mak
er', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Hangers',
 'Refrigerator', 'Essentials', 'Washer', 'BBQ grill', 'TV', 'Elevator']
 price                                       | 210
 minimum_nights                              | 7
 maximum_nights                              | 21
 minimum_minimum_nights                      | 7
 maximum_minimum_nights                      | 7
 minimum_maximum_nights                      | 21
 maximum_maximum_nights                      | 21
 minimum_nights_avg_ntm                      | 7.0
 maximum_nights_avg_ntm                      | 21.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 0
 availability_90                             | 0
 availability_365                            | 40
 instant_bookable                            | f
 calculated_host_listings_count              | 1
 calculated_host_listings_count_entire_homes | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms | 0
 since_year                                  | 2011
-RECORD 9-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------
```

```
 id                                          | 77157
 host_since                                  | 2011/02/27
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | within a few hours
 host_response_rate                          | 1
 host_acceptance_rate                        | 1
 host_is_superhost                           | f
 host_neighbourhood                          | Arbutus Ridge
 host_listings_count                         | 2
 host_total_listings_count                   | 2
 host_verifications                          | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'offline_
government_id', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                      | Arbutus Ridge
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.24675
 longitude                                   | -123.16711
 property_type                               | Private room in guest suite
 room_type                                   | Private room
 accommodates                                | 3
 bathrooms                                   | null
 bathrooms_text                              | 1 private bath
 bedrooms                                    | 1
 beds                                        | 2
 amenities                                   | ['Kitchen', 'Microwave', 'Wifi', 'Laundromat nearby', 'Drye
r', 'Hot water', 'Garden or backyard', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking bas
ics', 'Private entrance', 'Smoke alarm', 'Heating', 'Mini fridge', 'Single level home', 'Hair dryer', 'Lockb
ox', 'Dedicated workspace', 'First aid kit', 'Iron', 'Body soap', 'Free street parking', 'Coffee maker', 'Be
d linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Barbecue utensils', 'Shower gel', 'Hang
ers', 'Cable TV', 'Essentials', 'Washer', 'Conditioner', 'BBQ grill', 'Beach essentials', 'TV', 'Long term s
tays allowed']
 price                                       | 75
 minimum_nights                              | 3
 maximum_nights                              | 60
 minimum_minimum_nights                      | 3
 maximum_minimum_nights                      | 3
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 3.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 3
 availability_60                             | 3
 availability_90                             | 3
 availability_365                            | 267
 instant_bookable                            | t
 calculated_host_listings_count              | 2
```

```
 calculated_host_listings_count_entire_homes  | 0
 calculated_host_listings_count_private_rooms | 2
 calculated_host_listings_count_shared_rooms  | 0
 since_year                                   | 2011
-RECORD 10-------------------
```

```
 minimum_nights_avg_ntm                       | 3.0
 maximum_nights_avg_ntm                       | 365.0
 has_availability                             | t
 availability_30                              | 4
 availability_60                              | 34
 availability_90                              | 64
 availability_365                             | 339
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 since_year                                   | 2011
-RECORD 11-------------------
```

```
edicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Free parking on premises', 'Keypad', 'Shampo
o', 'Dishes and silverware', 'Patio or balcony', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Long
 term stays allowed', 'TV']
 price                                       | 75
 minimum_nights                              | 2
 maximum_nights                              | 60
 minimum_minimum_nights                      | 2
 maximum_minimum_nights                      | 2
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 2.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 0
 availability_60                             | 0
 availability_90                             | 2
 availability_365                            | 268
 instant_bookable                            | t
 calculated_host_listings_count              | 1
 calculated_host_listings_count_entire_homes | 0
 calculated_host_listings_count_private_rooms | 1
 calculated_host_listings_count_shared_rooms | 0
 since_year                                  | 2011
-RECORD 12-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
 id                                          | 129163
 host_since                                  | 2011/05/17
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | within a few hours
 host_response_rate                          | 1
 host_acceptance_rate                        | 1
 host_is_superhost                           | f
 host_neighbourhood                          | Renfrew-Collingwood
 host_listings_count                         | 2
 host_total_listings_count                   | 2
 host_verifications                          | ['email', 'phone', 'reviews', 'jumio', 'offline_government_i
d', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | null
 neighbourhood_cleansed                      | Renfrew-Collingwood
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.24507
```

```
 longitude                                     | -123.04941
 property_type                                 | Private room in house
 room_type                                     | Private room
 accommodates                                  | 1
 bathrooms                                     | null
 bathrooms_text                                | 1 shared bath
 bedrooms                                      | 1
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Hot water', 'Paid parking on premises', 'Lock o
n bedroom door', 'Free street parking', 'Cable TV', 'Smoke alarm', 'Wifi', 'Essentials', 'Heating', 'Washe
r', 'Dryer', 'Carbon monoxide alarm', 'Dishes and silverware', 'TV', 'Long term stays allowed', 'Host greets
you', 'Hair dryer', 'Shampoo', 'Fire extinguisher']
 price                                         | 45
 minimum_nights                                | 30
 maximum_nights                                | 280
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 280
 maximum_maximum_nights                        | 280
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 280.0
 has_availability                              | t
 availability_30                               | 30
 availability_60                               | 60
 availability_90                               | 90
 availability_365                              | 365
 instant_bookable                              | t
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 0
 calculated_host_listings_count_private_rooms  | 1
 calculated_host_listings_count_shared_rooms   | 0
 since_year                                    | 2011
-RECORD 13-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
 id                                            | 174135
 host_since                                    | 2011/07/17
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | within a few hours
 host_response_rate                            | 1
 host_acceptance_rate                          | 0
 host_is_superhost                             | f
 host_neighbourhood                            | Downtown Eastside
 host_listings_count                           | 1
```

```
 host_total_listings_count                    | 1
 host_verifications                           | ['email', 'phone', 'reviews', 'offline_government_id', 'self
ie', 'government_id', 'identity_manual']
 host_has_profile_pic                         | t
 host_identity_verified                       | t
 neighbourhood                                | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                       | Downtown Eastside
 neighbourhood_group_cleansed                 | null
 latitude                                     | 49.27782
 longitude                                    | -123.08985
 property_type                                | Entire guest suite
 room_type                                    | Entire home/apt
 accommodates                                 | 2
 bathrooms                                    | null
 bathrooms_text                               | 1 bath
 bedrooms                                     | 1
 beds                                         | 2
 amenities                                    | ['Kitchen', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Hot w
ater', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated wor
kspace', 'Iron', 'Free street parking', 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Hangers', 'Cable TV',
'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning']
 price                                        | 100
 minimum_nights                               | 30
 maximum_nights                               | 365
 minimum_minimum_nights                       | 30
 maximum_minimum_nights                       | 30
 minimum_maximum_nights                       | 365
 maximum_maximum_nights                       | 365
 minimum_nights_avg_ntm                       | 30.0
 maximum_nights_avg_ntm                       | 365.0
 has_availability                             | t
 availability_30                              | 0
 availability_60                              | 0
 availability_90                              | 0
 availability_365                             | 224
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 since_year                                   | 2011
-RECORD 14-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------
```

```
 id                                          | 200643
 host_since                                  | 2011/08/16
 host_location                               | Vancouver, British Columbia, Canada
 host_response_time                          | within an hour
 host_response_rate                          | 1
 host_acceptance_rate                        | 1
 host_is_superhost                           | f
 host_neighbourhood                          | Strathcona
 host_listings_count                         | 4
 host_total_listings_count                   | 4
 host_verifications                          | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                        | t
 host_identity_verified                      | t
 neighbourhood                               | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                      | Downtown Eastside
 neighbourhood_group_cleansed                | null
 latitude                                    | 49.27384
 longitude                                   | -123.08289
 property_type                               | Private room in house
 room_type                                   | Private room
 accommodates                                | 4
 bathrooms                                   | null
 bathrooms_text                              | 2 shared baths
 bedrooms                                    | 1
 beds                                        | 1
 amenities                                   | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wifi',
'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Garden or backyard', 'Dishwasher', 'Extra pillows a
nd blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Room-darkening shades', 'Smoke alarm', 'Heatin
g', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Free street parking', 'Coffee maker', 'Be
d linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Essentials',
'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
 price                                       | 87
 minimum_nights                              | 1
 maximum_nights                              | 1125
 minimum_minimum_nights                      | 1
 maximum_minimum_nights                      | 1
 minimum_maximum_nights                      | 1125
 maximum_maximum_nights                      | 1125
 minimum_nights_avg_ntm                      | 1.0
 maximum_nights_avg_ntm                      | 1125.0
 has_availability                            | t
 availability_30                             | 30
 availability_60                             | 60
 availability_90                             | 90
 availability_365                            | 365
 instant_bookable                            | t
 calculated_host_listings_count              | 4
 calculated_host_listings_count_entire_homes | 0
 calculated_host_listings_count_private_rooms | 4
```

```
 calculated_host_listings_count_shared_rooms  | 0
 since_year                                   | 2011
-RECORD 15-----------------------
```

```
 maximum_nights_avg_ntm                       | 180.0
 has_availability                             | t
 availability_30                              | 4
 availability_60                              | 4
 availability_90                              | 4
 availability_365                             | 113
 instant_bookable                             | f
 calculated_host_listings_count               | 1
 calculated_host_listings_count_entire_homes  | 1
 calculated_host_listings_count_private_rooms | 0
 calculated_host_listings_count_shared_rooms  | 0
 since_year                                   | 2011
-RECORD 16-------------------
```

```
 minimum_nights                                | 120
 maximum_nights                                | 365
 minimum_minimum_nights                        | 120
 maximum_minimum_nights                        | 120
 minimum_maximum_nights                        | 365
 maximum_maximum_nights                        | 365
 minimum_nights_avg_ntm                        | 120.0
 maximum_nights_avg_ntm                        | 365.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 11
 availability_90                               | 41
 availability_365                              | 316
 instant_bookable                              | f
 calculated_host_listings_count                | 1
 calculated_host_listings_count_entire_homes   | 1
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 since_year                                    | 2011
-RECORD 17-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                            | 283841
 host_since                                    | 2009/08/10
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | within a day
 host_response_rate                            | 0.93
 host_acceptance_rate                          | 0.88
 host_is_superhost                             | f
 host_neighbourhood                            | Coal Harbour
 host_listings_count                           | 49
 host_total_listings_count                     | 49
 host_verifications                            | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
 neighbourhood                                 | null
 neighbourhood_cleansed                        | Downtown
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.2872
 longitude                                     | -123.12352
 property_type                                 | Entire apartment
 room_type                                     | Entire home/apt
 accommodates                                  | 6
 bathrooms                                     | null
```

```
 bathrooms_text                                | 3 baths
 bedrooms                                      | 3
 beds                                          | 3
 amenities                                     | ['Kitchen', 'Hot tub', 'Oven', 'Pool', 'Gym', 'Indoor firepl
ace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Dishwasher', 'Smoke alar
m', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Coffee maker', 'Dis
hes and silverware', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowe
d', 'TV', 'Air conditioning']
 price                                         | 197
 minimum_nights                                | 90
 maximum_nights                                | 1125
 minimum_minimum_nights                        | 90
 maximum_minimum_nights                        | 90
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 90.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 0
 availability_60                               | 0
 availability_90                               | 0
 availability_365                              | 0
 instant_bookable                              | f
 calculated_host_listings_count                | 41
 calculated_host_listings_count_entire_homes   | 41
 calculated_host_listings_count_private_rooms  | 0
 calculated_host_listings_count_shared_rooms   | 0
 since_year                                    | 2009
-RECORD 18-------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------
 id                                            | 323090
 host_since                                    | 2012/01/01
 host_location                                 | Vancouver, British Columbia, Canada
 host_response_time                            | N/A
 host_response_rate                            | N/A
 host_acceptance_rate                          | 0.85
 host_is_superhost                             | t
 host_neighbourhood                            | Downtown Vancouver
 host_listings_count                           | 1
 host_total_listings_count                     | 1
 host_verifications                            | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                          | t
 host_identity_verified                        | t
```

```
 neighbourhood                                  | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                         | Downtown
 neighbourhood_group_cleansed                   | null
 latitude                                       | 49.28206
 longitude                                      | -123.12461
 property_type                                  | Entire apartment
 room_type                                      | Entire home/apt
 accommodates                                   | 2
 bathrooms                                      | null
 bathrooms_text                                 | 1 bath
 bedrooms                                       | 1
 beds                                           | 1
 amenities                                      | ['Kitchen', 'Cooking basics', 'Iron', 'Elevator', 'Gym', 'Ha
ngers', 'Cable TV', 'Smoke alarm', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Carbon monoxide alarm', 'Wif
i', 'TV', 'Long term stays allowed', 'Hair dryer', 'Shampoo', 'Dishes and silverware', 'Dedicated workspac
e']
 price                                          | 168
 minimum_nights                                 | 180
 maximum_nights                                 | 180
 minimum_minimum_nights                         | 180
 maximum_minimum_nights                         | 180
 minimum_maximum_nights                         | 180
 maximum_maximum_nights                         | 180
 minimum_nights_avg_ntm                         | 180.0
 maximum_nights_avg_ntm                         | 180.0
 has_availability                               | t
 availability_30                                | 30
 availability_60                                | 60
 availability_90                                | 90
 availability_365                               | 141
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
 since_year                                     | 2012
-RECORD 19-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
 id                                             | 359942
 host_since                                     | 2011/09/23
 host_location                                  | Vancouver, British Columbia, Canada
 host_response_time                             | within an hour
 host_response_rate                             | 1
```

```
 host_acceptance_rate                          | N/A
 host_is_superhost                             | f
 host_neighbourhood                            | Yaletown
 host_listings_count                           | 8
 host_total_listings_count                     | 8
 host_verifications                            | ['email', 'phone', 'facebook', 'reviews']
 host_has_profile_pic                          | t
 host_identity_verified                        | f
 neighbourhood                                 | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                        | Downtown
 neighbourhood_group_cleansed                  | null
 latitude                                      | 49.27825
 longitude                                     | -123.11865
 property_type                                 | Private room in apartment
 room_type                                     | Private room
 accommodates                                  | 2
 bathrooms                                     | null
 bathrooms_text                                | 1 shared bath
 bedrooms                                      | 2
 beds                                          | 1
 amenities                                     | ['Kitchen', 'Lake access', 'Oven', 'Paid parking on premise
s', 'Paid parking off premises', 'Gym', 'Microwave', 'Wifi', 'Dryer', 'Stove', 'Hot water', 'Dishwasher', 'E
xtra pillows and blankets', 'Cooking basics', 'Heating', 'Lock on bedroom door', 'Hair dryer', 'Iron', 'Coff
ee maker', 'Bed linens', 'Pocket wifi', 'Shampoo', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Esse
ntials', 'Washer', 'Long term stays allowed', 'Elevator']
 price                                         | 240
 minimum_nights                                | 30
 maximum_nights                                | 1125
 minimum_minimum_nights                        | 30
 maximum_minimum_nights                        | 30
 minimum_maximum_nights                        | 1125
 maximum_maximum_nights                        | 1125
 minimum_nights_avg_ntm                        | 30.0
 maximum_nights_avg_ntm                        | 1125.0
 has_availability                              | t
 availability_30                               | 29
 availability_60                               | 59
 availability_90                               | 89
 availability_365                              | 364
 instant_bookable                              | f
 calculated_host_listings_count                | 7
 calculated_host_listings_count_entire_homes   | 3
 calculated_host_listings_count_private_rooms  | 4
 calculated_host_listings_count_shared_rooms   | 0
 since_year                                    | 2011
only showing top 20 rows
```

EDA

We will now plot rows against target of our dataset transformed as we have so far.

In [32]:
```
pipeline = Pipeline(stages=stages)
df = pipeline.fit(dataset).transform(dataset)
```

In [33]:
```
import matplotlib.pyplot as plt
%matplotlib inline

prices = [ p['price']  for p in df.select("price").toLocalIterator() ]

prices = sorted(prices)

plt.plot(prices, label="prices", color='gray')
```

Out[33]:
```
[<matplotlib.lines.Line2D at 0x7f9f4502bf10>]
```

This curve can profit from flattening using a log to better understand it.

In [34]:
```
prices = list(map(math.log, prices))

plt.plot(prices, label="prices", color='gray')
```

Out[34]:
```
[<matplotlib.lines.Line2D at 0x7f9f44edfcd0>]
```

In [35]:

```
class LogTransformer(UnaryTransformer):

    def createTransformFunc(self):
        return lambda x: None if x is None else math.log(x)

    def outputDataType(self):
        return DoubleType()

    def validateInputType(self, inputType):
        assert inputType == IntegerType(), f'Expected IntegerType and found {inputType}'

log_transformer = LogTransformer().\
      setInputCol("price"). \
      setOutputCol("log_price")
stages.append(log_transformer)

pipeline = Pipeline(stages=stages)
df = pipeline.fit(dataset).transform(dataset)
```

Let us look at the types of the current columns.

In [36]:

```
for col in df.columns:
    print(df.schema[col].dataType.typeName(), ': ', col)
```

```
integer :  id
string :  host_since
string :  host_location
```

```
string :  host_response_time
string :  host_response_rate
string :  host_acceptance_rate
string :  host_is_superhost
string :  host_neighbourhood
integer :  host_listings_count
integer :  host_total_listings_count
string :  host_verifications
string :  host_has_profile_pic
string :  host_identity_verified
string :  neighbourhood
string :  neighbourhood_cleansed
string :  neighbourhood_group_cleansed
double :  latitude
double :  longitude
string :  property_type
string :  room_type
integer :  accommodates
string :  bathrooms
string :  bathrooms_text
integer :  bedrooms
integer :  beds
string :  amenities
integer :  price
integer :  minimum_nights
integer :  maximum_nights
integer :  minimum_minimum_nights
integer :  maximum_minimum_nights
integer :  minimum_maximum_nights
integer :  maximum_maximum_nights
double :  minimum_nights_avg_ntm
double :  maximum_nights_avg_ntm
string :  has_availability
integer :  availability_30
integer :  availability_60
integer :  availability_90
integer :  availability_365
string :  instant_bookable
integer :  calculated_host_listings_count
integer :  calculated_host_listings_count_entire_homes
integer :  calculated_host_listings_count_private_rooms
integer :  calculated_host_listings_count_shared_rooms
string :  neighbourhood_reduced
string :  property_reduced
string :  room_reduced
double :  neighbourhood_indexed
double :  property_indexed
double :  room_indexed
array :  amenities_array
```

```
array :  verifications_array
vectorudt :  amenities_vector
vectorudt :  verifications_vector
integer :  since_year
double :  log_price
```

We can now plot each of the numeric columns against the target (`log_price`).

In [37]:

```
plottable = 0

for col in df.columns:
    _type = df.schema[col].dataType.typeName()
    if (_type == 'double' or _type == 'integer') and (col != 'price'):
        plottable += 1

plt.rcParams['figure.figsize'] = [5, 5 * plottable]

current = 0
for col in df.columns:
    _type = df.schema[col].dataType.typeName()
    if (_type == 'double' or _type == 'integer') \
      and (col != 'price') and (col != 'log_price'):
        current += 1
        plt.subplot(plottable, 1, current)

        pairs = list()
        for p in df.select("log_price", col).toLocalIterator():
            pairs.append( (p['log_price'], p[col]) )
        pairs = sorted(pairs, key=lambda t:t[0])
        vals = [ t[1] for t in pairs if t[1] is not None ]
        minv = min(vals)
        maxv = max(vals)
        if minv != maxv:
            vals = list(map(lambda x: x if x is None else (x-minv)/(maxv-minv), vals))
        logp = [ t[0] for t in pairs if t[1] is not None ]

        plt.plot(logp, vals, '.', color='gray')

        plt.xlabel('log prices')
        plt.ylabel(col)
        plt.title(col.upper())

plt.savefig("eda.png", bbox_inches='tight', dpi=300)
plt
```

Out[37]: <module 'matplotlib.pyplot' from '/vol/scratch/virtualenv/tourism/lib/python3.7/site-packages/matplotlib/pyplot.py'>

host_to
0.2
0.0
3
4
5
6
7
log prices

LATITUDE
latitude
1.0
0.8
0.6
0.4
0.2
0.0
3
4
5
6
7
log prices

LONGITUDE
longitude
1.0
0.8
0.6
0.4
0.2
0.0
3
4
5
6
7
log prices

ACCOMMODATES
1.0
0.8

accommodates
0.0
0.2
0.4
0.6
3
4
5
6
7
log prices

BEDROOMS
bedrooms
0.0
0.2
0.4
0.6
0.8
1.0
3
4
5
6
7
log prices

BEDS
beds
0.0
0.2
0.4
0.6
0.8
1.0
3
4
5
6
7

log prices
MINIMUM_NIGHTS
minimum_nights
1.0
0.8
0.6
0.4
0.2
0.0
3
4
5
6
7
log prices

MAXIMUM_NIGHTS
maximum_nights
1.0
0.8
0.6
0.4
0.2
0.0
3
4
5
6
7
log prices

MINIMUM_MINIMUM_NIGHTS
inimum_minimum_nights
1.0
0.8
0.6
0.4

0.2
0.0
3
4
5
6
7
log prices

MAXIMUM_MINIMUM_NIGHTS
maximum_minimum_nights
1.0
0.8
0.6
0.4
0.2
0.0
3
4
5
6
7
log prices

MINIMUM_MAXIMUM_NIGHTS
minimum_maximum_nights
1.0
0.8
0.6
0.4
0.2
0.0
3
4
5
6
7
log prices

MAXIMUM_MAXIMUM_NIGHTS
1.0
0.8

MINIMUM_NIGHTS_AVG_NTM

MAXIMUM_NIGHTS_AVG_NTM

AVAILABILITY_30

availability_30
log prices

AVAILABILITY_60
availability_60
log prices

AVAILABILITY_90
availability_90

AVAILABILITY_365

CALCULATED_HOST_LISTINGS_COUNT

CALCULATED_HOST_LISTINGS_COUNT_ENTIRE_HOMES

CALCULATED_HOST_LISTINGS_COUNT_PRIVATE_ROOMS

CALCULATED_HOST_LISTINGS_COUNT_SHARED_ROOMS

NEIGHBOURHOOD_INDEXED

PROPERTY_INDEXED

ROOM_INDEXED

For categorical data, we can plot box plots for each of the prices

In [38]:

```
amenities_prices = [ list()  for a in fitted_amenities.vocabulary ]
for r in df.select("amenities_vector", "log_price").toLocalIterator():
    for idx in r[0].indices:
        amenities_prices[idx].append(r[1])

labels = [ label[:8] for label in fitted_amenities.vocabulary ]

plt.rcParams['figure.figsize'] = [20, 10]
plt.boxplot(amenities_prices, labels=labels)
plt
```

Out[38]:

```
<module 'matplotlib.pyplot' from '/vol/scratch/virtualenv/tourism/lib/python3.7/site-packages/matplotlib/pyp
lot.py'>
```

Having a TV pays better on average? Who would have thought.

In [39]:

```
verification_prices = [ list()  for a in fitted_verifications.vocabulary ]
for r in df.select("verifications_vector", "log_price").toLocalIterator():
    for idx in r[0].indices:
        verification_prices[idx].append(r[1])

plt.rcParams['figure.figsize'] = [10, 10]
plt.boxplot(verification_prices, labels=fitted_verifications.vocabulary)
plt
```

Out[39]:

```
<module 'matplotlib.pyplot' from '/vol/scratch/virtualenv/tourism/lib/python3.7/site-packages/matplotlib/pyp
lot.py'>
```

No noticeable difference.

In [40]:

```
n_prices = [ list()  for a in fitted_n_indexer.labels ]
for r in df.select("neighbourhood_indexed", "log_price").toLocalIterator():
    n_prices[int(r[0])].append(r[1])

labels = [ label[:8] for label in fitted_n_indexer.labels ]
```

```
plt.rcParams['figure.figsize'] = [20, 10]
plt.boxplot(n_prices, labels=labels)
plt
```

Out[40]: <module 'matplotlib.pyplot' from '/vol/scratch/virtualenv/tourism/lib/python3.7/site-packages/matplotlib/pyplot.py'>

From here, the following features look good:

- host_since
- host_listings_count
- host_total_listings_count
- accommodates
- bedrooms
- beds

- minimum_nights
- maximum_nights
- minimum_minimum_nights
- maximum_minimum_nights
- minimum_nights_avg_ntm
- availability_30
- availability_60
- availability_90
- availability_365
- calculated_host_listings_count
- calculated_host_listings_count_entire_homes
- calculated_host_listings_count_private_rooms
- neighbourhood_indexed
- property_type_indexed
- room_type_indexed

But `availability` cannot be computed for new properties, so we will not use it. Similarly with `since_years` .

In [41]:

```
columns = [ "host_listings_count", "host_total_listings_count", "accommodates",
            "bedrooms", "beds", "minimum_nights", "maximum_nights", "minimum_minimum_nights",
            "maximum_minimum_nights", "minimum_nights_avg_ntm",
            "calculated_host_listings_count",
            "calculated_host_listings_count_entire_homes", "calculated_host_listings_count_private_rooms",
            "neighbourhood_indexed", "property_indexed", "room_indexed" ]
```

Imputation

Finally, we need to deal with undefined values for each of the selected columns. We will replace them with the *mode*, the most common value for that column.

In [42]:

```
modes = { c: df.groupBy(c).\
  count().\
  sort("count", ascending=False).head()[c] for c in columns }
modes
```

```
{'host_listings_count': 1,
```

Out[42]:

```
'host_total_listings_count': 1,
'accommodates': 2,
'bedrooms': 1,
'beds': 1,
'minimum_nights': 30,
'maximum_nights': 1125,
'minimum_minimum_nights': 30,
'maximum_minimum_nights': 30,
'minimum_nights_avg_ntm': 30.0,
'calculated_host_listings_count': 1,
'calculated_host_listings_count_entire_homes': 1,
'calculated_host_listings_count_private_rooms': 0,
'neighbourhood_indexed': 0.0,
'property_indexed': 0.0,
'room_indexed': 0.0}
```

We can verify the code works as intended because the indexed columns return 0.0, as expected. We are now ready to impute. We will separate the imputers by type (integer vs. dobule).

In [43]:

```
class Imputer(UnaryTransformer):

    def __init__(self, value, outType):
        super(UnaryTransformer, self).__init__()
        self.value = value
        self.outType = outType

    def createTransformFunc(self):
        return lambda x: self.value if x is None else x

    def outputDataType(self):
        return self.outType

    def validateInputType(self, inputType):
        assert inputType == self.outType, f'Expected {outType.typeName()} and found {inputType}'

imp_columns = list()
for c in columns:
    c_imp = c + "_imp"
    imp_columns.append(c_imp)
    imputer = Imputer(modes[c], df.schema[c].dataType).\
      setInputCol(c). \
      setOutputCol(c_imp)
    stages.append(imputer)
```

Now we re-execute the full-pipeline to obtain the imputed values.

In [44]:
```
pipeline = Pipeline(stages=stages)
df = pipeline.fit(dataset).transform(dataset)
```

In [45]:
```
print("base")
for c in columns:
    print(" ",c, df.schema[c].dataType.typeName(), df.filter(c + " is null").count())
print("imputed")
for c in imp_columns:
    print(" ",c, df.schema[c].dataType.typeName(), df.filter(c + " is null").count())
print("amenities_vector", df.filter("amenities_vector is null").count())
```

```
base
  host_listings_count integer 1
  host_total_listings_count integer 1
  accommodates integer 0
  bedrooms integer 151
  beds integer 20
  minimum_nights integer 0
  maximum_nights integer 0
  minimum_minimum_nights integer 0
  maximum_minimum_nights integer 0
  minimum_nights_avg_ntm double 0
  calculated_host_listings_count integer 0
  calculated_host_listings_count_entire_homes integer 0
  calculated_host_listings_count_private_rooms integer 0
  neighbourhood_indexed double 0
  property_indexed double 0
  room_indexed double 0
imputed
  host_listings_count_imp integer 0
  host_total_listings_count_imp integer 0
  accommodates_imp integer 0
  bedrooms_imp integer 0
  beds_imp integer 0
  minimum_nights_imp integer 0
  maximum_nights_imp integer 0
  minimum_minimum_nights_imp integer 0
  maximum_minimum_nights_imp integer 0
  minimum_nights_avg_ntm_imp double 0
  calculated_host_listings_count_imp integer 0
  calculated_host_listings_count_entire_homes_imp integer 0
  calculated_host_listings_count_private_rooms_imp integer 0
  neighbourhood_indexed_imp double 0
  property_indexed_imp double 0
  room_indexed_imp double 0
amenities_vector 0
```

We can see all columns have been correctly imputed.

In [46]:
```
df.count()
```

Out[46]:
```
2167
```

In [47]:
```
assembler = VectorAssembler(
    inputCols=imp_columns + ["amenities_vector"],
    outputCol="features", handleInvalid="error")

output = assembler.transform(df)
output.count()
```

Out[47]:
```
2167
```

In [48]:
```
len(next(output.toLocalIterator())['features'])
```

Out[48]:
```
36
```

In [49]:
```
next(output.toLocalIterator())['features']
```

Out[49]:
```
DenseVector([1.0, 1.0, 2.0, 1.0, 1.0, 30.0, 60.0, 30.0, 30.0, 30.0, 1.0, 1.0, 0.0, 3.0, 1.0, 0.0, 1.0, 1.0,
1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.0, 1.0, 1.0])
```

In [50]:
```
rf = RandomForestRegressor(featuresCol="features", labelCol='log_price', seed=42)

model = rf.fit(output)
```

In [51]:
```
pipeline = Pipeline(stages=stages + [assembler])
df_devset = pipeline.fit(dataset).transform(devset)

predictions = model.transform(df_devset)

predictions.select("features","log_price", "prediction").show(5)
```

```
+--------------------+------------------+-----------------+
|            features|         log_price|       prediction|
+--------------------+------------------+-----------------+
```

```
|[49.0,49.0,5.0,2....|5.0106352940962555|5.111937128072566|
|[1.0,1.0,4.0,1.0,...| 5.293304824724492|4.783042564532028|
|[5.0,5.0,6.0,3.0,...| 5.501258210544727|5.281162445096307|
|[1.0,1.0,2.0,1.0,...| 5.225746673713202|4.725892380368114|
|[1.0,1.0,2.0,1.0,...| 4.852030263919617|4.628450415150274|
+--------------------+------------------+-----------------+
only showing top 5 rows
```

In [52]:

```
metrics = RegressionMetrics(predictions.select("prediction", "log_price").rdd\
                            .map(lambda r: (math.exp(r['prediction']), math.exp(r['log_price']))))
```

In [53]:

```
baseError = metrics.meanAbsoluteError
baseError
```

Out[53]:

```
56.83296858469796
```

In [54]:

```
print("Average price", sum(map(lambda x:x['price'],
                                df_devset.select("price").toLocalIterator())) / df_devset.count())
prices = [ p['price'] for p in df_devset.select("price").toLocalIterator() ]
prices = sorted(prices)
print("Median price", prices[len(prices)// 2])
```

```
Average price 150.7699704142012
Median price 115
```

So the current model put us about 50 dollars from the actual price, which is an error between 30% to 50% of the actual price. It is a poor model, but it is a start. Let us see which features are better.

Error Analysis

Let us reassemble the labels for our feature vector.

In [55]:

