Destination USA: Where to live as a new immigrant in USA

By: Ernest K. Kwegyir-Afful

INTRODUCTION/BUSINESS PROBLEM

- Every year a number of immigrants become new U.S citizens
- In the 2018 fiscal year 163,000 immigrants became new citizens
- The diversity visa program makes available 50,000 immigrant visas every year across the globe
- Some of these immigrants migrate to the U.S without having relatives that they can associate with.

Project Goal

The goal of this project is to develop an analysis that compares
neighborhoods in different U.S cities to identify those with a cultural
composition that will be attractive to a specific group of immigrants with a
similar cultural background.

Data

- Data was obtained for neighborhoods in Washington DC, Baltimore, Maryland and Des Moines, Iowa from the following sources
- Des Moines, Iowa: https://www.areavibes.com/des+moines-ia/neighborhoods/
- Baltimore, Maryland:
 https://en.wikipedia.org/wiki/List_of_Baltimore_neighborhoods
- Washington, DC: <u>https://en.wikipedia.org/wiki/Neighborhoods_in_Washington, D.C.</u>

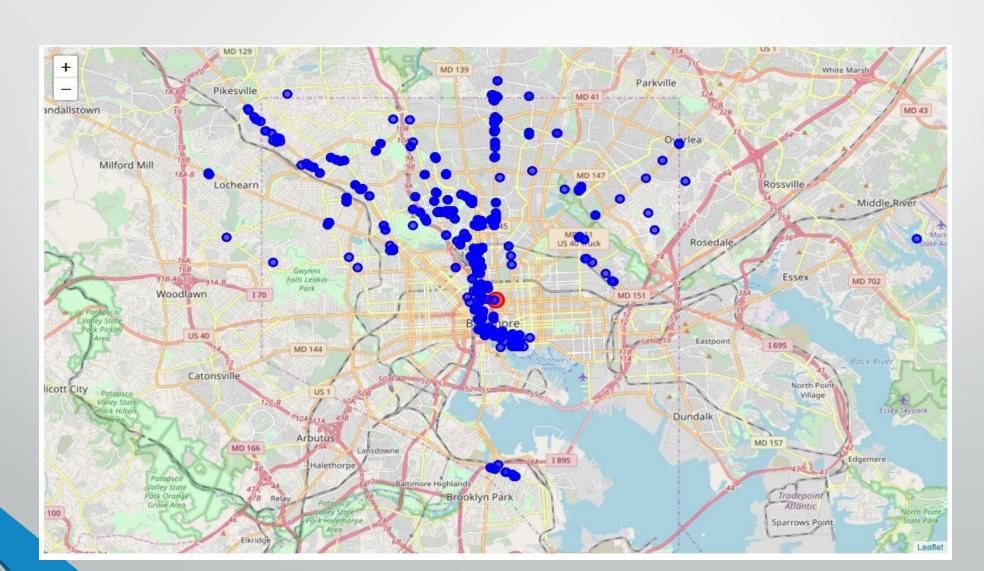
Neighborhood Data

- Baltimore has 308 neighborhoods, Des Moines has 53 neighborhoods and Washington DC has 138
- For Baltimore and Washington DC 100 neighborhoods each were used to minimize density of plots on the maps.
- The Python geocoder returned coordinates for 81 Baltimore neighborhoods, 24 neighborhoods for Des Moines and 98 neighborhoods for DC.
- These coordinates were then passed to the Foursquare API to return venues of interest.

