**Project contents:**

This project consists of 3 csv files and 3 Jupyter notebook files

The order to which they were created is:

1. Listings\_notebook.ipynb: contains an understanding of the data in the listings table and machine learning models to predict the Price
2. Calender\_nb.ipynb : contains an analysis of the data in calendar table and a subset of the listings table appears near the end
3. Reviews\_nb.ipynb : contains sentiment analyses of the reviews column found in the reviews dataset

**Installation:**

All packages are standard with python except the following two sentiment analysis libraries:

TextBlob: $ pip install -U textblob

NLTK: $ pip install nltk

**Project Motivation:**

This project’s requirements were to get business value out of the available datasets, after understanding the data 4 questions were raised:

1. To what extent can we predict the price of a listing based on all other parameters in this Dataset?
2. What parameters can a unit owner change to increase the price of a listing?
3. What are the best possible neighborhoods from a business aspect based on price only?
4. Can we quantify the sentiment of the ‘reviews’ column in the reviews table?

**Results:**

The first two questions are answered with a model and a point plot based on the model in the listing notebook

Third question is answered in the calendar notebook

Fourth question is answered by the reviews notebook

**Further Work:**

* A table with the bookings would be very beneficial as it is not mentioned if units are available as in the market or available as they could be booked through these dates. Also, the unavailability of the units is not clear if its due to being booked or its off the market during that time.
* Data for more years could be obtained and analyzed for seasonality and trends