```
labels = columns + list(map(lambda x: "amenity: " + x, fitted_amenities.vocabulary))
labels
```

Out[55]:

```
['host_listings_count',
 'host_total_listings_count',
 'accommodates',
 'bedrooms',
 'beds',
```

```
 'minimum_nights',
 'maximum_nights',
 'minimum_minimum_nights',
 'maximum_minimum_nights',
 'minimum_nights_avg_ntm',
 'calculated_host_listings_count',
 'calculated_host_listings_count_entire_homes',
 'calculated_host_listings_count_private_rooms',
 'neighbourhood_indexed',
 'property_indexed',
 'room_indexed',
 'amenity: Wifi',
 'amenity: Heating',
 'amenity: Smoke alarm',
 'amenity: Essentials',
 'amenity: Kitchen',
 'amenity: Hangers',
 'amenity: Long term stays allowed',
 'amenity: Washer',
 'amenity: Dryer',
 'amenity: Hair dryer',
 'amenity: TV',
 'amenity: Dedicated workspace',
 'amenity: Hot water',
 'amenity: Shampoo',
 'amenity: Iron',
 'amenity: Carbon monoxide alarm',
 'amenity: Dishes and silverware',
 'amenity: Fire extinguisher',
 'amenity: Refrigerator',
 'amenity: Cooking basics']
```

We can now check the random forest feature importance.

In [56]:

```
importances = [ (model.featureImportances[idx], lbl) for idx, lbl in enumerate(labels) ]
table = '<table><tr><th>Rank</th><th>Feature</th><th>Importance</th></tr>'
for idx, pair in enumerate(sorted(importances, reverse=True)):
    score, lbl =  pair
    table = '{}<tr><td>{}</td><td>{}</td><td style="text-align: left">{}</td></tr>'.format(table, idx + 1, l
table += '</table>'
display(HTML(table))
```

Rank	Feature	Importance
1	bedrooms	0.29759833711831807
2	accommodates	0.1427191216188187

3	room_indexed	0.1127149138402227
4	calculated_host_listings_count_private_rooms	0.10739723134374748
5	calculated_host_listings_count_entire_homes	0.08297496298387186
6	beds	0.05974627314563132
7	property_indexed	0.04717241092082373
8	neighbourhood_indexed	0.028542794641018582
9	minimum_nights	0.02128008411329211
10	minimum_minimum_nights	0.015979070148242725
11	calculated_host_listings_count	0.013996832881259188
12	host_listings_count	0.010204820776045778
13	maximum_minimum_nights	0.010034851731445702
14	amenity: TV	0.007056027176559329
15	minimum_nights_avg_ntm	0.006005886468711853
16	amenity: Washer	0.004455509972657118
17	maximum_nights	0.004358285051469576
18	host_total_listings_count	0.0037251114201160987
19	amenity: Shampoo	0.0036793946167604113
20	amenity: Wifi	0.0029486516717849845
21	amenity: Kitchen	0.0025464909525526004
22	amenity: Dryer	0.0022176197867234603
23	amenity: Dedicated workspace	0.002035125243326297
24	amenity: Fire extinguisher	0.001780232816892939
25	amenity: Iron	0.0015979303521886784
26	amenity: Essentials	0.0013465241889189868
27	amenity: Hair dryer	0.0010627095723940202
28	amenity: Heating	0.001036165720235597
29	amenity: Carbon monoxide alarm	0.0009801199944787862
30	amenity: Long term stays allowed	0.0007968870266094076

31	amenity: Hangers	0.0007259039580038782
32	amenity: Dishes and silverware	0.0004421557326424239
33	amenity: Hot water	0.0004196893223580955
34	amenity: Cooking basics	0.000379539732620414
35	amenity: Refrigerator	4.233395925725329e-05
36	amenity: Smoke alarm	0.0

We can see that many amenities are not important for price prediction. It might be because they are expected amenities (like "Dishes and silverware") or not important for the type of tourism received by Vancouver (e.g., "Dedicated workspace").

Feature selection: dropping amenities

We can keep the only two amenities that seem to matter: TV and Wifi by dropping the processing of amenities as StringIndexer and CountVectorizer and instead using a spotter, checking whether the original amenities string contains "TV" and "Wifi" to produce the two binary features.

It is also a good moment to clean the pipeline.

In [57]:

```
for idx, stage in enumerate(stages):
    print(idx, stage)
```

```
0 SetReducer_be1cd06abde5
1 SetReducer_1553fba51e0c
2 SetReducer_1a1f72ec4e63
3 StringIndexerModel: uid=StringIndexer_ee4a421f98bb, handleInvalid=error
4 StringIndexerModel: uid=StringIndexer_65923eaf74ab, handleInvalid=error
5 StringIndexerModel: uid=StringIndexer_41256a6da518, handleInvalid=error
6 FieldSplitter_9d732ace06c1
7 FieldSplitter_071896664d79
8 CountVectorizerModel: uid=CountVectorizer_6c0c861e16c7, vocabularySize=20
9 CountVectorizerModel: uid=CountVectorizer_7f37b47c4a36, vocabularySize=5
10 YearKeeper_44fe8355dd7e
11 LogTransformer_1ed73acf9ec7
12 Imputer_64cfe6164b2c
13 Imputer_cf0bd26a84de
14 Imputer_7f5907f4048f
15 Imputer_32bf93c9bbe9
16 Imputer_f8e3a806c6f5
17 Imputer_9e9a37076a45
```

```
18 Imputer_2d735b3489a3
19 Imputer_38e7adab229b
20 Imputer_adc3d2523088
21 Imputer_6f7dd177fa52
22 Imputer_bf8d9571070c
23 Imputer_f578f1561c41
24 Imputer_6e651794eb6b
25 Imputer_8d9ee1f455d7
26 Imputer_1f39c4d33256
27 Imputer_752707a53b3d
```

In [58]:

```
new_stages = stages[0:6] + stages[11:]
for idx, stage in enumerate(new_stages):
    print(idx, stage)
```

```
0 SetReducer_be1cd06abde5
1 SetReducer_1553fba51e0c
2 SetReducer_1a1f72ec4e63
3 StringIndexerModel: uid=StringIndexer_ee4a421f98bb, handleInvalid=error
4 StringIndexerModel: uid=StringIndexer_65923eaf74ab, handleInvalid=error
5 StringIndexerModel: uid=StringIndexer_41256a6da518, handleInvalid=error
6 LogTransformer_1ed73acf9ec7
7 Imputer_64cfe6164b2c
8 Imputer_cf0bd26a84de
9 Imputer_7f5907f4048f
10 Imputer_32bf93c9bbe9
11 Imputer_f8e3a806c6f5
12 Imputer_9e9a37076a45
13 Imputer_2d735b3489a3
14 Imputer_38e7adab229b
15 Imputer_adc3d2523088
16 Imputer_6f7dd177fa52
17 Imputer_bf8d9571070c
18 Imputer_f578f1561c41
19 Imputer_6e651794eb6b
20 Imputer_8d9ee1f455d7
21 Imputer_1f39c4d33256
22 Imputer_752707a53b3d
```

In [59]:

```
class AmenitySpotter(UnaryTransformer):
    def __init__(self, amenity):
        super(UnaryTransformer, self).__init__()
        self.amenity = amenity

    def createTransformFunc(self):
        return lambda x: self.amenity in x
```

```
    def outputDataType(self):
        return BooleanType()

    def validateInputType(self, inputType):
        assert inputType == StringType(), f'Expected StringType() and found {inputType}'

tv_spotter = AmenitySpotter("TV").\
  setInputCol("amenities"). \
  setOutputCol("has_tv")
wifi_spotter = AmenitySpotter("Wifi").\
  setInputCol("amenities"). \
  setOutputCol("has_wifi")
new_stages.append(tv_spotter)
new_stages.append(wifi_spotter)

tv_spotter.transform(df).show(vertical=True,truncate=False)
```

```
-RECORD 0-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                                  | 13358
 host_since                                          | 2009/11/07
 host_location                                       | Vancouver, British Columbia, Canada
 host_response_time                                  | within a day
 host_response_rate                                  | 0.71
 host_acceptance_rate                                | 0.71
 host_is_superhost                                   | t
 host_neighbourhood                                  | Downtown Vancouver
 host_listings_count                                 | 1
 host_total_listings_count                           | 1
 host_verifications                                  | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'offl
ine_government_id', 'selfie', 'government_id']
 host_has_profile_pic                                | f
 host_identity_verified                              | t
 neighbourhood                                       | null
 neighbourhood_cleansed                              | West End
 neighbourhood_group_cleansed                        | null
 latitude                                            | 49.28209
 longitude                                           | -123.12515
 property_type                                       | Entire condominium
 room_type                                           | Entire home/apt
 accommodates                                        | 2
 bathrooms                                           | null
```

```
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 1
 beds                                               | 1
 amenities                                          | ['Kitchen', 'Oven', 'Gym', 'Paid parking off premises',
'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Extra pillows and blankets',
'Dishwasher', 'Cooking basics', 'Hot water kettle', 'Smoke alarm', 'Heating', 'Single level home', 'Hair dry
er', 'Dedicated workspace', 'Iron', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Hange
rs', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditionin
g', 'Elevator']
 price                                              | 90
 minimum_nights                                     | 30
 maximum_nights                                     | 60
 minimum_minimum_nights                             | 30
 maximum_minimum_nights                             | 30
 minimum_maximum_nights                             | 60
 maximum_maximum_nights                             | 60
 minimum_nights_avg_ntm                             | 30.0
 maximum_nights_avg_ntm                             | 60.0
 has_availability                                   | t
 availability_30                                    | 11
 availability_60                                    | 41
 availability_90                                    | 70
 availability_365                                   | 70
 instant_bookable                                   | f
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 1
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | West End
 property_reduced                                   | Entire condominium
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 3.0
 property_indexed                                   | 1.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Oven, Gym, Paid parking off premises, Microwav
e, Wifi, Dryer, Carbon monoxide alarm, Stove, Hot water, Extra pillows and blankets, Dishwasher, Cooking bas
ics, Hot water kettle, Smoke alarm, Heating, Single level home, Hair dryer, Dedicated workspace, Iron, Coffe
e maker, Bed linens, Shampoo, Dishes and silverware, Hangers, Cable TV, Refrigerator, Essentials, Washer, Lo
ng term stays allowed, TV, Air conditioning, Elevator]
 verifications_array                                | [email, phone, facebook, reviews, jumio, offline_governm
ent_id, selfie, government_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,18,19],[1.
0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                         | 2009
 log_price                                          | 4.499809670330265
 host_listings_count_imp                            | 1
 host_total_listings_count_imp                      | 1
 accommodates_imp                                   | 2
```

```
 bedrooms_imp                                       | 1
 beds_imp                                           | 1
 minimum_nights_imp                                 | 30
 maximum_nights_imp                                 | 60
 minimum_minimum_nights_imp                         | 30
 maximum_minimum_nights_imp                         | 30
 minimum_nights_avg_ntm_imp                         | 30.0
 calculated_host_listings_count_imp                 | 1
 calculated_host_listings_count_entire_homes_imp    | 1
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 3.0
 property_indexed_imp                               | 1.0
 room_indexed_imp                                   | 0.0
 has_tv                                             | true
-RECORD 1-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                                 | 16254
 host_since                                         | 2009/12/15
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | N/A
 host_response_rate                                 | N/A
 host_acceptance_rate                               | N/A
 host_is_superhost                                  | f
 host_neighbourhood                                 | Hastings-Sunrise
 host_listings_count                                | 1
 host_total_listings_count                          | 1
 host_verifications                                 | ['phone', 'facebook', 'offline_government_id', 'selfie',
'government_id', 'identity_manual']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Hastings-Sunrise
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.27734
 longitude                                          | -123.04152
 property_type                                      | Entire guest suite
 room_type                                          | Entire home/apt
 accommodates                                       | 4
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 2
 beds                                               | 3
 amenities                                          | ['Kitchen', 'Oven', 'Microwave', 'Wifi', 'Dryer', 'Carbo
```

```
n monoxide alarm', 'Stove', 'Hot water', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking b
asics', 'Private entrance', 'Smoke alarm', 'Hair dryer', 'Lockbox', 'Dedicated workspace', 'Iron', 'Free str
eet parking', 'Free parking on premises', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware',
 'Fire extinguisher', 'Patio or balcony', 'Children\u2019s books and toys', 'Hangers', 'Refrigerator', 'Esse
ntials', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
 price                                          | 195
 minimum_nights                                 | 3
 maximum_nights                                 | 31
 minimum_minimum_nights                         | 3
 maximum_minimum_nights                         | 3
 minimum_maximum_nights                         | 1125
 maximum_maximum_nights                         | 1125
 minimum_nights_avg_ntm                         | 3.0
 maximum_nights_avg_ntm                         | 1125.0
 has_availability                               | t
 availability_30                                | 0
 availability_60                                | 0
 availability_90                                | 0
 availability_365                               | 228
 instant_bookable                               | t
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
 neighbourhood_reduced                          | Hastings-Sunrise
 property_reduced                               | Entire guest suite
 room_reduced                                   | Entire home/apt
 neighbourhood_indexed                          | 9.0
 property_indexed                               | 4.0
 room_indexed                                   | 0.0
 amenities_array                                | [Kitchen, Oven, Microwave, Wifi, Dryer, Carbon monoxide
 alarm, Stove, Hot water, Extra pillows and blankets, Luggage dropoff allowed, Cooking basics, Private entra
nce, Smoke alarm, Hair dryer, Lockbox, Dedicated workspace, Iron, Free street parking, Free parking on premi
ses, Coffee maker, Bed linens, Shampoo, Dishes and silverware, Fire extinguisher, Patio or balcony, Children
\u2019s books and toys, Hangers, Refrigerator, Essentials, Washer, Long term stays allowed, Pack \u2019n Pla
y/travel crib]
 verifications_array                            | [phone, facebook, offline_government_id, selfie, governm
ent_id, identity_manual]
 amenities_vector                               | (20,[0,2,3,4,5,6,7,8,9,11,12,13,14,15,16,17,18,19],[1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                           | (5,[0,2,3],[1.0,1.0,1.0])
 since_year                                     | 2009
 log_price                                      | 5.272999558563747
 host_listings_count_imp                        | 1
 host_total_listings_count_imp                  | 1
 accommodates_imp                               | 4
 bedrooms_imp                                   | 2
 beds_imp                                       | 3
 minimum_nights_imp                             | 3
```

```
 maximum_nights_imp                                | 31
 minimum_minimum_nights_imp                        | 3
 maximum_minimum_nights_imp                        | 3
 minimum_nights_avg_ntm_imp                        | 3.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 1
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 9.0
 property_indexed_imp                              | 4.0
 room_indexed_imp                                  | 0.0
 has_tv                                            | false
-RECORD 2-----------------------
```

```
r']
 price                                          | 94
 minimum_nights                                 | 3
 maximum_nights                                 | 60
 minimum_minimum_nights                         | 3
 maximum_minimum_nights                         | 3
 minimum_maximum_nights                         | 60
 maximum_maximum_nights                         | 60
 minimum_nights_avg_ntm                         | 3.0
 maximum_nights_avg_ntm                         | 60.0
 has_availability                               | t
 availability_30                                | 0
 availability_60                                | 0
 availability_90                                | 0
 availability_365                               | 0
 instant_bookable                               | f
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 1
 calculated_host_listings_count_private_rooms   | 0
 calculated_host_listings_count_shared_rooms    | 0
 neighbourhood_reduced                          | Mount Pleasant
 property_reduced                               | Entire apartment
 room_reduced                                   | Entire home/apt
 neighbourhood_indexed                          | 5.0
 property_indexed                               | 0.0
 room_indexed                                   | 0.0
 amenities_array                                | [Kitchen, Wifi, Host greets you, Dryer, Carbon monoxide
 alarm, Hot water, Luggage dropoff allowed, Smoke alarm, Heating, Hair dryer, Dedicated workspace, Iron, Fre
e street parking, Free parking on premises, Coffee maker, Shampoo, Dishes and silverware, Fire extinguisher,
Hangers, Essentials, Washer, Long term stays allowed, Elevator]
 verifications_array                            | [phone, facebook, reviews, jumio, offline_government_id,
selfie, government_id]
 amenities_vector                               | (20,[0,1,2,3,4,5,6,7,8,9,11,12,13,14,15,16,17],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                           | (5,[0,2,3,4],[1.0,1.0,1.0,1.0])
 since_year                                     | 2010
 log_price                                      | 4.543294782270004
 host_listings_count_imp                        | 1
 host_total_listings_count_imp                  | 1
 accommodates_imp                               | 2
 bedrooms_imp                                   | 1
 beds_imp                                       | 1
 minimum_nights_imp                             | 3
 maximum_nights_imp                             | 60
 minimum_minimum_nights_imp                     | 3
 maximum_minimum_nights_imp                     | 3
 minimum_nights_avg_ntm_imp                     | 3.0
 calculated_host_listings_count_imp             | 1
 calculated_host_listings_count_entire_homes_imp | 1
```

```
 calculated_host_listings_count_private_rooms_imp | 0
 neighbourhood_indexed_imp                        | 5.0
 property_indexed_imp                             | 0.0
 room_indexed_imp                                 | 0.0
 has_tv                                           | false
-RECORD 3-------------------
```

```
 has_availability                                  | t
 availability_30                                   | 9
 availability_60                                   | 20
 availability_90                                   | 50
 availability_365                                  | 325
 instant_bookable                                  | f
 calculated_host_listings_count                    | 3
 calculated_host_listings_count_entire_homes       | 0
 calculated_host_listings_count_private_rooms      | 3
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | West End
 property_reduced                                  | Private room in house
 room_reduced                                      | Private room
 neighbourhood_indexed                             | 3.0
 property_indexed                                  | 3.0
 room_indexed                                      | 1.0
 amenities_array                                   | [Kitchen, Cable TV, Wifi, Essentials, Heating, Washer, D
ryer, Long term stays allowed]
 verifications_array                               | [email, phone, reviews, jumio, government_id]
 amenities_vector                                  | (20,[0,1,3,4,6,7,8],[1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                        | 2010
 log_price                                         | 3.912023005428146
 host_listings_count_imp                           | 3
 host_total_listings_count_imp                     | 3
 accommodates_imp                                  | 2
 bedrooms_imp                                      | 1
 beds_imp                                          | 1
 minimum_nights_imp                                | 30
 maximum_nights_imp                                | 360
 minimum_minimum_nights_imp                        | 30
 maximum_minimum_nights_imp                        | 30
 minimum_nights_avg_ntm_imp                        | 30.0
 calculated_host_listings_count_imp                | 3
 calculated_host_listings_count_entire_homes_imp   | 0
 calculated_host_listings_count_private_rooms_imp  | 3
 neighbourhood_indexed_imp                         | 3.0
 property_indexed_imp                              | 3.0
 room_indexed_imp                                  | 1.0
 has_tv                                            | true
-RECORD 4-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                                | 21913
```

```
 host_since                                          | 2010/02/22
 host_location                                       | Vancouver, British Columbia, Canada
 host_response_time                                  | within an hour
 host_response_rate                                  | 1
 host_acceptance_rate                                | 1
 host_is_superhost                                   | t
 host_neighbourhood                                  | Renfrew-Collingwood
 host_listings_count                                 | 3
 host_total_listings_count                           | 3
 host_verifications                                  | ['email', 'phone', 'reviews', 'offline_government_id',
 'selfie', 'government_id']
 host_has_profile_pic                                | t
 host_identity_verified                              | t
 neighbourhood                                       | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                              | Renfrew-Collingwood
 neighbourhood_group_cleansed                        | null
 latitude                                            | 49.23427
 longitude                                           | -123.05309
 property_type                                       | Entire house
 room_type                                           | Entire home/apt
 accommodates                                        | 2
 bathrooms                                           | null
 bathrooms_text                                      | 1 bath
 bedrooms                                            | 1
 beds                                                | 1
 amenities                                           | ['Kitchen', 'Oven', 'Microwave', 'Wifi', 'Dryer', 'Carbo
n monoxide alarm', 'Stove', 'Hot water', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heat
ing', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Free parking
on premises', 'Coffee maker', 'Bed linens', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV']
 price                                               | 80
 minimum_nights                                      | 28
 maximum_nights                                      | 400
 minimum_minimum_nights                              | 28
 maximum_minimum_nights                              | 28
 minimum_maximum_nights                              | 1125
 maximum_maximum_nights                              | 1125
 minimum_nights_avg_ntm                              | 28.0
 maximum_nights_avg_ntm                              | 1125.0
 has_availability                                    | t
 availability_30                                     | 2
 availability_60                                     | 13
 availability_90                                     | 43
 availability_365                                    | 318
 instant_bookable                                    | f
 calculated_host_listings_count                      | 3
 calculated_host_listings_count_entire_homes         | 1
 calculated_host_listings_count_private_rooms        | 2
 calculated_host_listings_count_shared_rooms         | 0
```

```
 neighbourhood_reduced                              | Renfrew-Collingwood
 property_reduced                                   | Entire house
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 11.0
 property_indexed                                   | 2.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Oven, Microwave, Wifi, Dryer, Carbon monoxide
 alarm, Stove, Hot water, Luggage dropoff allowed, Private entrance, Smoke alarm, Heating, Single level hom
e, Hair dryer, Dedicated workspace, Iron, Free street parking, Free parking on premises, Coffee maker, Bed l
inens, Keypad, Shampoo, Dishes and silverware, Fire extinguisher, Hangers, Cable TV, Refrigerator, Essential
s, Washer, Long term stays allowed, TV]
 verifications_array                                | [email, phone, reviews, offline_government_id, selfie, g
overnment_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18],[1.
0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                         | 2010
 log_price                                          | 4.382026634673881
 host_listings_count_imp                            | 3
 host_total_listings_count_imp                      | 3
 accommodates_imp                                   | 2
 bedrooms_imp                                       | 1
 beds_imp                                           | 1
 minimum_nights_imp                                 | 28
 maximum_nights_imp                                 | 400
 minimum_minimum_nights_imp                         | 28
 maximum_minimum_nights_imp                         | 28
 minimum_nights_avg_ntm_imp                         | 28.0
 calculated_host_listings_count_imp                 | 3
 calculated_host_listings_count_entire_homes_imp    | 1
 calculated_host_listings_count_private_rooms_imp   | 2
 neighbourhood_indexed_imp                          | 11.0
 property_indexed_imp                               | 2.0
 room_indexed_imp                                   | 0.0
 has_tv                                             | true
-RECORD 5-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                                 | 24273
 host_since                                         | 2011/10/09
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within an hour
 host_response_rate                                 | 1
 host_acceptance_rate                               | 0.82
```

```
 host_is_superhost                                  | f
 host_neighbourhood                                 | Kitsilano
 host_listings_count                                | 2
 host_total_listings_count                          | 2
 host_verifications                                 | ['email', 'phone', 'reviews', 'jumio', 'offline_governme
nt_id', 'government_id']
 host_has_profile_pic                               | f
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Kitsilano
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.2626
 longitude                                          | -123.17308
 property_type                                      | Entire apartment
 room_type                                          | Entire home/apt
 accommodates                                       | 6
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 2
 beds                                               | 3
 amenities                                          | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wi
fi', 'Dryer', 'Fireplace guards', 'Carbon monoxide alarm', 'Stove', 'Crib', 'Hot water', 'Babysitter recomme
ndations', 'Garden or backyard', 'Extra pillows and blankets', 'Dishwasher', 'Luggage dropoff allowed', 'Coo
king basics', 'Room-darkening shades', 'Private entrance', 'Smoke alarm', 'Baking sheet', 'Heating', 'Single
level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Coffee maker',
 'Bed linens', 'Keypad', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Shower gel', 'Patio or ba
lcony', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials',
'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'TV', 'Ethernet connection']
 price                                              | 200
 minimum_nights                                     | 30
 maximum_nights                                     | 365
 minimum_minimum_nights                             | 30
 maximum_minimum_nights                             | 30
 minimum_maximum_nights                             | 365
 maximum_maximum_nights                             | 365
 minimum_nights_avg_ntm                             | 30.0
 maximum_nights_avg_ntm                             | 365.0
 has_availability                                   | t
 availability_30                                    | 0
 availability_60                                    | 0
 availability_90                                    | 10
 availability_365                                   | 223
 instant_bookable                                   | f
 calculated_host_listings_count                     | 2
 calculated_host_listings_count_entire_homes        | 2
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Kitsilano
 property_reduced                                   | Entire apartment
```

```
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 2.0
 property_indexed                                   | 0.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Oven, Indoor fireplace, Microwave, Wifi, Drye
r, Fireplace guards, Carbon monoxide alarm, Stove, Crib, Hot water, Babysitter recommendations, Garden or ba
ckyard, Extra pillows and blankets, Dishwasher, Luggage dropoff allowed, Cooking basics, Room-darkening shad
es, Private entrance, Smoke alarm, Baking sheet, Heating, Single level home, Hair dryer, Dedicated workspac
e, Iron, Free street parking, Bathtub, Coffee maker, Bed linens, Keypad, Shampoo, Fire extinguisher, Dishes
 and silverware, Shower gel, Patio or balcony, Children\u2019s books and toys, High chair, Hangers, Cable T
V, Refrigerator, Essentials, Washer, Long term stays allowed, Pack \u2019n Play/travel crib, TV, Ethernet co
nnection]
 verifications_array                                | [email, phone, reviews, jumio, offline_government_id, go
vernment_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19],
[1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                         | 2011
 log_price                                          | 5.298317366548036
 host_listings_count_imp                            | 2
 host_total_listings_count_imp                      | 2
 accommodates_imp                                   | 6
 bedrooms_imp                                       | 2
 beds_imp                                           | 3
 minimum_nights_imp                                 | 30
 maximum_nights_imp                                 | 365
 minimum_minimum_nights_imp                         | 30
 maximum_minimum_nights_imp                         | 30
 minimum_nights_avg_ntm_imp                         | 30.0
 calculated_host_listings_count_imp                 | 2
 calculated_host_listings_count_entire_homes_imp    | 2
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 2.0
 property_indexed_imp                               | 0.0
 room_indexed_imp                                   | 0.0
 has_tv                                             | true
-RECORD 6-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                                 | 42645
 host_since                                         | 2010/04/17
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within an hour
 host_response_rate                                 | 1
```

```
 host_acceptance_rate                                | 1
 host_is_superhost                                   | t
 host_neighbourhood                                  | Downtown Vancouver
 host_listings_count                                 | 4
 host_total_listings_count                           | 4
 host_verifications                                  | ['email', 'phone', 'facebook', 'reviews']
 host_has_profile_pic                                | t
 host_identity_verified                              | t
 neighbourhood                                       | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                              | Downtown
 neighbourhood_group_cleansed                        | null
 latitude                                            | 49.28001
 longitude                                           | -123.12475
 property_type                                       | Entire condominium
 room_type                                           | Entire home/apt
 accommodates                                        | 3
 bathrooms                                           | null
 bathrooms_text                                      | 1 bath
 bedrooms                                            | 1
 beds                                                | 2
 amenities                                           | ['Kitchen', 'Oven', 'Paid parking on premises', 'Paid pa
rking off premises', 'Gym', 'Microwave', 'Wifi', 'Host greets you', 'Dryer', 'Carbon monoxide alarm', 'Stov
e', 'Hot water', 'Dishwasher', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair
 dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Coffee maker', 'Shampoo', 'Fire extinguisher', 'Di
shes and silverware', 'Patio or balcony', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable
 TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'T
V', 'Elevator']
 price                                               | 159
 minimum_nights                                      | 32
 maximum_nights                                      | 365
 minimum_minimum_nights                              | 32
 maximum_minimum_nights                              | 32
 minimum_maximum_nights                              | 1125
 maximum_maximum_nights                              | 1125
 minimum_nights_avg_ntm                              | 32.0
 maximum_nights_avg_ntm                              | 1125.0
 has_availability                                    | t
 availability_30                                     | 0
 availability_60                                     | 29
 availability_90                                     | 59
 availability_365                                    | 333
 instant_bookable                                    | t
 calculated_host_listings_count                      | 4
 calculated_host_listings_count_entire_homes         | 4
 calculated_host_listings_count_private_rooms        | 0
 calculated_host_listings_count_shared_rooms         | 0
 neighbourhood_reduced                               | Downtown
 property_reduced                                    | Entire condominium
 room_reduced                                        | Entire home/apt
```

```
 neighbourhood_indexed                              | 0.0
 property_indexed                                   | 1.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Oven, Paid parking on premises, Paid parking o
ff premises, Gym, Microwave, Wifi, Host greets you, Dryer, Carbon monoxide alarm, Stove, Hot water, Dishwash
er, Luggage dropoff allowed, Cooking basics, Smoke alarm, Heating, Hair dryer, Dedicated workspace, First ai
d kit, Iron, Coffee maker, Shampoo, Fire extinguisher, Dishes and silverware, Patio or balcony, Children\u20
19s books and toys, High chair, Hangers, Cable TV, Refrigerator, Essentials, Washer, Long term stays allowe
d, Pack \u2019n Play/travel crib, TV, Elevator]
 verifications_array                                | [email, phone, facebook, reviews]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19],
[1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,4],[1.0,1.0,1.0])
 since_year                                         | 2010
 log_price                                          | 5.0689042022202315
 host_listings_count_imp                            | 4
 host_total_listings_count_imp                      | 4
 accommodates_imp                                   | 3
 bedrooms_imp                                       | 1
 beds_imp                                           | 2
 minimum_nights_imp                                 | 32
 maximum_nights_imp                                 | 365
 minimum_minimum_nights_imp                         | 32
 maximum_minimum_nights_imp                         | 32
 minimum_nights_avg_ntm_imp                         | 32.0
 calculated_host_listings_count_imp                 | 4
 calculated_host_listings_count_entire_homes_imp    | 4
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 0.0
 property_indexed_imp                               | 1.0
 room_indexed_imp                                   | 0.0
 has_tv                                             | true
-RECORD 7-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                                 | 54261
 host_since                                         | 2010/10/05
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | a few days or more
 host_response_rate                                 | 0
 host_acceptance_rate                               | N/A
 host_is_superhost                                  | f
 host_neighbourhood                                 | Downtown Vancouver
 host_listings_count                                | 2
```

```
 host_total_listings_count                          | 2
 host_verifications                                 | ['email', 'phone', 'reviews', 'offline_government_id',
 'selfie', 'government_id', 'identity_manual']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | null
 neighbourhood_cleansed                             | Downtown
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.27655
 longitude                                          | -123.12485
 property_type                                      | Entire condominium
 room_type                                          | Entire home/apt
 accommodates                                       | 4
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 1
 beds                                               | 1
 amenities                                          | ['Kitchen', 'Hot tub', 'Gym', 'Indoor fireplace', 'Wif
i', 'Host greets you', 'Dryer', 'Carbon monoxide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alar
m', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Shampoo', 'Fire ext
inguisher', 'Hangers', 'Cable TV', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditionin
g', 'Elevator']
 price                                              | 126
 minimum_nights                                     | 90
 maximum_nights                                     | 365
 minimum_minimum_nights                             | 90
 maximum_minimum_nights                             | 90
 minimum_maximum_nights                             | 365
 maximum_maximum_nights                             | 365
 minimum_nights_avg_ntm                             | 90.0
 maximum_nights_avg_ntm                             | 365.0
 has_availability                                   | t
 availability_30                                    | 30
 availability_60                                    | 60
 availability_90                                    | 90
 availability_365                                   | 365
 instant_bookable                                   | f
 calculated_host_listings_count                     | 2
 calculated_host_listings_count_entire_homes        | 2
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Downtown
 property_reduced                                   | Entire condominium
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 0.0
 property_indexed                                   | 1.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Hot tub, Gym, Indoor fireplace, Wifi, Host gre
ets you, Dryer, Carbon monoxide alarm, Hot water, Luggage dropoff allowed, Smoke alarm, Heating, Hair dryer,
```

