FINDING *SEATTLE* SUBSTITUTE IN THE WASHINGTON STATE

Fadhel Hariz Dzulfikar

Coursera Capstone Project

- *Washington is the eighteenth largest state, with an area of 71,362 square miles (184,827 km2), and the thirteenth most populous state, with more than 7.6 million people. Approximately 60 percent of Washington's residents live in the Seattle metropolitan area. Some people might think to live in another city (leave Seattle) where the city is not too densely populated, but similar to Seattle, in order to solve population distribution problem. They need some alternatives city to think of where to buy their new house and property.* So, are there nearby cities around Washington state that similar to Seattle to live in?

BUSSINESS PROBLEM

- *Wikipedia >>>>> List of cities in the Washington State*

- *MapQuest API >>>>> Locating each city in the Washington State*

- *FourSquare API >>>>> Listing top venues of each city*

DATA (PREPROCESSING)

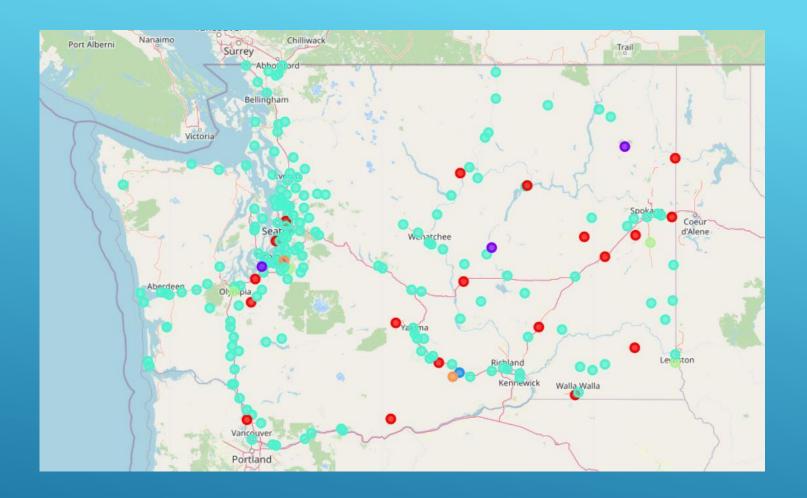K-Means clustering → Listing top 5 venues in each city → Implementing the cluster based on K-Means Clustering

METHODOLOGY (DATA PROCESSING)

RESULT AND DISCUSSION (1)

- *Cluster 0: 19 cities with mainly parks and farms avenue.*

- *Cluster 1: There are only four cities listed in cluster 1. Therefore, this cluster cannot be used as reference to find your city.*

- *Cluster 3: These cities on the list have many similarities to Seattle and might be people's alternative to find home where the city are less populated than Seattle but have Seattle vibes. Moreover, the cities on this cluster have many entertainment avenues as the top five avenues such bar, restaurant, cafe, coffe shop, and many more which is good for business.*

- *Cluster 4: It has only four cities listed on this cluster. The top avenues here are fishing spot and farmers market.*

- *Cluster 5: candy store and falafel restaurant making them unique in this dataset, but unreliable.*

RESULT AND DISCUSSION (2)

- *To solve the problem, we used KMeans clustering in order to group the cities into 6 clusters. However there are only four data that are reliable. This project needs to be improved whether using more reliable dataset or clustering method. Based on this project, we sure can answer the main problem stated in the beginning of this project that is "are there nearby city around Washington state that similar to Seattle to live in?". And the answer are listed on cluster 3.*

CONCLUSION