**Recommending Portland Neighborhoods for Renters: A Geospatial Analysis of Rent and Venues**

Jonathan Jettenberger-Burleson

April 27, 2020

**1. Introduction**

**1.1 Background**

The cost of a home is just one aspect that comes into play when calculating expenses. For many people the idea of buying a home can be either overwhelming or outside of their financial means. Renting is a good alternative that many decide is the optimal option for them. There are resources out there, such as Zillow that assist renters and home buyers alike in finding the right home. The monthly cost to rent can be affected by the unit size, location, number of bedrooms, and many other features. In addition, a unit’s proximity to venues such as restaurants, parks, night clubs, and museums may play a role in pricing and, more importantly, a renter’s interest in that location. Information relating monthly rent rates by neighborhood to nearby venues can be an important tool for renters.

**1.2 Problem**

A new home buyer that has decided that renting may be an option they wish to explore may weigh several criteria when looking for a place that is right for them. Other than the monthly rent, a location’s proximity to certain venues can play a role in a renter’s decision. This report will aim to recommend a neighborhood for a prospective renter based on venue type, venue proximity, and of course, monthly rent. In this report, we will focus on the Portland, Oregon area to minimize the scope.

**1.3 Interest**

This report is primarily of interest to both renters familiar with the Portland, Oregon area and those unfamiliar and new to the area. Home buyers and landlords looking to rent out their future property would also find this information of use in gauging the potential value of their investment based on proximity to venues. Those in real estate or city planning may also find this information of use when predicting rent, property value, or interest in an area. Business owners, especially small, local businesses, may find this report noteworthy in predicting their potential customer base.

**2. Data**

**2.1 Data sources**

In this report, I will be leveraging geospatial venue data gathered from [Foursquare](https://foursquare.com/) and rental data gathered from [Zillow](https://www.zillow.com/). The Foursquare API will mainly be utilized for venue types and their locations in the Portland area, where requests send back trending locations within the search radius. As the Zillow API does not allow for general searches for an area, information on rent from multiple properties and their locations will come from [here](https://www.zillow.com/research/data/). This data is periodically updated and is relevant at the time of this report. The Zillow service uses a proprietary algorithm to determine rent as an attribute they call Rent Zestimate. This rent value factors in a property’s characteristics, unique features, on-market data, and off-market data. We will be selecting the Zillow Rent Index (ZRI) data with neighborhood geography. Boundary data, found [here](https://gis-pdx.opendata.arcgis.com/datasets/1ef75e34b8504ab9b14bef0c26cade2c_3), will also be used for modeling of neighborhoods. To assist in the cleaning of data, information about neighborhood names was gathered from [PDX Listed](https://www.pdxlisted.com/neighborhoods/).

**2.2 Data cleaning and feature selection**

Data requested from Foursquare provides much information about venues in an area. Information can be gathered based on location, other users, or even just by venue category. Foursquare also provides data on tips, hours, menus, photos, and events. Much of this information would be useful depending on the needs of the user, but are out of scope for this report.

We will be focusing on the latitude and longitude location of each venue, and the categories the venue falls under. The names of the venues will also be included for labeling. To restrict the data size for this report, we will search for venues with Portland as a city value. Due to the 100-entry limit on returned data per request for the Foursquare API, five separate requests were made in an attempt to increase the dataset size. Each request was centered on one of the five sections of Portland (Table 1) and was searched in a radius of 8046 meters, or approximately 5 miles. Overlapping data points were also removed. This method yielded a dataset of 366 venues in the Portland area (Figure 1).

