Understanding Millennial Tastes

Millennials: An Overview

- They are the group born between 1981 and 1996
  - As of 2019 they are in the age range of 23 to 38
- On average they are:
  - Better Educated
  - Less likely to get married early
  - Less likely to have children early
  - Have more workplace diversity

A Tale of 2(Millennial) Cities

- Portland, Oregon and Seattle, Washington are two cities that have a very high net migration of millennials.
  - In Seattle, they make up almost 40% of the population
- With such a prominent demographic becoming the dominant force in society, how might we better grow to understand their unique set of needs and desires?

Data

Sources:

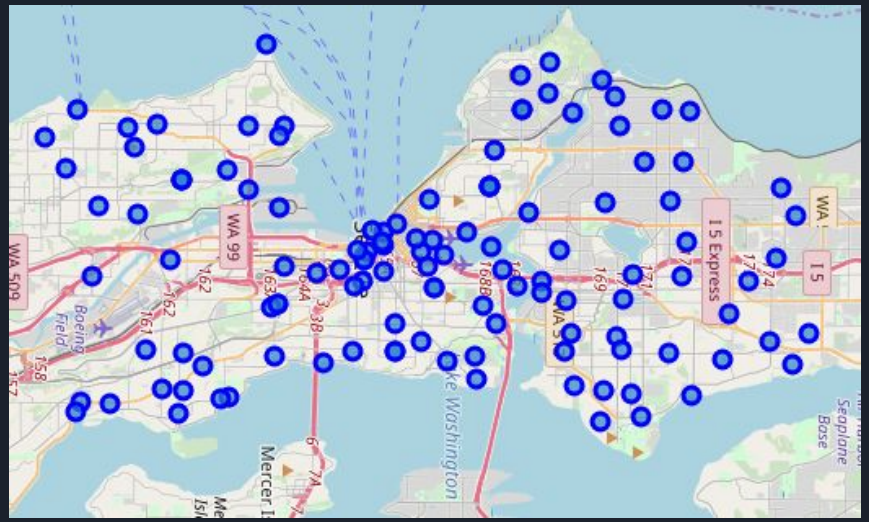
The data for this project was collected from the following places:

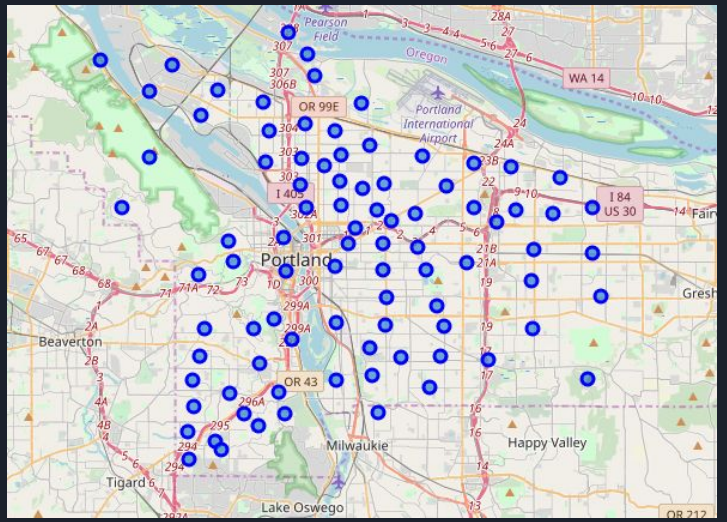
- A table of neighborhoods in Seattle was scraped from wikipedia
- Latitude and Longitude coordinates were collected for each neighborhood using the Nominatim geocoder
- Foursquare API was used in conjunction with the neighborhood names and the coordinates to obtain a list of venues with their respective categories
- A CSV of Portland neighborhood names was obtained http://gis-pdx.opendata.arcgis.com/datasets/neighborhoods-regions

Data Cleaning:

- The data was imported, superfluous columns were dropped, and the column of neighborhood names was iterated over so that it was in a format that could be called by the geocoder

Folium Maps of Portland and Seattle





How was the data used?

- A function was used to call Foursquare's api for each row of geocoded Latitude and Longitude coordinates
- This provided a list of venues near each neighborhood
- All of the venues in each city were aggregated, and a count was performed for each unique venue category.

