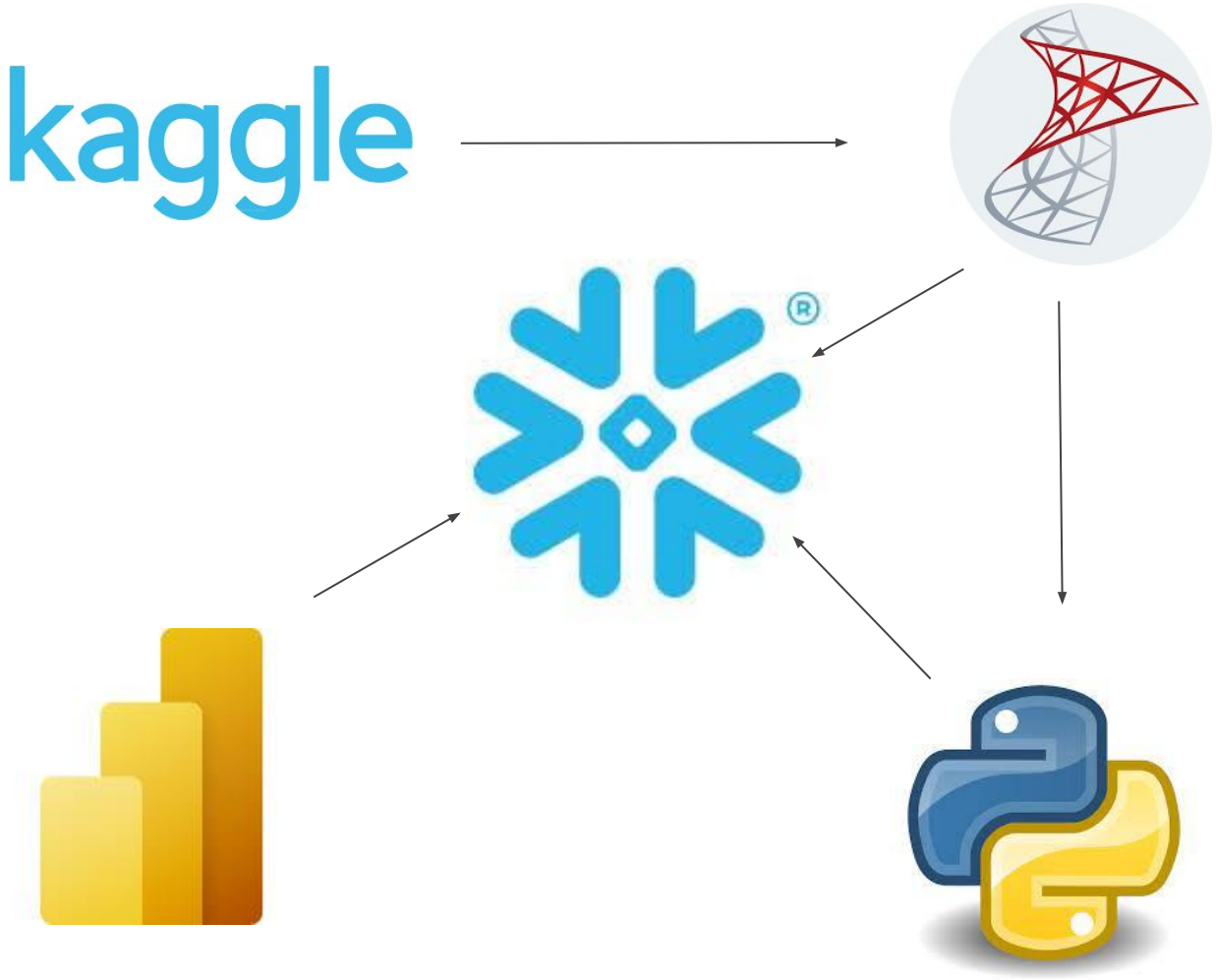
New York November 2021

- Bawkat Mazengya
- Gal Hadad
- Ziva Einav Grushka

Process

- Download data
- Wearhouse
- Analyze
- Dashboards
- Recommendations



EDA

```sql
select COLUMN_NAME from INFORMATION_SCHEMA.COLUMNS
where TABLE_NAME= 'JUNDATYSTPROJECT'
```

Results | Messages

COLUMN_NAME
id
listing_url
scrape_id
last_scraped
name
description
neighborhood_overview
picture_url
host_id
host_url
host_name
host_since
host_location
host_about
host_response_time
host_response_rate
host_acceptance_rate
host_is_superhost
host_thumbnail_url
host_picture_url
host_neighbourhood
host_total_listings_count
host_verifications