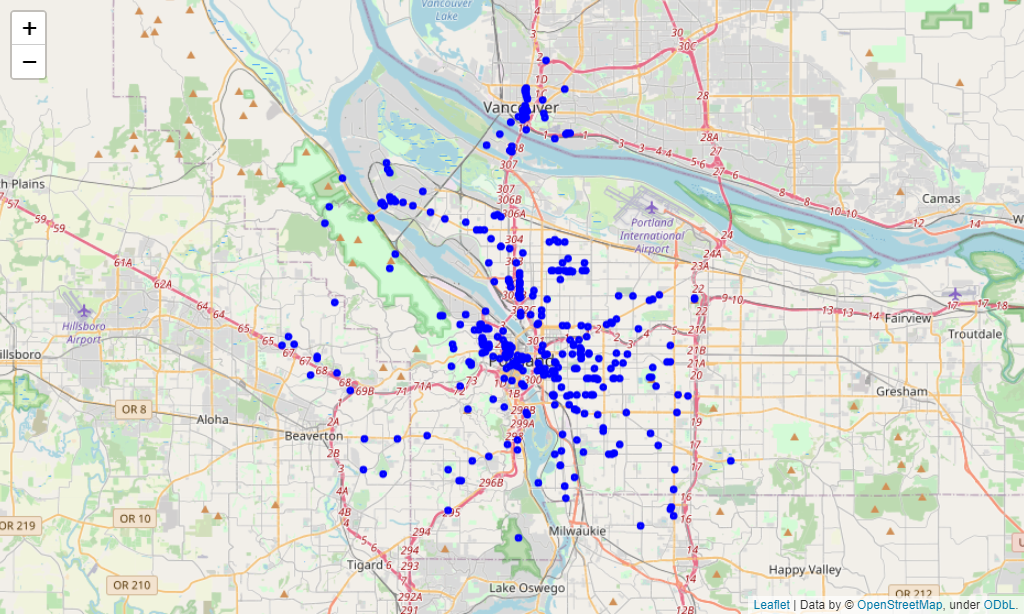


Figure 1. Map of Foursquare venue data points.

Table 1. Latitude and longitude data on the five sections of Portland, OR.

|  |  |
| --- | --- |
| Section | Location (lat, lng) |
| Northwest | 45.5586, -122.7609 |
| Northeast | 45.5676, -122.6179 |
| North | 45.6104, -122.7034 |
| Southwest | 45.4849, -122.7116 |
| Southeast | 45.4914, -122.5930 |

Zillow has detailed information on property characteristics and market data. The dataset collected has information from multiple locations. To refine the data, I began with selecting the subset focused on rental properties with the values City and State equaling Portland and OR, respectively. Due to the limited scope of this report, only the neighborhood names and the Rent Zestimate data will be examined. The column titles for this information was standardized to Neighborhood and Rent as well. In Portland, OR, many neighborhoods are represented by a home owners association or a league. This information was not represented between data sources. To match up the rent subset with the geospatial information gathered from the boundary data, 17 data entries for neighborhood names had to be cleaned because of this. There was one additional instance of a neighborhood being represented by its former name.

