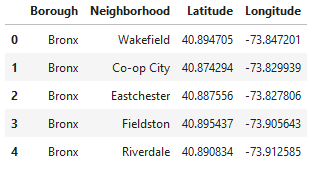
**Safeguarding the drunk New York City**

**Merrymel George**

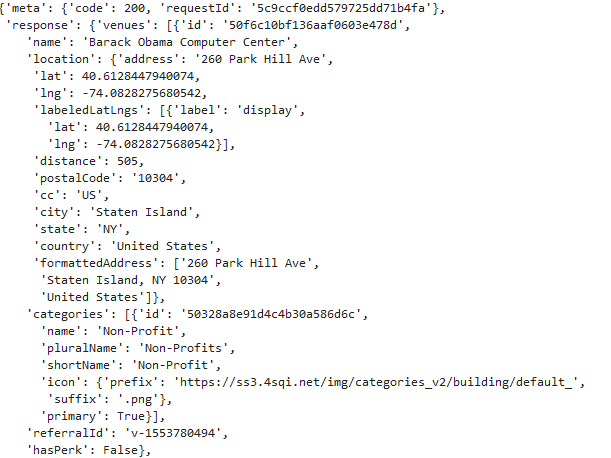
**March 28, 2019**

1. **Data Acquisition**
   1. **Data sources**

The data that is required for doing the project is acquired from two sources. First dataset, which gives the required data related New York, the different boroughs, neighborhoods, latitudes and longitudes are acquired from a data source which is freely available - <https://cocl.us/new_york_dataset> . A set of sample records of the acquired data is shown below:



Then, with the help of Foursquare API, the other dataset, which is data of the pubs, bars and nightclubs in New York are acquired. Required Foursquare credentials are created and is used to pull data with details like name of the pub/bar/nightclub, location details etc for each neighborhood of the boroughs of New York. A sample of the data that is pulled using the Foursquare API is shown below.



This data needs to be cleansed, wrangled and munged to the required data like pub/bar/nightclub name, location, latitude, longitude etc can be extracted. A sample set of extracted records are shown below.