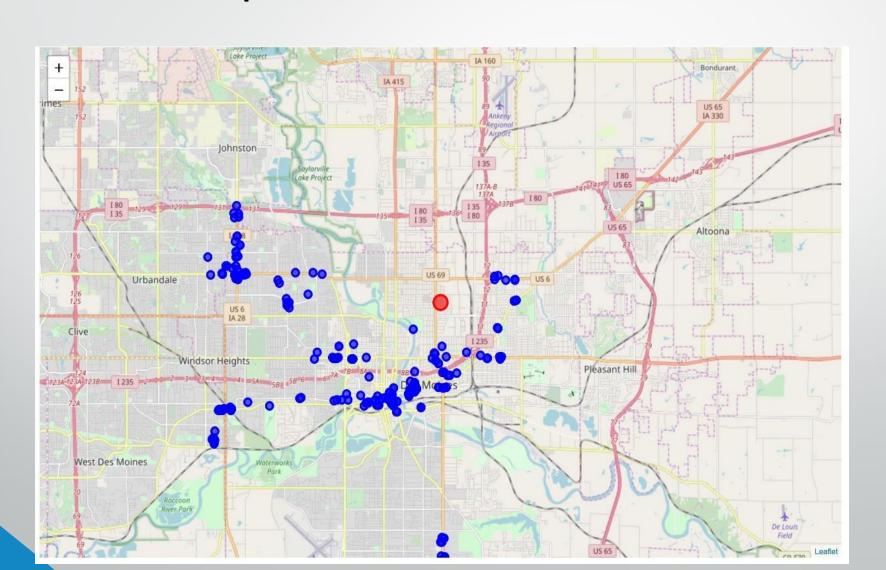
METHODOLOGY

- For our Foursquare API call, we looked within a radius of a 1000 meters from our location coordinates and returned 300 venues
- We filtered the data returned from Foursquare based on key words in the venue category that were identified
 - Venue Categories: Restaurant, Market, Pizza, Food and Sandwich
- For Baltimore, 249 number of unique categories were returned. 190 unique categories and 323 unique categories were returned for Des Moines and Washington DC respectively.

