

## Best Portland Neighborhoods for Non-Profit

### Introduction:

Portland, Oregon is the largest city in Oregon. It is known for many things. For example, it rains a lot. Portland averages 42 inches of rain per year and 154 days per year have measurable precipitation. The city is also known for its restaurants, breweries, and numerous coffee shops. Perhaps less known is the city's focus on service. In fact, Portland State University requires all students to complete a capstone project that includes community service.

A local Portland non-profit is wanting to expand its outreach to Portland teens. They are looking for neighborhoods with access to mass transit for small group gatherings. They would also like the neighborhoods to have venues such as coffee shops and pizza places that might be willing to host these gatherings.

Using available data we will create a list of the Portland area neighborhoods that satisfy these requirements.

### Data:

To complete this project, two data sources were used. First, a list of Portland neighborhoods was found on the [Portland City Government website](#). This list was then put into a CSV file. The other data source was the [Foursquare API](#). From Foursquare, we were able to create a data frame of neighborhoods and their top venues. These venues include places such as parks, restaurants, stores, mass transit, and more.

### Methodology:

The first step to analyzing the Portland, Oregon neighborhoods was to create a panda dataframe with the neighborhood names with their corresponding Latitude and Longitude coordinates. The neighborhood names were copied from the Portland City Government website and put into google sheets. This was then saved as a csv file and added to the project in Watson Studio. The Geocoder Python package was used to gather the coordinates of each neighborhood. The coordinates from one neighborhood could not be found. That row was dropped from the data frame. The resulting data frame had 92 neighborhoods.

	Neighborhood	Latitude	Longitude
0	Alameda	45.548631	-122.636481
1	Arbor Lodge	45.571794	-122.690152
2	Ardenwald/Johnson Creek	45.458516	-122.627539
3	Argay Terrace	45.552830	-122.523204
4	Arlington Heights	45.519496	-122.710667
5	Arnold Creek	45.441395	-122.699933

Figure 1: Portland Neighborhood Dataframe:  
First five entries

The folium python package was used to visualize the neighborhoods on a map.

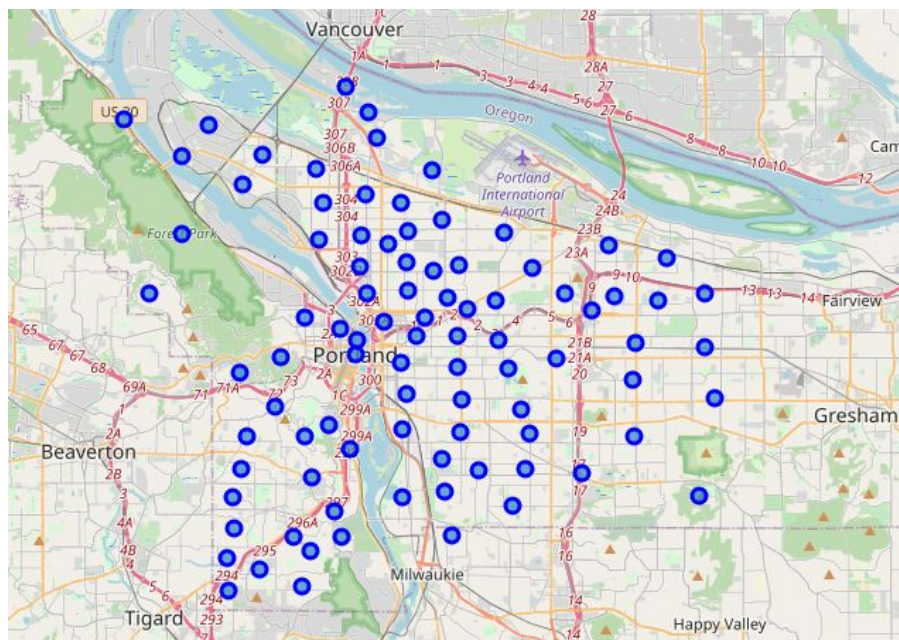


Figure 2: Map of Portland Neighborhoods

The Foursquare API was used to collect venue data for each neighborhood. The top venues for each neighborhood were collected within 600 meters of the latitude and longitude coordinates for the neighborhood. The amount of venues collected was limited to 100. This method will not collect any neighborhood venues outside the 600-meter radius of the coordinates. However, this data should still give a good view of the neighborhoods' top venues.