	Venues	Raw_Counts	Ratio_Counts
0	Coffee Shop	204	7.019959
1	Park	120	4.129387
2	Pizza Place	101	3.475568
3	Bar	70	2.408809
4	Mexican Restaurant	64	2.202340

How was the data used?[Part 2]

- These counts were divided by the total number of unique venues to give the ratio of those venue categories in each city.
- The 2 dataframes were merged, and a function looped over each row to perform a Z-test for 2 proportions.

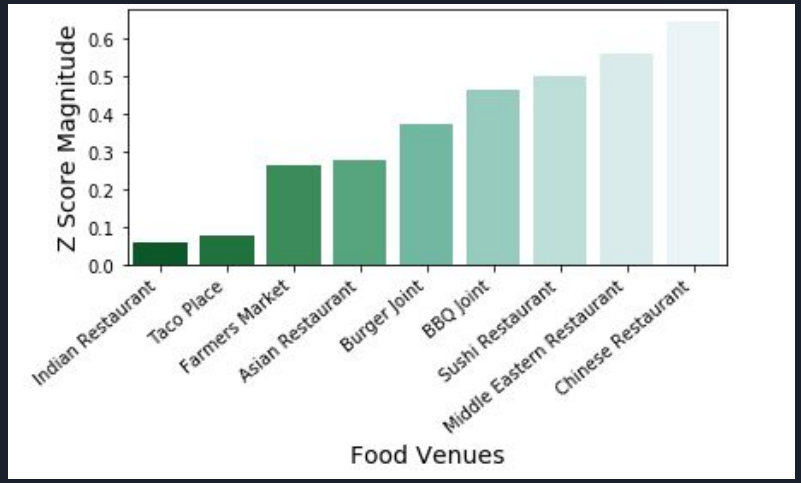
	Venues	Raw_Counts_Seattle	Ratio_Counts_Seattle	Raw_Counts_Portland	Ratio_Counts_Portland	Z_Score
12	Trail	38.0	1.307639	37.0	1.309271	-0.005473
305	Bed & Breakfast	1.0	0.034412	1.0	0.035386	-0.019880
282	Food Stand	1.0	0.034412	1.0	0.035386	-0.019880
277	Campground	1.0	0.034412	1.0	0.035386	-0.019880
270	Golf Driving Range	1.0	0.034412	1.0	0.035386	-0.019880