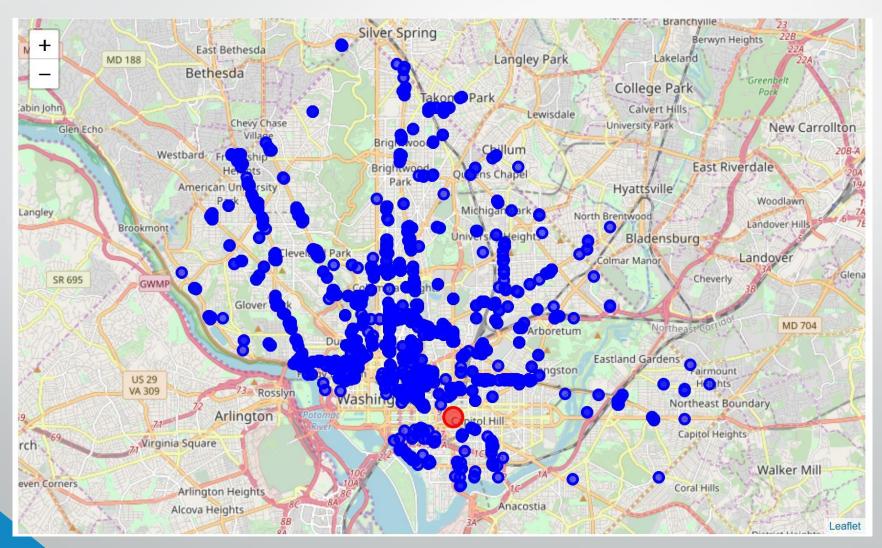
Map of Baltimore venues



Map of Des Moines Venues



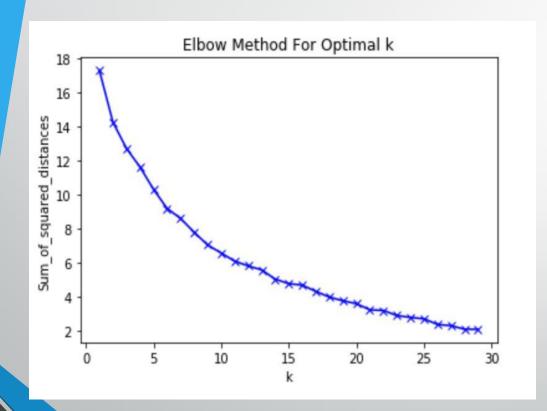
Map of Washington DC Venues

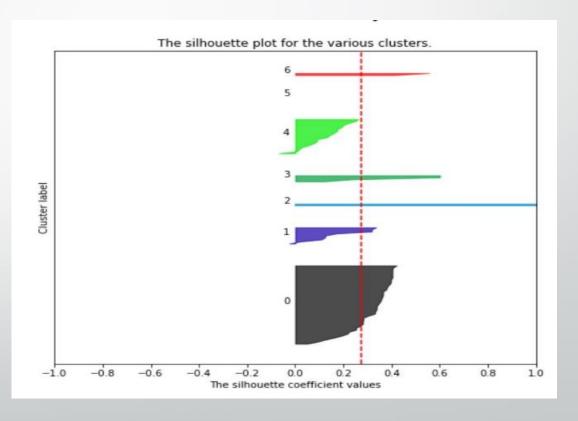


METHODOLOGY

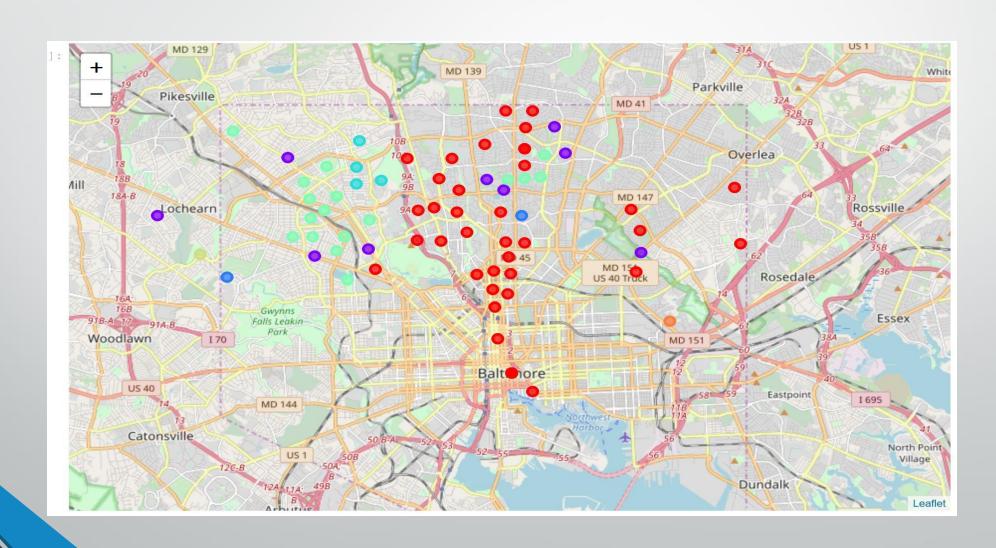
- K-means clustering was used to generate clusters of the different neighborhoods based on the venues returned by Foursquare.
- Clusters were thus based on restaurants or eateries in these neighborhoods
- To determine the optimum number of K we used:
 - Elbow method
 - Silhouette analysis
 - Calinski-Harabaz method
- We also explored Density Based Clustering to see if it eliminates the shortcomings of Kmeans

