

The background features abstract green geometric shapes. On the left, a solid green trapezoid points upwards. On the right, a complex arrangement of overlapping translucent green triangles and polygons creates a layered, crystalline effect. A thin, light gray line extends from the bottom left towards the right, passing through the lower part of the right-side geometric shapes.

Similarity of Cities

Goals

- ▶ The project is trying to establish which locations in the cities are similar to each other.
- ▶ To what extent does that similarity of neighborhoods in the city makes cities similar to one another
- ▶ To find what makes certain neighborhoods similar or dissimilar to others



Data acquisition and preparation

- ▶ The project uses data for three locations
 1. NYC from https://cocl.us/new_york_dataset
 2. Toronto from
 1. https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
 2. https://cocl.us/Geospatial_data
 3. Paris from https://en.wikipedia.org/wiki/Arrondissements_of_Paris and coordinates using geocoders
- ▶ We also use data related to venues from foursquare api.
- ▶ After getting data for different cities they are merged into same dataframe

Feature Selection

- ▶ In the project venues are categorized into 7 primary categories:
 1. Arts and entertainment
 2. Building
 3. Food
 4. Nightlife
 5. Parks and outdoors
 6. Shops
 7. Travel
- ▶ These categories are used to Cluster the neighborhood



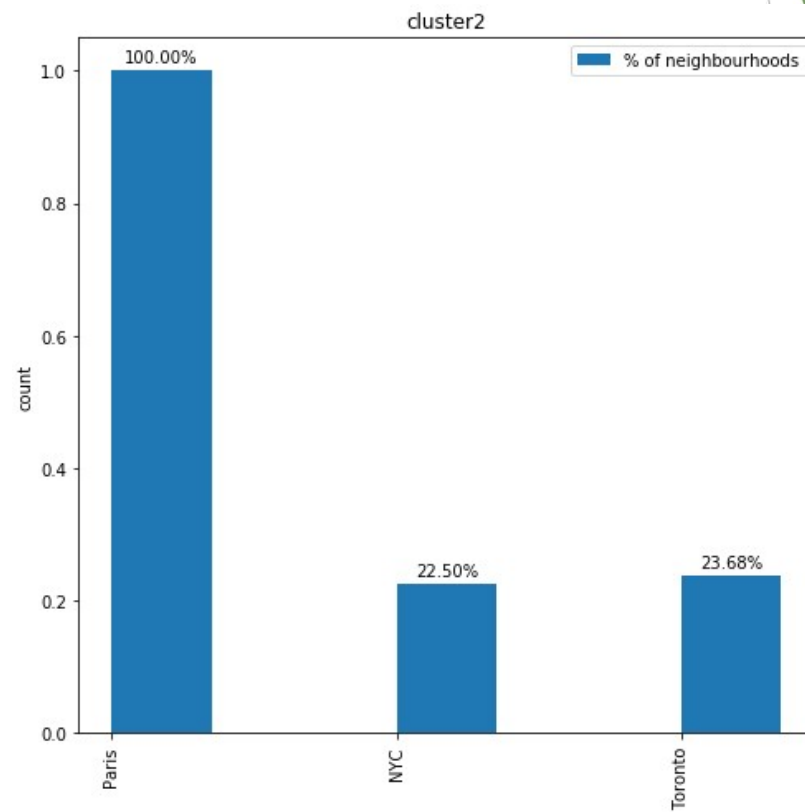
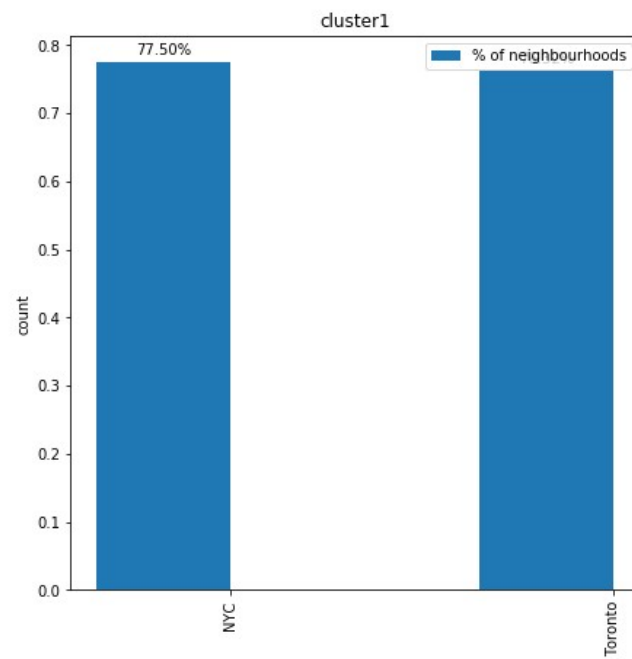
Predictive Model

- ▶ The project uses K- means to cluster the neighborhoods in the cities based on venues present near the locations.
- ▶ Number of clusters created are two.
- ▶ If two cities are similar most of their neighborhoods will lie in same cluster.
- ▶ Striking different properties of clusters indicate very clear difference in the cities.
- ▶ Having cluster analysis will inform about properties of neighborhoods in the cluster
- ▶ If a cities have similar kind of venues then such cities are clustered together.

Result

- ▶ It turns out that most neighborhoods of Toronto and Manhattan(NYC) are similar to each other when compared to neighborhoods of Paris.
- ▶ Most of the venues in cluster 1 which had mostly neighborhoods from Toronto and Manhattan fall under category food about 61% followed by venues related to shopping at around 15%.
- ▶ For cluster 2 most common category was parks and outdoors at 33% followed by food at 26% and shopping related venues at 22% respectively

Results



Results

