This project is to explore the variety of venues available to residents of certain Toronto neighbourhoods.

* **Premise**: There are many people in Toronto, who do not have cars. It is important to them to have a variety of diverse venues close by.
* **Stakeholders**: The residents of Toronto, future business owners
* **Problem**: Not all neighbourhoods have a good variety of venues to access near their homes such as a big variety of restaurants, malls, entertainment.
* **Data**: Venue data from foursquare as well as geographic data from Wiki will outline the neighbourhoods and expose the variety or lack thereof of venues in each area.

Data

* Data was scraped from Wikipedia to determine the postal codes of each neighbourhood in Toronto
* A list of venues in Toronto was obtained through the Foursquare API

Methodology

* Analyze the data to understand top 5 venues in each neighborhood.
* Grouped data per neighborhood
* Analyze results using Seaborn library
* Normalize the data to reduce impact on data spread and keep to scale

Analysis

* Used Machine learning algorithm KMean clustering to cluster all neighborhood areas based on venue category
* Used Folium Python library to create a Toronto map with clustered neighborhoods displayed

Results

K-mean clustering has shown

* + Cluster 3, 4 and 5 popular for cafes
  + Cluster 1 is popular for hotels
  + Cluster 2 is popular for a unique venue

Outcomes

* Report shows most popular venues in the City of Toronto
* Useful to attract new business to the area that are not currently in existence
* Residents and tourists can see which venues are close by
* Potential for expanded analysis