Project description

## DATA SOURCES

Inside Airbnb Amsterdam Snapshot (December 6th, 2018)

This dataset provides a comprehensive snapshot of Airbnb listings in Amsterdam, as of December 6th, 2018. The data is sourced from insideairbnb.com, which contains information on listings, reviews, and additional details.

## DATA COLLECTION

The data is collected through web scraping of Airbnb listings in Amsterdam. The snapshot is taken on December 6th, 2018, and it covers information on over 20,000 listings. Additionally, a shapefile of the neighborhoods in Amsterdam is included.

## DATA CONTENTS

The dataset is composed of the following files:

a. Calendar.csv: Contains availability and pricing information for each listing 365 days ahead.

b. Listings.csv: Provides basic information on all the Airbnb listings in Amsterdam on December 6th, 2018.

c. Reviews\_details.csv: Includes detailed information on each listing, such as descriptions, host details, and property characteristics.

d. Neighbourhoods.geojson: Contains the geographical shapefile of Amsterdam neighborhoods.

e. Reviews.csv: Contains the listing ID and the date of the reviews.

f. Listings\_details.csv: Contains detailed information on each listing, similar to reviews\_details.csv.

g. Neighbourhoods.csv: Includes the neighborhood group and neighborhood names in Amsterdam.

## DATA LIMITATIONS

The dataset is a snapshot taken on December 6th, 2018, which means that it may not accurately represent the current situation of Airbnb listings in Amsterdam. Additionally, the data is collected through web scraping, so some information may be incomplete or missing.

## DATA RELEVANCY

Project Objective: Analyze the Airbnb market in Amsterdam, identify trends, and explore factors affecting pricing and demand.

Relevant Hypotheses:

a. The location of a listing (neighborhood) has a significant impact on its price and demand.

b. The type of accommodation (room type, property type) influences pricing and customer preferences.

c. Host-related factors (e.g., response time, experience, and superhost status) may affect listing performance.

d. Analyzing reviews can provide insights into factors that influence customer satisfaction and contribute to a listing's success.

## HYPOTHETHIS

a. The location of a listing (neighborhood) has a significant impact on its price and demand.

* **Which neighborhoods have the highest and lowest average prices for listings?**
* Is there a relationship between neighborhood location and the number of reviews a listing receives?
* Do listings in central or popular areas have higher prices and demand compared to those in less popular areas?

b. The type of accommodation (room type, property type) influences pricing and customer preferences.

* **How do prices vary between different room types (e.g., private room, entire home, shared room)?**
* Are certain property types (e.g., apartments, houses, lofts) more popular among customers based on the number of reviews?
* Do specific room and property types have a higher price per night on average?

c. Host-related factors (e.g., response time, experience, and superhost status) may affect listing performance.

* Is there a relationship between host response time and the number of reviews a listing receives?
* **Do listings with superhosts have higher prices and demand compared to listings with non-superhosts?**
* How does host experience (e.g., the number of listings managed by the host) affect the price and demand for their listings?

d. Analyzing reviews can provide insights into factors that influence customer satisfaction and contribute to a listing's success.

* **Is there a correlation between review scores (e.g., rating, accuracy, value) and the price of a listing?**
* Do listings with higher review scores receive more reviews on average?
* Can certain review aspects, such as accuracy or value, be linked to higher customer satisfaction and, in turn, a more successful listing?

e. Seasonal and temporal factors influence listing prices and demand.

• Do prices fluctuate depending on the time of the year (e.g., peak seasons, holidays)?

• Are there specific months or days of the week when the demand for listings is higher based on the number of bookings or availability?

• Is there a relationship between the time of the year and the number of reviews a listing receives?