Selecting K for Baltimore venue data

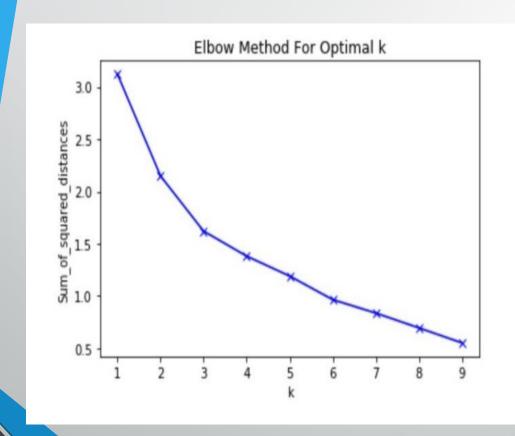


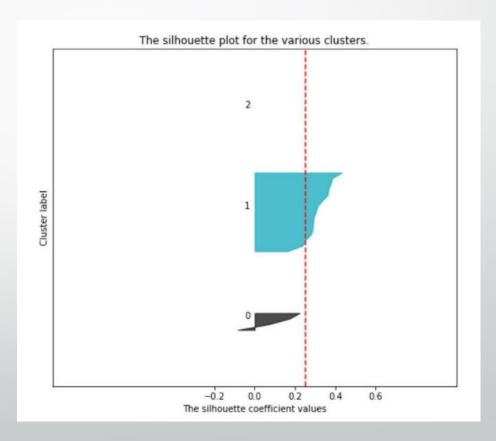


7 Neighborhood clusters in Baltimore



Selecting K for Des Moines

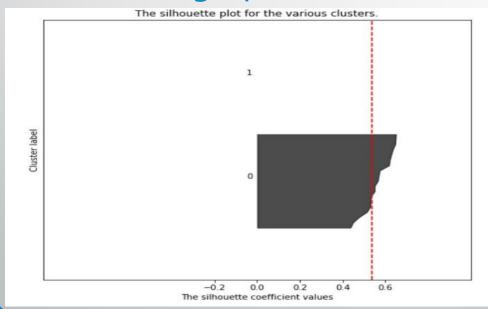




Silhouette coefficient: 3 clusters: 0.25

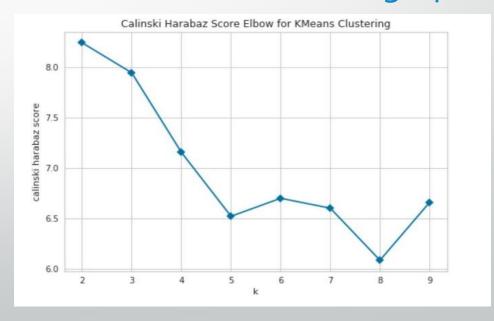
Selecting K for Des Moines

Silhouette graph for 2 clusters

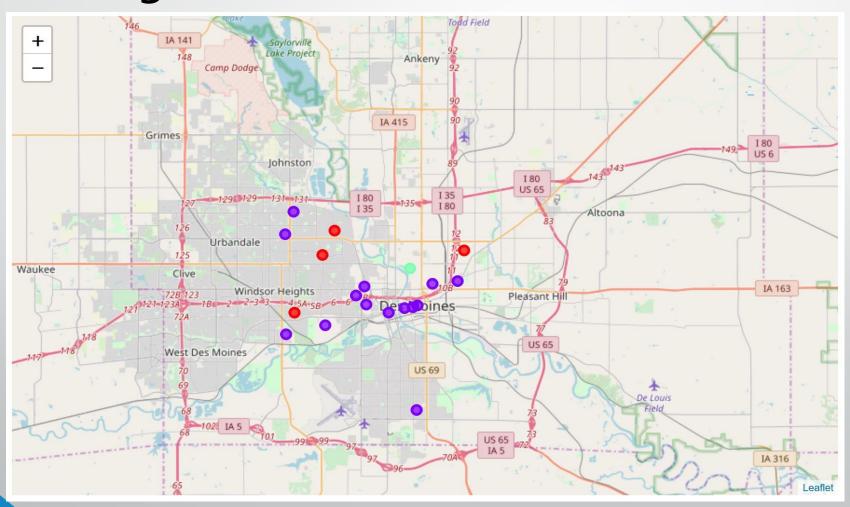


Silhouette coefficient: 2 clusters: 0.54

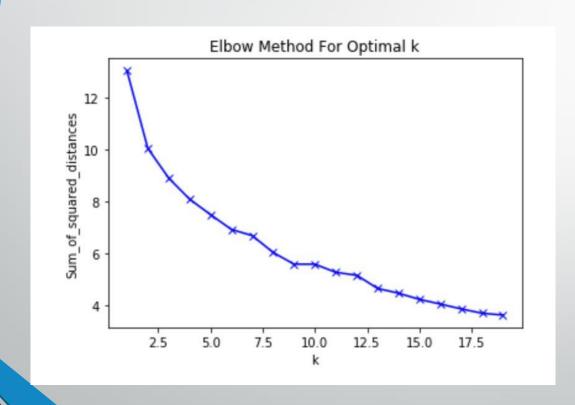
Calinski-Harabaz score graph

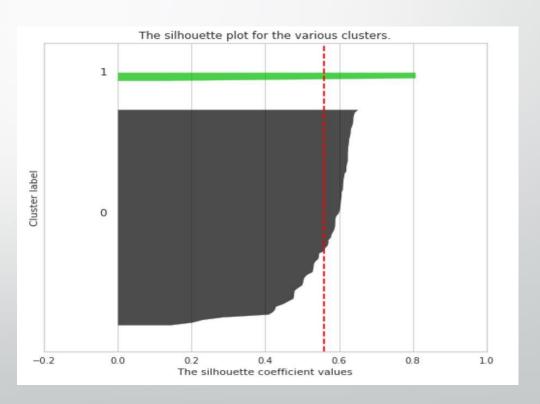


2 Neighborhood clusters in Des Moines

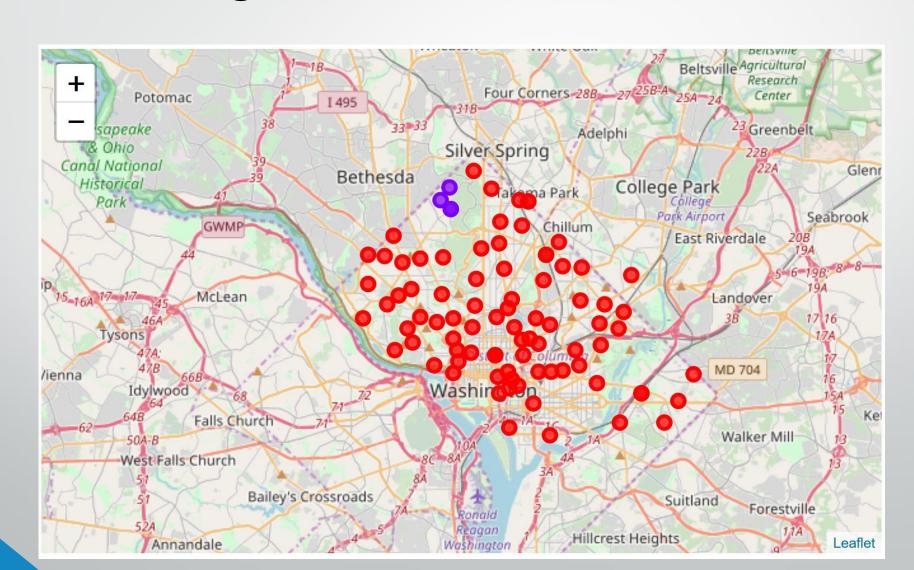


Selecting K for Washington DC venue data





2 Neighborhood clusters in DC



Cultural backgrounds represented in cities

- Washington DC
 - American
 - Asian
 - African
 - European
 - South American/Mexican
 - Mediterranean
 - Caribbean
 - Middle Eastern

- Baltimore
 - American
 - Asian
 - African
 - European
 - Mediterranean
 - Caribbean
 - Middle Eastern
 - Mexican/South American

- Des Moines
 - American
 - Asian
 - Mexican
 - European
 - Mediterranean
 - Middle Eastern

Conclusions

- While some of the same cultural backgrounds are represented in the other cities, the sheer density in Washington DC emphasizes the differences in diversity.
- For the three cities explored, in descending order with the city with the highest cultural diversity
 - Washington DC
 - Baltimore, MD
 - Des Moines, IA