Final Project Proposal – Ist 652

**Analyzing nyc airbnb data from 2019**

**Courtney Zimmer-Bartels**

DATA SET

Airbnb, Inc. has quickly become the go-to marketplace for vacation rentals in the United States and abroad. Founded in 2007, Airbnb now claims they have on average 2 million people staying on Airbnb per night in over 1,000 cities globally. The New York City Airbnb Open Dataset for 2019 contains over 19 columns, containing information such as the listing id, host id, neighbourhood, price, availability and reviews. Combined with a rental inventory dataset from StreetEasy, a NYC only rental and sales platform.

Data Source:

1. <https://www.kaggle.com/dgomonov/new-york-city-airbnb-open-data/data#>
2. <https://streeteasy.com/blog/data-dashboard/?agg=Total&metric=Inventory&type=Rentals&bedrooms=Any%20Bedrooms&property=Any%20Property%20Type&minDate=2010-01-01&maxDate=2020-06-01&area=Flatiron,Brooklyn%20Heights>

TOPIC

**Has in the introduction of Airbnb to NYC negatively affected the rental market?**

1. **Exploratory Analysis**

We will explore the Airbnb inventory in NYC by borough and more specifically neighbourhood. Look at the hotspots within each borough and as NYC as a whole. Try to understand if price is a factor in an Airbnb listing or if location is more of a factor. We also want to take a look at listings by room type to understand the percentage of listing that are entire house/apartments vs. shared rooms. Additionally, we will also look at the rental inventory of each borough and neighbourhood in NYC and see if there has been a noticeable decrease since the introduction of Airbnb into the marketplace.

PROJECT SCOPE

The analysis will include various tabular reports. These tabular reports will show the percentage of Airbnb listing by NYC boroughs and room type. We will also be looking at the average price of room types and regions, as well as the average listing availability of each region. The data sets I have identified are both structured data, I would like to explore adding some unstructured or text data, such as twitter data surrounding Airbnb rentals in NYC. An aspect of this project that creates hesitation is the lack of experience working in python and with unstructured data. Therefore the focus will be in importing and analyzing the structured datasets identified above and exploring social media unstructured data that could add value to the project afterwards.