This data returned 290 unique venues. Also, not all neighborhoods returned the full 100 max venues.

Neighborhood	
Alameda	7
Arbor Lodge	13
Ardenwald/Johnson Creek	9
Argay Terrace	6
Arlington Heights	30
Arnold Creek	1
Ashcreek	2
Beaumont-Wilshire	17
Boise	100
Brentwood/Darlington	4

Figure 3: Number of Venues:  
First ten entries

The neighborhoods with less than 10 venues were removed from the dataframe. The final number of neighborhoods being analyzed was 51.

A new data frame was created showing each neighborhood and it's 10 top venues.

	NeighborhoodName	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Arbor Lodge	Convenience Store	Mexican Restaurant	Bus Stop	Thai Restaurant	Coffee Shop	Gym	Park	Bar	Marijuana Dispensary	Video Store
1	Arlington Heights	Trail	Park	Garden	Playground	Gift Shop	Train Station	Botanical Garden	Fountain	Tennis Court	Field
2	Beaumont-Wilshire	Pizza Place	Breakfast Spot	Bar	Grocery Store	Bus Line	Flower Shop	Taco Place	Food & Drink Shop	Bagel Shop	Thai Restaurant
3	Boise	Food Truck	Coffee Shop	Pizza Place	Gift Shop	Bar	Cocktail Bar	Brewery	Yoga Studio	Beer Garden	Pet Store
4	Bridgeton	Harbor / Marina	American Restaurant	Hawaiian Restaurant	General Travel	Deli / Bodega	Motel	Thai Restaurant	Grocery Store	Bar	Ethiopian Restaurant

Figure 4: Top 10 Venues:  
First five entries

One hot encoding was used to transform our venue data into numerical values. A data frame was created with the neighborhood names and the frequency of each of the 290 unique venues in that neighborhood.

	NeighborhoodName	ATM	Accessories Store	African Restaurant	American Restaurant	Amphitheater	Antique Shop	Arcade	Argentinian Restaurant	Art Gallery	Arts & Crafts Store	Arts & Entertainment	Asian Restaurant	Athletics & Sports
0	Arbor Lodge	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000000
1	Arlington Heights	0.000000	0.000000	0.000	0.000000	0.033333	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000000
2	Beaumont-Wilshire	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000000
3	Boise	0.000000	0.000000	0.000	0.010000	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.010000	0.000000
4	Bridgeton	0.000000	0.000000	0.000	0.166667	0.000000	0.000000	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000000
5	Brooklyn	0.071429	0.000000	0.000	0.000000	0.000000	0.000000	0.000000	0.000	0.000000	0.071429	0.000000	0.000000	0.000000

Figure 5: Sample of dataframe used for clustering

This data was used to cluster (group together because of certain similarities) the neighborhoods based on their venues. To do that, K-means clustering was used. K-means is a





