Exercises

1. Exercise 1

You work as a data analyst in Airbnb.

The management wants to better understand the rental trends in New York City in 2019.

Explore the listings dataset to answer the following questions.

Download the AB\_NYC\_2019.csv file, to your computer, and read it into the notebook.

It is recommended to give the dataframe an informative name.

Before you answer the questions, get yourself acquainted with the dataframe (you can use methods like head, shape, columns)

* 1. Sort the data and display the three most expensive listings.
  2. How many listings are offered in each part of the city? ("neighbourhood\_group")
  3. What is the average price of listings in Manhattan? (use the .mean() method)
  4. Since the average might be influenced by extreme values, use the describe() method to display supplemental statistics for the prices in Manhattan.

1. Exercise 2

Reviews are an important part of Airbnb mechanism, as they increase customer trust, and encourage hosts to provide better service.

You've noticed that some properties have no reviews or that the reviews are out of date.

Create a new dataframe which contains only the problematic listings - where the last review date is missing, or before 2019.

Sort the new dataframe by id, and then by host\_id, so it would be easier for other departments to further investigate this issue.

1. Exercise 3

"Professional Hosts" are hosts that have 10 listings or more.

The company would like to track them, so they could offer them extra services.

Create a dataframe of "Professional Hosts" in Manhattan and Brooklyn.

The dataframe should include the host\_id, host name and number of listings, sorted by the number of listings (from the highest to the lowest)

finally, use the .drop\_duplicates() method to include only one row per host in the new dataframe.

ראש הטופס