	COLUMN_NAME
36	beds
37	amenities
38	price
39	minimum_nights
40	maximum_nights
41	minimum_minimum_nights
42	maximum_minimum_nights
43	minimum_maximum_nights
44	maximum_maximum_nights
45	minimum_nights_avg_ntm
46	maximum_nights_avg_ntm
47	has_availability
48	availability_30
49	availability_60
50	availability_90
51	availability_365
52	calendar_last_scraped
53	number_of_reviews
54	number_of_reviews_ltm
55	number_of_reviews_l30d
56	first_review
57	last_review
58	review_scores_rating
59	review_scores_accuracy
60	review_scores_cleanliness
61	review_scores_checkin
62	review_scores_communication
63	review_scores_location
64	review_scores_value
65	license
66	instant_bookable
67	calculated_host_listings_count
68	calculated_host_listings_count_entire_homes
69	calculated_host_listings_count_private_rooms
70	calculated_host_listings_count_shared_rooms
71	reviews_per_month

```sql
--- list records
select
count (id)
from JUNDATYSTPROJECT
-- 37,713

--hosts
select
count (distinct host_id)
from JUNDATYSTPROJECT
--25,912 hosts

--num lists not available
select count(id)num_id
from JUNDATYSTPROJECT
where availability_365 = 0
```

```python
try:
  import snowflake.connector
except ImportError:
  # install the required packages and dependencies
  !pip install --upgrade snowflake-connector-python
  !pip install --upgrade --force-reinstall pandas
  !pip install --upgrade --force-reinstall pyarrow
  !pip install --upgrade --force-reinstall snowflake-connector-python
  !pip install --upgrade --force-reinstall sqlalchemy
  !pip install --upgrade --force-reinstall snowflake-sqlalchemy

  print('Stopping RUNTIME! Please run again.')

  # restart runtime (required for using the installed packages and dependencies)
  # note: will crash the session so the next step should be executed manually after session restart
  import os
  os.kill(os.getpid(), 9)
```

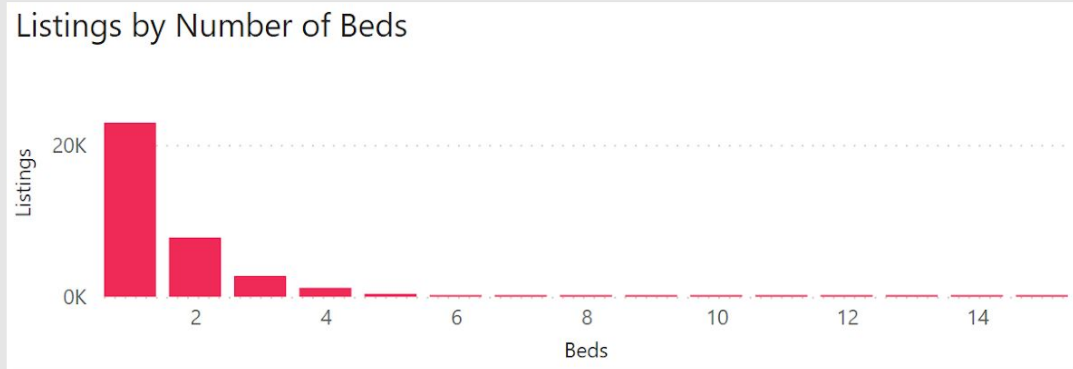
```
Users\User\anaconda3\lib\site-packages\snowflake\connector\options.py:96: UserWarning: You have an incomp
arrow' installed (7.0.0), please install a version that adheres to: 'pyarrow<6.1.0,>=6.0.0; extra == "pa
arn_incompatible_dep(
```

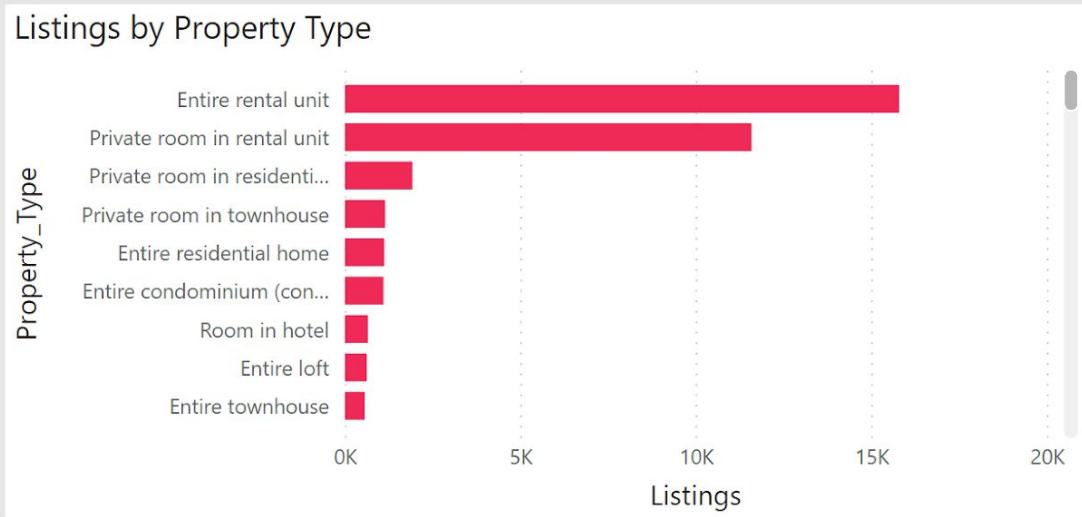
```python
import math
import datetime
import numpy as np
import pandas as pd
import snowflake.connector
from snowflake.sqlalchemy import URL
from sqlalchemy import create_engine
```

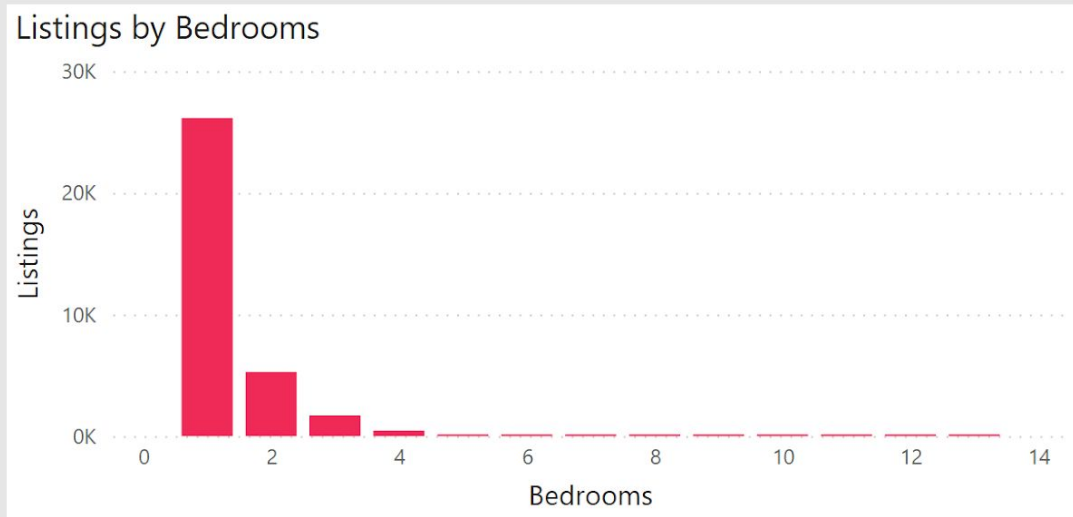
```python
engine = create_engine(URL(
    account = 'hz32324.uae-north.azure',
    user = 'BAWKAT',
    password = 'Bawkat123456',
    database = 'TST_DB',
    schema = 'PUBLIC',
    warehouse = 'COMPUTE_WH'#,
    #role = customer_name
```

37.713K	79	14.732K	7
Listings	Property Types	Listings Fully Booked	Average of BEDROOMS

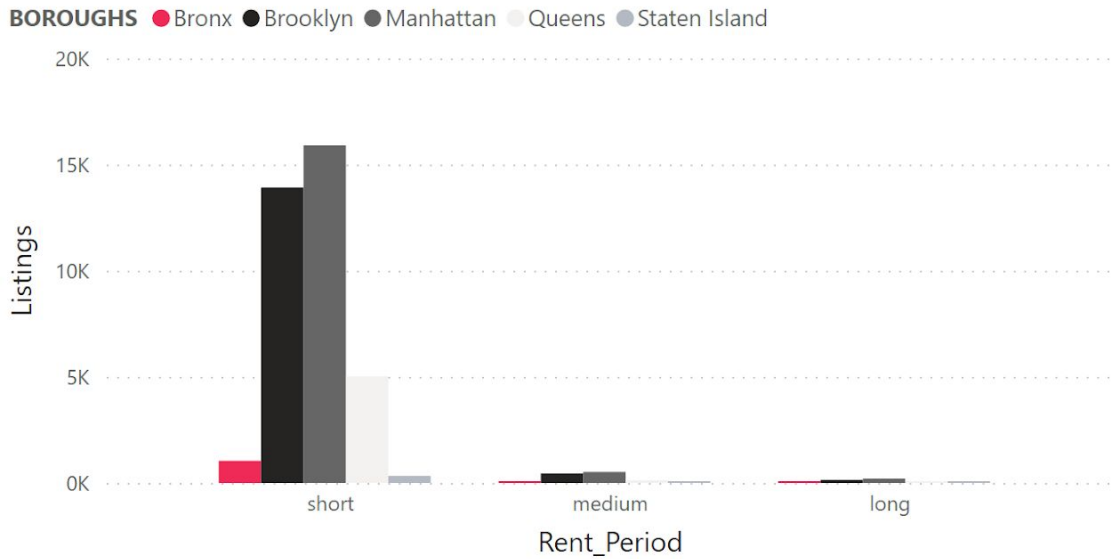
BOROUGH	NUM_LISTS	PCT_RUNNING_TOTAL	SUM_RUNNING_TOTAL	LIST_PERCENT
Manhattan	16,625.00	44.08	16,625.00	44.08
Brooklyn	14,512.00	82.56	31,137.00	38.48
Queens	5,178.00	96.29	36,315.00	13.73
Bronx	1,059.00	99.10	37,374.00	2.81
Staten Island	339.00	100.00	37,713.00	0.90
Total	**37,713.00**			**100.00**







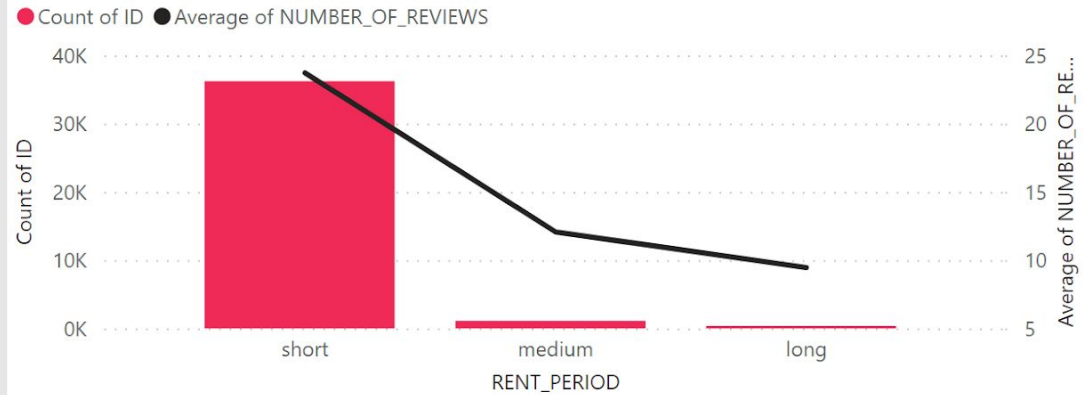
Listings by Rent_Period and BOROUGHS



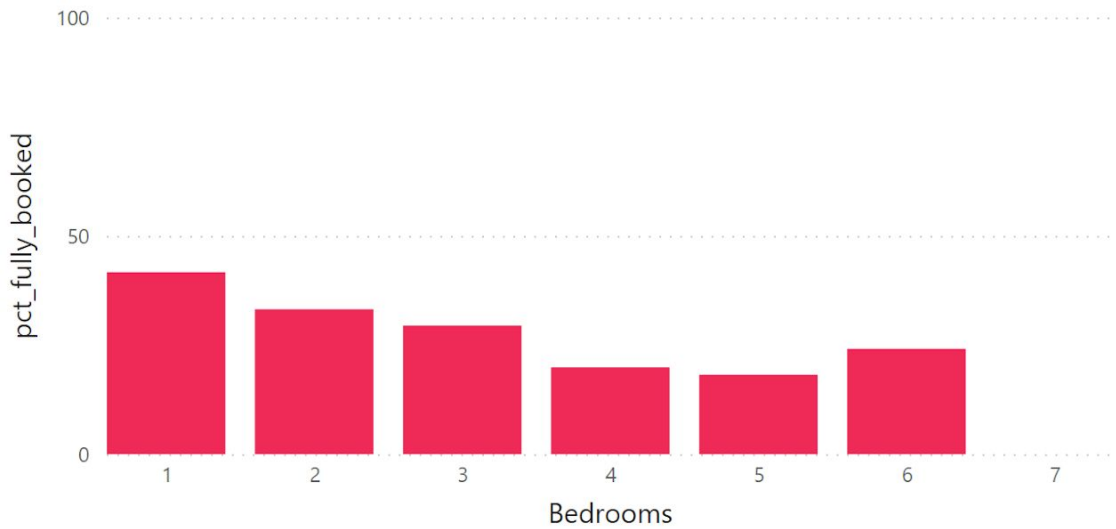
14.732K
Listings Fully Booked

7
AVG BEDROOMS

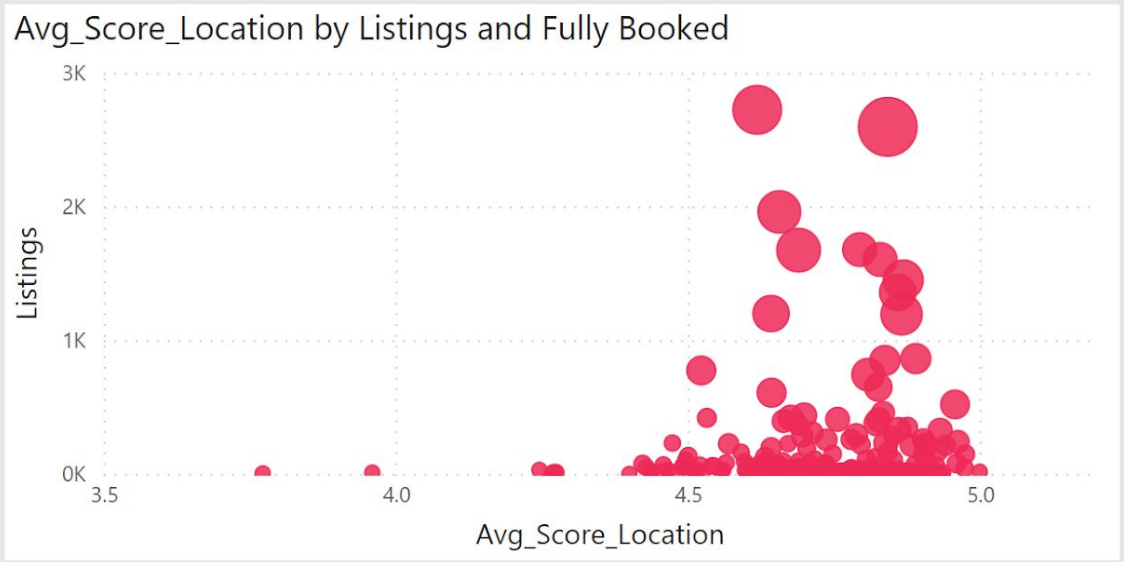
Num Lists by RENT_PERIOD



pct_fully_booked by Bedrooms



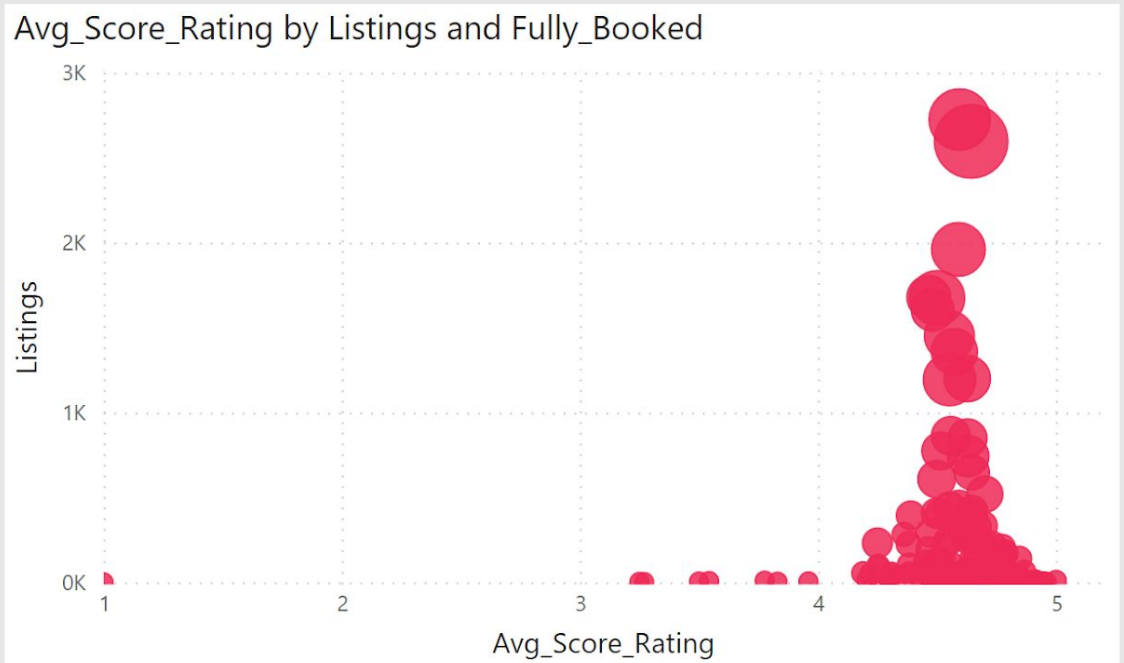
BOROUGHS	PROPERTY_TYPE	FULLY_BOOKED	NUM_LISTINGS
Manhattan	Entire rental unit	3,650.00	8,696.00
Brooklyn	Private room in rental unit	2,787.00	5,152.00
Brooklyn	Entire rental unit	2,522.00	5,520.00
Manhattan	Private room in rental unit	2,300.00	4,718.00
Queens	Private room in rental unit	479.00	1,372.00
Queens	Entire rental unit	398.00	1,211.00
Queens	Private room in residential home	204.00	905.00
Queens	Private room in townhouse	190.00	558.00
Brooklyn	Private room in residential home	185.00	699.00
Manhattan	Entire condominium (condo)	167.00	602.00
Brooklyn	Entire loft	142.00	349.00
Brooklyn	Entire residential home	140.00	530.00
Brooklyn	Entire condominium (condo)	138.00	394.00
Total		**14,732.00**	**37,713.00**

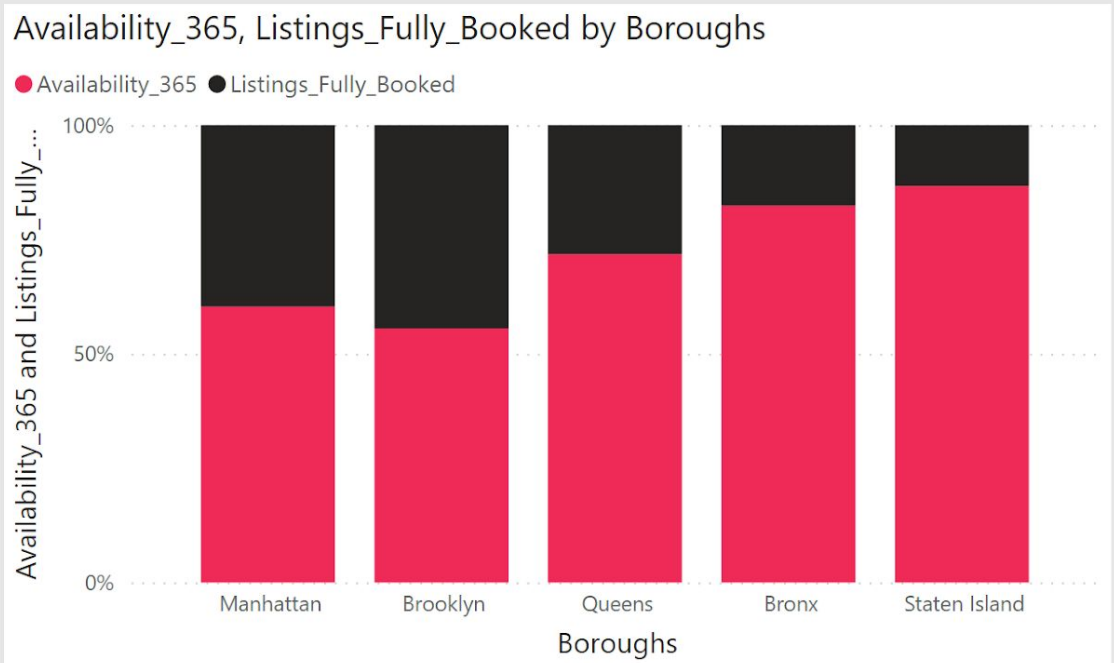


70.49
Avg Listings Fully Booked

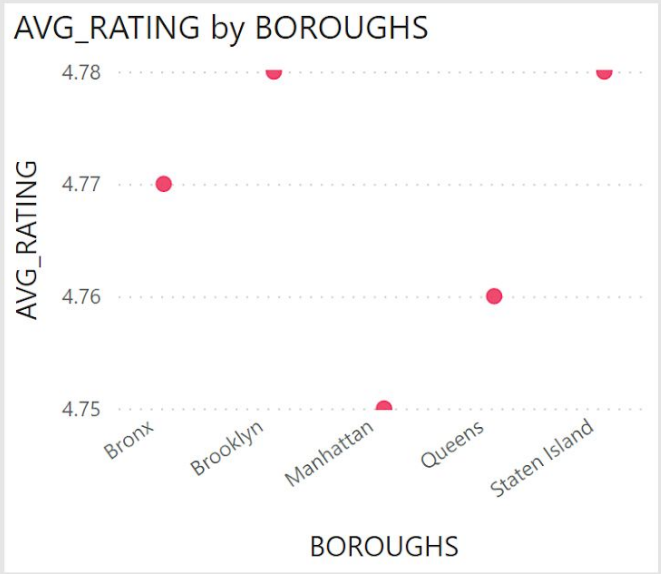
4.61
Avg Score Rating

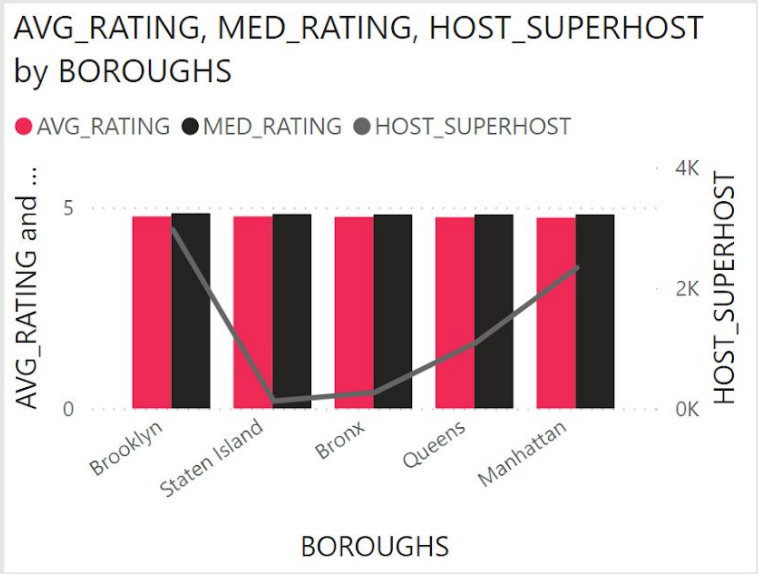
4.73
Avg Score Location





Listings with the highest total rating (between 4 and 5), broken-down by boroughs

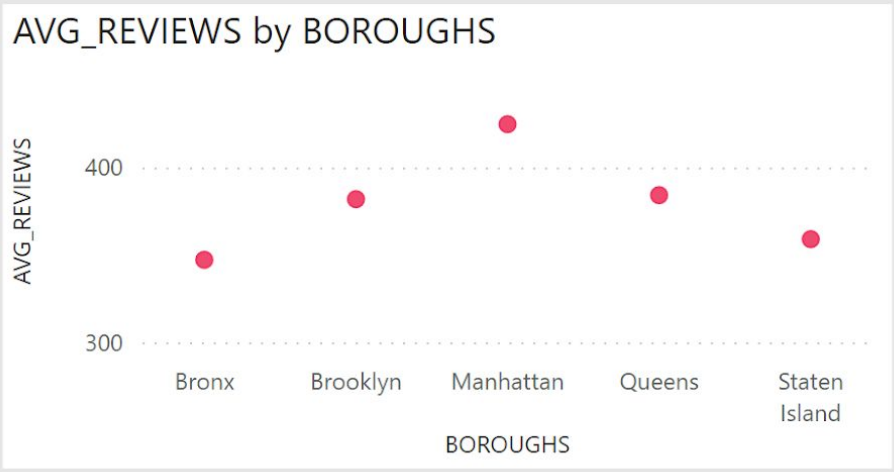


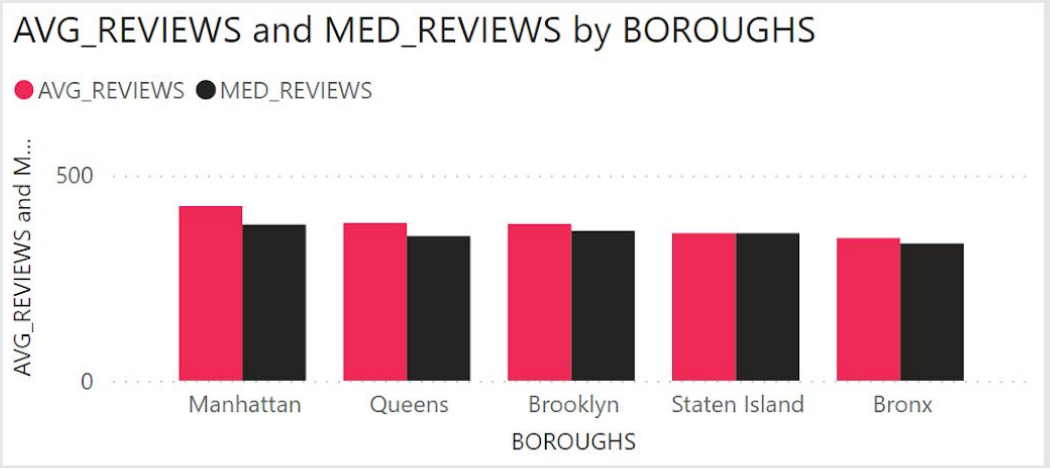


BOROUGHS	NUM_LISTINGS
Bronx	770.00
Brooklyn	10,835.00
Manhattan	11,114.00
Queens	3,726.00
Staten Island	279.00
Total	**26,724.00**

Listings without a rating score

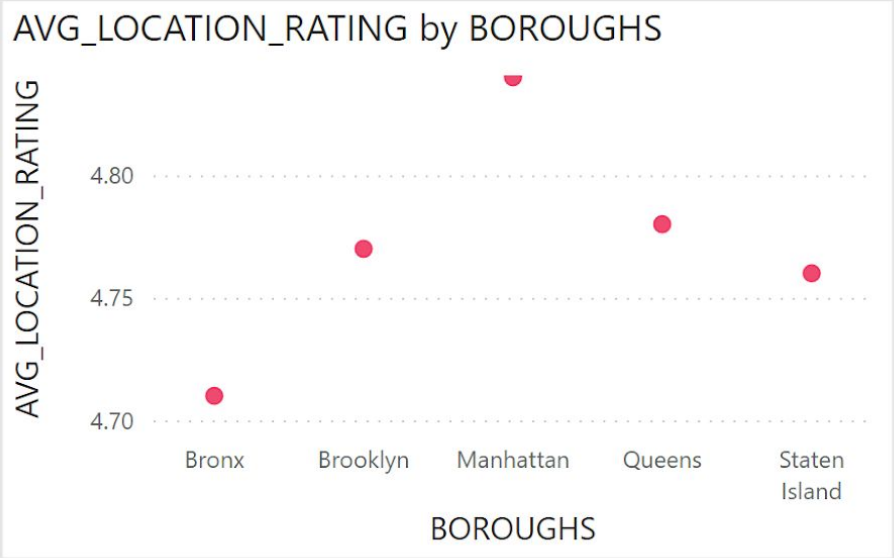
Listings with the most number of reviews (300 and above), broken-down by boroughs

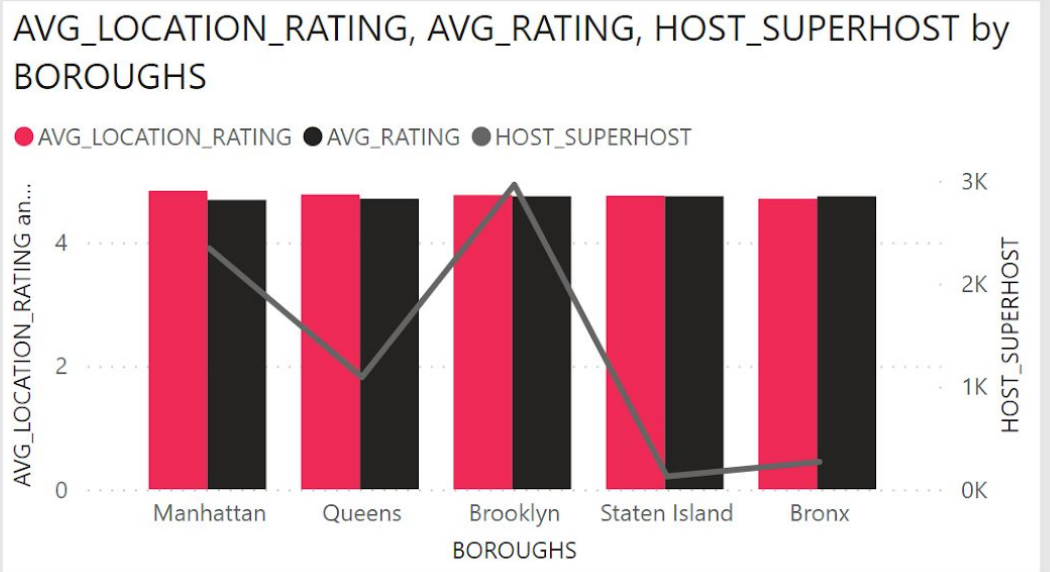




BOROUGHS	NUM_LISTINGS
Bronx	7.00
Brooklyn	65.00
Manhattan	67.00
Queens	41.00
Staten Island	1.00
Total	**181.00**

Listings with the highest location rating (between 4 and 5), broken-down by





BOROUGHS	NUM_LISTINGS
Staten Island	284.00
Queens	3,796.00
Manhattan	11,440.00
Brooklyn	10,895.00
Bronx	763.00
Total	**27,178.00**

Distribution of properties by neighborhoods

Bronx

1,059.00
NUM_LISTS

14.53
Average of AVAILABILITY_30

Brooklyn

14,512.00
NUM_LISTS

119.52
Average of AVAILABILITY_30

Manhattan

16,625.00
NUM_LISTS

220.06
Average of AVAILABILITY_30

Queens

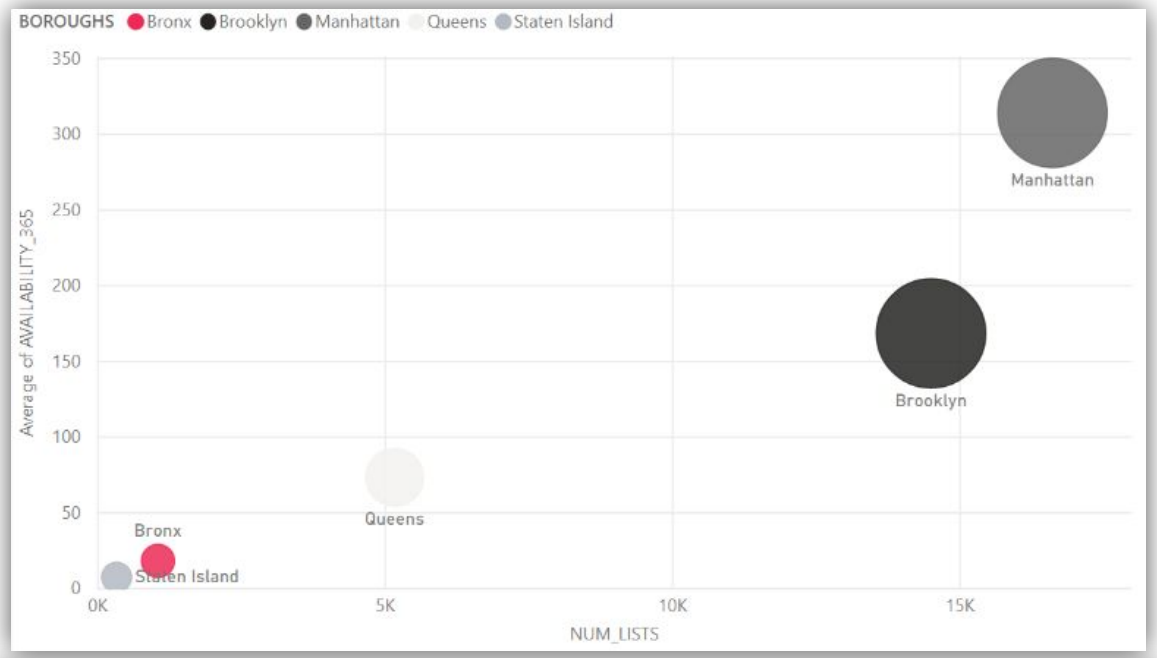
5,178.00
NUM_LISTS

58.22
Average of AVAILABILITY_30

Staten Island

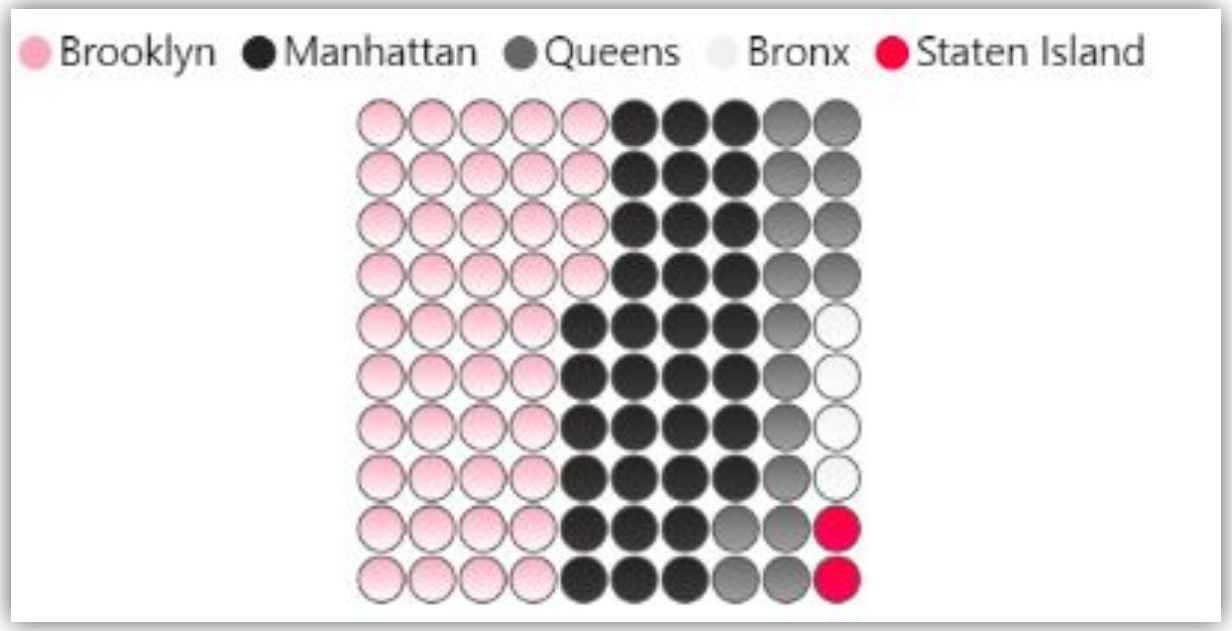
339.00
NUM_LISTS

5.21
Average of AVAILABILITY_30

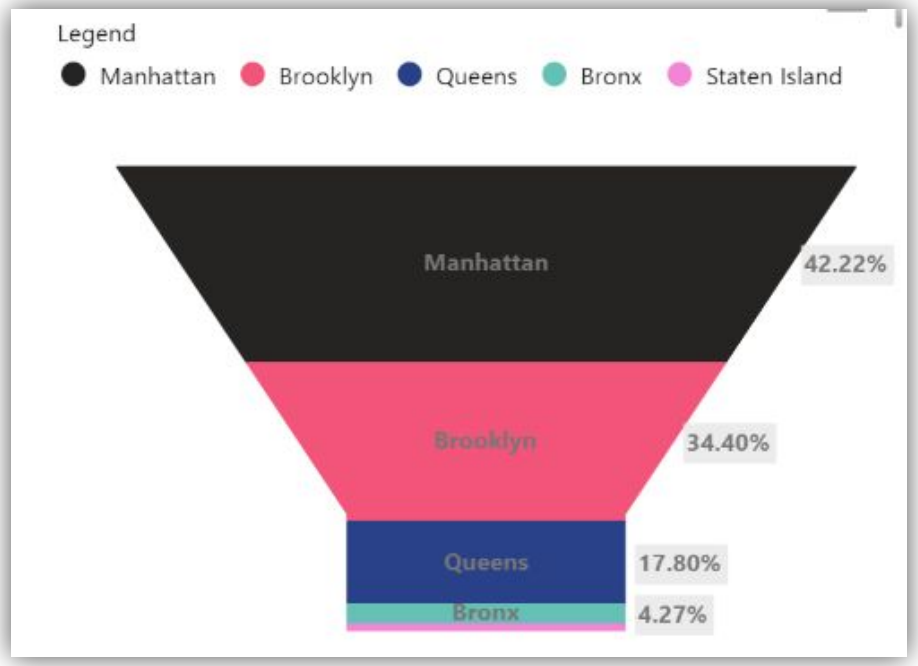


Neighborhoods	Number of listings	Average price for listing
5	37.71K	125.40

Superhost classification by neighborhood



Monthly availability by neighborhood



Thank You!