	NeighborhoodName	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Arbor Lodge	45.571794	-122.690152	6	Convenience Store	Mexican Restaurant	Bus Stop	Thai Restaurant	Coffee Shop	Gym	Park	Bar	Marijuana Dispensary	Video Store
1	Arlington Heights	45.519496	-122.710667	4	Trail	Park	Garden	Playground	Gift Shop	Train Station	Botanical Garden	Fountain	Tennis Court	Field
2	Beaumont-Wilshire	45.550391	-122.623694	0	Pizza Place	Breakfast Spot	Bar	Grocery Store	Bus Line	Flower Shop	Taco Place	Food & Drink Shop	Bagel Shop	Thai Restaurant
3	Boise	45.550159	-122.671878	1	Food Truck	Coffee Shop	Pizza Place	Gift Shop	Bar	Cocktail Bar	Brewery	Yoga Studio	Beer Garden	Pet Store
4	Bridgeton	45.602409	-122.668102	2	Harbor / Marina	American Restaurant	Hawaiian Restaurant	General Travel	Deli / Bodega	Motel	Thai Restaurant	Grocery Store	Bar	Ethiopian Restaurant
5	Brooklyn	45.494819	-122.651552	0	Bar	ATM	Light Rail Station	Grocery Store	Marijuana Dispensary	Park	Café	Arts & Crafts Store	Bus Stop	Mexican Restaurant

Figure 7: Neighborhoods with Top Ten and Cluster Label:  
First five entries

From this the top 10 venues in each cluster was determined.

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
Bar Grocery Store Pizza Place Coffee Shop Mexican Restaurant Flower Shop Breakfast Spot Brewery Italian Restaurant Bus Stop	Food Truck Bar Coffee Shop Pizza Place Asian Restaurant Brewery Burger Joint Gym Breakfast Spot Sandwich Place	American Restaurant Bar Deli / Bodega Ethiopian Restaurant General Travel Grocery Store Harbor / Marina Hawaiian Restaurant Motel Thai Restaurant	Coffee Shop Sandwich Place Hotel Pizza Place Park Mobile Phone Shop Yoga Studio Bus Stop Food Truck Clothing Store	Botanical Garden Field Fountain Garden Gift Shop Park Playground Tennis Court Trail Train Station	Coffee Shop Mexican Restaurant Bar Pizza Place Sandwich Place Food Truck Brewery Grocery Store Liquor Store Café	Coffee Shop Convenience Store Bar Park Grocery Store Japanese Restaurant Thai Restaurant Smoke Shop Sandwich Place Pizza Place

Figure 8: Top 10 Venues in Each Cluster

Results:

The analysis of the data yielded some helpful results. The clusters can be named by focussing on the differences of each cluster.

1. Food and drinks
2. Food trucks and restaurants
3. Eclectic
4. Coffee and mass transit
5. Recreation

6. Coffee and food
7. Coffee and parks

Also, by looking at the map it is shown we can look at whether the neighborhoods are on the East or West side and how far they are from downtown.

Cluster	East or West	Proximity to downtown
1	East	Central
2	East	Away from central
3	East	Very far away
4	East and West	Central
5	West	Somewhat away
6	East and West	Away from central
7	East and West	Very far away

Figure 9: Where the Clusters are on the Map

#### Discussion:

The purpose of this analysis was to help a Portland non-profit find the best neighborhoods to host get-togethers for local teens. Looking at the data, it can be seen that most neighborhoods will have appropriate venues to host the get togethers. Coffee shops show up in almost all the clusters. Therefore, we must look a little deeper. Cluster four has coffee shops, pizza places, and parks in their top venues list. All of these venues would work well for the non-profit's needs. Also, these neighborhoods have good access to mass transportation and are centrally located. This will make it easier for the teens to get there. In addition, cluster four neighborhoods are on both the east and west sides. This gives the non-profit more options. They may choose to alternate east and west locations for their gatherings.

Cluster Four Neighborhoods
Creston-Kenilworth
Eliot
Hayden Island
Lloyd District
Markham

Mt Tabor
Multnomah
Old Town
Pearl
South Portland

Figure 10: Recommended Neighborhoods

#### Conclusion:

The purpose of this project was to create a recommendation for a Portland, Oregon non profit. By combining a list of neighborhoods with their latitude and longitudes we were able to obtain venue data from the Foursquare API. With k-means clustering we were able to cluster the neighborhoods into meaningful groups. Each cluster had a slightly different make-up. By looking deeply at the results of the data analysis we found that cluster four most closely meets the non-profits needs.