

## **The Battle of the Neighborhoods**

- **Data Section**

- » Toronto Dataset

The dataset for Toronto is scrapped from Wikipedia page and it is stored in the dataframe. The dataset contains three columns namely postcode, borough and neighborhood. It has 288 rows with raw data which is refined by removing unassigned borough which leaves only 102 rows for further processing.

- » New York Dataset

New York dataset is taken from New York University Website. This dataset has a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood.