

index

sample text

```
library(tidyverse)
```

```
## Warning: package 'ggplot2' was built under R version 4.5.2  
## Warning: package 'readr' was built under R version 4.5.2  
## Warning: package 'purrr' was built under R version 4.5.2  
## Warning: package 'stringr' was built under R version 4.5.2  
  
## -- Attaching core tidyverse packages -----  
## * v dplyr     1.1.4     v readr     2.1.6  
## * vforcats   1.0.1     v stringr   1.6.0  
## * v ggplot2   4.0.1     v tibble    3.3.0  
## * v lubridate 1.9.4     v tidyR    1.3.1  
## * v purrr    1.2.0  
## -- Conflicts -----  
## * x dplyr::filter() masks stats::filter()  
## * x dplyr::lag()   masks stats::lag()  
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(patchwork)
```

```
## Warning: package 'patchwork' was built under R version 4.5.2
```

```
library(naniar)  
library(skimr)
```

```
## Warning: package 'skimr' was built under R version 4.5.2  
  
##  
## Attaching package: 'skimr'  
##  
## The following object is masked from 'package:naniar':  
##  
##     n_complete
```

```
library(visdat)  
library(mice)
```

```

## 
## Attaching package: 'mice'
##
## The following object is masked from 'package:stats':
## 
##     filter
##
## The following objects are masked from 'package:base':
## 
##     cbind, rbind

library(rapportools)

## Warning: package 'rapportools' was built under R version 4.5.2

## 
## Attaching package: 'rapportools'
##
## The following object is masked from 'package:dplyr':
## 
##     n
##
## The following objects are masked from 'package:stats':
## 
##     IQR, median, sd, var
##
## The following objects are masked from 'package:base':
## 
##     max, mean, min, range, sum

library(knitr)
source("data_pull.R")

## Warning: package 'rcompendium' was built under R version 4.5.2

head(df)

##           id          listing_url    scrape_id last_scraped      source
## 1 40824219 https://www.airbnb.com/rooms/40824219 2.0251e+13 2025-10-02    city scrape
## 2 40833186 https://www.airbnb.com/rooms/40833186 2.0251e+13 2025-10-02 previous scrape
## 3 40837137 https://www.airbnb.com/rooms/40837137 2.0251e+13 2025-10-02 previous scrape
## 4 40838018 https://www.airbnb.com/rooms/40838018 2.0251e+13 2025-10-02 previous scrape
## 5 40839416 https://www.airbnb.com/rooms/40839416 2.0251e+13 2025-10-02    city scrape
## 6 40843980 https://www.airbnb.com/rooms/40843980 2.0251e+13 2025-10-01    city scrape
## 
##           name
## 1 Room close to Manhattan for FEMALE guests
## 2 Soho LES East village private room downtown
## 3 Sunset Park - Quiet and close to subway!
## 4 Cozy One Bedroom in Clinton Hill
## 5 XL dojo shared green yogi palace apt
## 6 Cozy 2 Bedroom Spacious Apartment near Manhattan

```

```

## 1
## 2
## 3
## 4
## 5
## 6 This 2 bed. furnished apt on the 2nd fl. in Ozone Park, Queens, provides guests with the experience
## 1
## 2
## 3
## 4
## 5
## 6 The borough of Queens offers plenty of outdoor recreation space and easy access to Manhattan. For p
## 1 https://a0.muscache.com/pictures/hosting/Hosting-40824219/original/70
## 2 https://a0.muscache.com/pictures/
## 3 https://a0.muscache.com/pictures/
## 4 https://a0.muscache.com/pictures/
## 5 https://a0.muscache.com/pictures/hosting/Hosting-U3RheVN1cHBseUxpc3Rpcmc6NDA4Mzk0MTY%3D/original/20
## 6 https://a0.muscache.com/pictures/
## host_id host_url host_name host_since host_location
## 1 317540555 https://www.airbnb.com/users/show/317540555 Sevil 2019-12-14
## 2 68718914 https://www.airbnb.com/users/show/68718914 Patricio 2016-04-24
## 3 317770098 https://www.airbnb.com/users/show/317770098 Xuan 2019-12-15 New York, United States
## 4 17211451 https://www.airbnb.com/users/show/17211451 Jamila 2014-06-24 New York, NY
## 5 4765305 https://www.airbnb.com/users/show/4765305 Haffro 2013-01-19 New York, NY
## 6 295370107 https://www.airbnb.com/users/show/295370107 Dr. Tania 2019-09-16 New York, United States
## 1
## 2
## 3
## 4 Hello, I'm a global communications p
## 5
## 6 Hi, I'm Dr. Tania Sheikh. I'm a physician at a local NYC hospital and I'm a mother of three kids. I
## host_response_time host_response_rate host_acceptance_rate host_is_superhost
## 1 within a few hours 100% 54% t
## 2 N/A N/A N/A f
## 3 N/A N/A N/A f
## 4 N/A N/A N/A f
## 5 within a day 90% 69% t
## 6 within a few hours 75% 100% t
## host_thumb
## 1 https://a0.muscache.com/im/pictures/user/771f9d60-559e-43f8-90b6-acaf31f81875.jpg?aki_policy=profile
## 2 https://a0.muscache.com/im/pictures/user/0b4f989e-4e0a-45ee-8a7f-6b3a8930e7ce.jpg?aki_policy=profile
## 3 https://a0.muscache.com/im/pictures/user/2c4b164f-513b-4c06-8500-10c8fe079528.jpg?aki_policy=profile
## 4 https://a0.muscache.com/im/users/17211451/profile_pic/1403641887/original.jpg?aki_policy=profile
## 5 https://a0.muscache.com/im/pictures/user/9565915b-93d8-4c72-96e4-d2538a8ff72c.jpg?aki_policy=profile
## 6 https://a0.muscache.com/im/pictures/user/5d78b12e-948e-4fc1-9663-e1d823dfef92.jpg?aki_policy=profile
## host_
## 1 https://a0.muscache.com/im/pictures/user/771f9d60-559e-43f8-90b6-acaf31f81875.jpg?aki_policy=profile
## 2 https://a0.muscache.com/im/pictures/user/0b4f989e-4e0a-45ee-8a7f-6b3a8930e7ce.jpg?aki_policy=profile
## 3 https://a0.muscache.com/im/pictures/user/2c4b164f-513b-4c06-8500-10c8fe079528.jpg?aki_policy=profile
## 4 https://a0.muscache.com/im/users/17211451/profile_pic/1403641887/original.jpg?aki_policy=profile

```

```

## 5 https://a0.muscache.com/im/pictures/user/9565915b-93d8-4c72-96e4-d2538a8ff72c.jpg?aki_policy=profile
## 6 https://a0.muscache.com/im/pictures/user/5d78b12e-948e-4fc1-9663-e1d823dfef92.jpg?aki_policy=profile
##   host_listings_count host_total_listings_count host_verifications host_has_profile_pic host_identity
## 1                   3                           3 ['email', 'phone']          t
## 2                   1                           1 ['phone']                  t
## 3                   1                           1 ['email', 'phone']          t
## 4                   1                           3 ['email', 'phone']          t
## 5                   8                          16 ['email', 'phone']          t
## 6                   2                           2 ['email', 'phone']          t
##               neighbourhood neighbourhood_cleansed neighbourhood_group_cleansed latitude longitude
## 1 Neighborhood highlights           Sunnyside                    Queens 40.74698 -73.91763
## 2                               Nolita                    Manhattan 40.72314 -73.99323 Pri
## 3 Neighborhood highlights           Sunset Park                 Brooklyn 40.64607 -74.00552
## 4 Neighborhood highlights           Clinton Hill                Brooklyn 40.68370 -73.96115
## 5 Neighborhood highlights           East Village                Manhattan 40.72147 -73.98270
## 6 Neighborhood highlights           Ozone Park                  Queens 40.68230 -73.84545
##   room_type accommodates bathrooms bathrooms_text bedrooms beds
## 1 Private room        1       1 shared bath      1   1
## 2 Private room        2       NA 1 shared bath     1   NA
## 3 Private room        1       NA 1 shared bath    NA   NA
## 4 Entire home/apt     1       NA             1 bath     1   NA
## 5 Private room        1       2 2 shared baths    1   1
## 6 Entire home/apt     6       1             1 bath      2   3
##
## 1                                         ["Extra pillows"
## 2
## 3
## 4
## 5 ["Dedicated workspace", "Microwave", "Elevator", "Dryer", "Iron", "Private living room", "Washer",
## 6
##   price minimum_nights maximum_nights minimum_minimum_nights maximum_maximum_nights minimum_maximum
## 1 $66.00            30          365          30          30
## 2                  30          1125         30          30
## 3                  30          180          30          30
## 4                  90          90          90          90
## 5 $76.00            30          900          30          30
## 6 $97.00            30          365          30          30
##   minimum_nights_avg_ntm maximum_nights_avg_ntm calendar_updated has_availability availability_30 ava
## 1                  30          365          NA          t          16
## 2                  30          1125         NA          t          0
## 3                  30          180          NA          t          0
## 4                  90          1125         NA          t          0
## 5                  30          900          NA          t          26
## 6                  30          1125         NA          t          29
##   availability_365 calendar_last_scraped number_of_reviews number_of_reviews_ltm number_of_reviews_l
## 1                  77 2025-10-02          16          2
## 2                  0 2025-10-02          0          0
## 3                  0 2025-10-02          1          0
## 4                  0 2025-10-02          1          0
## 5                 168 2025-10-02          20          0
## 6                 364 2025-10-01          93          1
##   number_of_reviews_ly estimated_occupancy_1365d estimated_revenue_1365d first_review last_review re
## 1                  0          120          7920 2019-12-25 2025-08-06
## 2                  0              0          NA

```

```

## 3          0          0          NA 2020-01-19 2020-01-19
## 4          0          0          NA 2020-03-01 2020-03-01
## 5          1          0          0  2021-09-10 2024-09-01
## 6          0         60          5820 2020-07-06 2025-01-25
##   review_scores_accuracy review_scores_cleanliness review_scores_checkin review_scores_communication
## 1              4.81            4.88            4.81           4.88
## 2              NA             NA             NA            NA
## 3              5.00            4.00            5.00           5.00
## 4              5.00            5.00            5.00           5.00
## 5              4.80            4.85            5.00           5.00
## 6              4.26            4.09            4.45           4.06
##   review_scores_value license instant_bookable calculated_host_listings_count calculated_host_listing
## 1              4.69            f               3
## 2              NA             t               1
## 3              5.00            f               1
## 4              5.00            t               1
## 5              4.95            f               8
## 6              4.00            f               2
##   calculated_host_listings_count_private_rooms calculated_host_listings_count_shared_rooms reviews_p
## 1                           3                         0
## 2                           1                         0
## 3                           1                         0
## 4                           0                         0
## 5                           8                         0
## 6                           0                         0

```