Amit Kamat

I. Introduction

One of the most important tasks someone moving to a new city is to identify accommodation. The aim of this project is to aid recent graduates, moving to a new city in selecting a suitable property to rent.

In this project, we explore the neighborhoods of Atlanta in order to extract the correlation between the real estate value and its surrounding venues. We shall employ the Foursquare location data to compare various neighborhoods in the city of Atlanta.

It is common for owners or agents to advertise their properties being in the proximity of some interesting venues like supermarkets, restaurants, businesses etc; displaying convenience of the location in order to raise the property's value. We shall be analyzing this correlation and derive some highlights.

This project could be used as a buyers' guide by the following:

- Students moving to a new city, interested to live in a happening neighborhood.
- Property agents who can optimize the value of their properties.
- Real estate planners, aiding them to decide what kind of venues around their products can help them earn a big profit.
- Other budding Data Scientists who are curious about exploring trends in the real estate market and conduct their own studies.

The inspiration for this project came from my ordeal of moving to a new city for work. I was unable to find a source of information that was concise and lucid in order to make a well-informed decision when selecting a property to rent. Hence, I decided to try to undertake my own study and construct a frame to base future studies upon.

Amit Kamat

II. Data Description

Atlanta was chosen as the pilot city for this project. The real estate prices by neighborhoods were scrapped from the following website:

https://www.rentcafe.com/average-rent-market-trends/us/ga/atlanta/

The data covers all the neighborhoods in the city of Atlanta. Note that this data describes the average rent for a studio apartment in a given neighborhood. This data is one of the driving factors in the study. Along with this data, we shall be using geodata of Atlanta pulled from a file stored in cloud:

https://www.dropbox.com/s/edecety6jpkbtfv/atlanta-postal.csv?dl=0

Due to the lack of open data available I decided to construct a couple of dictionaries mapping the latitudes and longitudes to their neighborhoods.

Finally, in order to explore venues in and around the neighborhoods in question, we use the Foursquare API.

The process of cleaning the data and processing it includes the following steps:

1. Scrapping the rentcafe website fo the average rent prices for a condo in Atlanta by neighborhood.
2. Pull the geographic data of the neighborhoods from the cloud file and concatenating the two dataframes.
3. Construct two dictionaries for the neighborhoods, storing their coordinates and then appending the dictionaries to the previous dataframe.
4. The coordinates of each neighborhood are then passed to the Foursquare API . We then obtain a list of venues in the pre-determined radius.
5. We then count the number of occurrences of each type of venue in a neighborhood. We turn each venue type into a column with their occurrence as the value, using one hot encoding.
6. We then standardize the average price by removing the mean and scaling to unit variance.