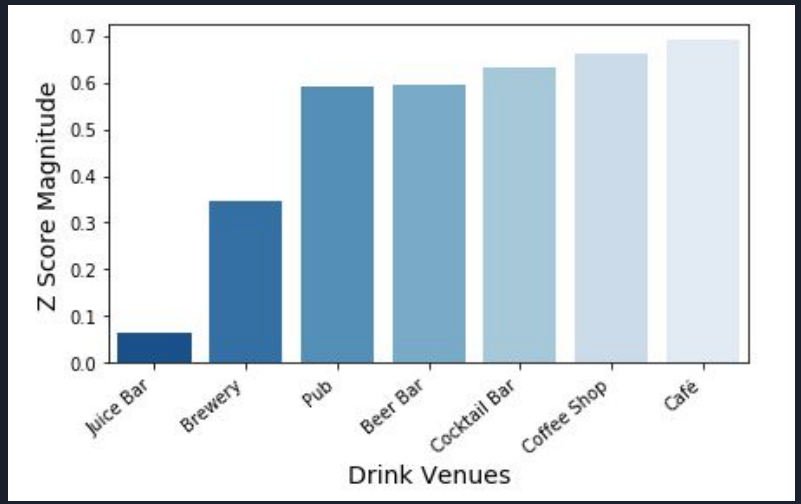
Data Analysis and Categorization

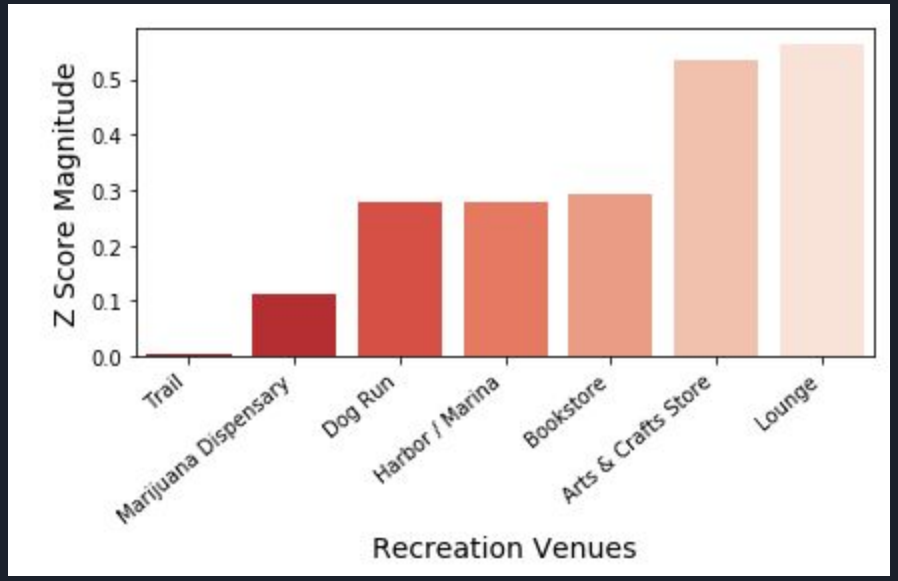
- An alpha value of 0.05 was selected to divide the 398 unique venue categories into 2 tables labelled ‘differences’ for Z scores with a magnitude greater than 1.96, and ‘similarities’ for Z scores with a magnitude smaller than 1.96.
- The similarity dataframe was further subsetted into values with especially small Z Score magnitudes, leaving 32 total venue categories

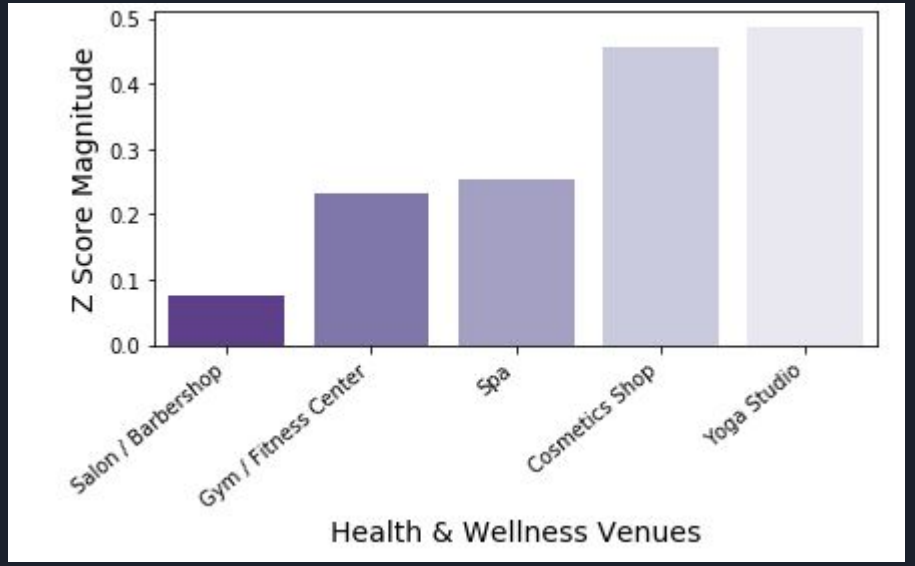
Data Categorization

- These 32 venues were grouped into 5 different groups: Food, Drink, Recreation, Wellness, City Essentials.
- The first 4 groups were graphically represented to highlight the venue categories that possessed a very similar representation of that respective venue category






Z Score Magnitude
0.0
0.1
0.2
0.3
0.4
0.5
Trail
Marijuana Dispensary
Dog Run
Harbor / Marina
Bookstore
Arts & Crafts Store
Lounge
Recreation Venues


Z Score Magnitude
0.0
0.1
0.2
0.3
0.4
0.5
Salon / Barbershop
Gym / Fitness Center
Spa
Cosmetics Shop
Yoga Studio
Health & Wellness Venues

Although further research should be done into the topic, we can at least say with some degree of certainty that of the venues returned from foursquare's api in both the cities of Seattle and Portland, the venues that have the most similar representation are international restaurants, venues that serve physiologically altering substances(caffeine, alcohol, cannabis), nature trails, and wellness centers(spas, yoga studios, gyms).