The Battle of the Neighborhoods

IBM Data Science Professional Certificate Final Presentation

Introduction

- A company notices the high growth rates of the number of young people in Milwaukee, WI, and wants to capitalize on it
- They are unfamiliar with Milwaukee and need to make an informed decision
- Ideal location qualities:
    - High foot traffic
    - Lots of nightlife
    - Convenient to get to

Data

Sources:

- Wikipedia
  - Neighborhood names and districts
- Foursquare
  - Venue data
- Google/Google maps
  - Latitude and longitude

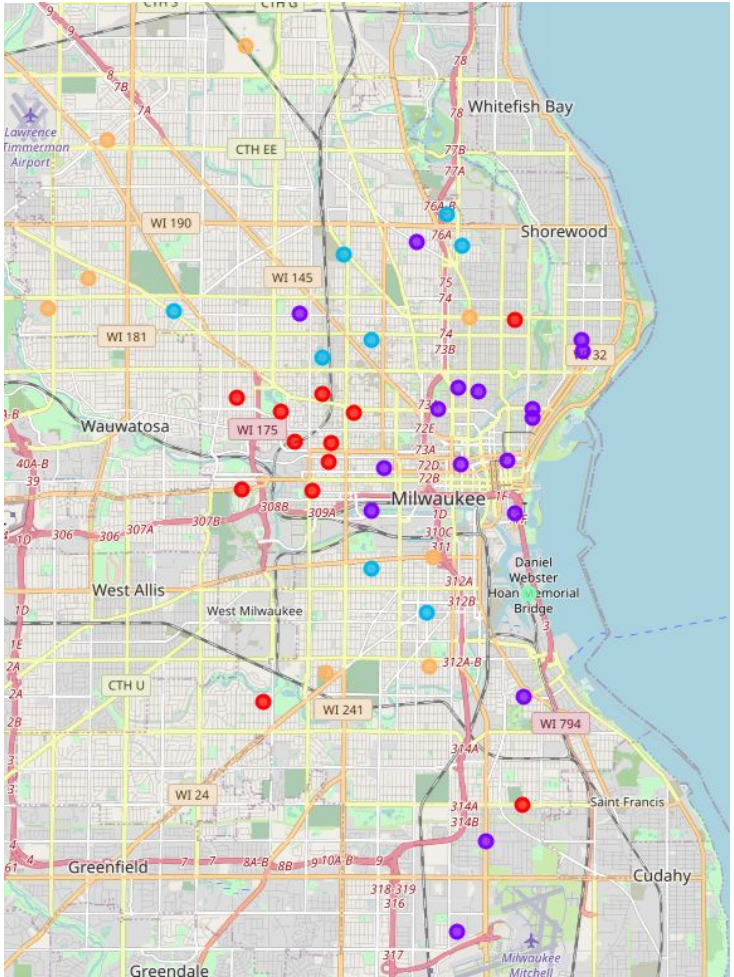
Methodology

- Foursquare
  - 10 most common out of 100 closest venues
- Create Milwaukee CSV
- Clustering Algorithm: K-Means
  - Unlabelled data set
  - Multiple iterations
  - 5 clusters

Results

Cluster:
1
2
3
4
5

- Cluster 1
  - Bars, parks, activities
- Cluster 2
  - Bars, restaurants, coffee shops
- Cluster 3
  - Gyms, grocery stores, parks
- Cluster 4
  - Marina
- Cluster 5
  - Mexican restaurants, gas stations



Conclusion

- Depends on type of venue
- Cluster 1 more nightlife
- Cluster 2 more daytime attractions
    - University of Milwaukee