```
Dedicated workspace, Iron, Free parking on premises, Shampoo, Fire extinguisher, Hangers, Cable TV, Essentia
ls, Washer, Long term stays allowed, TV, Air conditioning, Elevator]
 verifications_array                                | [email, phone, reviews, offline_government_id, selfie, g
overnment_id, identity_manual]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,17],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                         | 2010
 log_price                                          | 4.836281906951478
 host_listings_count_imp                            | 2
 host_total_listings_count_imp                      | 2
 accommodates_imp                                   | 4
 bedrooms_imp                                       | 1
 beds_imp                                           | 1
 minimum_nights_imp                                 | 90
 maximum_nights_imp                                 | 365
 minimum_minimum_nights_imp                         | 90
 maximum_minimum_nights_imp                         | 90
 minimum_nights_avg_ntm_imp                         | 90.0
 calculated_host_listings_count_imp                 | 2
 calculated_host_listings_count_entire_homes_imp    | 2
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 0.0
 property_indexed_imp                               | 1.0
 room_indexed_imp                                   | 0.0
 has_tv                                             | true
-RECORD 8-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------
 id                                                 | 72659
 host_since                                         | 2011/02/03
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | N/A
 host_response_rate                                 | N/A
 host_acceptance_rate                               | 0
 host_is_superhost                                  | f
 host_neighbourhood                                 | Commercial Drive
 host_listings_count                                | 1
 host_total_listings_count                          | 1
 host_verifications                                 | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | null
 neighbourhood_cleansed                             | Grandview-Woodland
```

```
 neighbourhood_group_cleansed                     | null
 latitude                                         | 49.27357
 longitude                                        | -123.07024
 property_type                                    | Entire condominium
 room_type                                        | Entire home/apt
 accommodates                                     | 4
 bathrooms                                        | null
 bathrooms_text                                   | 1 bath
 bedrooms                                         | 2
 beds                                             | 2
 amenities                                        | ['Kitchen', 'Lake access', 'Oven', 'Paid parking off pre
mises', 'Gym', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot wate
r', 'Dishwasher', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm',
 'Heating', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Coffee
maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Hanger
s', 'Refrigerator', 'Essentials', 'Washer', 'BBQ grill', 'TV', 'Elevator']
 price                                            | 210
 minimum_nights                                   | 7
 maximum_nights                                   | 21
 minimum_minimum_nights                           | 7
 maximum_minimum_nights                           | 7
 minimum_maximum_nights                           | 21
 maximum_maximum_nights                           | 21
 minimum_nights_avg_ntm                           | 7.0
 maximum_nights_avg_ntm                           | 21.0
 has_availability                                 | t
 availability_30                                  | 0
 availability_60                                  | 0
 availability_90                                  | 0
 availability_365                                 | 40
 instant_bookable                                 | f
 calculated_host_listings_count                   | 1
 calculated_host_listings_count_entire_homes      | 1
 calculated_host_listings_count_private_rooms     | 0
 calculated_host_listings_count_shared_rooms      | 0
 neighbourhood_reduced                            | Grandview-Woodland
 property_reduced                                 | Entire condominium
 room_reduced                                     | Entire home/apt
 neighbourhood_indexed                            | 8.0
 property_indexed                                 | 1.0
 room_indexed                                     | 0.0
 amenities_array                                  | [Kitchen, Lake access, Oven, Paid parking off premises,
 Gym, Indoor fireplace, Microwave, Wifi, Dryer, Carbon monoxide alarm, Stove, Hot water, Dishwasher, Extra p
illows and blankets, Luggage dropoff allowed, Cooking basics, Smoke alarm, Heating, Single level home, Hair
 dryer, Dedicated workspace, Iron, Free street parking, Coffee maker, Bed linens, Shampoo, Dishes and silver
ware, Fire extinguisher, Patio or balcony, Hangers, Refrigerator, Essentials, Washer, BBQ grill, TV, Elevato
r]
 verifications_array                              | [email, phone, reviews, jumio, government_id]
 amenities_vector                                 | (20,[0,1,2,3,4,5,7,8,9,10,11,12,13,14,15,16,17,18,19],
```

```
[1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                        | 2011
 log_price                                         | 5.3471075307174685
 host_listings_count_imp                           | 1
 host_total_listings_count_imp                     | 1
 accommodates_imp                                  | 4
 bedrooms_imp                                      | 2
 beds_imp                                          | 2
 minimum_nights_imp                                | 7
 maximum_nights_imp                                | 21
 minimum_minimum_nights_imp                        | 7
 maximum_minimum_nights_imp                        | 7
 minimum_nights_avg_ntm_imp                        | 7.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 1
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 8.0
 property_indexed_imp                              | 1.0
 room_indexed_imp                                  | 0.0
 has_tv                                            | true
-RECORD 9-----------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                                | 77157
 host_since                                        | 2011/02/27
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | within a few hours
 host_response_rate                                | 1
 host_acceptance_rate                              | 1
 host_is_superhost                                 | f
 host_neighbourhood                                | Arbutus Ridge
 host_listings_count                               | 2
 host_total_listings_count                         | 2
 host_verifications                                | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'offl
ine_government_id', 'government_id']
 host_has_profile_pic                              | t
 host_identity_verified                            | t
 neighbourhood                                     | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                            | Arbutus Ridge
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.24675
 longitude                                         | -123.16711
 property_type                                     | Private room in guest suite
```

```
 room_type                                          | Private room
 accommodates                                       | 3
 bathrooms                                          | null
 bathrooms_text                                     | 1 private bath
 bedrooms                                           | 1
 beds                                               | 2
 amenities                                          | ['Kitchen', 'Microwave', 'Wifi', 'Laundromat nearby', 'D
ryer', 'Hot water', 'Garden or backyard', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking
 basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Mini fridge', 'Single level home', 'Hair dryer', 'L
ockbox', 'Dedicated workspace', 'First aid kit', 'Iron', 'Body soap', 'Free street parking', 'Coffee maker',
'Bed linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Barbecue utensils', 'Shower gel', 'H
angers', 'Cable TV', 'Essentials', 'Washer', 'Conditioner', 'BBQ grill', 'Beach essentials', 'TV', 'Long ter
m stays allowed']
 price                                              | 75
 minimum_nights                                     | 3
 maximum_nights                                     | 60
 minimum_minimum_nights                             | 3
 maximum_minimum_nights                             | 3
 minimum_maximum_nights                             | 1125
 maximum_maximum_nights                             | 1125
 minimum_nights_avg_ntm                             | 3.0
 maximum_nights_avg_ntm                             | 1125.0
 has_availability                                   | t
 availability_30                                    | 3
 availability_60                                    | 3
 availability_90                                    | 3
 availability_365                                   | 267
 instant_bookable                                   | t
 calculated_host_listings_count                     | 2
 calculated_host_listings_count_entire_homes        | 0
 calculated_host_listings_count_private_rooms       | 2
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | OTHER
 property_reduced                                   | OTHER
 room_reduced                                       | Private room
 neighbourhood_indexed                              | 1.0
 property_indexed                                   | 5.0
 room_indexed                                       | 1.0
 amenities_array                                    | [Kitchen, Microwave, Wifi, Laundromat nearby, Dryer, Hot
water, Garden or backyard, Extra pillows and blankets, Luggage dropoff allowed, Cooking basics, Private entr
ance, Smoke alarm, Heating, Mini fridge, Single level home, Hair dryer, Lockbox, Dedicated workspace, First
 aid kit, Iron, Body soap, Free street parking, Coffee maker, Bed linens, Shampoo, Fire extinguisher, Dishes
and silverware, Barbecue utensils, Shower gel, Hangers, Cable TV, Essentials, Washer, Conditioner, BBQ gril
l, Beach essentials, TV, Long term stays allowed]
 verifications_array                                | [email, phone, facebook, reviews, jumio, offline_governm
ent_id, government_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,16,17,19],[1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
```

```
 since_year                                          | 2011
 log_price                                           | 4.31748811353631
 host_listings_count_imp                             | 2
 host_total_listings_count_imp                       | 2
 accommodates_imp                                    | 3
 bedrooms_imp                                        | 1
 beds_imp                                            | 2
 minimum_nights_imp                                  | 3
 maximum_nights_imp                                  | 60
 minimum_minimum_nights_imp                          | 3
 maximum_minimum_nights_imp                          | 3
 minimum_nights_avg_ntm_imp                          | 3.0
 calculated_host_listings_count_imp                  | 2
 calculated_host_listings_count_entire_homes_imp     | 0
 calculated_host_listings_count_private_rooms_imp    | 2
 neighbourhood_indexed_imp                           | 1.0
 property_indexed_imp                                | 5.0
 room_indexed_imp                                    | 1.0
 has_tv                                              | true
-RECORD 10-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                                  | 111077
 host_since                                          | 2011/05/09
 host_location                                       | North Vancouver, British Columbia, Canada
 host_response_time                                  | within a few hours
 host_response_rate                                  | 1
 host_acceptance_rate                                | 0.33
 host_is_superhost                                   | f
 host_neighbourhood                                  | Grandview-Woodland
 host_listings_count                                 | 1
 host_total_listings_count                           | 1
 host_verifications                                  | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'gove
rnment_id']
 host_has_profile_pic                                | t
 host_identity_verified                              | t
 neighbourhood                                       | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                              | Grandview-Woodland
 neighbourhood_group_cleansed                        | null
 latitude                                            | 49.26726
 longitude                                           | -123.05969
 property_type                                       | Entire guest suite
 room_type                                           | Entire home/apt
 accommodates                                        | 3
```

```
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 2
 beds                                               | 2
 amenities                                          | ['Kitchen', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dr
yer', 'Stove', 'Hot water', 'Cooking basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'I
ron', 'Free street parking', 'Coffee maker', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed']
 price                                              | 125
 minimum_nights                                     | 3
 maximum_nights                                     | 365
 minimum_minimum_nights                             | 3
 maximum_minimum_nights                             | 3
 minimum_maximum_nights                             | 365
 maximum_maximum_nights                             | 365
 minimum_nights_avg_ntm                             | 3.0
 maximum_nights_avg_ntm                             | 365.0
 has_availability                                   | t
 availability_30                                    | 4
 availability_60                                    | 34
 availability_90                                    | 64
 availability_365                                   | 339
 instant_bookable                                   | f
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 1
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Grandview-Woodland
 property_reduced                                   | Entire guest suite
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 8.0
 property_indexed                                   | 4.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Indoor fireplace, Microwave, Wifi, Dryer, Stov
e, Hot water, Cooking basics, Private entrance, Smoke alarm, Heating, Hair dryer, Iron, Free street parking,
Coffee maker, Keypad, Shampoo, Dishes and silverware, Fire extinguisher, Hangers, Refrigerator, Essentials,
 Washer, Long term stays allowed]
 verifications_array                                | [email, phone, facebook, reviews, jumio, government_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,12,13,14,16,17,18,19],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                         | 2011
 log_price                                          | 4.8283137373023015
 host_listings_count_imp                            | 1
 host_total_listings_count_imp                      | 1
 accommodates_imp                                   | 3
 bedrooms_imp                                       | 2
 beds_imp                                           | 2
 minimum_nights_imp                                 | 3
```

```
 maximum_nights_imp                                | 365
 minimum_minimum_nights_imp                        | 3
 maximum_minimum_nights_imp                        | 3
 minimum_nights_avg_ntm_imp                        | 3.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 1
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 8.0
 property_indexed_imp                              | 4.0
 room_indexed_imp                                  | 0.0
 has_tv                                            | false
-RECORD 11-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
 id                                                | 125176
 host_since                                        | 2011/05/24
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | within an hour
 host_response_rate                                | 1
 host_acceptance_rate                              | 1
 host_is_superhost                                 | f
 host_neighbourhood                                | Riley Park
 host_listings_count                               | 1
 host_total_listings_count                         | 1
 host_verifications                                | ['email', 'phone', 'facebook', 'reviews', 'offline_gover
nment_id', 'government_id', 'work_email']
 host_has_profile_pic                              | t
 host_identity_verified                            | t
 neighbourhood                                     | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                            | Riley Park
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.24648
 longitude                                         | -123.09241
 property_type                                     | Private room in house
 room_type                                         | Private room
 accommodates                                      | 3
 bathrooms                                         | null
 bathrooms_text                                    | 1 private bath
 bedrooms                                          | 1
 beds                                              | 0
 amenities                                         | ['Kitchen', 'Hot tub', 'Oven', 'Wifi', 'Carbon monoxide
 alarm', 'Stove', 'Hot water', 'Garden or backyard', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair drye
r', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Free parking on premises', 'Keypad',
 'Shampoo', 'Dishes and silverware', 'Patio or balcony', 'Hangers', 'Cable TV', 'Refrigerator', 'Essential
```

```
s', 'Long term stays allowed', 'TV']
 price                                          | 75
 minimum_nights                                 | 2
 maximum_nights                                 | 60
 minimum_minimum_nights                         | 2
 maximum_minimum_nights                         | 2
 minimum_maximum_nights                         | 1125
 maximum_maximum_nights                         | 1125
 minimum_nights_avg_ntm                         | 2.0
 maximum_nights_avg_ntm                         | 1125.0
 has_availability                               | t
 availability_30                                | 0
 availability_60                                | 0
 availability_90                                | 2
 availability_365                               | 268
 instant_bookable                               | t
 calculated_host_listings_count                 | 1
 calculated_host_listings_count_entire_homes    | 0
 calculated_host_listings_count_private_rooms   | 1
 calculated_host_listings_count_shared_rooms    | 0
 neighbourhood_reduced                          | Riley Park
 property_reduced                               | Private room in house
 room_reduced                                   | Private room
 neighbourhood_indexed                          | 7.0
 property_indexed                               | 3.0
 room_indexed                                   | 1.0
 amenities_array                                | [Kitchen, Hot tub, Oven, Wifi, Carbon monoxide alarm, St
ove, Hot water, Garden or backyard, Cooking basics, Smoke alarm, Heating, Hair dryer, Dedicated workspace, I
ron, Free street parking, Bathtub, Free parking on premises, Keypad, Shampoo, Dishes and silverware, Patio o
r balcony, Hangers, Cable TV, Refrigerator, Essentials, Long term stays allowed, TV]
 verifications_array                            | [email, phone, facebook, reviews, offline_government_id,
government_id, work_email]
 amenities_vector                               | (20,[0,1,2,3,4,5,6,9,10,11,12,13,14,15,16,18,19],[1.0,1.
0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                           | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                     | 2011
 log_price                                      | 4.31748811353631
 host_listings_count_imp                        | 1
 host_total_listings_count_imp                  | 1
 accommodates_imp                               | 3
 bedrooms_imp                                   | 1
 beds_imp                                       | 0
 minimum_nights_imp                             | 2
 maximum_nights_imp                             | 60
 minimum_minimum_nights_imp                     | 2
 maximum_minimum_nights_imp                     | 2
 minimum_nights_avg_ntm_imp                     | 2.0
 calculated_host_listings_count_imp             | 1
 calculated_host_listings_count_entire_homes_imp | 0
```

```
 calculated_host_listings_count_private_rooms_imp | 1
 neighbourhood_indexed_imp                        | 7.0
 property_indexed_imp                             | 3.0
 room_indexed_imp                                 | 1.0
 has_tv                                           | true
-RECORD 12-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
 id                                               | 129163
 host_since                                       | 2011/05/17
 host_location                                    | Vancouver, British Columbia, Canada
 host_response_time                               | within a few hours
 host_response_rate                               | 1
 host_acceptance_rate                             | 1
 host_is_superhost                                | f
 host_neighbourhood                               | Renfrew-Collingwood
 host_listings_count                              | 2
 host_total_listings_count                        | 2
 host_verifications                               | ['email', 'phone', 'reviews', 'jumio', 'offline_governme
nt_id', 'government_id']
 host_has_profile_pic                             | t
 host_identity_verified                           | t
 neighbourhood                                    | null
 neighbourhood_cleansed                           | Renfrew-Collingwood
 neighbourhood_group_cleansed                     | null
 latitude                                         | 49.24507
 longitude                                        | -123.04941
 property_type                                    | Private room in house
 room_type                                        | Private room
 accommodates                                     | 1
 bathrooms                                        | null
 bathrooms_text                                   | 1 shared bath
 bedrooms                                         | 1
 beds                                             | 1
 amenities                                        | ['Kitchen', 'Hot water', 'Paid parking on premises', 'Lo
ck on bedroom door', 'Free street parking', 'Cable TV', 'Smoke alarm', 'Wifi', 'Essentials', 'Heating', 'Was
her', 'Dryer', 'Carbon monoxide alarm', 'Dishes and silverware', 'TV', 'Long term stays allowed', 'Host gree
ts you', 'Hair dryer', 'Shampoo', 'Fire extinguisher']
 price                                            | 45
 minimum_nights                                   | 30
 maximum_nights                                   | 280
 minimum_minimum_nights                           | 30
 maximum_minimum_nights                           | 30
 minimum_maximum_nights                           | 280
```

```
 maximum_maximum_nights                            | 280
 minimum_nights_avg_ntm                            | 30.0
 maximum_nights_avg_ntm                            | 280.0
 has_availability                                  | t
 availability_30                                   | 30
 availability_60                                   | 60
 availability_90                                   | 90
 availability_365                                  | 365
 instant_bookable                                  | t
 calculated_host_listings_count                    | 1
 calculated_host_listings_count_entire_homes       | 0
 calculated_host_listings_count_private_rooms      | 1
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | Renfrew-Collingwood
 property_reduced                                  | Private room in house
 room_reduced                                      | Private room
 neighbourhood_indexed                             | 11.0
 property_indexed                                  | 3.0
 room_indexed                                      | 1.0
 amenities_array                                   | [Kitchen, Hot water, Paid parking on premises, Lock on b
edroom door, Free street parking, Cable TV, Smoke alarm, Wifi, Essentials, Heating, Washer, Dryer, Carbon mo
noxide alarm, Dishes and silverware, TV, Long term stays allowed, Host greets you, Hair dryer, Shampoo, Fire
extinguisher]
 verifications_array                               | [email, phone, reviews, jumio, offline_government_id, go
vernment_id]
 amenities_vector                                  | (20,[0,1,2,3,4,6,7,8,9,10,12,13,15,16,17],[1.0,1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                        | 2011
 log_price                                         | 3.8066624897703196
 host_listings_count_imp                           | 2
 host_total_listings_count_imp                     | 2
 accommodates_imp                                  | 1
 bedrooms_imp                                      | 1
 beds_imp                                          | 1
 minimum_nights_imp                                | 30
 maximum_nights_imp                                | 280
 minimum_minimum_nights_imp                        | 30
 maximum_minimum_nights_imp                        | 30
 minimum_nights_avg_ntm_imp                        | 30.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 0
 calculated_host_listings_count_private_rooms_imp  | 1
 neighbourhood_indexed_imp                         | 11.0
 property_indexed_imp                              | 3.0
 room_indexed_imp                                  | 1.0
 has_tv                                            | true
-RECORD 13-------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
```

```
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
 id                                                  | 174135
 host_since                                          | 2011/07/17
 host_location                                       | Vancouver, British Columbia, Canada
 host_response_time                                  | within a few hours
 host_response_rate                                  | 1
 host_acceptance_rate                                | 0
 host_is_superhost                                   | f
 host_neighbourhood                                  | Downtown Eastside
 host_listings_count                                 | 1
 host_total_listings_count                           | 1
 host_verifications                                  | ['email', 'phone', 'reviews', 'offline_government_id',
 'selfie', 'government_id', 'identity_manual']
 host_has_profile_pic                                | t
 host_identity_verified                              | t
 neighbourhood                                       | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                              | Downtown Eastside
 neighbourhood_group_cleansed                        | null
 latitude                                            | 49.27782
 longitude                                           | -123.08985
 property_type                                       | Entire guest suite
 room_type                                           | Entire home/apt
 accommodates                                        | 2
 bathrooms                                           | null
 bathrooms_text                                      | 1 bath
 bedrooms                                            | 1
 beds                                                | 2
 amenities                                           | ['Kitchen', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'H
ot water', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated
workspace', 'Iron', 'Free street parking', 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Hangers', 'Cable T
V', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning']
 price                                               | 100
 minimum_nights                                      | 30
 maximum_nights                                      | 365
 minimum_minimum_nights                              | 30
 maximum_minimum_nights                              | 30
 minimum_maximum_nights                              | 365
 maximum_maximum_nights                              | 365
 minimum_nights_avg_ntm                              | 30.0
 maximum_nights_avg_ntm                              | 365.0
 has_availability                                    | t
 availability_30                                     | 0
 availability_60                                     | 0
 availability_90                                     | 0
```

```
 availability_365                                  | 224
 instant_bookable                                  | f
 calculated_host_listings_count                    | 1
 calculated_host_listings_count_entire_homes       | 1
 calculated_host_listings_count_private_rooms      | 0
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | Downtown Eastside
 property_reduced                                  | Entire guest suite
 room_reduced                                      | Entire home/apt
 neighbourhood_indexed                             | 6.0
 property_indexed                                  | 4.0
 room_indexed                                      | 0.0
 amenities_array                                   | [Kitchen, Wifi, Dryer, Carbon monoxide alarm, Hot water,
Luggage dropoff allowed, Private entrance, Smoke alarm, Heating, Hair dryer, Dedicated workspace, Iron, Free
street parking, Bed linens, Shampoo, Fire extinguisher, Hangers, Cable TV, Essentials, Washer, Long term sta
ys allowed, TV, Air conditioning]
 verifications_array                               | [email, phone, reviews, offline_government_id, selfie, g
overnment_id, identity_manual]
 amenities_vector                                  | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,17],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                        | 2011
 log_price                                         | 4.605170185988092
 host_listings_count_imp                           | 1
 host_total_listings_count_imp                     | 1
 accommodates_imp                                  | 2
 bedrooms_imp                                      | 1
 beds_imp                                          | 2
 minimum_nights_imp                                | 30
 maximum_nights_imp                                | 365
 minimum_minimum_nights_imp                        | 30
 maximum_minimum_nights_imp                        | 30
 minimum_nights_avg_ntm_imp                        | 30.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 1
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 6.0
 property_indexed_imp                              | 4.0
 room_indexed_imp                                  | 0.0
 has_tv                                            | true
-RECORD 14-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
 id                                                | 200643
```

```
 host_since                                          | 2011/08/16
 host_location                                       | Vancouver, British Columbia, Canada
 host_response_time                                  | within an hour
 host_response_rate                                  | 1
 host_acceptance_rate                                | 1
 host_is_superhost                                   | f
 host_neighbourhood                                  | Strathcona
 host_listings_count                                 | 4
 host_total_listings_count                           | 4
 host_verifications                                  | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                                | t
 host_identity_verified                              | t
 neighbourhood                                       | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                              | Downtown Eastside
 neighbourhood_group_cleansed                        | null
 latitude                                            | 49.27384
 longitude                                           | -123.08289
 property_type                                       | Private room in house
 room_type                                           | Private room
 accommodates                                        | 4
 bathrooms                                           | null
 bathrooms_text                                      | 2 shared baths
 bedrooms                                            | 1
 beds                                                | 1
 amenities                                           | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wi
fi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Garden or backyard', 'Dishwasher', 'Extra pill
ows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Room-darkening shades', 'Smoke alarm', 'Hea
ting', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Free street parking', 'Coffee maker',
 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Essentia
ls', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
 price                                               | 87
 minimum_nights                                      | 1
 maximum_nights                                      | 1125
 minimum_minimum_nights                              | 1
 maximum_minimum_nights                              | 1
 minimum_maximum_nights                              | 1125
 maximum_maximum_nights                              | 1125
 minimum_nights_avg_ntm                              | 1.0
 maximum_nights_avg_ntm                              | 1125.0
 has_availability                                    | t
 availability_30                                     | 30
 availability_60                                     | 60
 availability_90                                     | 90
 availability_365                                    | 365
 instant_bookable                                    | t
 calculated_host_listings_count                      | 4
 calculated_host_listings_count_entire_homes         | 0
 calculated_host_listings_count_private_rooms        | 4
 calculated_host_listings_count_shared_rooms         | 0
```

```
 neighbourhood_reduced                              | Downtown Eastside
 property_reduced                                   | Private room in house
 room_reduced                                       | Private room
 neighbourhood_indexed                              | 6.0
 property_indexed                                   | 3.0
 room_indexed                                       | 1.0
 amenities_array                                    | [Kitchen, Oven, Indoor fireplace, Microwave, Wifi, Drye
r, Carbon monoxide alarm, Stove, Hot water, Garden or backyard, Dishwasher, Extra pillows and blankets, Lugg
age dropoff allowed, Cooking basics, Room-darkening shades, Smoke alarm, Heating, Hair dryer, Dedicated work
space, First aid kit, Iron, Free street parking, Coffee maker, Bed linens, Shampoo, Fire extinguisher, Dishe
s and silverware, Hangers, Refrigerator, Essentials, Washer, Long term stays allowed, Pack \u2019n Play/trav
el crib]
 verifications_array                                | [email, phone, reviews, jumio, government_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,11,12,13,14,15,16,17,18,19],[1.
0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                         | 2011
 log_price                                          | 4.465908118654584
 host_listings_count_imp                            | 4
 host_total_listings_count_imp                      | 4
 accommodates_imp                                   | 4
 bedrooms_imp                                       | 1
 beds_imp                                           | 1
 minimum_nights_imp                                 | 1
 maximum_nights_imp                                 | 1125
 minimum_minimum_nights_imp                         | 1
 maximum_minimum_nights_imp                         | 1
 minimum_nights_avg_ntm_imp                         | 1.0
 calculated_host_listings_count_imp                 | 4
 calculated_host_listings_count_entire_homes_imp    | 0
 calculated_host_listings_count_private_rooms_imp   | 4
 neighbourhood_indexed_imp                          | 6.0
 property_indexed_imp                               | 3.0
 room_indexed_imp                                   | 1.0
 has_tv                                             | false
-RECORD 15-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                                 | 248014
 host_since                                         | 2011/10/16
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within a day
 host_response_rate                                 | 0.5
 host_acceptance_rate                               | N/A
```

```
 host_is_superhost                                  | f
 host_neighbourhood                                 | Grandview-Woodland
 host_listings_count                                | 1
 host_total_listings_count                          | 1
 host_verifications                                 | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Grandview-Woodland
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.29103
 longitude                                          | -123.05782
 property_type                                      | Entire house
 room_type                                          | Entire home/apt
 accommodates                                       | 3
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 2
 beds                                               | 0
 amenities                                          | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wi
fi', 'Host greets you', 'Dryer', 'Stove', 'Carbon monoxide alarm', 'Hot water', 'Garden or backyard', 'Dishw
asher', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'I
ron', 'Free street parking', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Cable TV',
 'Refrigerator', 'Hangers', 'Essentials', 'Washer', 'Long term stays allowed', 'BBQ grill', 'TV', 'Cleaning
 before checkout']
 price                                              | 100
 minimum_nights                                     | 31
 maximum_nights                                     | 180
 minimum_minimum_nights                             | 31
 maximum_minimum_nights                             | 31
 minimum_maximum_nights                             | 180
 maximum_maximum_nights                             | 180
 minimum_nights_avg_ntm                             | 31.0
 maximum_nights_avg_ntm                             | 180.0
 has_availability                                   | t
 availability_30                                    | 4
 availability_60                                    | 4
 availability_90                                    | 4
 availability_365                                   | 113
 instant_bookable                                   | f
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 1
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Grandview-Woodland
 property_reduced                                   | Entire house
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 8.0
 property_indexed                                   | 2.0
```

```
 room_indexed                                      | 0.0
 amenities_array                                   | [Kitchen, Oven, Indoor fireplace, Microwave, Wifi, Host
 greets you, Dryer, Stove, Carbon monoxide alarm, Hot water, Garden or backyard, Dishwasher, Cooking basics,
Smoke alarm, Heating, Hair dryer, Dedicated workspace, First aid kit, Iron, Free street parking, Dishes and
 silverware, Fire extinguisher, Patio or balcony, Cable TV, Refrigerator, Hangers, Essentials, Washer, Long
 term stays allowed, BBQ grill, TV, Cleaning before checkout]
 verifications_array                               | [email, phone, reviews, jumio, government_id]
 amenities_vector                                  | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,14,15,16,17,18,19],[1.
0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                        | 2011
 log_price                                         | 4.605170185988092
 host_listings_count_imp                           | 1
 host_total_listings_count_imp                     | 1
 accommodates_imp                                  | 3
 bedrooms_imp                                      | 2
 beds_imp                                          | 0
 minimum_nights_imp                                | 31
 maximum_nights_imp                                | 180
 minimum_minimum_nights_imp                        | 31
 maximum_minimum_nights_imp                        | 31
 minimum_nights_avg_ntm_imp                        | 31.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 1
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 8.0
 property_indexed_imp                              | 2.0
 room_indexed_imp                                  | 0.0
 has_tv                                            | true
-RECORD 16-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                                | 258636
 host_since                                        | 2011/08/09
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | N/A
 host_response_rate                                | N/A
 host_acceptance_rate                              | 1
 host_is_superhost                                 | f
 host_neighbourhood                                | Downtown Vancouver
 host_listings_count                               | 1
 host_total_listings_count                         | 1
 host_verifications                                | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                              | t
```

```
 host_identity_verified                          | t
 neighbourhood                                   | null
 neighbourhood_cleansed                          | Downtown
 neighbourhood_group_cleansed                    | null
 latitude                                        | 49.27796
 longitude                                       | -123.12671
 property_type                                   | Entire apartment
 room_type                                       | Entire home/apt
 accommodates                                    | 4
 bathrooms                                       | null
 bathrooms_text                                  | 1 bath
 bedrooms                                        | 1
 beds                                            | 2
 amenities                                       | ['Kitchen', 'Iron', 'Elevator', 'Gym', 'Hangers', 'Smoke
alarm', 'Wifi', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Free parking on premises', 'Long term stays all
owed', 'Hair dryer', 'Shampoo', 'Dedicated workspace']
 price                                           | 99
 minimum_nights                                  | 120
 maximum_nights                                  | 365
 minimum_minimum_nights                          | 120
 maximum_minimum_nights                          | 120
 minimum_maximum_nights                          | 365
 maximum_maximum_nights                          | 365
 minimum_nights_avg_ntm                          | 120.0
 maximum_nights_avg_ntm                          | 365.0
 has_availability                                | t
 availability_30                                 | 0
 availability_60                                 | 11
 availability_90                                 | 41
 availability_365                                | 316
 instant_bookable                                | f
 calculated_host_listings_count                  | 1
 calculated_host_listings_count_entire_homes     | 1
 calculated_host_listings_count_private_rooms    | 0
 calculated_host_listings_count_shared_rooms     | 0
 neighbourhood_reduced                           | Downtown
 property_reduced                                | Entire apartment
 room_reduced                                    | Entire home/apt
 neighbourhood_indexed                           | 0.0
 property_indexed                                | 0.0
 room_indexed                                    | 0.0
 amenities_array                                 | [Kitchen, Iron, Elevator, Gym, Hangers, Smoke alarm, Wif
i, Essentials, Heating, Washer, Dryer, Free parking on premises, Long term stays allowed, Hair dryer, Shampo
o, Dedicated workspace]
 verifications_array                             | [email, phone, reviews, jumio, government_id]
 amenities_vector                                | (20,[0,1,2,3,4,5,6,7,8,9,11,13,14],[1.0,1.0,1.0,1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                            | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                      | 2011
```

