

A study of Amsterdam Airbnb rentals

Business Problem

Amsterdam is the capital and most populous city of Netherlands. Colloquially referred to as the "Venice of the North", attributed by the large number of canals which form a UNESCO World Heritage Site, it is also home to numerous tourist attractions. As a result, it is one of the most visited places in Europe, receiving more than 4.63 million international visitors annually. This results in a demand for places for a short term stay.

Airbnb, Inc. is an online marketplace for arranging or offering lodging, primarily homestays, or tourism experiences. The company does not own any of the real estate listings, nor does it host events; it acts as a broker, receiving commissions from each booking.

We will examine available data to answer these questions:

- Which part of the city has the highest price of rentals?
- Which areas have the highest number of rentals?
- What are the factors that drive the listing price?
- What are the factors that drive the popularity of a rental?
- When are more Airbnbs likely to be available?

The answers should be beneficial to anyone looking to host a short term vacation rental in Amsterdam by telling them which are the most desirable places to put up listings and what type of listings are in demand. Any traveler traveling to the city should also be able to make an informed decision about where to stay based on their budget and interests.

Data

Data Collection

The data for this project has been sourced from mainly 2 sources:

- Foursquare API
 1. Venue Categories - Hierarchical list of categories applied to venues. This data has been used to group venues in each neighborhood
 2. Venue Search - List of venues near the current location. All venues have been rolled-up to these categories for analysis - Shop & Service, Food, Outdoors & Recreation, Travel & Transport, Nightlife Spot, Arts & Entertainment, Professional & Other Places
- Publicly available Airbnb listing data that includes:
 1. Information and metrics for listings in Amsterdam
 2. Detailed Calendar Data for listings in Amsterdam
 3. GeoJSON file of neighbourhoods of Amsterdam

Here are the relevant details of the listings dataset:

Name	Description
listing_id	Id of the airbnb listing
neighbourhood	Neighborhood of the listing
latitude/longitude	Coordinates of the listing
room_type	The type of listing. Private room, Shared room etc.
price	Price per night
minimum_nights	Minimum number of nights required to book
number_of_reviews	Number of reviews available for the listing
last_review	Last review date
reviews_per_month	Average number for reviews posted per month
availability_365	Number of days available in the next 365 days

Listing calendar dataset contains:

Name	Description
listing_id	Id of the airbnb listing
date	Date
available	Flag to indicate if available for booking
price	Listing price
adjusted_price	Booking price
minimum_nights	Minimum number of nights required to book
maximum_nights	Maximum number of nights the rental can be booked for

All sources of data and other information have been mentioned in the Reference section.

Data Cleaning

The following data cleaning measures were taken before analysis:

- The focus of this study will be on short term rentals. So, all rentals requiring more than 6 months of minimum stay have not been considered for the analysis.
- The listing type of 'hotel rooms' were removed as hotels are not considered as a traditional Airbnb hosting. Note that hotel listings only accounts for only 1.3% of the original data set.
- The listing data contained some \$0 listings that have been removed as erroneous entries.

In addition to the above steps, some of the missing values were imputed. For example, the rentals with value NA 'reviews_per_month' have been updated to 0. This column is a measure of popularity of a listing. Finally, during data analysis and modeling, outliers have been removed from the relevant columns wherever necessary, so that the analysis results do not get skewed.

References

- <https://en.wikipedia.org>
- <https://www.airbnb.com/>
- <http://insideairbnb.com/>
- https://matplotlib.org/3.2.1/api/pyplot_summary.html
- <https://pandas.pydata.org/docs/reference/index.html>
- <https://python-visualization.github.io/folium/>
- <https://developer.foursquare.com/docs/places-api/>