Access to Parks and Fitness Facilities in Manhattan

Racine Nassau

September 8, 2020

IBM Data Science Specialization Capstone Project

Exercise and Green Space

- Green space improves physical and mental health
- > 80% of US adults do not meet guidelines for physical activity
- **Many Americans live in neighborhoods that lack access to parks and fitness facilities**

Project Aims

- Characterize Manhattan neighborhoods based on access to parks and sports/fitness facilities
- Explore the relation between socio-economic factors such as income and the level of access to these facilities across different neighborhoods

Data Acquisition

- Demographic/socio-economic data from this Kaggle dataset
- 2010 Census Tract to 2010 Neighborhood Tabulation Area Equivalency CSV file to group the above data into Manhattan neighborhoods
- GeoJSON file of Neighborhood Tabulation Areas from the NYC Department of City Planning website
- A list of nearby parks and sports/fitness facilities for each neighborhood collected using the Foursquare API

Gathering and Cleaning Foursquare Location Data

Search for parks and sports & athletics venues within 500 m of the neighborhood's latitude and longitude point

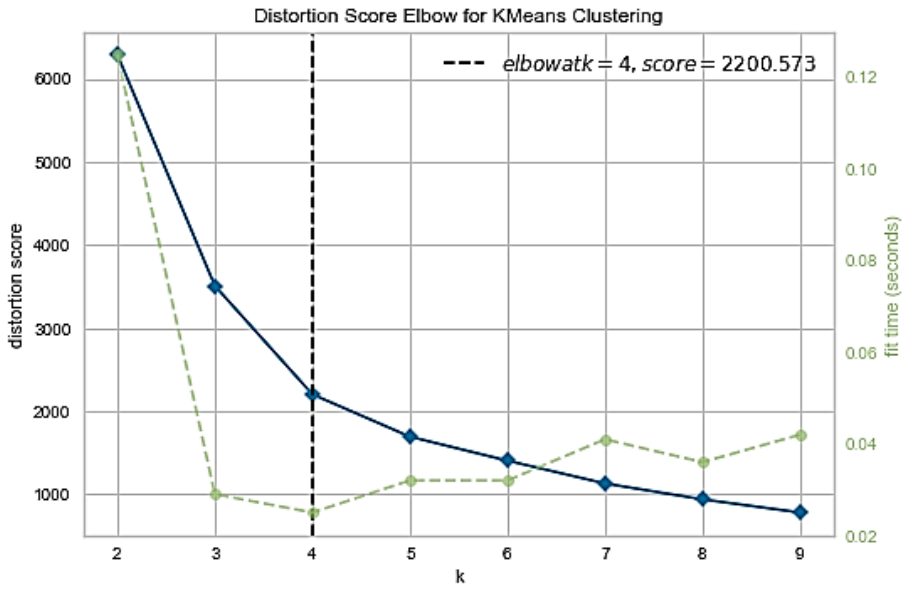
Count the number of venues falling into each of these categories:

Neighborhood	Fitness Class	Gym	Park	Sports Facility
Central Harlem South	6	9	3	5

Repeat for other neighborhoods

Clustering the Neighborhoods

- **Which neighborhoods are most similar in terms of park and fitness access?**
  - Group the neighborhoods into clusters using k-means
- **How many clusters?**
  - Use the elbow method to find the optimal number



Group neighborhoods into 4 clusters

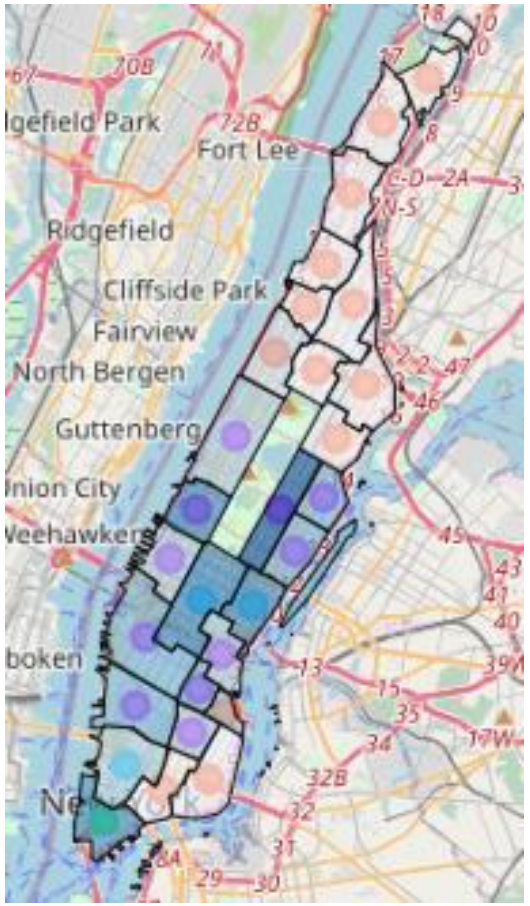
Clustering Results

- **Orange cluster**
  - Clustered around northern Manhattan and the Lower East Side
  - Overlaps with the poorest areas of Manhattan
- **Blue and green clusters**
  - Some of Manhattan's wealthiest neighborhoods

Manhattan Socio-economic Data

Population Density

Income



Poverty Rate

Socio-economic Data by Cluster

Cluster Label	Fitness Class	Gym	Park	Sports Facility	Population Density (people/ sq mi)	Income ($)	Poverty Rate
Orange	2.6	5.9	6.6	6.0	8202	45,241	27.3%
Purple	8.6	26.4	9.5	7.6	9050	106,227	9.9%
Blue	18.0	57.0	15.0	10.0	4728	107,442	9.7%
Green	2.0	92.0	29.0	2.0	6050	135,073	7.4%

- **Orange Cluster**
  - Low income, high poverty rates
  - Low numbers of parks and fitness facilities
- **Blue and Green Clusters**
  - High income, relatively low poverty rates
  - High numbers of parks and fitness facilities

Differences in income and poverty rate correspond to differences in the numbers of parks and fitness facilities

Conclusions & Recommendations

- Low income, high poverty rates → low access to parks and fitness facilities
- Parks and fitness facilities are **NOT** distributed according to population density
- Current distribution of parks and fitness facilities is **inequitable**
  - Low income residents have less access to a healthy lifestyle
- Focusing future investments in parks and rec into low income neighborhoods could help reduce these gaps