Airbnb, Inc. is an American company that operates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. It acts as an intermediate between those who want to rent out space and those who are looking for rent.

San Diego is a city on the pacific coast of California, known for its beaches, parks and warm climate. It’s one of the most popular tourist spots in California. There are lot of Airbnb rental homes.

The aim of the project is to create a statistical modeling to predict the rental price and thereby allowing the host to fix the rent of the houses in San Diego.

Factors considered during the rental price prediction:

1. Property type
2. House type
3. Number of bedrooms
4. Number of bathrooms
5. Accommodates
6. Beds
7. Amenities

**Data Sources:**

<http://insideairbnb.com/get-the-data.html>

This is the link for the San Diego Airbnb data set. This contains Airbnb listing for all the cities.

Each dataset from Inside Airbnb contains several data of interest:

* *listings.csv.gz* - Detailed listings data, including various attributes (features) of each listing such as number of bedrooms, bathrooms, location, etc.
* *calendar.csv.gz* - Information about the availability calendar for each listing
* *reviews.csv.gz* - Individual reviews for listings
* *neighborhoods and GeoJSON files* - Mappings and detailed information for each neighborhood within a city

In this project we have used the data for the city of San Diego.

We have used the data from listings.csv.gz, as highlighted.

Graphical user interface, application

Description automatically generated

**Technologies Used:**

1. Jupyter Notebook
2. Matplotlib
3. Numpy
4. Pandas
5. Seaborn
6. Regression Models

**Steps:**

1. Import data from csv into the Jupyter notebook.
2. Data Wrangling
3. Exploratory data analysis