```
 log_price                                          | 4.59511985013459
 host_listings_count_imp                            | 1
 host_total_listings_count_imp                      | 1
 accommodates_imp                                   | 4
 bedrooms_imp                                       | 1
 beds_imp                                           | 2
 minimum_nights_imp                                 | 120
 maximum_nights_imp                                 | 365
 minimum_minimum_nights_imp                         | 120
 maximum_minimum_nights_imp                         | 120
 minimum_nights_avg_ntm_imp                         | 120.0
 calculated_host_listings_count_imp                 | 1
 calculated_host_listings_count_entire_homes_imp    | 1
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 0.0
 property_indexed_imp                               | 0.0
 room_indexed_imp                                   | 0.0
 has_tv                                             | false
-RECORD 17-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                                 | 283841
 host_since                                         | 2009/08/10
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within a day
 host_response_rate                                 | 0.93
 host_acceptance_rate                               | 0.88
 host_is_superhost                                  | f
 host_neighbourhood                                 | Coal Harbour
 host_listings_count                                | 49
 host_total_listings_count                          | 49
 host_verifications                                 | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | null
 neighbourhood_cleansed                             | Downtown
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.2872
 longitude                                          | -123.12352
 property_type                                      | Entire apartment
 room_type                                          | Entire home/apt
 accommodates                                       | 6
 bathrooms                                          | null
 bathrooms_text                                     | 3 baths
```

```
 bedrooms                                          | 3
 beds                                              | 3
 amenities                                         | ['Kitchen', 'Hot tub', 'Oven', 'Pool', 'Gym', 'Indoor fi
replace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Dishwasher', 'Smoke
 alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Coffee maker',
'Dishes and silverware', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays all
owed', 'TV', 'Air conditioning']
 price                                             | 197
 minimum_nights                                    | 90
 maximum_nights                                    | 1125
 minimum_minimum_nights                            | 90
 maximum_minimum_nights                            | 90
 minimum_maximum_nights                            | 1125
 maximum_maximum_nights                            | 1125
 minimum_nights_avg_ntm                            | 90.0
 maximum_nights_avg_ntm                            | 1125.0
 has_availability                                  | t
 availability_30                                   | 0
 availability_60                                   | 0
 availability_90                                   | 0
 availability_365                                  | 0
 instant_bookable                                  | f
 calculated_host_listings_count                    | 41
 calculated_host_listings_count_entire_homes       | 41
 calculated_host_listings_count_private_rooms      | 0
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | Downtown
 property_reduced                                  | Entire apartment
 room_reduced                                      | Entire home/apt
 neighbourhood_indexed                             | 0.0
 property_indexed                                  | 0.0
 room_indexed                                      | 0.0
 amenities_array                                   | [Kitchen, Hot tub, Oven, Pool, Gym, Indoor fireplace, Mi
crowave, Wifi, Dryer, Carbon monoxide alarm, Stove, Hot water, Dishwasher, Smoke alarm, Heating, Hair dryer,
Dedicated workspace, Iron, Free parking on premises, Coffee maker, Dishes and silverware, Hangers, Cable TV,
Refrigerator, Essentials, Washer, Long term stays allowed, TV, Air conditioning]
 verifications_array                               | [email, phone, reviews, jumio, government_id]
 amenities_vector                                  | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,14,15,16,18],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                        | 2009
 log_price                                         | 5.2832037287379885
 host_listings_count_imp                           | 49
 host_total_listings_count_imp                     | 49
 accommodates_imp                                  | 6
 bedrooms_imp                                      | 3
 beds_imp                                          | 3
 minimum_nights_imp                                | 90
 maximum_nights_imp                                | 1125
```

```
 minimum_minimum_nights_imp                       | 90
 maximum_minimum_nights_imp                       | 90
 minimum_nights_avg_ntm_imp                       | 90.0
 calculated_host_listings_count_imp               | 41
 calculated_host_listings_count_entire_homes_imp  | 41
 calculated_host_listings_count_private_rooms_imp | 0
 neighbourhood_indexed_imp                        | 0.0
 property_indexed_imp                             | 0.0
 room_indexed_imp                                 | 0.0
 has_tv                                           | true
-RECORD 18-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
 id                                               | 323090
 host_since                                       | 2012/01/01
 host_location                                    | Vancouver, British Columbia, Canada
 host_response_time                               | N/A
 host_response_rate                               | N/A
 host_acceptance_rate                             | 0.85
 host_is_superhost                                | t
 host_neighbourhood                               | Downtown Vancouver
 host_listings_count                              | 1
 host_total_listings_count                        | 1
 host_verifications                               | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                             | t
 host_identity_verified                           | t
 neighbourhood                                    | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                           | Downtown
 neighbourhood_group_cleansed                     | null
 latitude                                         | 49.28206
 longitude                                        | -123.12461
 property_type                                    | Entire apartment
 room_type                                        | Entire home/apt
 accommodates                                     | 2
 bathrooms                                        | null
 bathrooms_text                                   | 1 bath
 bedrooms                                         | 1
 beds                                             | 1
 amenities                                        | ['Kitchen', 'Cooking basics', 'Iron', 'Elevator', 'Gym',
'Hangers', 'Cable TV', 'Smoke alarm', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Carbon monoxide alarm',
 'Wifi', 'TV', 'Long term stays allowed', 'Hair dryer', 'Shampoo', 'Dishes and silverware', 'Dedicated works
pace']
 price                                            | 168
 minimum_nights                                   | 180
```

maximum_nights	180
minimum_minimum_nights	180
maximum_minimum_nights	180
minimum_maximum_nights	180
maximum_maximum_nights	180
minimum_nights_avg_ntm	180.0
maximum_nights_avg_ntm	180.0
has_availability	t
availability_30	30
availability_60	60
availability_90	90
availability_365	141
instant_bookable	f
calculated_host_listings_count	1
calculated_host_listings_count_entire_homes	1
calculated_host_listings_count_private_rooms	0
calculated_host_listings_count_shared_rooms	0
neighbourhood_reduced	Downtown
property_reduced	Entire apartment
room_reduced	Entire home/apt
neighbourhood_indexed	0.0
property_indexed	0.0
room_indexed	0.0
amenities_array	[Kitchen, Cooking basics, Iron, Elevator, Gym, Hangers, Cable TV, Smoke alarm, Essentials, Heating, Washer, Dryer, Carbon monoxide alarm, Wifi, TV, Long term stays allowed, Hair dryer, Shampoo, Dishes and silverware, Dedicated workspace]
verifications_array	[email, phone, reviews, jumio, government_id]
amenities_vector	(20,[0,1,2,3,4,5,6,7,8,9,10,11,13,14,15,16,19],[1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
verifications_vector	(5,[0,1,2,4],[1.0,1.0,1.0,1.0])
since_year	2012
log_price	5.123963979403259
host_listings_count_imp	1
host_total_listings_count_imp	1
accommodates_imp	2
bedrooms_imp	1
beds_imp	1
minimum_nights_imp	180
maximum_nights_imp	180
minimum_minimum_nights_imp	180
maximum_minimum_nights_imp	180
minimum_nights_avg_ntm_imp	180.0
calculated_host_listings_count_imp	1
calculated_host_listings_count_entire_homes_imp	1
calculated_host_listings_count_private_rooms_imp	0
neighbourhood_indexed_imp	0.0
property_indexed_imp	0.0
room_indexed_imp	0.0
has_tv	true

```
-RECORD 19-----------------------
```

```
 availability_60                                    | 59
 availability_90                                    | 89
 availability_365                                   | 364
 instant_bookable                                   | f
 calculated_host_listings_count                     | 7
 calculated_host_listings_count_entire_homes        | 3
 calculated_host_listings_count_private_rooms       | 4
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Downtown
 property_reduced                                   | Private room in apartment
 room_reduced                                       | Private room
 neighbourhood_indexed                              | 0.0
 property_indexed                                   | 6.0
 room_indexed                                       | 1.0
 amenities_array                                    | [Kitchen, Lake access, Oven, Paid parking on premises, P
aid parking off premises, Gym, Microwave, Wifi, Dryer, Stove, Hot water, Dishwasher, Extra pillows and blank
ets, Cooking basics, Heating, Lock on bedroom door, Hair dryer, Iron, Coffee maker, Bed linens, Pocket wifi,
Shampoo, Dishes and silverware, Hangers, Refrigerator, Essentials, Washer, Long term stays allowed, Elevato
r]
 verifications_array                                | [email, phone, facebook, reviews]
 amenities_vector                                   | (20,[0,1,3,4,5,6,7,8,9,12,13,14,16,18,19],[1.0,1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,4],[1.0,1.0,1.0])
 since_year                                         | 2011
 log_price                                          | 5.480638923341991
 host_listings_count_imp                            | 8
 host_total_listings_count_imp                      | 8
 accommodates_imp                                   | 2
 bedrooms_imp                                       | 2
 beds_imp                                           | 1
 minimum_nights_imp                                 | 30
 maximum_nights_imp                                 | 1125
 minimum_minimum_nights_imp                         | 30
 maximum_minimum_nights_imp                         | 30
 minimum_nights_avg_ntm_imp                         | 30.0
 calculated_host_listings_count_imp                 | 7
 calculated_host_listings_count_entire_homes_imp    | 3
 calculated_host_listings_count_private_rooms_imp   | 4
 neighbourhood_indexed_imp                          | 0.0
 property_indexed_imp                               | 6.0
 room_indexed_imp                                   | 1.0
 has_tv                                             | false
only showing top 20 rows
```

In [60]:

```
assembler = VectorAssembler(
    inputCols=imp_columns + ["has_tv", "has_wifi"],
```

```
    outputCol="features", handleInvalid="error")

pipeline = Pipeline(stages=new_stages + [assembler]).fit(dataset)
df = pipeline.transform(dataset)

rf = RandomForestRegressor(featuresCol="features", labelCol='log_price', seed=42)
model = rf.fit(df)

df_devset = pipeline.transform(devset)
predictions = model.transform(df_devset)

predictions.select("features","log_price", "prediction").show(5)

metrics = RegressionMetrics(predictions.select("prediction", "log_price").rdd\
                            .map(lambda r: (math.exp(r['prediction']), math.exp(r['log_price']))))

print("base:", baseError)
selectedError = metrics.meanAbsoluteError
selectedError
```

```
+--------------------+------------------+-----------------+
|            features|         log_price|       prediction|
+--------------------+------------------+-----------------+
|[49.0,49.0,5.0,2....|5.0106352940962555|5.060099291695992|
|[1.0,1.0,4.0,1.0,...| 5.293304824724492|4.865733059150112|
|[5.0,5.0,6.0,3.0,...| 5.501258210544727|5.348326732154461|
|[1.0,1.0,2.0,1.0,...| 5.225746673713202|4.724549844833268|
|[1.0,1.0,2.0,1.0,...| 4.852030263919617|4.604958242838359|
+--------------------+------------------+-----------------+
only showing top 5 rows

base: 56.83296858469796
```

Out[60]:

```
57.18225268896161
```

That didn't work that well. Random forests are very robust, so there is no boost in the signal by dropping the amenities.

Let us take a look at the feature importance again.

In [61]:

```
labels = columns + ["has_tv", "has_wifi"]

importances = [ (model.featureImportances[idx], lbl) for idx, lbl in enumerate(labels) ]
table = '<table><tr><th>Rank</th><th>Feature</th><th>Importance</th></tr>'
for idx, pair in enumerate(sorted(importances, reverse=True)):
    score, lbl = pair
    table = '{}<tr><td>{}</td><td>{}</td><td style="text-align: left">{}</td></tr>'.format(table, idx+1, lbl
```

```
table += '</table>'
display(HTML(table))
```

Rank	Feature	Importance
1	bedrooms	0.20267493002987597
2	room_indexed	0.19087475710575744
3	accommodates	0.18591641207704024
4	beds	0.13877602324509647
5	calculated_host_listings_count_private_rooms	0.05974810554598172
6	calculated_host_listings_count_entire_homes	0.04394849805976533
7	property_indexed	0.03464265882956363
8	neighbourhood_indexed	0.02871239631489586
9	minimum_nights	0.01828477397852655
10	has_wifi	0.016757682036745353
11	minimum_nights_avg_ntm	0.015937450249146064
12	maximum_minimum_nights	0.01383781614482204
13	host_listings_count	0.012508879206775864
14	minimum_minimum_nights	0.010560372225586371
15	calculated_host_listings_count	0.010178819411908919
16	host_total_listings_count	0.006227356047715361
17	has_tv	0.005643951966945448
18	maximum_nights	0.004769117523851434

Now, the neighbourhood where a property resides seems to be as important as having Wifi. This is intuitively wrong but we know most of our training properties are downtown. It is possible the neighbourhood is not being as useful as we would like because the value of a location is regarding its distance, for example to SkyTrain stations (Vancouver's elevated subway alternative).

Expanding using external data: SkyTrain stations

For that, we will have to go back to the latitude and longitude columns in the raw data and compute a rough aproximation of the distance to the closest SkyTrain station for them, using OpenStreet map data.

In [62]:

```
stops = list()
with open("stops.tsv") as s:
    for row in s:
        stops.append(tuple(map(float,row.split("\t"))))
stops
```

Out[62]:

```
[(49.2591619, -122.9639466),
 (49.2533813, -122.9181633),
 (49.2484049, -122.8968898),
 (49.1895232, -122.847929),
 (49.1827396, -122.84468),
 (49.2297489, -123.0127361),
 (49.2483091, -123.0559279),
 (49.2441271, -123.045876),
 (49.2257767, -123.0038962),
 (49.2200544, -122.9885247),
 (49.2121895, -122.9593092),
 (49.2000809, -122.9488866),
 (49.2013864, -122.9126807),
 (49.2047284, -122.9062802),
 (49.2534613, -122.918123),
 (49.2793597, -123.1093495),
 (49.2861459, -123.1120211),
 (49.2484747, -122.8968053),
 (49.2618723, -123.069195),
 (49.2629063, -123.0685311),
 (49.2649587, -123.013462),
 (49.1956136, -123.1260625),
 (49.1967024, -123.1464206),
 (49.1839068, -123.136362),
 (49.1747215, -123.1365025),
 (49.2611963, -122.8901154),
 (49.1930946, -123.1579202),
 (49.2043891, -122.8741208),
 (49.224476, -122.8895687),
 (49.2384152, -123.0319039),
 (49.2657805, -123.078985),
 (49.2730906, -123.1003413),
 (49.2333355, -122.8827927),
 (49.2545929, -122.9390127),
 (49.2647152, -122.98211),
 (49.2663735, -123.001671),
 (49.2607565, -123.0328558),
 (49.258881, -123.045543),
 (49.2747473, -122.8004037),
 (49.1681288, -123.136343),
```

```
 (49.2778955, -122.8449182),
 (49.2771345, -122.8277649),
 (49.2803533, -122.7940182),
 (49.2853142, -122.7917096),
 (49.2096206, -123.117002),
 (49.2264759, -123.1165034),
 (49.2330037, -123.1162801),
 (49.2489773, -123.1154122),
 (49.2629029, -123.1148243),
 (49.2666036, -123.1154558),
 (49.2742425, -123.1217911),
 (49.2820749, -123.1190169),
 (49.2856579, -123.1134773),
 (49.2829007, -123.1151647),
 (49.2852909, -123.1189344),
 (49.1942258, -123.1784147),
 (49.1989111, -122.8506765)]
```

In [63]:

```
gps_point = VectorAssembler(inputCols=["latitude", "longitude" ], outputCol="gps")
stages.append(gps_point)
```

In [64]:

```
gps_point.transform(df).show(truncate=False,vertical=True)
```

```
-RECORD 0-----------------------
```

```
 neighbourhood_cleansed                          | West End
 neighbourhood_group_cleansed                    | null
 latitude                                        | 49.28209
 longitude                                       | -123.12515
 property_type                                   | Entire condominium
 room_type                                       | Entire home/apt
 accommodates                                    | 2
 bathrooms                                       | null
 bathrooms_text                                  | 1 bath
 bedrooms                                        | 1
 beds                                            | 1
 amenities                                       | ['Kitchen', 'Oven', 'Gym', 'Paid parking off premises',
'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Extra pillows and blankets',
'Dishwasher', 'Cooking basics', 'Hot water kettle', 'Smoke alarm', 'Heating', 'Single level home', 'Hair dry
er', 'Dedicated workspace', 'Iron', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Hange
rs', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditionin
g', 'Elevator']
 price                                           | 90
 minimum_nights                                  | 30
 maximum_nights                                  | 60
 minimum_minimum_nights                          | 30
 maximum_minimum_nights                          | 30
 minimum_maximum_nights                          | 60
 maximum_maximum_nights                          | 60
 minimum_nights_avg_ntm                          | 30.0
 maximum_nights_avg_ntm                          | 60.0
 has_availability                                | t
 availability_30                                 | 11
 availability_60                                 | 41
 availability_90                                 | 70
 availability_365                                | 70
 instant_bookable                                | f
 calculated_host_listings_count                  | 1
 calculated_host_listings_count_entire_homes     | 1
 calculated_host_listings_count_private_rooms    | 0
 calculated_host_listings_count_shared_rooms     | 0
 neighbourhood_reduced                           | West End
 property_reduced                                | Entire condominium
 room_reduced                                    | Entire home/apt
 neighbourhood_indexed                           | 3.0
 property_indexed                                | 1.0
 room_indexed                                    | 0.0
 log_price                                       | 4.499809670330265
 host_listings_count_imp                         | 1
 host_total_listings_count_imp                   | 1
 accommodates_imp                                | 2
 bedrooms_imp                                    | 1
 beds_imp                                        | 1
 minimum_nights_imp                              | 30
```

```
 maximum_nights_imp                                | 60
 minimum_minimum_nights_imp                        | 30
 maximum_minimum_nights_imp                        | 30
 minimum_nights_avg_ntm_imp                        | 30.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 1
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 3.0
 property_indexed_imp                              | 1.0
 room_indexed_imp                                  | 0.0
 has_tv                                            | true
 has_wifi                                          | true
 features                                          | [1.0,1.0,2.0,1.0,1.0,30.0,60.0,30.0,30.0,30.0,1.0,1.0,0.
0,3.0,1.0,0.0,1.0,1.0]
 gps                                               | [49.28209,-123.12515]
-RECORD 1-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------
 id                                                | 16254
 host_since                                        | 2009/12/15
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | N/A
 host_response_rate                                | N/A
 host_acceptance_rate                              | N/A
 host_is_superhost                                 | f
 host_neighbourhood                                | Hastings-Sunrise
 host_listings_count                               | 1
 host_total_listings_count                         | 1
 host_verifications                                | ['phone', 'facebook', 'offline_government_id', 'selfie',
'government_id', 'identity_manual']
 host_has_profile_pic                              | t
 host_identity_verified                            | t
 neighbourhood                                     | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                            | Hastings-Sunrise
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.27734
 longitude                                         | -123.04152
 property_type                                     | Entire guest suite
 room_type                                         | Entire home/apt
 accommodates                                      | 4
 bathrooms                                         | null
 bathrooms_text                                    | 1 bath
 bedrooms                                          | 2
 beds                                              | 3
```

amenities	['Kitchen', 'Oven', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Private entrance', 'Smoke alarm', 'Hair dryer', 'Lockbox', 'Dedicated workspace', 'Iron', 'Free street parking', 'Free parking on premises', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Children\u2019s books and toys', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
price	195
minimum_nights	3
maximum_nights	31
minimum_minimum_nights	3
maximum_minimum_nights	3
minimum_maximum_nights	1125
maximum_maximum_nights	1125
minimum_nights_avg_ntm	3.0
maximum_nights_avg_ntm	1125.0
has_availability	t
availability_30	0
availability_60	0
availability_90	0
availability_365	228
instant_bookable	t
calculated_host_listings_count	1
calculated_host_listings_count_entire_homes	1
calculated_host_listings_count_private_rooms	0
calculated_host_listings_count_shared_rooms	0
neighbourhood_reduced	Hastings-Sunrise
property_reduced	Entire guest suite
room_reduced	Entire home/apt
neighbourhood_indexed	9.0
property_indexed	4.0
room_indexed	0.0
log_price	5.272999558563747
host_listings_count_imp	1
host_total_listings_count_imp	1
accommodates_imp	4
bedrooms_imp	2
beds_imp	3
minimum_nights_imp	3
maximum_nights_imp	31
minimum_minimum_nights_imp	3
maximum_minimum_nights_imp	3
minimum_nights_avg_ntm_imp	3.0
calculated_host_listings_count_imp	1
calculated_host_listings_count_entire_homes_imp	1
calculated_host_listings_count_private_rooms_imp	0
neighbourhood_indexed_imp	9.0
property_indexed_imp	4.0
room_indexed_imp	0.0
has_tv	false

```
 has_wifi                                          | true
 features                                          | [1.0,1.0,4.0,2.0,3.0,3.0,31.0,3.0,3.0,3.0,1.0,1.0,0.0,9.
0,4.0,0.0,0.0,1.0]
 gps                                               | [49.27734,-123.04152]
-RECORD 2-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                                | 17765
 host_since                                        | 2010/01/07
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | N/A
 host_response_rate                                | N/A
 host_acceptance_rate                              | 1
 host_is_superhost                                 | f
 host_neighbourhood                                | Mount Pleasant
 host_listings_count                               | 1
 host_total_listings_count                         | 1
 host_verifications                                | ['phone', 'facebook', 'reviews', 'jumio', 'offline_gover
nment_id', 'selfie', 'government_id']
 host_has_profile_pic                              | t
 host_identity_verified                            | t
 neighbourhood                                     | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                            | Mount Pleasant
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.26181
 longitude                                         | -123.10905
 property_type                                     | Entire apartment
 room_type                                         | Entire home/apt
 accommodates                                      | 2
 bathrooms                                         | null
 bathrooms_text                                    | 1 bath
 bedrooms                                          | 1
 beds                                              | 1
 amenities                                         | ['Kitchen', 'Wifi', 'Host greets you', 'Dryer', 'Carbon
 monoxide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated
workspace', 'Iron', 'Free street parking', 'Free parking on premises', 'Coffee maker', 'Shampoo', 'Dishes an
d silverware', 'Fire extinguisher', 'Hangers', 'Essentials', 'Washer', 'Long term stays allowed', 'Elevato
r']
 price                                             | 94
 minimum_nights                                    | 3
 maximum_nights                                    | 60
 minimum_minimum_nights                            | 3
 maximum_minimum_nights                            | 3
 minimum_maximum_nights                            | 60
```

```
 maximum_maximum_nights                            | 60
 minimum_nights_avg_ntm                            | 3.0
 maximum_nights_avg_ntm                            | 60.0
 has_availability                                  | t
 availability_30                                   | 0
 availability_60                                   | 0
 availability_90                                   | 0
 availability_365                                  | 0
 instant_bookable                                  | f
 calculated_host_listings_count                    | 1
 calculated_host_listings_count_entire_homes       | 1
 calculated_host_listings_count_private_rooms      | 0
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | Mount Pleasant
 property_reduced                                  | Entire apartment
 room_reduced                                      | Entire home/apt
 neighbourhood_indexed                             | 5.0
 property_indexed                                  | 0.0
 room_indexed                                      | 0.0
 log_price                                         | 4.543294782270004
 host_listings_count_imp                           | 1
 host_total_listings_count_imp                     | 1
 accommodates_imp                                  | 2
 bedrooms_imp                                      | 1
 beds_imp                                          | 1
 minimum_nights_imp                                | 3
 maximum_nights_imp                                | 60
 minimum_minimum_nights_imp                        | 3
 maximum_minimum_nights_imp                        | 3
 minimum_nights_avg_ntm_imp                        | 3.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 1
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 5.0
 property_indexed_imp                              | 0.0
 room_indexed_imp                                  | 0.0
 has_tv                                            | false
 has_wifi                                          | true
 features                                          | [1.0,1.0,2.0,1.0,1.0,3.0,60.0,3.0,3.0,3.0,1.0,1.0,0.0,5.
0,0.0,0.0,0.0,1.0]
 gps                                               | [49.26181,-123.10905]
-RECORD 3-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
```

id	18795
host_since	2010/01/13
host_location	Vancouver, British Columbia, Canada
host_response_time	within a day
host_response_rate	0.7
host_acceptance_rate	0.4
host_is_superhost	f
host_neighbourhood	West End
host_listings_count	3
host_total_listings_count	3
host_verifications	['email', 'phone', 'reviews', 'jumio', 'government_id']
host_has_profile_pic	t
host_identity_verified	t
neighbourhood	null
neighbourhood_cleansed	West End
neighbourhood_group_cleansed	null
latitude	49.28251
longitude	-123.12827
property_type	Private room in house
room_type	Private room
accommodates	2
bathrooms	null
bathrooms_text	1 private bath
bedrooms	1
beds	1
amenities	['Kitchen', 'Cable TV', 'Wifi', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Long term stays allowed']
price	50
minimum_nights	30
maximum_nights	360
minimum_minimum_nights	30
maximum_minimum_nights	30
minimum_maximum_nights	360
maximum_maximum_nights	360
minimum_nights_avg_ntm	30.0
maximum_nights_avg_ntm	360.0
has_availability	t
availability_30	9
availability_60	20
availability_90	50
availability_365	325
instant_bookable	f
calculated_host_listings_count	3
calculated_host_listings_count_entire_homes	0
calculated_host_listings_count_private_rooms	3
calculated_host_listings_count_shared_rooms	0
neighbourhood_reduced	West End
property_reduced	Private room in house
room_reduced	Private room

```
 neighbourhood_indexed                              | 3.0
 property_indexed                                   | 3.0
 room_indexed                                       | 1.0
 log_price                                          | 3.912023005428146
 host_listings_count_imp                            | 3
 host_total_listings_count_imp                      | 3
 accommodates_imp                                   | 2
 bedrooms_imp                                       | 1
 beds_imp                                           | 1
 minimum_nights_imp                                 | 30
 maximum_nights_imp                                 | 360
 minimum_minimum_nights_imp                         | 30
 maximum_minimum_nights_imp                         | 30
 minimum_nights_avg_ntm_imp                         | 30.0
 calculated_host_listings_count_imp                 | 3
 calculated_host_listings_count_entire_homes_imp    | 0
 calculated_host_listings_count_private_rooms_imp   | 3
 neighbourhood_indexed_imp                          | 3.0
 property_indexed_imp                               | 3.0
 room_indexed_imp                                   | 1.0
 has_tv                                             | true
 has_wifi                                           | true
 features                                           | [3.0,3.0,2.0,1.0,1.0,30.0,360.0,30.0,30.0,30.0,3.0,0.0,
3.0,3.0,3.0,1.0,1.0,1.0]
 gps                                                | [49.28251,-123.12827]
-RECORD 4-----------------------
```

```
 neighbourhood_group_cleansed                     | null
 latitude                                         | 49.23427
 longitude                                        | -123.05309
 property_type                                    | Entire house
 room_type                                        | Entire home/apt
 accommodates                                     | 2
 bathrooms                                        | null
 bathrooms_text                                   | 1 bath
 bedrooms                                         | 1
 beds                                             | 1
 amenities                                        | ['Kitchen', 'Oven', 'Microwave', 'Wifi', 'Dryer', 'Carbo
n monoxide alarm', 'Stove', 'Hot water', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heat
ing', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Free parking
on premises', 'Coffee maker', 'Bed linens', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV']
 price                                            | 80
 minimum_nights                                   | 28
 maximum_nights                                   | 400
 minimum_minimum_nights                           | 28
 maximum_minimum_nights                           | 28
 minimum_maximum_nights                           | 1125
 maximum_maximum_nights                           | 1125
 minimum_nights_avg_ntm                           | 28.0
 maximum_nights_avg_ntm                           | 1125.0
 has_availability                                 | t
 availability_30                                  | 2
 availability_60                                  | 13
 availability_90                                  | 43
 availability_365                                 | 318
 instant_bookable                                 | f
 calculated_host_listings_count                   | 3
 calculated_host_listings_count_entire_homes      | 1
 calculated_host_listings_count_private_rooms     | 2
 calculated_host_listings_count_shared_rooms      | 0
 neighbourhood_reduced                            | Renfrew-Collingwood
 property_reduced                                 | Entire house
 room_reduced                                     | Entire home/apt
 neighbourhood_indexed                            | 11.0
 property_indexed                                 | 2.0
 room_indexed                                     | 0.0
 log_price                                        | 4.382026634673881
 host_listings_count_imp                          | 3
 host_total_listings_count_imp                    | 3
 accommodates_imp                                 | 2
 bedrooms_imp                                     | 1
 beds_imp                                         | 1
 minimum_nights_imp                               | 28
 maximum_nights_imp                               | 400
 minimum_minimum_nights_imp                       | 28
```

```
 maximum_minimum_nights_imp                       | 28
 minimum_nights_avg_ntm_imp                       | 28.0
 calculated_host_listings_count_imp               | 3
 calculated_host_listings_count_entire_homes_imp  | 1
 calculated_host_listings_count_private_rooms_imp | 2
 neighbourhood_indexed_imp                        | 11.0
 property_indexed_imp                             | 2.0
 room_indexed_imp                                 | 0.0
 has_tv                                           | true
 has_wifi                                         | true
 features                                         | [3.0,3.0,2.0,1.0,1.0,28.0,400.0,28.0,28.0,28.0,3.0,1.0,
2.0,11.0,2.0,0.0,1.0,1.0]
 gps                                              | [49.23427,-123.05309]
-RECORD 5-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                               | 24273
 host_since                                       | 2011/10/09
 host_location                                    | Vancouver, British Columbia, Canada
 host_response_time                               | within an hour
 host_response_rate                               | 1
 host_acceptance_rate                             | 0.82
 host_is_superhost                                | f
 host_neighbourhood                               | Kitsilano
 host_listings_count                              | 2
 host_total_listings_count                        | 2
 host_verifications                               | ['email', 'phone', 'reviews', 'jumio', 'offline_governme
nt_id', 'government_id']
 host_has_profile_pic                             | f
 host_identity_verified                           | t
 neighbourhood                                    | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                           | Kitsilano
 neighbourhood_group_cleansed                     | null
 latitude                                         | 49.2626
 longitude                                        | -123.17308
 property_type                                    | Entire apartment
 room_type                                        | Entire home/apt
 accommodates                                     | 6
 bathrooms                                        | null
 bathrooms_text                                   | 1 bath
 bedrooms                                         | 2
 beds                                             | 3
 amenities                                        | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wi
fi', 'Dryer', 'Fireplace guards', 'Carbon monoxide alarm', 'Stove', 'Crib', 'Hot water', 'Babysitter recomme
```

ndations', 'Garden or backyard', 'Extra pillows and blankets', 'Dishwasher', 'Luggage dropoff allowed', 'Coo
king basics', 'Room-darkening shades', 'Private entrance', 'Smoke alarm', 'Baking sheet', 'Heating', 'Single
level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Coffee maker',
 'Bed linens', 'Keypad', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Shower gel', 'Patio or ba
lcony', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials',
'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'TV', 'Ethernet connection']

price	200
minimum_nights	30
maximum_nights	365
minimum_minimum_nights	30
maximum_minimum_nights	30
minimum_maximum_nights	365
maximum_maximum_nights	365
minimum_nights_avg_ntm	30.0
maximum_nights_avg_ntm	365.0
has_availability	t
availability_30	0
availability_60	0
availability_90	10
availability_365	223
instant_bookable	f
calculated_host_listings_count	2
calculated_host_listings_count_entire_homes	2
calculated_host_listings_count_private_rooms	0
calculated_host_listings_count_shared_rooms	0
neighbourhood_reduced	Kitsilano
property_reduced	Entire apartment
room_reduced	Entire home/apt
neighbourhood_indexed	2.0
property_indexed	0.0
room_indexed	0.0
log_price	5.298317366548036
host_listings_count_imp	2
host_total_listings_count_imp	2
accommodates_imp	6
bedrooms_imp	2
beds_imp	3
minimum_nights_imp	30
maximum_nights_imp	365
minimum_minimum_nights_imp	30
maximum_minimum_nights_imp	30
minimum_nights_avg_ntm_imp	30.0
calculated_host_listings_count_imp	2
calculated_host_listings_count_entire_homes_imp	2
calculated_host_listings_count_private_rooms_imp	0
neighbourhood_indexed_imp	2.0
property_indexed_imp	0.0
room_indexed_imp	0.0
has_tv	true