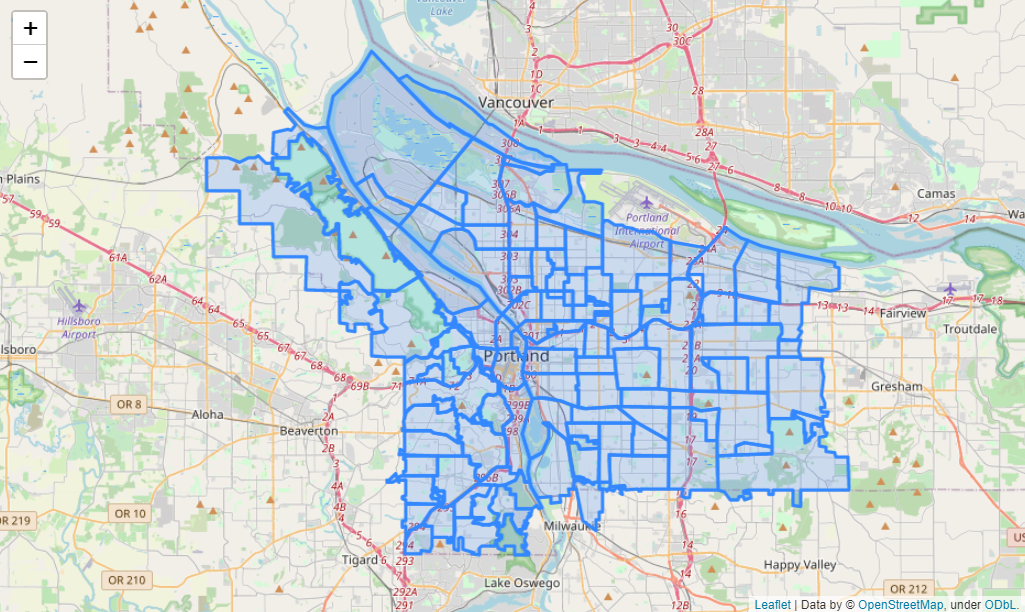


Figure 2. Boundary data for the neighborhoods of Portland, OR.

The neighborhood boundary data includes 25 entries that represent overlapping of the neighborhoods. In these instances, available rental data was averaged for the overlapping areas. From the rental data from Zillow, only 93 of the 130 areas on Figure 3 were filled. This still left some areas of the map with no rental data. A few of these areas represent unclaimed neighborhoods that are not claimed by any association. The other areas simply have no rental data on Zillow. This could be due to a lack of rental information posted on Zillow, or even no rental properties to speak of.

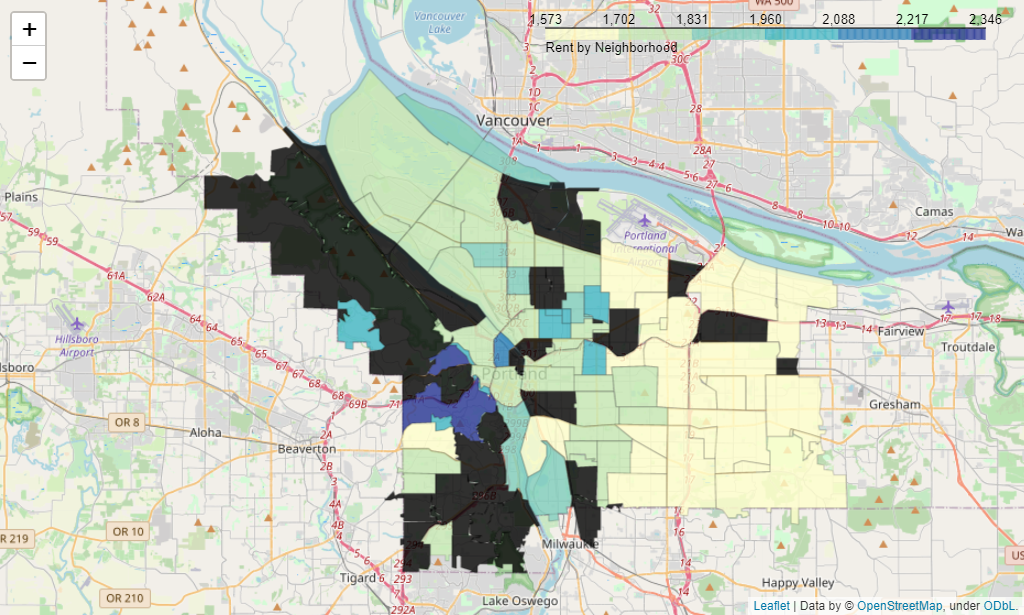


Figure 3. Heatmap representing average rent for neighborhoods in Portland, OR.

**3. Methodology**

**3.1 Exploratory data analysis**

**3.2 Merging data**

**3.2 Clustering**

**4. Results**

**4.1 Relationship between rent and entertainment**

**4.2 Relationship between rent and restaurants**

**4.3 Relationship between rent and parks**

**4. Discussion**

**4.1 Observations**

**4.2 Recommendations**

**5. Conclusions**

**6. Future directions**