**Readme file**

**Analysis of Zillow and Airbnb Data**

Considering 2 datasets for analysis i.e. of Zillow and Airbnb to find out zipcodes that are suitable to be bought by a real estate agency targeting investment in 2 bedroom apartments.

Here is my consideration and assumption points and approach.

My submission contains an ipython notebook, a readme file and a data visualization report.

Entire code, analysis and visualization has been written in Python. After cleaning and summarizing the data, here are my observations.

There are 6497, 2-bedroom apartments in total in the entire dataset of 2019 Airbnb.

There are 25 zipcodes in the NY city area that have recordings of property price rates for 2 bedroom from the Zillow dataset.

After performing time series analysis on the Zillow dataset from all over the years, we get the forecasted value of 2\_bedroom\_price as the current price value for the property.

The average ratings of particular zipcode has been evaluated as a weighted average of number\_of\_reviews and review\_scores\_rating.

We have a total of 377 unique zipcodes in the Airbnb dataset. Out of which 315 zipcodes offer 2-bedroom sites.

We are calculating the average rent of a property over a particular zipcode out of the mentioned 25 zipcodes.

Since we see fluctuation from time series and stationary analysis the value of property can increase and decrease wrt seasonality, we preferred to take the median of the final year as a reference.

For similar comparison of property values and rent rates we take the scaled measures (0 to 1 range) for both the parameters and find the ratio of rent value/property price.

As per the calculation these zipcodes have value >1 which means the rent value is much more profitable than the purchase price in the year of consideration:

Assuming that the value of rent remains the same throughout the year, I would recommend these zipcodes to be considered for investing:

10305, 10306, 10308, 11234, 11434

**What we could do next:** Given the data for Zillow for anywhere near 2019, we could complete the time series analysis to predict a value for the property prices in that year. Considering factors like nearby location and perform analysis on the nearby transit facilities available as that will add to the locality demand and help understand the cause of prices increase and decrease.