```
 has_wifi                                          | true
 features                                          | [2.0,2.0,6.0,2.0,3.0,30.0,365.0,30.0,30.0,30.0,2.0,2.0,
0.0,2.0,0.0,0.0,1.0,1.0]
 gps                                               | [49.2626,-123.17308]
-RECORD 6-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                                | 42645
 host_since                                        | 2010/04/17
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | within an hour
 host_response_rate                                | 1
 host_acceptance_rate                              | 1
 host_is_superhost                                 | t
 host_neighbourhood                                | Downtown Vancouver
 host_listings_count                               | 4
 host_total_listings_count                         | 4
 host_verifications                                | ['email', 'phone', 'facebook', 'reviews']
 host_has_profile_pic                              | t
 host_identity_verified                            | t
 neighbourhood                                     | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                            | Downtown
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.28001
 longitude                                         | -123.12475
 property_type                                     | Entire condominium
 room_type                                         | Entire home/apt
 accommodates                                      | 3
 bathrooms                                         | null
 bathrooms_text                                    | 1 bath
 bedrooms                                          | 1
 beds                                              | 2
 amenities                                         | ['Kitchen', 'Oven', 'Paid parking on premises', 'Paid pa
rking off premises', 'Gym', 'Microwave', 'Wifi', 'Host greets you', 'Dryer', 'Carbon monoxide alarm', 'Stov
e', 'Hot water', 'Dishwasher', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair
 dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Coffee maker', 'Shampoo', 'Fire extinguisher', 'Di
shes and silverware', 'Patio or balcony', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable
 TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'T
V', 'Elevator']
 price                                             | 159
 minimum_nights                                    | 32
 maximum_nights                                    | 365
 minimum_minimum_nights                            | 32
 maximum_minimum_nights                            | 32
```

```
 minimum_maximum_nights                            | 1125
 maximum_maximum_nights                            | 1125
 minimum_nights_avg_ntm                            | 32.0
 maximum_nights_avg_ntm                            | 1125.0
 has_availability                                  | t
 availability_30                                   | 0
 availability_60                                   | 29
 availability_90                                   | 59
 availability_365                                  | 333
 instant_bookable                                  | t
 calculated_host_listings_count                    | 4
 calculated_host_listings_count_entire_homes       | 4
 calculated_host_listings_count_private_rooms      | 0
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | Downtown
 property_reduced                                  | Entire condominium
 room_reduced                                      | Entire home/apt
 neighbourhood_indexed                             | 0.0
 property_indexed                                  | 1.0
 room_indexed                                      | 0.0
 log_price                                         | 5.0689042022202315
 host_listings_count_imp                           | 4
 host_total_listings_count_imp                     | 4
 accommodates_imp                                  | 3
 bedrooms_imp                                      | 1
 beds_imp                                          | 2
 minimum_nights_imp                                | 32
 maximum_nights_imp                                | 365
 minimum_minimum_nights_imp                        | 32
 maximum_minimum_nights_imp                        | 32
 minimum_nights_avg_ntm_imp                        | 32.0
 calculated_host_listings_count_imp                | 4
 calculated_host_listings_count_entire_homes_imp   | 4
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 0.0
 property_indexed_imp                              | 1.0
 room_indexed_imp                                  | 0.0
 has_tv                                            | true
 has_wifi                                          | true
 features                                          | [4.0,4.0,3.0,1.0,2.0,32.0,365.0,32.0,32.0,32.0,4.0,4.0,
0.0,0.0,1.0,0.0,1.0,1.0]
 gps                                               | [49.28001,-123.12475]
-RECORD 7-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
```

```
-----------------------------------------------------------------------------------------------------
 id                                                | 54261
 host_since                                        | 2010/10/05
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | a few days or more
 host_response_rate                                | 0
 host_acceptance_rate                              | N/A
 host_is_superhost                                 | f
 host_neighbourhood                                | Downtown Vancouver
 host_listings_count                               | 2
 host_total_listings_count                         | 2
 host_verifications                                | ['email', 'phone', 'reviews', 'offline_government_id',
 'selfie', 'government_id', 'identity_manual']
 host_has_profile_pic                              | t
 host_identity_verified                            | t
 neighbourhood                                     | null
 neighbourhood_cleansed                            | Downtown
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.27655
 longitude                                         | -123.12485
 property_type                                     | Entire condominium
 room_type                                         | Entire home/apt
 accommodates                                      | 4
 bathrooms                                         | null
 bathrooms_text                                    | 1 bath
 bedrooms                                          | 1
 beds                                              | 1
 amenities                                         | ['Kitchen', 'Hot tub', 'Gym', 'Indoor fireplace', 'Wif
i', 'Host greets you', 'Dryer', 'Carbon monoxide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alar
m', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Shampoo', 'Fire ext
inguisher', 'Hangers', 'Cable TV', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditionin
g', 'Elevator']
 price                                             | 126
 minimum_nights                                    | 90
 maximum_nights                                    | 365
 minimum_minimum_nights                            | 90
 maximum_minimum_nights                            | 90
 minimum_maximum_nights                            | 365
 maximum_maximum_nights                            | 365
 minimum_nights_avg_ntm                            | 90.0
 maximum_nights_avg_ntm                            | 365.0
 has_availability                                  | t
 availability_30                                   | 30
 availability_60                                   | 60
 availability_90                                   | 90
 availability_365                                  | 365
 instant_bookable                                  | f
 calculated_host_listings_count                    | 2
 calculated_host_listings_count_entire_homes       | 2
```

```
 calculated_host_listings_count_private_rooms     | 0
 calculated_host_listings_count_shared_rooms      | 0
 neighbourhood_reduced                            | Downtown
 property_reduced                                 | Entire condominium
 room_reduced                                     | Entire home/apt
 neighbourhood_indexed                            | 0.0
 property_indexed                                 | 1.0
 room_indexed                                     | 0.0
 log_price                                        | 4.836281906951478
 host_listings_count_imp                          | 2
 host_total_listings_count_imp                    | 2
 accommodates_imp                                 | 4
 bedrooms_imp                                     | 1
 beds_imp                                         | 1
 minimum_nights_imp                               | 90
 maximum_nights_imp                               | 365
 minimum_minimum_nights_imp                       | 90
 maximum_minimum_nights_imp                       | 90
 minimum_nights_avg_ntm_imp                       | 90.0
 calculated_host_listings_count_imp               | 2
 calculated_host_listings_count_entire_homes_imp  | 2
 calculated_host_listings_count_private_rooms_imp | 0
 neighbourhood_indexed_imp                        | 0.0
 property_indexed_imp                             | 1.0
 room_indexed_imp                                 | 0.0
 has_tv                                           | true
 has_wifi                                         | true
 features                                         | [2.0,2.0,4.0,1.0,1.0,90.0,365.0,90.0,90.0,90.0,2.0,2.0,
0.0,0.0,1.0,0.0,1.0,1.0]
 gps                                              | [49.27655,-123.12485]
-RECORD 8-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                               | 72659
 host_since                                       | 2011/02/03
 host_location                                    | Vancouver, British Columbia, Canada
 host_response_time                               | N/A
 host_response_rate                               | N/A
 host_acceptance_rate                             | 0
 host_is_superhost                                | f
 host_neighbourhood                               | Commercial Drive
 host_listings_count                              | 1
 host_total_listings_count                        | 1
 host_verifications                               | ['email', 'phone', 'reviews', 'jumio', 'government_id']
```

host_has_profile_pic	t
host_identity_verified	t
neighbourhood	null
neighbourhood_cleansed	Grandview-Woodland
neighbourhood_group_cleansed	null
latitude	49.27357
longitude	-123.07024
property_type	Entire condominium
room_type	Entire home/apt
accommodates	4
bathrooms	null
bathrooms_text	1 bath
bedrooms	2
beds	2
amenities	['Kitchen', 'Lake access', 'Oven', 'Paid parking off premises', 'Gym', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Dishwasher', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm', 'Heating', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'BBQ grill', 'TV', 'Elevator']
price	210
minimum_nights	7
maximum_nights	21
minimum_minimum_nights	7
maximum_minimum_nights	7
minimum_maximum_nights	21
maximum_maximum_nights	21
minimum_nights_avg_ntm	7.0
maximum_nights_avg_ntm	21.0
has_availability	t
availability_30	0
availability_60	0
availability_90	0
availability_365	40
instant_bookable	f
calculated_host_listings_count	1
calculated_host_listings_count_entire_homes	1
calculated_host_listings_count_private_rooms	0
calculated_host_listings_count_shared_rooms	0
neighbourhood_reduced	Grandview-Woodland
property_reduced	Entire condominium
room_reduced	Entire home/apt
neighbourhood_indexed	8.0
property_indexed	1.0
room_indexed	0.0
log_price	5.3471075307174685
host_listings_count_imp	1
host_total_listings_count_imp	1
accommodates_imp	4

```
 bedrooms_imp                                       | 2
 beds_imp                                           | 2
 minimum_nights_imp                                 | 7
 maximum_nights_imp                                 | 21
 minimum_minimum_nights_imp                         | 7
 maximum_minimum_nights_imp                         | 7
 minimum_nights_avg_ntm_imp                         | 7.0
 calculated_host_listings_count_imp                 | 1
 calculated_host_listings_count_entire_homes_imp    | 1
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 8.0
 property_indexed_imp                               | 1.0
 room_indexed_imp                                   | 0.0
 has_tv                                             | true
 has_wifi                                           | true
 features                                           | [1.0,1.0,4.0,2.0,2.0,7.0,21.0,7.0,7.0,7.0,1.0,1.0,0.0,8.
0,1.0,0.0,1.0,1.0]
 gps                                                | [49.27357,-123.07024]
-RECORD 9-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                                 | 77157
 host_since                                         | 2011/02/27
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within a few hours
 host_response_rate                                 | 1
 host_acceptance_rate                               | 1
 host_is_superhost                                  | f
 host_neighbourhood                                 | Arbutus Ridge
 host_listings_count                                | 2
 host_total_listings_count                          | 2
 host_verifications                                 | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'offl
ine_government_id', 'government_id']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Arbutus Ridge
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.24675
 longitude                                          | -123.16711
 property_type                                      | Private room in guest suite
 room_type                                          | Private room
 accommodates                                       | 3
 bathrooms                                          | null
```

```
 bathrooms_text                                      | 1 private bath
 bedrooms                                            | 1
 beds                                                | 2
 amenities                                           | ['Kitchen', 'Microwave', 'Wifi', 'Laundromat nearby', 'D
ryer', 'Hot water', 'Garden or backyard', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking
 basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Mini fridge', 'Single level home', 'Hair dryer', 'L
ockbox', 'Dedicated workspace', 'First aid kit', 'Iron', 'Body soap', 'Free street parking', 'Coffee maker',
'Bed linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Barbecue utensils', 'Shower gel', 'H
angers', 'Cable TV', 'Essentials', 'Washer', 'Conditioner', 'BBQ grill', 'Beach essentials', 'TV', 'Long ter
m stays allowed']
 price                                               | 75
 minimum_nights                                      | 3
 maximum_nights                                      | 60
 minimum_minimum_nights                              | 3
 maximum_minimum_nights                              | 3
 minimum_maximum_nights                              | 1125
 maximum_maximum_nights                              | 1125
 minimum_nights_avg_ntm                              | 3.0
 maximum_nights_avg_ntm                              | 1125.0
 has_availability                                    | t
 availability_30                                     | 3
 availability_60                                     | 3
 availability_90                                     | 3
 availability_365                                    | 267
 instant_bookable                                    | t
 calculated_host_listings_count                      | 2
 calculated_host_listings_count_entire_homes         | 0
 calculated_host_listings_count_private_rooms        | 2
 calculated_host_listings_count_shared_rooms         | 0
 neighbourhood_reduced                               | OTHER
 property_reduced                                    | OTHER
 room_reduced                                        | Private room
 neighbourhood_indexed                               | 1.0
 property_indexed                                    | 5.0
 room_indexed                                        | 1.0
 log_price                                           | 4.31748811353631
 host_listings_count_imp                             | 2
 host_total_listings_count_imp                       | 2
 accommodates_imp                                    | 3
 bedrooms_imp                                        | 1
 beds_imp                                            | 2
 minimum_nights_imp                                  | 3
 maximum_nights_imp                                  | 60
 minimum_minimum_nights_imp                          | 3
 maximum_minimum_nights_imp                          | 3
 minimum_nights_avg_ntm_imp                          | 3.0
 calculated_host_listings_count_imp                  | 2
 calculated_host_listings_count_entire_homes_imp     | 0
 calculated_host_listings_count_private_rooms_imp    | 2
```

```
 neighbourhood_indexed_imp                       | 1.0
 property_indexed_imp                            | 5.0
 room_indexed_imp                                | 1.0
 has_tv                                          | true
 has_wifi                                        | true
 features                                        | [2.0,2.0,3.0,1.0,2.0,3.0,60.0,3.0,3.0,3.0,2.0,0.0,2.0,1.
0,5.0,1.0,1.0,1.0]
 gps                                             | [49.24675,-123.16711]
-RECORD 10-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                              | 111077
 host_since                                      | 2011/05/09
 host_location                                   | North Vancouver, British Columbia, Canada
 host_response_time                              | within a few hours
 host_response_rate                              | 1
 host_acceptance_rate                            | 0.33
 host_is_superhost                               | f
 host_neighbourhood                              | Grandview-Woodland
 host_listings_count                             | 1
 host_total_listings_count                       | 1
 host_verifications                              | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'gove
rnment_id']
 host_has_profile_pic                            | t
 host_identity_verified                          | t
 neighbourhood                                   | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                          | Grandview-Woodland
 neighbourhood_group_cleansed                    | null
 latitude                                        | 49.26726
 longitude                                       | -123.05969
 property_type                                   | Entire guest suite
 room_type                                       | Entire home/apt
 accommodates                                    | 3
 bathrooms                                       | null
 bathrooms_text                                  | 1 bath
 bedrooms                                        | 2
 beds                                            | 2
 amenities                                       | ['Kitchen', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dr
yer', 'Stove', 'Hot water', 'Cooking basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'I
ron', 'Free street parking', 'Coffee maker', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed']
 price                                           | 125
 minimum_nights                                  | 3
 maximum_nights                                  | 365
```

```
 minimum_minimum_nights                          | 3
 maximum_minimum_nights                          | 3
 minimum_maximum_nights                          | 365
 maximum_maximum_nights                          | 365
 minimum_nights_avg_ntm                          | 3.0
 maximum_nights_avg_ntm                          | 365.0
 has_availability                                | t
 availability_30                                 | 4
 availability_60                                 | 34
 availability_90                                 | 64
 availability_365                                | 339
 instant_bookable                                | f
 calculated_host_listings_count                  | 1
 calculated_host_listings_count_entire_homes     | 1
 calculated_host_listings_count_private_rooms    | 0
 calculated_host_listings_count_shared_rooms     | 0
 neighbourhood_reduced                           | Grandview-Woodland
 property_reduced                                | Entire guest suite
 room_reduced                                    | Entire home/apt
 neighbourhood_indexed                           | 8.0
 property_indexed                                | 4.0
 room_indexed                                    | 0.0
 log_price                                       | 4.8283137373023015
 host_listings_count_imp                         | 1
 host_total_listings_count_imp                   | 1
 accommodates_imp                                | 3
 bedrooms_imp                                    | 2
 beds_imp                                        | 2
 minimum_nights_imp                              | 3
 maximum_nights_imp                              | 365
 minimum_minimum_nights_imp                      | 3
 maximum_minimum_nights_imp                      | 3
 minimum_nights_avg_ntm_imp                      | 3.0
 calculated_host_listings_count_imp              | 1
 calculated_host_listings_count_entire_homes_imp | 1
 calculated_host_listings_count_private_rooms_imp | 0
 neighbourhood_indexed_imp                       | 8.0
 property_indexed_imp                            | 4.0
 room_indexed_imp                                | 0.0
 has_tv                                          | false
 has_wifi                                        | true
 features                                        | [1.0,1.0,3.0,2.0,2.0,3.0,365.0,3.0,3.0,3.0,1.0,1.0,0.0,
8.0,4.0,0.0,0.0,1.0]
 gps                                             | [49.26726,-123.05969]
-RECORD 11-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 calculated_host_listings_count                   | 1
 calculated_host_listings_count_entire_homes      | 0
 calculated_host_listings_count_private_rooms     | 1
 calculated_host_listings_count_shared_rooms      | 0
 neighbourhood_reduced                            | Riley Park
 property_reduced                                 | Private room in house
 room_reduced                                     | Private room
 neighbourhood_indexed                            | 7.0
 property_indexed                                 | 3.0
 room_indexed                                     | 1.0
 log_price                                        | 4.31748811353631
 host_listings_count_imp                          | 1
 host_total_listings_count_imp                    | 1
 accommodates_imp                                 | 3
 bedrooms_imp                                     | 1
 beds_imp                                         | 0
 minimum_nights_imp                               | 2
 maximum_nights_imp                               | 60
 minimum_minimum_nights_imp                       | 2
 maximum_minimum_nights_imp                       | 2
 minimum_nights_avg_ntm_imp                       | 2.0
 calculated_host_listings_count_imp               | 1
 calculated_host_listings_count_entire_homes_imp  | 0
 calculated_host_listings_count_private_rooms_imp | 1
 neighbourhood_indexed_imp                        | 7.0
 property_indexed_imp                             | 3.0
 room_indexed_imp                                 | 1.0
 has_tv                                           | true
 has_wifi                                         | true
 features                                         | [1.0,1.0,3.0,1.0,0.0,2.0,60.0,2.0,2.0,2.0,1.0,0.0,1.0,7.
0,3.0,1.0,1.0,1.0]
 gps                                              | [49.24648,-123.09241]
-RECORD 12-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                               | 129163
 host_since                                       | 2011/05/17
 host_location                                    | Vancouver, British Columbia, Canada
 host_response_time                               | within a few hours
 host_response_rate                               | 1
 host_acceptance_rate                             | 1
 host_is_superhost                                | f
 host_neighbourhood                               | Renfrew-Collingwood
 host_listings_count                              | 2
```

```
 host_total_listings_count                        | 2
 host_verifications                               | ['email', 'phone', 'reviews', 'jumio', 'offline_governme
nt_id', 'government_id']
 host_has_profile_pic                             | t
 host_identity_verified                           | t
 neighbourhood                                    | null
 neighbourhood_cleansed                           | Renfrew-Collingwood
 neighbourhood_group_cleansed                     | null
 latitude                                         | 49.24507
 longitude                                        | -123.04941
 property_type                                    | Private room in house
 room_type                                        | Private room
 accommodates                                     | 1
 bathrooms                                        | null
 bathrooms_text                                   | 1 shared bath
 bedrooms                                         | 1
 beds                                             | 1
 amenities                                        | ['Kitchen', 'Hot water', 'Paid parking on premises', 'Lo
ck on bedroom door', 'Free street parking', 'Cable TV', 'Smoke alarm', 'Wifi', 'Essentials', 'Heating', 'Was
her', 'Dryer', 'Carbon monoxide alarm', 'Dishes and silverware', 'TV', 'Long term stays allowed', 'Host gree
ts you', 'Hair dryer', 'Shampoo', 'Fire extinguisher']
 price                                            | 45
 minimum_nights                                   | 30
 maximum_nights                                   | 280
 minimum_minimum_nights                           | 30
 maximum_minimum_nights                           | 30
 minimum_maximum_nights                           | 280
 maximum_maximum_nights                           | 280
 minimum_nights_avg_ntm                           | 30.0
 maximum_nights_avg_ntm                           | 280.0
 has_availability                                 | t
 availability_30                                  | 30
 availability_60                                  | 60
 availability_90                                  | 90
 availability_365                                 | 365
 instant_bookable                                 | t
 calculated_host_listings_count                   | 1
 calculated_host_listings_count_entire_homes      | 0
 calculated_host_listings_count_private_rooms     | 1
 calculated_host_listings_count_shared_rooms      | 0
 neighbourhood_reduced                            | Renfrew-Collingwood
 property_reduced                                 | Private room in house
 room_reduced                                     | Private room
 neighbourhood_indexed                            | 11.0
 property_indexed                                 | 3.0
 room_indexed                                     | 1.0
 log_price                                        | 3.8066624897703196
 host_listings_count_imp                          | 2
 host_total_listings_count_imp                    | 2
```

```
 accommodates_imp                                   | 1
 bedrooms_imp                                       | 1
 beds_imp                                           | 1
 minimum_nights_imp                                 | 30
 maximum_nights_imp                                 | 280
 minimum_minimum_nights_imp                         | 30
 maximum_minimum_nights_imp                         | 30
 minimum_nights_avg_ntm_imp                         | 30.0
 calculated_host_listings_count_imp                 | 1
 calculated_host_listings_count_entire_homes_imp    | 0
 calculated_host_listings_count_private_rooms_imp   | 1
 neighbourhood_indexed_imp                          | 11.0
 property_indexed_imp                               | 3.0
 room_indexed_imp                                   | 1.0
 has_tv                                             | true
 has_wifi                                           | true
 features                                           | [2.0,2.0,1.0,1.0,1.0,30.0,280.0,30.0,30.0,30.0,1.0,0.0,
1.0,11.0,3.0,1.0,1.0,1.0]
 gps                                                | [49.24507,-123.04941]
-RECORD 13-----------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                                 | 174135
 host_since                                         | 2011/07/17
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within a few hours
 host_response_rate                                 | 1
 host_acceptance_rate                               | 0
 host_is_superhost                                  | f
 host_neighbourhood                                 | Downtown Eastside
 host_listings_count                                | 1
 host_total_listings_count                          | 1
 host_verifications                                 | ['email', 'phone', 'reviews', 'offline_government_id',
 'selfie', 'government_id', 'identity_manual']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Downtown Eastside
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.27782
 longitude                                          | -123.08985
 property_type                                      | Entire guest suite
 room_type                                          | Entire home/apt
 accommodates                                       | 2
```

bathrooms	null
bathrooms_text	1 bath
bedrooms	1
beds	2
amenities	['Kitchen', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Hot water', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Hangers', 'Cable TV', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning']
price	100
minimum_nights	30
maximum_nights	365
minimum_minimum_nights	30
maximum_minimum_nights	30
minimum_maximum_nights	365
maximum_maximum_nights	365
minimum_nights_avg_ntm	30.0
maximum_nights_avg_ntm	365.0
has_availability	t
availability_30	0
availability_60	0
availability_90	0
availability_365	224
instant_bookable	f
calculated_host_listings_count	1
calculated_host_listings_count_entire_homes	1
calculated_host_listings_count_private_rooms	0
calculated_host_listings_count_shared_rooms	0
neighbourhood_reduced	Downtown Eastside
property_reduced	Entire guest suite
room_reduced	Entire home/apt
neighbourhood_indexed	6.0
property_indexed	4.0
room_indexed	0.0
log_price	4.605170185988092
host_listings_count_imp	1
host_total_listings_count_imp	1
accommodates_imp	2
bedrooms_imp	1
beds_imp	2
minimum_nights_imp	30
maximum_nights_imp	365
minimum_minimum_nights_imp	30
maximum_minimum_nights_imp	30
minimum_nights_avg_ntm_imp	30.0
calculated_host_listings_count_imp	1
calculated_host_listings_count_entire_homes_imp	1
calculated_host_listings_count_private_rooms_imp	0
neighbourhood_indexed_imp	6.0
property_indexed_imp	4.0

```
 room_indexed_imp                                   | 0.0
 has_tv                                             | true
 has_wifi                                           | true
 features                                           | [1.0,1.0,2.0,1.0,2.0,30.0,365.0,30.0,30.0,30.0,1.0,1.0,
0.0,6.0,4.0,0.0,1.0,1.0]
 gps                                                | [49.27782,-123.08985]
-RECORD 14-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------
 id                                                 | 200643
 host_since                                         | 2011/08/16
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within an hour
 host_response_rate                                 | 1
 host_acceptance_rate                               | 1
 host_is_superhost                                  | f
 host_neighbourhood                                 | Strathcona
 host_listings_count                                | 4
 host_total_listings_count                          | 4
 host_verifications                                 | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Downtown Eastside
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.27384
 longitude                                          | -123.08289
 property_type                                      | Private room in house
 room_type                                          | Private room
 accommodates                                       | 4
 bathrooms                                          | null
 bathrooms_text                                     | 2 shared baths
 bedrooms                                           | 1
 beds                                               | 1
 amenities                                          | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wi
fi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Garden or backyard', 'Dishwasher', 'Extra pill
ows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Room-darkening shades', 'Smoke alarm', 'Hea
ting', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Free street parking', 'Coffee maker',
 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Essentia
ls', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
 price                                              | 87
 minimum_nights                                     | 1
 maximum_nights                                     | 1125
 minimum_minimum_nights                             | 1
```

```
 maximum_minimum_nights                          | 1
 minimum_maximum_nights                          | 1125
 maximum_maximum_nights                          | 1125
 minimum_nights_avg_ntm                          | 1.0
 maximum_nights_avg_ntm                          | 1125.0
 has_availability                                | t
 availability_30                                 | 30
 availability_60                                 | 60
 availability_90                                 | 90
 availability_365                                | 365
 instant_bookable                                | t
 calculated_host_listings_count                  | 4
 calculated_host_listings_count_entire_homes     | 0
 calculated_host_listings_count_private_rooms    | 4
 calculated_host_listings_count_shared_rooms     | 0
 neighbourhood_reduced                           | Downtown Eastside
 property_reduced                                | Private room in house
 room_reduced                                    | Private room
 neighbourhood_indexed                           | 6.0
 property_indexed                                | 3.0
 room_indexed                                    | 1.0
 log_price                                       | 4.465908118654584
 host_listings_count_imp                         | 4
 host_total_listings_count_imp                   | 4
 accommodates_imp                                | 4
 bedrooms_imp                                    | 1
 beds_imp                                        | 1
 minimum_nights_imp                              | 1
 maximum_nights_imp                              | 1125
 minimum_minimum_nights_imp                      | 1
 maximum_minimum_nights_imp                      | 1
 minimum_nights_avg_ntm_imp                      | 1.0
 calculated_host_listings_count_imp              | 4
 calculated_host_listings_count_entire_homes_imp | 0
 calculated_host_listings_count_private_rooms_imp | 4
 neighbourhood_indexed_imp                       | 6.0
 property_indexed_imp                            | 3.0
 room_indexed_imp                                | 1.0
 has_tv                                          | false
 has_wifi                                        | true
 features                                        | [4.0,4.0,4.0,1.0,1.0,1.0,1125.0,1.0,1.0,1.0,4.0,0.0,4.0,
6.0,3.0,1.0,0.0,1.0]
 gps                                             | [49.27384,-123.08289]
-RECORD 15-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 calculated_host_listings_count_entire_homes      | 1
 calculated_host_listings_count_private_rooms     | 0
 calculated_host_listings_count_shared_rooms      | 0
 neighbourhood_reduced                            | Grandview-Woodland
 property_reduced                                 | Entire house
 room_reduced                                     | Entire home/apt
 neighbourhood_indexed                            | 8.0
 property_indexed                                 | 2.0
 room_indexed                                     | 0.0
 log_price                                        | 4.605170185988092
 host_listings_count_imp                          | 1
 host_total_listings_count_imp                    | 1
 accommodates_imp                                 | 3
 bedrooms_imp                                     | 2
 beds_imp                                         | 0
 minimum_nights_imp                               | 31
 maximum_nights_imp                               | 180
 minimum_minimum_nights_imp                       | 31
 maximum_minimum_nights_imp                       | 31
 minimum_nights_avg_ntm_imp                       | 31.0
 calculated_host_listings_count_imp               | 1
 calculated_host_listings_count_entire_homes_imp  | 1
 calculated_host_listings_count_private_rooms_imp | 0
 neighbourhood_indexed_imp                        | 8.0
 property_indexed_imp                             | 2.0
 room_indexed_imp                                 | 0.0
 has_tv                                           | true
 has_wifi                                         | true
 features                                         | [1.0,1.0,3.0,2.0,0.0,31.0,180.0,31.0,31.0,31.0,1.0,1.0,
0.0,8.0,2.0,0.0,1.0,1.0]
 gps                                              | [49.29103,-123.05782]
-RECORD 16-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                               | 258636
 host_since                                       | 2011/08/09
 host_location                                    | Vancouver, British Columbia, Canada
 host_response_time                               | N/A
 host_response_rate                               | N/A
 host_acceptance_rate                             | 1
 host_is_superhost                                | f
 host_neighbourhood                               | Downtown Vancouver
 host_listings_count                              | 1
 host_total_listings_count                        | 1
```

host_verifications	['email', 'phone', 'reviews', 'jumio', 'government_id']
host_has_profile_pic	t
host_identity_verified	t
neighbourhood	null
neighbourhood_cleansed	Downtown
neighbourhood_group_cleansed	null
latitude	49.27796
longitude	-123.12671
property_type	Entire apartment
room_type	Entire home/apt
accommodates	4
bathrooms	null
bathrooms_text	1 bath
bedrooms	1
beds	2
amenities	['Kitchen', 'Iron', 'Elevator', 'Gym', 'Hangers', 'Smoke alarm', 'Wifi', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Free parking on premises', 'Long term stays allowed', 'Hair dryer', 'Shampoo', 'Dedicated workspace']
price	99
minimum_nights	120
maximum_nights	365
minimum_minimum_nights	120
maximum_minimum_nights	120
minimum_maximum_nights	365
maximum_maximum_nights	365
minimum_nights_avg_ntm	120.0
maximum_nights_avg_ntm	365.0
has_availability	t
availability_30	0
availability_60	11
availability_90	41
availability_365	316
instant_bookable	f
calculated_host_listings_count	1
calculated_host_listings_count_entire_homes	1
calculated_host_listings_count_private_rooms	0
calculated_host_listings_count_shared_rooms	0
neighbourhood_reduced	Downtown
property_reduced	Entire apartment
room_reduced	Entire home/apt
neighbourhood_indexed	0.0
property_indexed	0.0
room_indexed	0.0
log_price	4.59511985013459
host_listings_count_imp	1
host_total_listings_count_imp	1
accommodates_imp	4
bedrooms_imp	1
beds_imp	2

```
 minimum_nights_imp                                | 120
 maximum_nights_imp                                | 365
 minimum_minimum_nights_imp                        | 120
 maximum_minimum_nights_imp                        | 120
 minimum_nights_avg_ntm_imp                        | 120.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 1
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 0.0
 property_indexed_imp                              | 0.0
 room_indexed_imp                                  | 0.0
 has_tv                                            | false
 has_wifi                                          | true
 features                                          | [1.0,1.0,4.0,1.0,2.0,120.0,365.0,120.0,120.0,120.0,1.0,
1.0,0.0,0.0,0.0,0.0,0.0,1.0]
 gps                                               | [49.27796,-123.12671]
-RECORD 17-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                                | 283841
 host_since                                        | 2009/08/10
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | within a day
 host_response_rate                                | 0.93
 host_acceptance_rate                              | 0.88
 host_is_superhost                                 | f
 host_neighbourhood                                | Coal Harbour
 host_listings_count                               | 49
 host_total_listings_count                         | 49
 host_verifications                                | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                              | t
 host_identity_verified                            | t
 neighbourhood                                     | null
 neighbourhood_cleansed                            | Downtown
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.2872
 longitude                                         | -123.12352
 property_type                                     | Entire apartment
 room_type                                         | Entire home/apt
 accommodates                                      | 6
 bathrooms                                         | null
 bathrooms_text                                    | 3 baths
 bedrooms                                          | 3
 beds                                              | 3
```

```
 amenities                                          | ['Kitchen', 'Hot tub', 'Oven', 'Pool', 'Gym', 'Indoor fi
replace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Dishwasher', 'Smoke
 alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Coffee maker',
'Dishes and silverware', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays all
owed', 'TV', 'Air conditioning']
 price                                              | 197
 minimum_nights                                     | 90
 maximum_nights                                     | 1125
 minimum_minimum_nights                             | 90
 maximum_minimum_nights                             | 90
 minimum_maximum_nights                             | 1125
 maximum_maximum_nights                             | 1125
 minimum_nights_avg_ntm                             | 90.0
 maximum_nights_avg_ntm                             | 1125.0
 has_availability                                   | t
 availability_30                                    | 0
 availability_60                                    | 0
 availability_90                                    | 0
 availability_365                                   | 0
 instant_bookable                                   | f
 calculated_host_listings_count                     | 41
 calculated_host_listings_count_entire_homes        | 41
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Downtown
 property_reduced                                   | Entire apartment
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 0.0
 property_indexed                                   | 0.0
 room_indexed                                       | 0.0
 log_price                                          | 5.2832037287379885
 host_listings_count_imp                            | 49
 host_total_listings_count_imp                      | 49
 accommodates_imp                                   | 6
 bedrooms_imp                                       | 3
 beds_imp                                           | 3
 minimum_nights_imp                                 | 90
 maximum_nights_imp                                 | 1125
 minimum_minimum_nights_imp                         | 90
 maximum_minimum_nights_imp                         | 90
 minimum_nights_avg_ntm_imp                         | 90.0
 calculated_host_listings_count_imp                 | 41
 calculated_host_listings_count_entire_homes_imp    | 41
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 0.0
 property_indexed_imp                               | 0.0
 room_indexed_imp                                   | 0.0
 has_tv                                             | true
 has_wifi                                           | true
```

```
 features                                          | [49.0,49.0,6.0,3.0,3.0,90.0,1125.0,90.0,90.0,90.0,41.0,4
1.0,0.0,0.0,0.0,0.0,1.0,1.0]
 gps                                               | [49.2872,-123.12352]
-RECORD 18-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                                | 323090
 host_since                                        | 2012/01/01
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | N/A
 host_response_rate                                | N/A
 host_acceptance_rate                              | 0.85
 host_is_superhost                                 | t
 host_neighbourhood                                | Downtown Vancouver
 host_listings_count                               | 1
 host_total_listings_count                         | 1
 host_verifications                                | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                              | t
 host_identity_verified                            | t
 neighbourhood                                     | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                            | Downtown
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.28206
 longitude                                         | -123.12461
 property_type                                     | Entire apartment
 room_type                                         | Entire home/apt
 accommodates                                      | 2
 bathrooms                                         | null
 bathrooms_text                                    | 1 bath
 bedrooms                                          | 1
 beds                                              | 1
 amenities                                         | ['Kitchen', 'Cooking basics', 'Iron', 'Elevator', 'Gym',
'Hangers', 'Cable TV', 'Smoke alarm', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Carbon monoxide alarm',
 'Wifi', 'TV', 'Long term stays allowed', 'Hair dryer', 'Shampoo', 'Dishes and silverware', 'Dedicated works
pace']
 price                                             | 168
 minimum_nights                                    | 180
 maximum_nights                                    | 180
 minimum_minimum_nights                            | 180
 maximum_minimum_nights                            | 180
 minimum_maximum_nights                            | 180
 maximum_maximum_nights                            | 180
 minimum_nights_avg_ntm                            | 180.0
 maximum_nights_avg_ntm                            | 180.0
```

```
 has_availability                                    | t
 availability_30                                     | 30
 availability_60                                     | 60
 availability_90                                     | 90
 availability_365                                    | 141
 instant_bookable                                    | f
 calculated_host_listings_count                      | 1
 calculated_host_listings_count_entire_homes         | 1
 calculated_host_listings_count_private_rooms        | 0
 calculated_host_listings_count_shared_rooms         | 0
 neighbourhood_reduced                               | Downtown
 property_reduced                                    | Entire apartment
 room_reduced                                        | Entire home/apt
 neighbourhood_indexed                               | 0.0
 property_indexed                                    | 0.0
 room_indexed                                        | 0.0
 log_price                                           | 5.123963979403259
 host_listings_count_imp                             | 1
 host_total_listings_count_imp                       | 1
 accommodates_imp                                    | 2
 bedrooms_imp                                        | 1
 beds_imp                                            | 1
 minimum_nights_imp                                  | 180
 maximum_nights_imp                                  | 180
 minimum_minimum_nights_imp                          | 180
 maximum_minimum_nights_imp                          | 180
 minimum_nights_avg_ntm_imp                          | 180.0
 calculated_host_listings_count_imp                  | 1
 calculated_host_listings_count_entire_homes_imp     | 1
 calculated_host_listings_count_private_rooms_imp    | 0
 neighbourhood_indexed_imp                           | 0.0
 property_indexed_imp                                | 0.0
 room_indexed_imp                                    | 0.0
 has_tv                                              | true
 has_wifi                                            | true
 features                                            | [1.0,1.0,2.0,1.0,1.0,180.0,180.0,180.0,180.0,180.0,1.0,
1.0,0.0,0.0,0.0,0.0,1.0,1.0]
 gps                                                 | [49.28206,-123.12461]
-RECORD 19-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                                  | 359942
 host_since                                          | 2011/09/23
 host_location                                       | Vancouver, British Columbia, Canada
```

```
 host_response_time                                | within an hour
 host_response_rate                                | 1
 host_acceptance_rate                              | N/A
 host_is_superhost                                 | f
 host_neighbourhood                                | Yaletown
 host_listings_count                               | 8
 host_total_listings_count                         | 8
 host_verifications                                | ['email', 'phone', 'facebook', 'reviews']
 host_has_profile_pic                              | t
 host_identity_verified                            | f
 neighbourhood                                     | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                            | Downtown
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.27825
 longitude                                         | -123.11865
 property_type                                     | Private room in apartment
 room_type                                         | Private room
 accommodates                                      | 2
 bathrooms                                         | null
 bathrooms_text                                    | 1 shared bath
 bedrooms                                          | 2
 beds                                              | 1
 amenities                                         | ['Kitchen', 'Lake access', 'Oven', 'Paid parking on prem
ises', 'Paid parking off premises', 'Gym', 'Microwave', 'Wifi', 'Dryer', 'Stove', 'Hot water', 'Dishwasher',
'Extra pillows and blankets', 'Cooking basics', 'Heating', 'Lock on bedroom door', 'Hair dryer', 'Iron', 'Co
ffee maker', 'Bed linens', 'Pocket wifi', 'Shampoo', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Es
sentials', 'Washer', 'Long term stays allowed', 'Elevator']
 price                                             | 240
 minimum_nights                                    | 30
 maximum_nights                                    | 1125
 minimum_minimum_nights                            | 30
 maximum_minimum_nights                            | 30
 minimum_maximum_nights                            | 1125
 maximum_maximum_nights                            | 1125
 minimum_nights_avg_ntm                            | 30.0
 maximum_nights_avg_ntm                            | 1125.0
 has_availability                                  | t
 availability_30                                   | 29
 availability_60                                   | 59
 availability_90                                   | 89
 availability_365                                  | 364
 instant_bookable                                  | f
 calculated_host_listings_count                    | 7
 calculated_host_listings_count_entire_homes       | 3
 calculated_host_listings_count_private_rooms      | 4
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | Downtown
 property_reduced                                  | Private room in apartment
 room_reduced                                      | Private room
```

```
 neighbourhood_indexed                              | 0.0
 property_indexed                                   | 6.0
 room_indexed                                       | 1.0
 log_price                                          | 5.480638923341991
 host_listings_count_imp                            | 8
 host_total_listings_count_imp                      | 8
 accommodates_imp                                   | 2
 bedrooms_imp                                       | 2
 beds_imp                                           | 1
 minimum_nights_imp                                 | 30
 maximum_nights_imp                                 | 1125
 minimum_minimum_nights_imp                         | 30
 maximum_minimum_nights_imp                         | 30
 minimum_nights_avg_ntm_imp                         | 30.0
 calculated_host_listings_count_imp                 | 7
 calculated_host_listings_count_entire_homes_imp    | 3
 calculated_host_listings_count_private_rooms_imp   | 4
 neighbourhood_indexed_imp                          | 0.0
 property_indexed_imp                               | 6.0
 room_indexed_imp                                   | 1.0
 has_tv                                             | false
 has_wifi                                           | true
 features                                           | [8.0,8.0,2.0,2.0,1.0,30.0,1125.0,30.0,30.0,30.0,7.0,3.0,
4.0,0.0,6.0,1.0,0.0,1.0]
 gps                                                | [49.27825,-123.11865]
only showing top 20 rows
```

Now we can calculate the distance to the closest SkyTrain station. To get more accurate results, the curvature of the earth needs to be taken into account using a GIS library to compute distances. We will ignore this issue in this example. That means that differences in longitude will be stronger than differences in latitude, as Vancouver is rather North.

In [65]:

```
class SkyTrainDistance(UnaryTransformer):
    def __init__(self, stops):
        super(UnaryTransformer, self).__init__()
        self.stops = stops

    def createTransformFunc(self):
        def closest(x):
            latx = float(x[0])
            longx = float(x[1])
            closest_dist = None
            for lat, long in self.stops:
                dist = (latx-lat)**2 + (longx-long)**2
                if closest_dist is None or dist < closest_dist:
```

```
                    closest_dist = dist
            return closest_dist
        return closest

    def outputDataType(self):
        return DoubleType()

    def validateInputType(self, inputType):
        assert inputType == VectorUDT(), f'Expected VectorUDT() and found {inputType}'

skytrainer = SkyTrainDistance(stops).\
  setInputCol("gps"). \
  setOutputCol("sky_distance")
stages.append(skytrainer)
```

In [66]:

```
pipeline = Pipeline(stages=stages)
df = pipeline.fit(dataset).transform(dataset)

df.show(vertical=True,truncate=False)
```

```
-RECORD 0-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                                  | 13358
 host_since                                          | 2009/11/07
 host_location                                       | Vancouver, British Columbia, Canada
 host_response_time                                  | within a day
 host_response_rate                                  | 0.71
 host_acceptance_rate                                | 0.71
 host_is_superhost                                   | t
 host_neighbourhood                                  | Downtown Vancouver
 host_listings_count                                 | 1
 host_total_listings_count                           | 1
 host_verifications                                  | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'offl
ine_government_id', 'selfie', 'government_id']
 host_has_profile_pic                                | f
 host_identity_verified                              | t
 neighbourhood                                       | null
 neighbourhood_cleansed                              | West End
 neighbourhood_group_cleansed                        | null
 latitude                                            | 49.28209
```

```
 longitude                                          | -123.12515
 property_type                                      | Entire condominium
 room_type                                          | Entire home/apt
 accommodates                                       | 2
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 1
 beds                                               | 1
 amenities                                          | ['Kitchen', 'Oven', 'Gym', 'Paid parking off premises',
'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Extra pillows and blankets',
'Dishwasher', 'Cooking basics', 'Hot water kettle', 'Smoke alarm', 'Heating', 'Single level home', 'Hair dry
er', 'Dedicated workspace', 'Iron', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Hange
rs', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditionin
g', 'Elevator']
 price                                              | 90
 minimum_nights                                     | 30
 maximum_nights                                     | 60
 minimum_minimum_nights                             | 30
 maximum_minimum_nights                             | 30
 minimum_maximum_nights                             | 60
 maximum_maximum_nights                             | 60
 minimum_nights_avg_ntm                             | 30.0
 maximum_nights_avg_ntm                             | 60.0
 has_availability                                   | t
 availability_30                                    | 11
 availability_60                                    | 41
 availability_90                                    | 70
 availability_365                                   | 70
 instant_bookable                                   | f
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 1
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | West End
 property_reduced                                   | Entire condominium
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 3.0
 property_indexed                                   | 1.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Oven, Gym, Paid parking off premises, Microwav
e, Wifi, Dryer, Carbon monoxide alarm, Stove, Hot water, Extra pillows and blankets, Dishwasher, Cooking bas
ics, Hot water kettle, Smoke alarm, Heating, Single level home, Hair dryer, Dedicated workspace, Iron, Coffe
e maker, Bed linens, Shampoo, Dishes and silverware, Hangers, Cable TV, Refrigerator, Essentials, Washer, Lo
ng term stays allowed, TV, Air conditioning, Elevator]
 verifications_array                                | [email, phone, facebook, reviews, jumio, offline_governm
ent_id, selfie, government_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,18,19],[1.
0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
```

```
 since_year                                          | 2009
 log_price                                           | 4.499809670330265
 host_listings_count_imp                             | 1
 host_total_listings_count_imp                       | 1
 accommodates_imp                                    | 2
 bedrooms_imp                                        | 1
 beds_imp                                            | 1
 minimum_nights_imp                                  | 30
 maximum_nights_imp                                  | 60
 minimum_minimum_nights_imp                          | 30
 maximum_minimum_nights_imp                          | 30
 minimum_nights_avg_ntm_imp                          | 30.0
 calculated_host_listings_count_imp                  | 1
 calculated_host_listings_count_entire_homes_imp     | 1
 calculated_host_listings_count_private_rooms_imp    | 0
 neighbourhood_indexed_imp                           | 3.0
 property_indexed_imp                                | 1.0
 room_indexed_imp                                    | 0.0
 gps                                                 | [49.28209,-123.12515]
 sky_distance                                        | 3.761514361999455E-5
-RECORD 1-----------------------
```

```
 accommodates                                       | 4
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 2
 beds                                               | 3
 amenities                                          | ['Kitchen', 'Oven', 'Microwave', 'Wifi', 'Dryer', 'Carbo
n monoxide alarm', 'Stove', 'Hot water', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking b
asics', 'Private entrance', 'Smoke alarm', 'Hair dryer', 'Lockbox', 'Dedicated workspace', 'Iron', 'Free str
eet parking', 'Free parking on premises', 'Coffee maker', 'Bed linens', 'Shampoo', 'Dishes and silverware',
 'Fire extinguisher', 'Patio or balcony', 'Children\u2019s books and toys', 'Hangers', 'Refrigerator', 'Esse
ntials', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
 price                                              | 195
 minimum_nights                                     | 3
 maximum_nights                                     | 31
 minimum_minimum_nights                             | 3
 maximum_minimum_nights                             | 3
 minimum_maximum_nights                             | 1125
 maximum_maximum_nights                             | 1125
 minimum_nights_avg_ntm                             | 3.0
 maximum_nights_avg_ntm                             | 1125.0
 has_availability                                   | t
 availability_30                                    | 0
 availability_60                                    | 0
 availability_90                                    | 0
 availability_365                                   | 228
 instant_bookable                                   | t
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 1
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Hastings-Sunrise
 property_reduced                                   | Entire guest suite
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 9.0
 property_indexed                                   | 4.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Oven, Microwave, Wifi, Dryer, Carbon monoxide
 alarm, Stove, Hot water, Extra pillows and blankets, Luggage dropoff allowed, Cooking basics, Private entra
nce, Smoke alarm, Hair dryer, Lockbox, Dedicated workspace, Iron, Free street parking, Free parking on premi
ses, Coffee maker, Bed linens, Shampoo, Dishes and silverware, Fire extinguisher, Patio or balcony, Children
\u2019s books and toys, Hangers, Refrigerator, Essentials, Washer, Long term stays allowed, Pack \u2019n Pla
y/travel crib]
 verifications_array                                | [phone, facebook, offline_government_id, selfie, governm
ent_id, identity_manual]
 amenities_vector                                   | (20,[0,2,3,4,5,6,7,8,9,11,12,13,14,15,16,17,18,19],[1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,2,3],[1.0,1.0,1.0])
 since_year                                         | 2009
 log_price                                          | 5.272999558563747
```

```
 host_listings_count_imp                          | 1
 host_total_listings_count_imp                    | 1
 accommodates_imp                                 | 4
 bedrooms_imp                                     | 2
 beds_imp                                         | 3
 minimum_nights_imp                               | 3
 maximum_nights_imp                               | 31
 minimum_minimum_nights_imp                       | 3
 maximum_minimum_nights_imp                       | 3
 minimum_nights_avg_ntm_imp                       | 3.0
 calculated_host_listings_count_imp               | 1
 calculated_host_listings_count_entire_homes_imp  | 1
 calculated_host_listings_count_private_rooms_imp | 0
 neighbourhood_indexed_imp                        | 9.0
 property_indexed_imp                             | 4.0
 room_indexed_imp                                 | 0.0
 gps                                              | [49.27734,-123.04152]
 sky_distance                                     | 3.5008083389031613E-4
-RECORD 2-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                               | 17765
 host_since                                       | 2010/01/07
 host_location                                    | Vancouver, British Columbia, Canada
 host_response_time                               | N/A
 host_response_rate                               | N/A
 host_acceptance_rate                             | 1
 host_is_superhost                                | f
 host_neighbourhood                               | Mount Pleasant
 host_listings_count                              | 1
 host_total_listings_count                        | 1
 host_verifications                               | ['phone', 'facebook', 'reviews', 'jumio', 'offline_gover
nment_id', 'selfie', 'government_id']
 host_has_profile_pic                             | t
 host_identity_verified                           | t
 neighbourhood                                    | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                           | Mount Pleasant
 neighbourhood_group_cleansed                     | null
 latitude                                         | 49.26181
 longitude                                        | -123.10905
 property_type                                    | Entire apartment
 room_type                                        | Entire home/apt
 accommodates                                     | 2
 bathrooms                                        | null
```

```
 bathrooms_text                                    | 1 bath
 bedrooms                                          | 1
 beds                                              | 1
 amenities                                         | ['Kitchen', 'Wifi', 'Host greets you', 'Dryer', 'Carbon
 monoxide alarm', 'Hot water', 'Luggage dropoff allowed', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated
workspace', 'Iron', 'Free street parking', 'Free parking on premises', 'Coffee maker', 'Shampoo', 'Dishes an
d silverware', 'Fire extinguisher', 'Hangers', 'Essentials', 'Washer', 'Long term stays allowed', 'Elevato
r']
 price                                             | 94
 minimum_nights                                    | 3
 maximum_nights                                    | 60
 minimum_minimum_nights                            | 3
 maximum_minimum_nights                            | 3
 minimum_maximum_nights                            | 60
 maximum_maximum_nights                            | 60
 minimum_nights_avg_ntm                            | 3.0
 maximum_nights_avg_ntm                            | 60.0
 has_availability                                  | t
 availability_30                                   | 0
 availability_60                                   | 0
 availability_90                                   | 0
 availability_365                                  | 0
 instant_bookable                                  | f
 calculated_host_listings_count                    | 1
 calculated_host_listings_count_entire_homes       | 1
 calculated_host_listings_count_private_rooms      | 0
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | Mount Pleasant
 property_reduced                                  | Entire apartment
 room_reduced                                      | Entire home/apt
 neighbourhood_indexed                             | 5.0
 property_indexed                                  | 0.0
 room_indexed                                      | 0.0
 amenities_array                                   | [Kitchen, Wifi, Host greets you, Dryer, Carbon monoxide
 alarm, Hot water, Luggage dropoff allowed, Smoke alarm, Heating, Hair dryer, Dedicated workspace, Iron, Fre
e street parking, Free parking on premises, Coffee maker, Shampoo, Dishes and silverware, Fire extinguisher,
Hangers, Essentials, Washer, Long term stays allowed, Elevator]
 verifications_array                               | [phone, facebook, reviews, jumio, offline_government_id,
selfie, government_id]
 amenities_vector                                  | (20,[0,1,2,3,4,5,6,7,8,9,11,12,13,14,15,16,17],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,2,3,4],[1.0,1.0,1.0,1.0])
 since_year                                        | 2010
 log_price                                         | 4.543294782270004
 host_listings_count_imp                           | 1
 host_total_listings_count_imp                     | 1
 accommodates_imp                                  | 2
 bedrooms_imp                                      | 1
 beds_imp                                          | 1
```

```
 minimum_nights_imp                                | 3
 maximum_nights_imp                                | 60
 minimum_minimum_nights_imp                        | 3
 maximum_minimum_nights_imp                        | 3
 minimum_nights_avg_ntm_imp                        | 3.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 1
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 5.0
 property_indexed_imp                              | 0.0
 room_indexed_imp                                  | 0.0
 gps                                               | [49.26181,-123.10905]
 sky_distance                                      | 3.453697089999319E-5
-RECORD 3-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                                | 18795
 host_since                                        | 2010/01/13
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | within a day
 host_response_rate                                | 0.7
 host_acceptance_rate                              | 0.4
 host_is_superhost                                 | f
 host_neighbourhood                                | West End
 host_listings_count                               | 3
 host_total_listings_count                         | 3
 host_verifications                                | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                              | t
 host_identity_verified                            | t
 neighbourhood                                     | null
 neighbourhood_cleansed                            | West End
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.28251
 longitude                                         | -123.12827
 property_type                                     | Private room in house
 room_type                                         | Private room
 accommodates                                      | 2
 bathrooms                                         | null
 bathrooms_text                                    | 1 private bath
 bedrooms                                          | 1
 beds                                              | 1
 amenities                                         | ['Kitchen', 'Cable TV', 'Wifi', 'Essentials', 'Heating',
'Washer', 'Dryer', 'Long term stays allowed']
 price                                             | 50
```

```
 minimum_nights                                    | 30
 maximum_nights                                    | 360
 minimum_minimum_nights                            | 30
 maximum_minimum_nights                            | 30
 minimum_maximum_nights                            | 360
 maximum_maximum_nights                            | 360
 minimum_nights_avg_ntm                            | 30.0
 maximum_nights_avg_ntm                            | 360.0
 has_availability                                  | t
 availability_30                                   | 9
 availability_60                                   | 20
 availability_90                                   | 50
 availability_365                                  | 325
 instant_bookable                                  | f
 calculated_host_listings_count                    | 3
 calculated_host_listings_count_entire_homes       | 0
 calculated_host_listings_count_private_rooms      | 3
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | West End
 property_reduced                                  | Private room in house
 room_reduced                                      | Private room
 neighbourhood_indexed                             | 3.0
 property_indexed                                  | 3.0
 room_indexed                                      | 1.0
 amenities_array                                   | [Kitchen, Cable TV, Wifi, Essentials, Heating, Washer, D
ryer, Long term stays allowed]
 verifications_array                               | [email, phone, reviews, jumio, government_id]
 amenities_vector                                  | (20,[0,1,3,4,6,7,8],[1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                        | 2010
 log_price                                         | 3.912023005428146
 host_listings_count_imp                           | 3
 host_total_listings_count_imp                     | 3
 accommodates_imp                                  | 2
 bedrooms_imp                                      | 1
 beds_imp                                          | 1
 minimum_nights_imp                                | 30
 maximum_nights_imp                                | 360
 minimum_minimum_nights_imp                        | 30
 maximum_minimum_nights_imp                        | 30
 minimum_nights_avg_ntm_imp                        | 30.0
 calculated_host_listings_count_imp                | 3
 calculated_host_listings_count_entire_homes_imp   | 0
 calculated_host_listings_count_private_rooms_imp  | 3
 neighbourhood_indexed_imp                         | 3.0
 property_indexed_imp                              | 3.0
 room_indexed_imp                                  | 1.0
 gps                                               | [49.28251,-123.12827]
 sky_distance                                      | 8.580917161991354E-5
```

```
-RECORD 4-----------------------
```

```
 availability_30                                    | 2
 availability_60                                    | 13
 availability_90                                    | 43
 availability_365                                   | 318
 instant_bookable                                   | f
 calculated_host_listings_count                     | 3
 calculated_host_listings_count_entire_homes        | 1
 calculated_host_listings_count_private_rooms       | 2
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Renfrew-Collingwood
 property_reduced                                   | Entire house
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 11.0
 property_indexed                                   | 2.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Oven, Microwave, Wifi, Dryer, Carbon monoxide
 alarm, Stove, Hot water, Luggage dropoff allowed, Private entrance, Smoke alarm, Heating, Single level hom
e, Hair dryer, Dedicated workspace, Iron, Free street parking, Free parking on premises, Coffee maker, Bed l
inens, Keypad, Shampoo, Dishes and silverware, Fire extinguisher, Hangers, Cable TV, Refrigerator, Essential
s, Washer, Long term stays allowed, TV]
 verifications_array                                | [email, phone, reviews, offline_government_id, selfie, g
overnment_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18],[1.
0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                         | 2010
 log_price                                          | 4.382026634673881
 host_listings_count_imp                            | 3
 host_total_listings_count_imp                      | 3
 accommodates_imp                                   | 2
 bedrooms_imp                                       | 1
 beds_imp                                           | 1
 minimum_nights_imp                                 | 28
 maximum_nights_imp                                 | 400
 minimum_minimum_nights_imp                         | 28
 maximum_minimum_nights_imp                         | 28
 minimum_nights_avg_ntm_imp                         | 28.0
 calculated_host_listings_count_imp                 | 3
 calculated_host_listings_count_entire_homes_imp    | 1
 calculated_host_listings_count_private_rooms_imp   | 2
 neighbourhood_indexed_imp                          | 11.0
 property_indexed_imp                               | 2.0
 room_indexed_imp                                   | 0.0
 gps                                                | [49.23427,-123.05309]
 sky_distance                                       | 1.492042164098205E-4
-RECORD 5-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
```

```
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
 id                                                | 24273
 host_since                                        | 2011/10/09
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | within an hour
 host_response_rate                                | 1
 host_acceptance_rate                              | 0.82
 host_is_superhost                                 | f
 host_neighbourhood                                | Kitsilano
 host_listings_count                               | 2
 host_total_listings_count                         | 2
 host_verifications                                | ['email', 'phone', 'reviews', 'jumio', 'offline_governme
nt_id', 'government_id']
 host_has_profile_pic                              | f
 host_identity_verified                            | t
 neighbourhood                                     | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                            | Kitsilano
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.2626
 longitude                                         | -123.17308
 property_type                                     | Entire apartment
 room_type                                         | Entire home/apt
 accommodates                                      | 6
 bathrooms                                         | null
 bathrooms_text                                    | 1 bath
 bedrooms                                          | 2
 beds                                              | 3
 amenities                                         | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wi
fi', 'Dryer', 'Fireplace guards', 'Carbon monoxide alarm', 'Stove', 'Crib', 'Hot water', 'Babysitter recomme
ndations', 'Garden or backyard', 'Extra pillows and blankets', 'Dishwasher', 'Luggage dropoff allowed', 'Coo
king basics', 'Room-darkening shades', 'Private entrance', 'Smoke alarm', 'Baking sheet', 'Heating', 'Single
level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Coffee maker',
 'Bed linens', 'Keypad', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Shower gel', 'Patio or ba
lcony', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials',
'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'TV', 'Ethernet connection']
 price                                             | 200
 minimum_nights                                    | 30
 maximum_nights                                    | 365
 minimum_minimum_nights                            | 30
 maximum_minimum_nights                            | 30
 minimum_maximum_nights                            | 365
 maximum_maximum_nights                            | 365
 minimum_nights_avg_ntm                            | 30.0
 maximum_nights_avg_ntm                            | 365.0
 has_availability                                  | t
 availability_30                                   | 0
```

```
 availability_60                                    | 0
 availability_90                                    | 10
 availability_365                                   | 223
 instant_bookable                                   | f
 calculated_host_listings_count                     | 2
 calculated_host_listings_count_entire_homes        | 2
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Kitsilano
 property_reduced                                   | Entire apartment
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 2.0
 property_indexed                                   | 0.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Oven, Indoor fireplace, Microwave, Wifi, Drye
r, Fireplace guards, Carbon monoxide alarm, Stove, Crib, Hot water, Babysitter recommendations, Garden or ba
ckyard, Extra pillows and blankets, Dishwasher, Luggage dropoff allowed, Cooking basics, Room-darkening shad
es, Private entrance, Smoke alarm, Baking sheet, Heating, Single level home, Hair dryer, Dedicated workspac
e, Iron, Free street parking, Bathtub, Coffee maker, Bed linens, Keypad, Shampoo, Fire extinguisher, Dishes
 and silverware, Shower gel, Patio or balcony, Children\u2019s books and toys, High chair, Hangers, Cable T
V, Refrigerator, Essentials, Washer, Long term stays allowed, Pack \u2019n Play/travel crib, TV, Ethernet co
nnection]
 verifications_array                                | [email, phone, reviews, jumio, offline_government_id, go
vernment_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19],
[1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                         | 2011
 log_price                                          | 5.298317366548036
 host_listings_count_imp                            | 2
 host_total_listings_count_imp                      | 2
 accommodates_imp                                   | 6
 bedrooms_imp                                       | 2
 beds_imp                                           | 3
 minimum_nights_imp                                 | 30
 maximum_nights_imp                                 | 365
 minimum_minimum_nights_imp                         | 30
 maximum_minimum_nights_imp                         | 30
 minimum_nights_avg_ntm_imp                         | 30.0
 calculated_host_listings_count_imp                 | 2
 calculated_host_listings_count_entire_homes_imp    | 2
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 2.0
 property_indexed_imp                               | 0.0
 room_indexed_imp                                   | 0.0
 gps                                                | [49.2626,-123.17308]
 sky_distance                                       | 0.0027660990694603196
-RECORD 6-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
```

```
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------
 id                                                 | 42645
 host_since                                         | 2010/04/17
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within an hour
 host_response_rate                                 | 1
 host_acceptance_rate                               | 1
 host_is_superhost                                  | t
 host_neighbourhood                                 | Downtown Vancouver
 host_listings_count                                | 4
 host_total_listings_count                          | 4
 host_verifications                                 | ['email', 'phone', 'facebook', 'reviews']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Downtown
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.28001
 longitude                                          | -123.12475
 property_type                                      | Entire condominium
 room_type                                          | Entire home/apt
 accommodates                                       | 3
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 1
 beds                                               | 2
 amenities                                          | ['Kitchen', 'Oven', 'Paid parking on premises', 'Paid pa
rking off premises', 'Gym', 'Microwave', 'Wifi', 'Host greets you', 'Dryer', 'Carbon monoxide alarm', 'Stov
e', 'Hot water', 'Dishwasher', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair
 dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Coffee maker', 'Shampoo', 'Fire extinguisher', 'Di
shes and silverware', 'Patio or balcony', 'Children\u2019s books and toys', 'High chair', 'Hangers', 'Cable
 TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib', 'T
V', 'Elevator']
 price                                              | 159
 minimum_nights                                     | 32
 maximum_nights                                     | 365
 minimum_minimum_nights                             | 32
 maximum_minimum_nights                             | 32
 minimum_maximum_nights                             | 1125
 maximum_maximum_nights                             | 1125
 minimum_nights_avg_ntm                             | 32.0
 maximum_nights_avg_ntm                             | 1125.0
 has_availability                                   | t
 availability_30                                    | 0
```

```
 availability_60                                   | 29
 availability_90                                   | 59
 availability_365                                  | 333
 instant_bookable                                  | t
 calculated_host_listings_count                    | 4
 calculated_host_listings_count_entire_homes       | 4
 calculated_host_listings_count_private_rooms      | 0
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | Downtown
 property_reduced                                  | Entire condominium
 room_reduced                                      | Entire home/apt
 neighbourhood_indexed                             | 0.0
 property_indexed                                  | 1.0
 room_indexed                                      | 0.0
 amenities_array                                   | [Kitchen, Oven, Paid parking on premises, Paid parking o
ff premises, Gym, Microwave, Wifi, Host greets you, Dryer, Carbon monoxide alarm, Stove, Hot water, Dishwash
er, Luggage dropoff allowed, Cooking basics, Smoke alarm, Heating, Hair dryer, Dedicated workspace, First ai
d kit, Iron, Coffee maker, Shampoo, Fire extinguisher, Dishes and silverware, Patio or balcony, Children\u20
19s books and toys, High chair, Hangers, Cable TV, Refrigerator, Essentials, Washer, Long term stays allowe
d, Pack \u2019n Play/travel crib, TV, Elevator]
 verifications_array                               | [email, phone, facebook, reviews]
 amenities_vector                                  | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19],
[1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,1,4],[1.0,1.0,1.0])
 since_year                                        | 2010
 log_price                                         | 5.0689042022202315
 host_listings_count_imp                           | 4
 host_total_listings_count_imp                     | 4
 accommodates_imp                                  | 3
 bedrooms_imp                                      | 1
 beds_imp                                          | 2
 minimum_nights_imp                                | 32
 maximum_nights_imp                                | 365
 minimum_minimum_nights_imp                        | 32
 maximum_minimum_nights_imp                        | 32
 minimum_nights_avg_ntm_imp                        | 32.0
 calculated_host_listings_count_imp                | 4
 calculated_host_listings_count_entire_homes_imp   | 4
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 0.0
 property_indexed_imp                              | 1.0
 room_indexed_imp                                  | 0.0
 gps                                               | [49.28001,-123.12475]
 sky_distance                                      | 3.7132247620008065E-5
-RECORD 7-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
```

```
-------------------
```

```
 calculated_host_listings_count                   | 2
 calculated_host_listings_count_entire_homes      | 2
 calculated_host_listings_count_private_rooms     | 0
 calculated_host_listings_count_shared_rooms      | 0
 neighbourhood_reduced                            | Downtown
 property_reduced                                 | Entire condominium
 room_reduced                                     | Entire home/apt
 neighbourhood_indexed                            | 0.0
 property_indexed                                 | 1.0
 room_indexed                                     | 0.0
 amenities_array                                  | [Kitchen, Hot tub, Gym, Indoor fireplace, Wifi, Host gre
ets you, Dryer, Carbon monoxide alarm, Hot water, Luggage dropoff allowed, Smoke alarm, Heating, Hair dryer,
Dedicated workspace, Iron, Free parking on premises, Shampoo, Fire extinguisher, Hangers, Cable TV, Essentia
ls, Washer, Long term stays allowed, TV, Air conditioning, Elevator]
 verifications_array                              | [email, phone, reviews, offline_government_id, selfie, g
overnment_id, identity_manual]
 amenities_vector                                 | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,17],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                             | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                       | 2010
 log_price                                        | 4.836281906951478
 host_listings_count_imp                          | 2
 host_total_listings_count_imp                    | 2
 accommodates_imp                                 | 4
 bedrooms_imp                                     | 1
 beds_imp                                         | 1
 minimum_nights_imp                               | 90
 maximum_nights_imp                               | 365
 minimum_minimum_nights_imp                       | 90
 maximum_minimum_nights_imp                       | 90
 minimum_nights_avg_ntm_imp                       | 90.0
 calculated_host_listings_count_imp               | 2
 calculated_host_listings_count_entire_homes_imp  | 2
 calculated_host_listings_count_private_rooms_imp | 0
 neighbourhood_indexed_imp                        | 0.0
 property_indexed_imp                             | 1.0
 room_indexed_imp                                 | 0.0
 gps                                              | [49.27655,-123.12485]
 sky_distance                                     | 1.468142545999774E-5
-RECORD 8-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                               | 72659
 host_since                                       | 2011/02/03
```

```
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | N/A
 host_response_rate                                 | N/A
 host_acceptance_rate                               | 0
 host_is_superhost                                  | f
 host_neighbourhood                                 | Commercial Drive
 host_listings_count                                | 1
 host_total_listings_count                          | 1
 host_verifications                                 | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | null
 neighbourhood_cleansed                             | Grandview-Woodland
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.27357
 longitude                                          | -123.07024
 property_type                                      | Entire condominium
 room_type                                          | Entire home/apt
 accommodates                                       | 4
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 2
 beds                                               | 2
 amenities                                          | ['Kitchen', 'Lake access', 'Oven', 'Paid parking off pre
mises', 'Gym', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot wate
r', 'Dishwasher', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Smoke alarm',
 'Heating', 'Single level home', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free street parking', 'Coffee
maker', 'Bed linens', 'Shampoo', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Hanger
s', 'Refrigerator', 'Essentials', 'Washer', 'BBQ grill', 'TV', 'Elevator']
 price                                              | 210
 minimum_nights                                     | 7
 maximum_nights                                     | 21
 minimum_minimum_nights                             | 7
 maximum_minimum_nights                             | 7
 minimum_maximum_nights                             | 21
 maximum_maximum_nights                             | 21
 minimum_nights_avg_ntm                             | 7.0
 maximum_nights_avg_ntm                             | 21.0
 has_availability                                   | t
 availability_30                                    | 0
 availability_60                                    | 0
 availability_90                                    | 0
 availability_365                                   | 40
 instant_bookable                                   | f
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 1
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Grandview-Woodland
```

```
 property_reduced                                   | Entire condominium
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 8.0
 property_indexed                                   | 1.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Lake access, Oven, Paid parking off premises,
 Gym, Indoor fireplace, Microwave, Wifi, Dryer, Carbon monoxide alarm, Stove, Hot water, Dishwasher, Extra p
illows and blankets, Luggage dropoff allowed, Cooking basics, Smoke alarm, Heating, Single level home, Hair
 dryer, Dedicated workspace, Iron, Free street parking, Coffee maker, Bed linens, Shampoo, Dishes and silver
ware, Fire extinguisher, Patio or balcony, Hangers, Refrigerator, Essentials, Washer, BBQ grill, TV, Elevato
r]
 verifications_array                                | [email, phone, reviews, jumio, government_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,7,8,9,10,11,12,13,14,15,16,17,18,19],
[1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                         | 2011
 log_price                                          | 5.3471075307174685
 host_listings_count_imp                            | 1
 host_total_listings_count_imp                      | 1
 accommodates_imp                                   | 4
 bedrooms_imp                                       | 2
 beds_imp                                           | 2
 minimum_nights_imp                                 | 7
 maximum_nights_imp                                 | 21
 minimum_minimum_nights_imp                         | 7
 maximum_minimum_nights_imp                         | 7
 minimum_nights_avg_ntm_imp                         | 7.0
 calculated_host_listings_count_imp                 | 1
 calculated_host_listings_count_entire_homes_imp    | 1
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 8.0
 property_indexed_imp                               | 1.0
 room_indexed_imp                                   | 0.0
 gps                                                | [49.27357,-123.07024]
 sky_distance                                       | 1.1663483690003356E-4
-RECORD 9-----------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                                 | 77157
 host_since                                         | 2011/02/27
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within a few hours
 host_response_rate                                 | 1
 host_acceptance_rate                               | 1
```

```
 host_is_superhost                                  | f
 host_neighbourhood                                 | Arbutus Ridge
 host_listings_count                                | 2
 host_total_listings_count                          | 2
 host_verifications                                 | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'offl
ine_government_id', 'government_id']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Arbutus Ridge
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.24675
 longitude                                          | -123.16711
 property_type                                      | Private room in guest suite
 room_type                                          | Private room
 accommodates                                       | 3
 bathrooms                                          | null
 bathrooms_text                                     | 1 private bath
 bedrooms                                           | 1
 beds                                               | 2
 amenities                                          | ['Kitchen', 'Microwave', 'Wifi', 'Laundromat nearby', 'D
ryer', 'Hot water', 'Garden or backyard', 'Extra pillows and blankets', 'Luggage dropoff allowed', 'Cooking
 basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Mini fridge', 'Single level home', 'Hair dryer', 'L
ockbox', 'Dedicated workspace', 'First aid kit', 'Iron', 'Body soap', 'Free street parking', 'Coffee maker',
'Bed linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Barbecue utensils', 'Shower gel', 'H
angers', 'Cable TV', 'Essentials', 'Washer', 'Conditioner', 'BBQ grill', 'Beach essentials', 'TV', 'Long ter
m stays allowed']
 price                                              | 75
 minimum_nights                                     | 3
 maximum_nights                                     | 60
 minimum_minimum_nights                             | 3
 maximum_minimum_nights                             | 3
 minimum_maximum_nights                             | 1125
 maximum_maximum_nights                             | 1125
 minimum_nights_avg_ntm                             | 3.0
 maximum_nights_avg_ntm                             | 1125.0
 has_availability                                   | t
 availability_30                                    | 3
 availability_60                                    | 3
 availability_90                                    | 3
 availability_365                                   | 267
 instant_bookable                                   | t
 calculated_host_listings_count                     | 2
 calculated_host_listings_count_entire_homes        | 0
 calculated_host_listings_count_private_rooms       | 2
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | OTHER
 property_reduced                                   | OTHER
 room_reduced                                       | Private room
```

```
 neighbourhood_indexed                              | 1.0
 property_indexed                                   | 5.0
 room_indexed                                       | 1.0
 amenities_array                                    | [Kitchen, Microwave, Wifi, Laundromat nearby, Dryer, Hot
water, Garden or backyard, Extra pillows and blankets, Luggage dropoff allowed, Cooking basics, Private entr
ance, Smoke alarm, Heating, Mini fridge, Single level home, Hair dryer, Lockbox, Dedicated workspace, First
 aid kit, Iron, Body soap, Free street parking, Coffee maker, Bed linens, Shampoo, Fire extinguisher, Dishes
and silverware, Barbecue utensils, Shower gel, Hangers, Cable TV, Essentials, Washer, Conditioner, BBQ gril
l, Beach essentials, TV, Long term stays allowed]
 verifications_array                                | [email, phone, facebook, reviews, jumio, offline_governm
ent_id, government_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,16,17,19],[1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                         | 2011
 log_price                                          | 4.31748811353631
 host_listings_count_imp                            | 2
 host_total_listings_count_imp                      | 2
 accommodates_imp                                   | 3
 bedrooms_imp                                       | 1
 beds_imp                                           | 2
 minimum_nights_imp                                 | 3
 maximum_nights_imp                                 | 60
 minimum_minimum_nights_imp                         | 3
 maximum_minimum_nights_imp                         | 3
 minimum_nights_avg_ntm_imp                         | 3.0
 calculated_host_listings_count_imp                 | 2
 calculated_host_listings_count_entire_homes_imp    | 0
 calculated_host_listings_count_private_rooms_imp   | 2
 neighbourhood_indexed_imp                          | 1.0
 property_indexed_imp                               | 5.0
 room_indexed_imp                                   | 1.0
 gps                                                | [49.24675,-123.16711]
 sky_distance                                       | 0.0026776233901299497
-RECORD 10-------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                                 | 111077
 host_since                                         | 2011/05/09
 host_location                                      | North Vancouver, British Columbia, Canada
 host_response_time                                 | within a few hours
 host_response_rate                                 | 1
 host_acceptance_rate                               | 0.33
 host_is_superhost                                  | f
```

```
 host_neighbourhood                                 | Grandview-Woodland
 host_listings_count                                | 1
 host_total_listings_count                          | 1
 host_verifications                                 | ['email', 'phone', 'facebook', 'reviews', 'jumio', 'gove
rnment_id']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Grandview-Woodland
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.26726
 longitude                                          | -123.05969
 property_type                                      | Entire guest suite
 room_type                                          | Entire home/apt
 accommodates                                       | 3
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 2
 beds                                               | 2
 amenities                                          | ['Kitchen', 'Indoor fireplace', 'Microwave', 'Wifi', 'Dr
yer', 'Stove', 'Hot water', 'Cooking basics', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'I
ron', 'Free street parking', 'Coffee maker', 'Keypad', 'Shampoo', 'Dishes and silverware', 'Fire extinguishe
r', 'Hangers', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays allowed']
 price                                              | 125
 minimum_nights                                     | 3
 maximum_nights                                     | 365
 minimum_minimum_nights                             | 3
 maximum_minimum_nights                             | 3
 minimum_maximum_nights                             | 365
 maximum_maximum_nights                             | 365
 minimum_nights_avg_ntm                             | 3.0
 maximum_nights_avg_ntm                             | 365.0
 has_availability                                   | t
 availability_30                                    | 4
 availability_60                                    | 34
 availability_90                                    | 64
 availability_365                                   | 339
 instant_bookable                                   | f
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 1
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Grandview-Woodland
 property_reduced                                   | Entire guest suite
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 8.0
 property_indexed                                   | 4.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Indoor fireplace, Microwave, Wifi, Dryer, Stov
```

```
e, Hot water, Cooking basics, Private entrance, Smoke alarm, Heating, Hair dryer, Iron, Free street parking,
Coffee maker, Keypad, Shampoo, Dishes and silverware, Fire extinguisher, Hangers, Refrigerator, Essentials,
 Washer, Long term stays allowed]
 verifications_array                              | [email, phone, facebook, reviews, jumio, government_id]
 amenities_vector                                 | (20,[0,1,2,3,4,5,6,7,8,9,12,13,14,16,17,18,19],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                             | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                       | 2011
 log_price                                        | 4.8283137373023015
 host_listings_count_imp                          | 1
 host_total_listings_count_imp                    | 1
 accommodates_imp                                 | 3
 bedrooms_imp                                     | 2
 beds_imp                                         | 2
 minimum_nights_imp                               | 3
 maximum_nights_imp                               | 365
 minimum_minimum_nights_imp                       | 3
 maximum_minimum_nights_imp                       | 3
 minimum_nights_avg_ntm_imp                       | 3.0
 calculated_host_listings_count_imp               | 1
 calculated_host_listings_count_entire_homes_imp  | 1
 calculated_host_listings_count_private_rooms_imp | 0
 neighbourhood_indexed_imp                        | 8.0
 property_indexed_imp                             | 4.0
 room_indexed_imp                                 | 0.0
 gps                                              | [49.26726,-123.05969]
 sky_distance                                     | 9.711975289998879E-5
-RECORD 11-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                               | 125176
 host_since                                       | 2011/05/24
 host_location                                    | Vancouver, British Columbia, Canada
 host_response_time                               | within an hour
 host_response_rate                               | 1
 host_acceptance_rate                             | 1
 host_is_superhost                                | f
 host_neighbourhood                               | Riley Park
 host_listings_count                              | 1
 host_total_listings_count                        | 1
 host_verifications                               | ['email', 'phone', 'facebook', 'reviews', 'offline_gover
nment_id', 'government_id', 'work_email']
 host_has_profile_pic                             | t
 host_identity_verified                           | t
```

```
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Riley Park
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.24648
 longitude                                          | -123.09241
 property_type                                      | Private room in house
 room_type                                          | Private room
 accommodates                                       | 3
 bathrooms                                          | null
 bathrooms_text                                     | 1 private bath
 bedrooms                                           | 1
 beds                                               | 0
 amenities                                          | ['Kitchen', 'Hot tub', 'Oven', 'Wifi', 'Carbon monoxide
 alarm', 'Stove', 'Hot water', 'Garden or backyard', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair drye
r', 'Dedicated workspace', 'Iron', 'Free street parking', 'Bathtub', 'Free parking on premises', 'Keypad',
 'Shampoo', 'Dishes and silverware', 'Patio or balcony', 'Hangers', 'Cable TV', 'Refrigerator', 'Essential
s', 'Long term stays allowed', 'TV']
 price                                              | 75
 minimum_nights                                     | 2
 maximum_nights                                     | 60
 minimum_minimum_nights                             | 2
 maximum_minimum_nights                             | 2
 minimum_maximum_nights                             | 1125
 maximum_maximum_nights                             | 1125
 minimum_nights_avg_ntm                             | 2.0
 maximum_nights_avg_ntm                             | 1125.0
 has_availability                                   | t
 availability_30                                    | 0
 availability_60                                    | 0
 availability_90                                    | 2
 availability_365                                   | 268
 instant_bookable                                   | t
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 0
 calculated_host_listings_count_private_rooms       | 1
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Riley Park
 property_reduced                                   | Private room in house
 room_reduced                                       | Private room
 neighbourhood_indexed                              | 7.0
 property_indexed                                   | 3.0
 room_indexed                                       | 1.0
 amenities_array                                    | [Kitchen, Hot tub, Oven, Wifi, Carbon monoxide alarm, St
ove, Hot water, Garden or backyard, Cooking basics, Smoke alarm, Heating, Hair dryer, Dedicated workspace, I
ron, Free street parking, Bathtub, Free parking on premises, Keypad, Shampoo, Dishes and silverware, Patio o
r balcony, Hangers, Cable TV, Refrigerator, Essentials, Long term stays allowed, TV]
 verifications_array                                | [email, phone, facebook, reviews, offline_government_id,
government_id, work_email]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,9,10,11,12,13,14,15,16,18,19],[1.0,1.
```

```
0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                        | 2011
 log_price                                         | 4.31748811353631
 host_listings_count_imp                           | 1
 host_total_listings_count_imp                     | 1
 accommodates_imp                                  | 3
 bedrooms_imp                                      | 1
 beds_imp                                          | 0
 minimum_nights_imp                                | 2
 maximum_nights_imp                                | 60
 minimum_minimum_nights_imp                        | 2
 maximum_minimum_nights_imp                        | 2
 minimum_nights_avg_ntm_imp                        | 2.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 0
 calculated_host_listings_count_private_rooms_imp  | 1
 neighbourhood_indexed_imp                         | 7.0
 property_indexed_imp                              | 3.0
 room_indexed_imp                                  | 1.0
 gps                                               | [49.24648,-123.09241]
 sky_distance                                      | 5.353377121297059E-4
-RECORD 12-----------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                                | 129163
 host_since                                        | 2011/05/17
 host_location                                     | Vancouver, British Columbia, Canada
 host_response_time                                | within a few hours
 host_response_rate                                | 1
 host_acceptance_rate                              | 1
 host_is_superhost                                 | f
 host_neighbourhood                                | Renfrew-Collingwood
 host_listings_count                               | 2
 host_total_listings_count                         | 2
 host_verifications                                | ['email', 'phone', 'reviews', 'jumio', 'offline_governme
nt_id', 'government_id']
 host_has_profile_pic                              | t
 host_identity_verified                            | t
 neighbourhood                                     | null
 neighbourhood_cleansed                            | Renfrew-Collingwood
 neighbourhood_group_cleansed                      | null
 latitude                                          | 49.24507
 longitude                                         | -123.04941
```

```
 property_type                                      | Private room in house
 room_type                                          | Private room
 accommodates                                       | 1
 bathrooms                                          | null
 bathrooms_text                                     | 1 shared bath
 bedrooms                                           | 1
 beds                                               | 1
 amenities                                          | ['Kitchen', 'Hot water', 'Paid parking on premises', 'Lo
ck on bedroom door', 'Free street parking', 'Cable TV', 'Smoke alarm', 'Wifi', 'Essentials', 'Heating', 'Was
her', 'Dryer', 'Carbon monoxide alarm', 'Dishes and silverware', 'TV', 'Long term stays allowed', 'Host gree
ts you', 'Hair dryer', 'Shampoo', 'Fire extinguisher']
 price                                              | 45
 minimum_nights                                     | 30
 maximum_nights                                     | 280
 minimum_minimum_nights                             | 30
 maximum_minimum_nights                             | 30
 minimum_maximum_nights                             | 280
 maximum_maximum_nights                             | 280
 minimum_nights_avg_ntm                             | 30.0
 maximum_nights_avg_ntm                             | 280.0
 has_availability                                   | t
 availability_30                                    | 30
 availability_60                                    | 60
 availability_90                                    | 90
 availability_365                                   | 365
 instant_bookable                                   | t
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 0
 calculated_host_listings_count_private_rooms       | 1
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Renfrew-Collingwood
 property_reduced                                   | Private room in house
 room_reduced                                       | Private room
 neighbourhood_indexed                              | 11.0
 property_indexed                                   | 3.0
 room_indexed                                       | 1.0
 amenities_array                                    | [Kitchen, Hot water, Paid parking on premises, Lock on b
edroom door, Free street parking, Cable TV, Smoke alarm, Wifi, Essentials, Heating, Washer, Dryer, Carbon mo
noxide alarm, Dishes and silverware, TV, Long term stays allowed, Host greets you, Hair dryer, Shampoo, Fire
extinguisher]
 verifications_array                                | [email, phone, reviews, jumio, offline_government_id, go
vernment_id]
 amenities_vector                                   | (20,[0,1,2,3,4,6,7,8,9,10,12,13,15,16,17],[1.0,1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                         | 2011
 log_price                                          | 3.8066624897703196
 host_listings_count_imp                            | 2
 host_total_listings_count_imp                      | 2
```

```
 accommodates_imp                                   | 1
 bedrooms_imp                                       | 1
 beds_imp                                           | 1
 minimum_nights_imp                                 | 30
 maximum_nights_imp                                 | 280
 minimum_minimum_nights_imp                         | 30
 maximum_minimum_nights_imp                         | 30
 minimum_nights_avg_ntm_imp                         | 30.0
 calculated_host_listings_count_imp                 | 1
 calculated_host_listings_count_entire_homes_imp    | 0
 calculated_host_listings_count_private_rooms_imp   | 1
 neighbourhood_indexed_imp                          | 11.0
 property_indexed_imp                               | 3.0
 room_indexed_imp                                   | 1.0
 gps                                                | [49.24507,-123.04941]
 sky_distance                                       | 1.337821640990996E-5
-RECORD 13-------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------------------
 id                                                 | 174135
 host_since                                         | 2011/07/17
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within a few hours
 host_response_rate                                 | 1
 host_acceptance_rate                               | 0
 host_is_superhost                                  | f
 host_neighbourhood                                 | Downtown Eastside
 host_listings_count                                | 1
 host_total_listings_count                          | 1
 host_verifications                                 | ['email', 'phone', 'reviews', 'offline_government_id',
 'selfie', 'government_id', 'identity_manual']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Downtown Eastside
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.27782
 longitude                                          | -123.08985
 property_type                                      | Entire guest suite
 room_type                                          | Entire home/apt
 accommodates                                       | 2
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 1
```

```
 beds                                                | 2
 amenities                                           | ['Kitchen', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'H
ot water', 'Luggage dropoff allowed', 'Private entrance', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated
workspace', 'Iron', 'Free street parking', 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Hangers', 'Cable T
V', 'Essentials', 'Washer', 'Long term stays allowed', 'TV', 'Air conditioning']
 price                                               | 100
 minimum_nights                                      | 30
 maximum_nights                                      | 365
 minimum_minimum_nights                              | 30
 maximum_minimum_nights                              | 30
 minimum_maximum_nights                              | 365
 maximum_maximum_nights                              | 365
 minimum_nights_avg_ntm                              | 30.0
 maximum_nights_avg_ntm                              | 365.0
 has_availability                                    | t
 availability_30                                     | 0
 availability_60                                     | 0
 availability_90                                     | 0
 availability_365                                    | 224
 instant_bookable                                    | f
 calculated_host_listings_count                      | 1
 calculated_host_listings_count_entire_homes         | 1
 calculated_host_listings_count_private_rooms        | 0
 calculated_host_listings_count_shared_rooms         | 0
 neighbourhood_reduced                               | Downtown Eastside
 property_reduced                                    | Entire guest suite
 room_reduced                                        | Entire home/apt
 neighbourhood_indexed                               | 6.0
 property_indexed                                    | 4.0
 room_indexed                                        | 0.0
 amenities_array                                     | [Kitchen, Wifi, Dryer, Carbon monoxide alarm, Hot water,
Luggage dropoff allowed, Private entrance, Smoke alarm, Heating, Hair dryer, Dedicated workspace, Iron, Free
street parking, Bed linens, Shampoo, Fire extinguisher, Hangers, Cable TV, Essentials, Washer, Long term sta
ys allowed, TV, Air conditioning]
 verifications_array                                 | [email, phone, reviews, offline_government_id, selfie, g
overnment_id, identity_manual]
 amenities_vector                                    | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,17],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                                | (5,[0,1,2,3,4],[1.0,1.0,1.0,1.0,1.0])
 since_year                                          | 2011
 log_price                                           | 4.605170185988092
 host_listings_count_imp                             | 1
 host_total_listings_count_imp                       | 1
 accommodates_imp                                    | 2
 bedrooms_imp                                        | 1
 beds_imp                                            | 2
 minimum_nights_imp                                  | 30
 maximum_nights_imp                                  | 365
 minimum_minimum_nights_imp                          | 30
```

```
 maximum_minimum_nights_imp                       | 30
 minimum_nights_avg_ntm_imp                       | 30.0
 calculated_host_listings_count_imp               | 1
 calculated_host_listings_count_entire_homes_imp  | 1
 calculated_host_listings_count_private_rooms_imp | 0
 neighbourhood_indexed_imp                        | 6.0
 property_indexed_imp                             | 4.0
 room_indexed_imp                                 | 0.0
 gps                                              | [49.27782,-123.08985]
 sky_distance                                     | 1.3243460004991607E-4
-RECORD 14-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
 id                                               | 200643
 host_since                                       | 2011/08/16
 host_location                                    | Vancouver, British Columbia, Canada
 host_response_time                               | within an hour
 host_response_rate                               | 1
 host_acceptance_rate                             | 1
 host_is_superhost                                | f
 host_neighbourhood                               | Strathcona
 host_listings_count                              | 4
 host_total_listings_count                        | 4
 host_verifications                               | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                             | t
 host_identity_verified                           | t
 neighbourhood                                    | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                           | Downtown Eastside
 neighbourhood_group_cleansed                     | null
 latitude                                         | 49.27384
 longitude                                        | -123.08289
 property_type                                    | Private room in house
 room_type                                        | Private room
 accommodates                                     | 4
 bathrooms                                        | null
 bathrooms_text                                   | 2 shared baths
 bedrooms                                         | 1
 beds                                             | 1
 amenities                                        | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wi
fi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Garden or backyard', 'Dishwasher', 'Extra pill
ows and blankets', 'Luggage dropoff allowed', 'Cooking basics', 'Room-darkening shades', 'Smoke alarm', 'Hea
ting', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'Iron', 'Free street parking', 'Coffee maker',
 'Bed linens', 'Shampoo', 'Fire extinguisher', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Essentia
ls', 'Washer', 'Long term stays allowed', 'Pack \u2019n Play/travel crib']
```

```
 price                                          | 87
 minimum_nights                                 | 1
 maximum_nights                                 | 1125
 minimum_minimum_nights                         | 1
 maximum_minimum_nights                         | 1
 minimum_maximum_nights                         | 1125
 maximum_maximum_nights                         | 1125
 minimum_nights_avg_ntm                         | 1.0
 maximum_nights_avg_ntm                         | 1125.0
 has_availability                               | t
 availability_30                                | 30
 availability_60                                | 60
 availability_90                                | 90
 availability_365                               | 365
 instant_bookable                               | t
 calculated_host_listings_count                 | 4
 calculated_host_listings_count_entire_homes    | 0
 calculated_host_listings_count_private_rooms   | 4
 calculated_host_listings_count_shared_rooms    | 0
 neighbourhood_reduced                          | Downtown Eastside
 property_reduced                               | Private room in house
 room_reduced                                   | Private room
 neighbourhood_indexed                          | 6.0
 property_indexed                               | 3.0
 room_indexed                                   | 1.0
 amenities_array                                | [Kitchen, Oven, Indoor fireplace, Microwave, Wifi, Drye
r, Carbon monoxide alarm, Stove, Hot water, Garden or backyard, Dishwasher, Extra pillows and blankets, Lugg
age dropoff allowed, Cooking basics, Room-darkening shades, Smoke alarm, Heating, Hair dryer, Dedicated work
space, First aid kit, Iron, Free street parking, Coffee maker, Bed linens, Shampoo, Fire extinguisher, Dishe
s and silverware, Hangers, Refrigerator, Essentials, Washer, Long term stays allowed, Pack \u2019n Play/trav
el crib]
 verifications_array                            | [email, phone, reviews, jumio, government_id]
 amenities_vector                               | (20,[0,1,2,3,4,5,6,7,8,9,11,12,13,14,15,16,17,18,19],[1.
0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                           | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                     | 2011
 log_price                                      | 4.465908118654584
 host_listings_count_imp                        | 4
 host_total_listings_count_imp                  | 4
 accommodates_imp                               | 4
 bedrooms_imp                                   | 1
 beds_imp                                       | 1
 minimum_nights_imp                             | 1
 maximum_nights_imp                             | 1125
 minimum_minimum_nights_imp                     | 1
 maximum_minimum_nights_imp                     | 1
 minimum_nights_avg_ntm_imp                     | 1.0
 calculated_host_listings_count_imp             | 4
 calculated_host_listings_count_entire_homes_imp | 0
```

```
 calculated_host_listings_count_private_rooms_imp | 4
 neighbourhood_indexed_imp                        | 6.0
 property_indexed_imp                             | 3.0
 room_indexed_imp                                 | 1.0
 gps                                              | [49.27384,-123.08289]
 sky_distance                                     | 8.020456525005144E-5
-RECORD 15-------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------
 id                                               | 248014
 host_since                                       | 2011/10/16
 host_location                                    | Vancouver, British Columbia, Canada
 host_response_time                               | within a day
 host_response_rate                               | 0.5
 host_acceptance_rate                             | N/A
 host_is_superhost                                | f
 host_neighbourhood                               | Grandview-Woodland
 host_listings_count                              | 1
 host_total_listings_count                        | 1
 host_verifications                               | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                             | t
 host_identity_verified                           | t
 neighbourhood                                    | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                           | Grandview-Woodland
 neighbourhood_group_cleansed                     | null
 latitude                                         | 49.29103
 longitude                                        | -123.05782
 property_type                                    | Entire house
 room_type                                        | Entire home/apt
 accommodates                                     | 3
 bathrooms                                        | null
 bathrooms_text                                   | 1 bath
 bedrooms                                         | 2
 beds                                             | 0
 amenities                                        | ['Kitchen', 'Oven', 'Indoor fireplace', 'Microwave', 'Wi
fi', 'Host greets you', 'Dryer', 'Stove', 'Carbon monoxide alarm', 'Hot water', 'Garden or backyard', 'Dishw
asher', 'Cooking basics', 'Smoke alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'First aid kit', 'I
ron', 'Free street parking', 'Dishes and silverware', 'Fire extinguisher', 'Patio or balcony', 'Cable TV',
 'Refrigerator', 'Hangers', 'Essentials', 'Washer', 'Long term stays allowed', 'BBQ grill', 'TV', 'Cleaning
 before checkout']
 price                                            | 100
 minimum_nights                                   | 31
 maximum_nights                                   | 180
 minimum_minimum_nights                           | 31
```

```
 maximum_minimum_nights                            | 31
 minimum_maximum_nights                            | 180
 maximum_maximum_nights                            | 180
 minimum_nights_avg_ntm                            | 31.0
 maximum_nights_avg_ntm                            | 180.0
 has_availability                                  | t
 availability_30                                   | 4
 availability_60                                   | 4
 availability_90                                   | 4
 availability_365                                  | 113
 instant_bookable                                  | f
 calculated_host_listings_count                    | 1
 calculated_host_listings_count_entire_homes       | 1
 calculated_host_listings_count_private_rooms      | 0
 calculated_host_listings_count_shared_rooms       | 0
 neighbourhood_reduced                             | Grandview-Woodland
 property_reduced                                  | Entire house
 room_reduced                                      | Entire home/apt
 neighbourhood_indexed                             | 8.0
 property_indexed                                  | 2.0
 room_indexed                                      | 0.0
 amenities_array                                   | [Kitchen, Oven, Indoor fireplace, Microwave, Wifi, Host
 greets you, Dryer, Stove, Carbon monoxide alarm, Hot water, Garden or backyard, Dishwasher, Cooking basics,
Smoke alarm, Heating, Hair dryer, Dedicated workspace, First aid kit, Iron, Free street parking, Dishes and
 silverware, Fire extinguisher, Patio or balcony, Cable TV, Refrigerator, Hangers, Essentials, Washer, Long
 term stays allowed, BBQ grill, TV, Cleaning before checkout]
 verifications_array                               | [email, phone, reviews, jumio, government_id]
 amenities_vector                                  | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,14,15,16,17,18,19],[1.
0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                              | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                        | 2011
 log_price                                         | 4.605170185988092
 host_listings_count_imp                           | 1
 host_total_listings_count_imp                     | 1
 accommodates_imp                                  | 3
 bedrooms_imp                                      | 2
 beds_imp                                          | 0
 minimum_nights_imp                                | 31
 maximum_nights_imp                                | 180
 minimum_minimum_nights_imp                        | 31
 maximum_minimum_nights_imp                        | 31
 minimum_nights_avg_ntm_imp                        | 31.0
 calculated_host_listings_count_imp                | 1
 calculated_host_listings_count_entire_homes_imp   | 1
 calculated_host_listings_count_private_rooms_imp  | 0
 neighbourhood_indexed_imp                         | 8.0
 property_indexed_imp                              | 2.0
 room_indexed_imp                                  | 0.0
 gps                                               | [49.29103,-123.05782]
```

```
 sky_distance                                       | 9.0567016489999E-4
-RECORD 16-----------------------
```

```
 availability_90                                    | 41
 availability_365                                   | 316
 instant_bookable                                   | f
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 1
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Downtown
 property_reduced                                   | Entire apartment
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 0.0
 property_indexed                                   | 0.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Iron, Elevator, Gym, Hangers, Smoke alarm, Wif
i, Essentials, Heating, Washer, Dryer, Free parking on premises, Long term stays allowed, Hair dryer, Shampo
o, Dedicated workspace]
 verifications_array                                | [email, phone, reviews, jumio, government_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,11,13,14],[1.0,1.0,1.0,1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                         | 2011
 log_price                                          | 4.59511985013459
 host_listings_count_imp                            | 1
 host_total_listings_count_imp                      | 1
 accommodates_imp                                   | 4
 bedrooms_imp                                       | 1
 beds_imp                                           | 2
 minimum_nights_imp                                 | 120
 maximum_nights_imp                                 | 365
 minimum_minimum_nights_imp                         | 120
 maximum_minimum_nights_imp                         | 120
 minimum_nights_avg_ntm_imp                         | 120.0
 calculated_host_listings_count_imp                 | 1
 calculated_host_listings_count_entire_homes_imp    | 1
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 0.0
 property_indexed_imp                               | 0.0
 room_indexed_imp                                   | 0.0
 gps                                                | [49.27796,-123.12671]
 sky_distance                                       | 3.8015383460071904E-5
-RECORD 17-------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------------------
 id                                                 | 283841
```

```
 host_since                                          | 2009/08/10
 host_location                                       | Vancouver, British Columbia, Canada
 host_response_time                                  | within a day
 host_response_rate                                  | 0.93
 host_acceptance_rate                                | 0.88
 host_is_superhost                                   | f
 host_neighbourhood                                  | Coal Harbour
 host_listings_count                                 | 49
 host_total_listings_count                           | 49
 host_verifications                                  | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                                | t
 host_identity_verified                              | t
 neighbourhood                                       | null
 neighbourhood_cleansed                              | Downtown
 neighbourhood_group_cleansed                        | null
 latitude                                            | 49.2872
 longitude                                           | -123.12352
 property_type                                       | Entire apartment
 room_type                                           | Entire home/apt
 accommodates                                        | 6
 bathrooms                                           | null
 bathrooms_text                                      | 3 baths
 bedrooms                                            | 3
 beds                                                | 3
 amenities                                           | ['Kitchen', 'Hot tub', 'Oven', 'Pool', 'Gym', 'Indoor fi
replace', 'Microwave', 'Wifi', 'Dryer', 'Carbon monoxide alarm', 'Stove', 'Hot water', 'Dishwasher', 'Smoke
 alarm', 'Heating', 'Hair dryer', 'Dedicated workspace', 'Iron', 'Free parking on premises', 'Coffee maker',
'Dishes and silverware', 'Hangers', 'Cable TV', 'Refrigerator', 'Essentials', 'Washer', 'Long term stays all
owed', 'TV', 'Air conditioning']
 price                                               | 197
 minimum_nights                                      | 90
 maximum_nights                                      | 1125
 minimum_minimum_nights                              | 90
 maximum_minimum_nights                              | 90
 minimum_maximum_nights                              | 1125
 maximum_maximum_nights                              | 1125
 minimum_nights_avg_ntm                              | 90.0
 maximum_nights_avg_ntm                              | 1125.0
 has_availability                                    | t
 availability_30                                     | 0
 availability_60                                     | 0
 availability_90                                     | 0
 availability_365                                    | 0
 instant_bookable                                    | f
 calculated_host_listings_count                      | 41
 calculated_host_listings_count_entire_homes         | 41
 calculated_host_listings_count_private_rooms        | 0
 calculated_host_listings_count_shared_rooms         | 0
 neighbourhood_reduced                               | Downtown
```

```
 property_reduced                                    | Entire apartment
 room_reduced                                        | Entire home/apt
 neighbourhood_indexed                               | 0.0
 property_indexed                                    | 0.0
 room_indexed                                        | 0.0
 amenities_array                                     | [Kitchen, Hot tub, Oven, Pool, Gym, Indoor fireplace, Mi
crowave, Wifi, Dryer, Carbon monoxide alarm, Stove, Hot water, Dishwasher, Smoke alarm, Heating, Hair dryer,
Dedicated workspace, Iron, Free parking on premises, Coffee maker, Dishes and silverware, Hangers, Cable TV,
Refrigerator, Essentials, Washer, Long term stays allowed, TV, Air conditioning]
 verifications_array                                 | [email, phone, reviews, jumio, government_id]
 amenities_vector                                    | (20,[0,1,2,3,4,5,6,7,8,9,10,11,12,14,15,16,18],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                                | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                          | 2009
 log_price                                           | 5.2832037287379885
 host_listings_count_imp                             | 49
 host_total_listings_count_imp                       | 49
 accommodates_imp                                    | 6
 bedrooms_imp                                        | 3
 beds_imp                                            | 3
 minimum_nights_imp                                  | 90
 maximum_nights_imp                                  | 1125
 minimum_minimum_nights_imp                          | 90
 maximum_minimum_nights_imp                          | 90
 minimum_nights_avg_ntm_imp                          | 90.0
 calculated_host_listings_count_imp                  | 41
 calculated_host_listings_count_entire_homes_imp     | 41
 calculated_host_listings_count_private_rooms_imp    | 0
 neighbourhood_indexed_imp                           | 0.0
 property_indexed_imp                                | 0.0
 room_indexed_imp                                    | 0.0
 gps                                                 | [49.2872,-123.12352]
 sky_distance                                        | 2.4672390169983233E-5
-RECORD 18-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------
 id                                                  | 323090
 host_since                                          | 2012/01/01
 host_location                                       | Vancouver, British Columbia, Canada
 host_response_time                                  | N/A
 host_response_rate                                  | N/A
 host_acceptance_rate                                | 0.85
 host_is_superhost                                   | t
 host_neighbourhood                                  | Downtown Vancouver
```

```
 host_listings_count                                | 1
 host_total_listings_count                          | 1
 host_verifications                                 | ['email', 'phone', 'reviews', 'jumio', 'government_id']
 host_has_profile_pic                               | t
 host_identity_verified                             | t
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Downtown
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.28206
 longitude                                          | -123.12461
 property_type                                      | Entire apartment
 room_type                                          | Entire home/apt
 accommodates                                       | 2
 bathrooms                                          | null
 bathrooms_text                                     | 1 bath
 bedrooms                                           | 1
 beds                                               | 1
 amenities                                          | ['Kitchen', 'Cooking basics', 'Iron', 'Elevator', 'Gym',
'Hangers', 'Cable TV', 'Smoke alarm', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Carbon monoxide alarm',
 'Wifi', 'TV', 'Long term stays allowed', 'Hair dryer', 'Shampoo', 'Dishes and silverware', 'Dedicated works
pace']
 price                                              | 168
 minimum_nights                                     | 180
 maximum_nights                                     | 180
 minimum_minimum_nights                             | 180
 maximum_minimum_nights                             | 180
 minimum_maximum_nights                             | 180
 maximum_maximum_nights                             | 180
 minimum_nights_avg_ntm                             | 180.0
 maximum_nights_avg_ntm                             | 180.0
 has_availability                                   | t
 availability_30                                    | 30
 availability_60                                    | 60
 availability_90                                    | 90
 availability_365                                   | 141
 instant_bookable                                   | f
 calculated_host_listings_count                     | 1
 calculated_host_listings_count_entire_homes        | 1
 calculated_host_listings_count_private_rooms       | 0
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Downtown
 property_reduced                                   | Entire apartment
 room_reduced                                       | Entire home/apt
 neighbourhood_indexed                              | 0.0
 property_indexed                                   | 0.0
 room_indexed                                       | 0.0
 amenities_array                                    | [Kitchen, Cooking basics, Iron, Elevator, Gym, Hangers,
 Cable TV, Smoke alarm, Essentials, Heating, Washer, Dryer, Carbon monoxide alarm, Wifi, TV, Long term stays
allowed, Hair dryer, Shampoo, Dishes and silverware, Dedicated workspace]
```

```
 verifications_array                                | [email, phone, reviews, jumio, government_id]
 amenities_vector                                   | (20,[0,1,2,3,4,5,6,7,8,9,10,11,13,14,15,16,19],[1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,2,4],[1.0,1.0,1.0,1.0])
 since_year                                         | 2012
 log_price                                          | 5.123963979403259
 host_listings_count_imp                            | 1
 host_total_listings_count_imp                      | 1
 accommodates_imp                                   | 2
 bedrooms_imp                                       | 1
 beds_imp                                           | 1
 minimum_nights_imp                                 | 180
 maximum_nights_imp                                 | 180
 minimum_minimum_nights_imp                         | 180
 maximum_minimum_nights_imp                         | 180
 minimum_nights_avg_ntm_imp                         | 180.0
 calculated_host_listings_count_imp                 | 1
 calculated_host_listings_count_entire_homes_imp    | 1
 calculated_host_listings_count_private_rooms_imp   | 0
 neighbourhood_indexed_imp                          | 0.0
 property_indexed_imp                               | 0.0
 room_indexed_imp                                   | 0.0
 gps                                                | [49.28206,-123.12461]
 sky_distance                                       | 3.128298961998519E-5
-RECORD 19-------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------
 id                                                 | 359942
 host_since                                         | 2011/09/23
 host_location                                      | Vancouver, British Columbia, Canada
 host_response_time                                 | within an hour
 host_response_rate                                 | 1
 host_acceptance_rate                               | N/A
 host_is_superhost                                  | f
 host_neighbourhood                                 | Yaletown
 host_listings_count                                | 8
 host_total_listings_count                          | 8
 host_verifications                                 | ['email', 'phone', 'facebook', 'reviews']
 host_has_profile_pic                               | t
 host_identity_verified                             | f
 neighbourhood                                      | Vancouver, British Columbia, Canada
 neighbourhood_cleansed                             | Downtown
 neighbourhood_group_cleansed                       | null
 latitude                                           | 49.27825
```

```
 longitude                                          | -123.11865
 property_type                                      | Private room in apartment
 room_type                                          | Private room
 accommodates                                       | 2
 bathrooms                                          | null
 bathrooms_text                                     | 1 shared bath
 bedrooms                                           | 2
 beds                                               | 1
 amenities                                          | ['Kitchen', 'Lake access', 'Oven', 'Paid parking on prem
ises', 'Paid parking off premises', 'Gym', 'Microwave', 'Wifi', 'Dryer', 'Stove', 'Hot water', 'Dishwasher',
'Extra pillows and blankets', 'Cooking basics', 'Heating', 'Lock on bedroom door', 'Hair dryer', 'Iron', 'Co
ffee maker', 'Bed linens', 'Pocket wifi', 'Shampoo', 'Dishes and silverware', 'Hangers', 'Refrigerator', 'Es
sentials', 'Washer', 'Long term stays allowed', 'Elevator']
 price                                              | 240
 minimum_nights                                     | 30
 maximum_nights                                     | 1125
 minimum_minimum_nights                             | 30
 maximum_minimum_nights                             | 30
 minimum_maximum_nights                             | 1125
 maximum_maximum_nights                             | 1125
 minimum_nights_avg_ntm                             | 30.0
 maximum_nights_avg_ntm                             | 1125.0
 has_availability                                   | t
 availability_30                                    | 29
 availability_60                                    | 59
 availability_90                                    | 89
 availability_365                                   | 364
 instant_bookable                                   | f
 calculated_host_listings_count                     | 7
 calculated_host_listings_count_entire_homes        | 3
 calculated_host_listings_count_private_rooms       | 4
 calculated_host_listings_count_shared_rooms        | 0
 neighbourhood_reduced                              | Downtown
 property_reduced                                   | Private room in apartment
 room_reduced                                       | Private room
 neighbourhood_indexed                              | 0.0
 property_indexed                                   | 6.0
 room_indexed                                       | 1.0
 amenities_array                                    | [Kitchen, Lake access, Oven, Paid parking on premises, P
aid parking off premises, Gym, Microwave, Wifi, Dryer, Stove, Hot water, Dishwasher, Extra pillows and blank
ets, Cooking basics, Heating, Lock on bedroom door, Hair dryer, Iron, Coffee maker, Bed linens, Pocket wifi,
Shampoo, Dishes and silverware, Hangers, Refrigerator, Essentials, Washer, Long term stays allowed, Elevato
r]
 verifications_array                                | [email, phone, facebook, reviews]
 amenities_vector                                   | (20,[0,1,3,4,5,6,7,8,9,12,13,14,16,18,19],[1.0,1.0,1.0,
1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0])
 verifications_vector                               | (5,[0,1,4],[1.0,1.0,1.0])
 since_year                                         | 2011
 log_price                                          | 5.480638923341991
```

```
 host_listings_count_imp                          | 8
 host_total_listings_count_imp                    | 8
 accommodates_imp                                 | 2
 bedrooms_imp                                     | 2
 beds_imp                                         | 1
 minimum_nights_imp                               | 30
 maximum_nights_imp                               | 1125
 minimum_minimum_nights_imp                       | 30
 maximum_minimum_nights_imp                       | 30
 minimum_nights_avg_ntm_imp                       | 30.0
 calculated_host_listings_count_imp               | 7
 calculated_host_listings_count_entire_homes_imp  | 3
 calculated_host_listings_count_private_rooms_imp | 4
 neighbourhood_indexed_imp                        | 0.0
 property_indexed_imp                             | 6.0
 room_indexed_imp                                 | 1.0
 gps                                              | [49.27825,-123.11865]
 sky_distance                                     | 1.4764475619986258E-5
only showing top 20 rows
```

In [67]:

```python
plt.rcParams['figure.figsize'] = [5, 5]
pairs = list()

for p in df.select("log_price", "sky_distance").toLocalIterator():
    pairs.append( (p['log_price'], p["sky_distance"]) )
pairs = sorted(pairs, key=lambda t:t[0])
vals = [ t[1] for t in pairs if t[1] is not None ]
minv = min(vals)
maxv = max(vals)
if minv != maxv:
    vals = list(map(lambda x: x if x is None else (x-minv)/(maxv-minv), vals))
logp = [ t[0] for t in pairs if t[1] is not None ]

plt.plot(logp, vals, '.', color='gray')

plt.xlabel('log prices')
plt.ylabel("sky_distance")
plt
```

Out[67]:

```
<module 'matplotlib.pyplot' from '/vol/scratch/virtualenv/tourism/lib/python3.7/site-packages/matplotlib/pyp
lot.py'>
```

Doesn't look too peachy, let us see if the Random Forest can do something combining it with other features.

In [68]:

```
assembler = VectorAssembler(
    inputCols=imp_columns + ["amenities_vector", "sky_distance"],
    outputCol="features", handleInvalid="error")

pipeline = Pipeline(stages=stages + [assembler]).fit(dataset)
df = pipeline.transform(dataset)

rf = RandomForestRegressor(featuresCol="features", labelCol='log_price', seed=42)
model = rf.fit(df)

df_devset = pipeline.transform(devset)
predictions = model.transform(df_devset)

predictions.select("features","log_price", "prediction").show(5)

metrics = RegressionMetrics(predictions.select("prediction", "log_price").rdd\
                            .map(lambda r: (math.exp(r['prediction']), math.exp(r['log_price']))))

skyError = metrics.meanAbsoluteError
print("base:", baseError)
print("selected:", selectedError)
skyError
```

```
+--------------------+------------------+-----------------+
|            features|         log_price|       prediction|
+--------------------+------------------+-----------------+
|[49.0,49.0,5.0,2....|5.0106352940962555|5.013525997312197|
|[1.0,1.0,4.0,1.0,...| 5.293304824724492| 4.85490285833901|
|[5.0,5.0,6.0,3.0,...| 5.501258210544727| 5.25627305221444|
|[1.0,1.0,2.0,1.0,...| 5.225746673713202|4.729864780112891|
|[1.0,1.0,2.0,1.0,...| 4.852030263919617|4.571939227455491|
+--------------------+------------------+-----------------+
only showing top 5 rows

base: 56.83296858469796
selected: 57.18225268896161
```

Out[68]: 55.96113276013822

This reduced the error by almost one dollar. It seems a step in the right direction.

In [69]:

```
labels = columns + fitted_amenities.vocabulary + ["sky_distance"]

importances = [ (model.featureImportances[idx], lbl) for idx, lbl in enumerate(labels) ]
table = '<table><tr><th>Rank</th><th>Feature</th><th>Importance</th></tr>'
for idx, pair in enumerate(sorted(importances, reverse=True)):
    score, lbl = pair
    table = '{}<tr><td>{}</td><td>{}</td><td style="text-align: left">{}</td></tr>'.format(table, idx+1, lbl
table += '</table>'
display(HTML(table))
```

Rank	Feature	Importance
1	accommodates	0.22187044618185991
2	room_indexed	0.16042925153483512
3	bedrooms	0.15753887462516586
4	beds	0.09617709414703925
5	calculated_host_listings_count_private_rooms	0.08326651242819044
6	property_indexed	0.06305114161756117
7	calculated_host_listings_count_entire_homes	0.039874411329568475
8	neighbourhood_indexed	0.026023230790762075
9	maximum_minimum_nights	0.01733420753863666
10	sky_distance	0.01635666588260814

11	minimum_minimum_nights	0.014376545213620941
12	minimum_nights_avg_ntm	0.012871496835122972
13	minimum_nights	0.012789049203439106
14	calculated_host_listings_count	0.011107371382699986
15	TV	0.010880192489694209
16	host_total_listings_count	0.008993051954861515
17	host_listings_count	0.008692792952374965
18	Wifi	0.005767208516026559
19	Shampoo	0.0056203813370562395
20	maximum_nights	0.005197869716296374
21	Dryer	0.004171763488977607
22	Dedicated workspace	0.0021903817195076877
23	Fire extinguisher	0.0020779157444683212
24	Hot water	0.0020308292774049302
25	Heating	0.0018927212478844265
26	Hair dryer	0.0018409091690042355
27	Washer	0.001429516560734415
28	Kitchen	0.0012165706350505275
29	Long term stays allowed	0.0010487060359317692
30	Carbon monoxide alarm	0.0009642622782109762
31	Iron	0.0007966223561908031
32	Dishes and silverware	0.000594624278084045
33	Hangers	0.0005932677150737754
34	Essentials	0.0004091349124547263
35	Cooking basics	0.00040466766356220203
36	Refrigerator	0.00012031124003961051
37	Smoke alarm	0.0

The new feature is more important than neighbourhood. To conclude, we can do a drill-down on the instances with highest absolute

error using SHAP values. We first initialize a SHAP explainer using our model.

In [70]:
```
import shap
import numpy as np

explainer = shap.TreeExplainer(model)
```

We now identify some rows that contribute the most to the absolute error.

In [71]:
```
test_features = np.array(predictions.select("features").collect())
pred_actual = predictions.select("log_price", "prediction", "price").collect()
abs_err = list(map(lambda x:abs(x[0] - x[1]), pred_actual))
```

In [72]:
```
abs_err = list(map(lambda t:(t[1], t[0]), enumerate(abs_err)))
abs_err = sorted(abs_err, reverse=True)
abs_err[:10]
```

Out[72]:
```
[(3.3246071806661863, 862),
 (2.8965173585297146, 1312),
 (1.79262901646341, 664),
 (1.7715506281028377, 754),
 (1.6875898593949783, 695),
 (1.5658884091267078, 551),
 (1.5656351380182016, 995),
 (1.4835914017434364, 311),
 (1.4788509033722352, 1062),
 (1.4699892689246497, 649)]
```

In [73]:
```
shap.initjs()
err, idx = abs_err[0]
row_data = test_features[idx].reshape(-1)
print("Expected / prediction", pred_actual[idx], "err", err, "price", pred_actual[idx]["price"], "predicted"
shap_values = explainer.shap_values(row_data, check_additivity=False)
print(devset.take(idx+1)[-1])
shap.force_plot(explainer.expected_value[0], shap_values, row_data, feature_names=labels)
```

```
Expected / prediction Row(log_price=8.086410275323782, prediction=4.761803094657596, price=3250) err 3.32460
71806661863 price 3250 predicted 116.95661974766784
Row(id=37181228, host_since='2019/05/18', host_location='Surrey, British Columbia, Canada', host_response_ti
me='N/A', host_response_rate='N/A', host_acceptance_rate='N/A', host_is_superhost='f', host_neighbourhood='D
```

```
owntown Vancouver', host_listings_count=0, host_total_listings_count=0, host_verifications="['email', 'phon
e']", host_has_profile_pic='t', host_identity_verified='f', neighbourhood='Vancouver, British Columbia, Cana
da', neighbourhood_cleansed='Downtown', neighbourhood_group_cleansed=None, latitude=49.27673, longitude=-12
3.12571, property_type='Entire apartment', room_type='Entire home/apt', accommodates=2, bathrooms=None, bath
rooms_text='1 bath', bedrooms=1, beds=1, amenities="['Kitchen', 'Hot tub', 'Private entrance', 'Pool', 'Elev
ator', 'Gym', 'Hangers', 'Smoke alarm', 'Wifi', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Free parking on
premises', 'Long term stays allowed', 'TV', 'Hair dryer', 'Shampoo']", price=3250, minimum_nights=180, maxim
um_nights=1125, minimum_minimum_nights=180, maximum_minimum_nights=180, minimum_maximum_nights=1125, maximum
_maximum_nights=1125, minimum_nights_avg_ntm=180.0, maximum_nights_avg_ntm=1125.0, has_availability='t', ava
ilability_30=29, availability_60=59, availability_90=89, availability_365=364, instant_bookable='t', calcula
ted_host_listings_count=1, calculated_host_listings_count_entire_homes=1, calculated_host_listings_count_pri
vate_rooms=0, calculated_host_listings_count_shared_rooms=0)
```

Out[73]:

Well, that listing is new and the host made a mistake pricing it per month (or multiple months) while AirBnB is daily. The price set as $3,250 a night is meaningless. Let's look at the second entry.

In [74]:

```
shap.initjs()
err, idx = abs_err[1]
row_data = test_features[idx].reshape(-1)
print("Expected / prediction", pred_actual[idx], "err", err, "price", pred_actual[idx]["price"], "predicted"
shap_values = explainer.shap_values(row_data, check_additivity=False)
print(devset.take(idx+1)[-1])
shap.force_plot(explainer.expected_value[0], shap_values, row_data, feature_names=labels)
```

```
Expected / prediction Row(log_price=6.856461984594587, prediction=3.959944626064872, price=950) err 2.896517
3585297146 price 950 predicted 52.45442126095771
Row(id=47405647, host_since='2020/10/06', host_location='Vancouver, British Columbia, Canada', host_response
_time='N/A', host_response_rate='N/A', host_acceptance_rate='N/A', host_is_superhost='f', host_neighbourhood
=None, host_listings_count=4, host_total_listings_count=4, host_verifications="['email', 'phone', 'offline_g
overnment_id', 'selfie', 'government_id', 'identity_manual']", host_has_profile_pic='t', host_identity_verif
ied='t', neighbourhood=None, neighbourhood_cleansed='Renfrew-Collingwood', neighbourhood_group_cleansed=Non
e, latitude=49.24738, longitude=-123.02397, property_type='Private room in townhouse', room_type='Private ro
om', accommodates=1, bathrooms=None, bathrooms_text='Shared half-bath', bedrooms=None, beds=1, amenities="
['Kitchen', 'Hot tub', 'Iron', 'Private entrance', 'Lock on bedroom door', 'Indoor fireplace', 'Smoke alar
m', 'Essentials', 'Heating', 'Washer', 'Dryer', 'Long term stays allowed', 'Air conditioning', 'Hot wate
```

```
r']", price=950, minimum_nights=30, maximum_nights=150, minimum_minimum_nights=30, maximum_minimum_nights=3
0, minimum_maximum_nights=150, maximum_maximum_nights=150, minimum_nights_avg_ntm=30.0, maximum_nights_avg_n
tm=150.0, has_availability='t', availability_30=29, availability_60=59, availability_90=89, availability_365
=179, instant_bookable='t', calculated_host_listings_count=1, calculated_host_listings_count_entire_homes=0,
calculated_host_listings_count_private_rooms=1, calculated_host_listings_count_shared_rooms=0)
```

Out[74]:

Same error here, $950 seems to be the price per month because the minimum stay is 30 days. Again, a new listing.

In [75]:

```
shap.initjs()
err, idx = abs_err[2]
row_data = test_features[idx].reshape(-1)
print("Expected / prediction", pred_actual[idx], "err", err, "price", pred_actual[idx]["price"], "predicted"
shap_values = explainer.shap_values(row_data, check_additivity=False)
print(devset.take(idx+1)[-1])
shap.force_plot(explainer.expected_value[0], shap_values, row_data, feature_names=labels)
```

```
Expected / prediction Row(log_price=6.902742737158593, prediction=5.110113720695183, price=995) err 1.792629
01646341 price 995 predicted 165.68919609295506
Row(id=30352025, host_since='2016/07/26', host_location='West Vancouver, British Columbia, Canada', host_res
ponse_time='N/A', host_response_rate='N/A', host_acceptance_rate='N/A', host_is_superhost='f', host_neighbou
rhood='Fairview', host_listings_count=1, host_total_listings_count=1, host_verifications="['phone', 'faceboo
k', 'jumio', 'offline_government_id', 'selfie', 'government_id', 'identity_manual']", host_has_profile_pic
='t', host_identity_verified='t', neighbourhood=None, neighbourhood_cleansed='Fairview', neighbourhood_group
_cleansed=None, latitude=49.26657, longitude=-123.13922, property_type='Entire condominium', room_type='Enti
re home/apt', accommodates=6, bathrooms=None, bathrooms_text='3 baths', bedrooms=2, beds=3, amenities="['Hot
tub', 'Kitchen', 'First aid kit', 'Hangers', 'Wifi', 'Breakfast', 'Essentials', 'Heating', 'Washer', 'Drye
r', 'Long term stays allowed', 'Smoke alarm', 'TV', 'Air conditioning', 'Elevator', 'Shampoo']", price=995,
minimum_nights=3, maximum_nights=1125, minimum_minimum_nights=3, maximum_minimum_nights=3, minimum_maximum_n
ights=1125, maximum_maximum_nights=1125, minimum_nights_avg_ntm=3.0, maximum_nights_avg_ntm=1125.0, has_avai
lability='t', availability_30=0, availability_60=0, availability_90=0, availability_365=0, instant_bookable
='t', calculated_host_listings_count=1, calculated_host_listings_count_entire_homes=1, calculated_host_listi
ngs_count_private_rooms=0, calculated_host_listings_count_shared_rooms=0)
```

Out[75]:

This is an entire house that accommodates 6 people. Going back to the source data, it is described as a "luxury condo with an outdoor hot tub". We can see the neighborhoud and distance to skytrain went against the pricing but the description says it is very close to Granville Island and its shops and restaurants. Adding more POIs (points-of-interest) should help here. And of course, NLP on the descriptions.

Where to go from here? An obvious expansion is to download POIs ("points-of-interest") from OpenStreetMap and add a feature of the number of POIs within 1Km radius. Another interesting feature is the distance to the closest SkyTrain station of the Canada Line that connects the city directly to the airport. We can surmise that a direct airport connection might command higher prices.

Finally, it might be possible to train two different models, one for downtown, another for the rest of the city and use them to predict a price, then leave it to the final model to use that information for the final prediction.