# THE BATTLE OF NEIGHBORHOODS

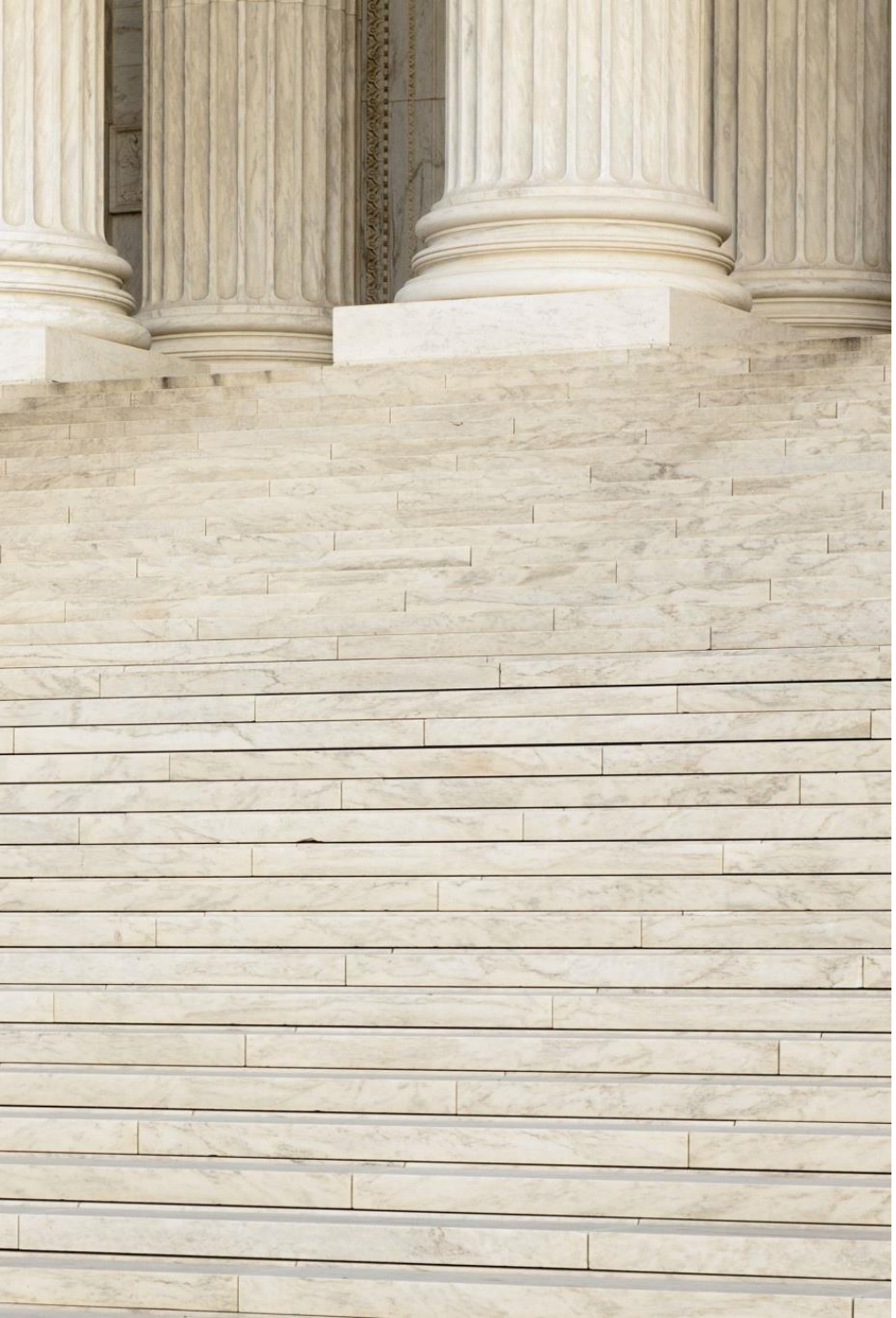
FINAL PRESENTATION FOR

APPLIED DATA SCIENCE CAPSTONE PROJECT

By

CORELLETA KASEKE

# Table of Contents



# Introduction

- Confectionery Limited is a Multi-National Company that specializes in confectionery products such as assorted ice-creams, chocolates, cookies and candies.
- The company has succeeded in the various locations that it is currently operating and wishes to make great returns from a new shop in Manhattan, New York, USA.

# Business Problem

- The company has been presented with two investment opportunities to open a new shop in one of two locations.
- As a result, the company must choose the location that offers a possibility of greater revenue and thus profits.
- Due to the nature of its products, the company is looking to locate near schools as young people are their greatest customers.
- The aim of this study will therefore be to find the best area to open this new venture for Confectionery Ltd based on the most valued attributes.

# Data

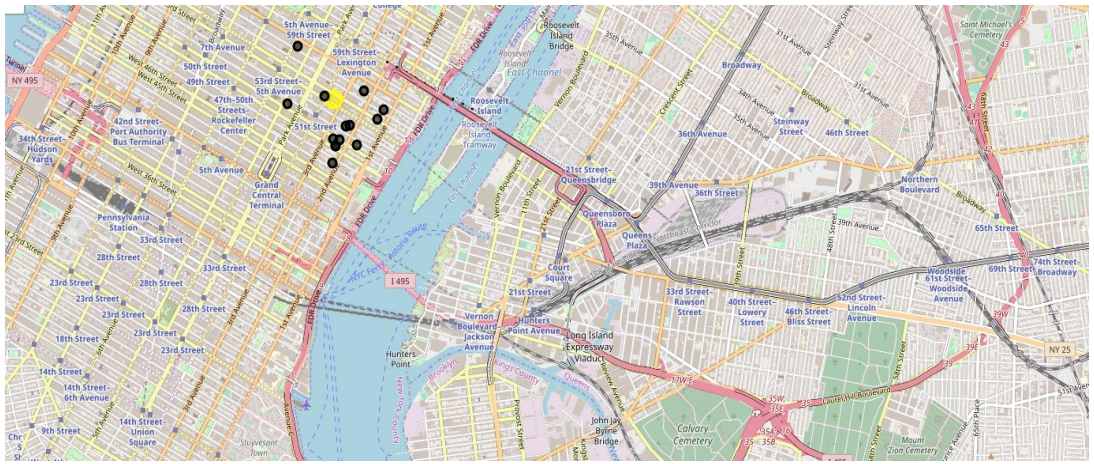
- The analysis will make use of Foursquare data together with geo-data of New York, Manhattan downloaded here: https://cocl.us/new_york_dataset
- The variable that was mostly of interest was the **availability of customers** in the two locations provided. This was mostly **the number of schools** that were in the vicinity. However, the study also had to include the **number of trending venues** so as to filter down the best location from the two.
- Foursquare API was used to get data for number of schools and trending venues in the vicinity.


# Methodology

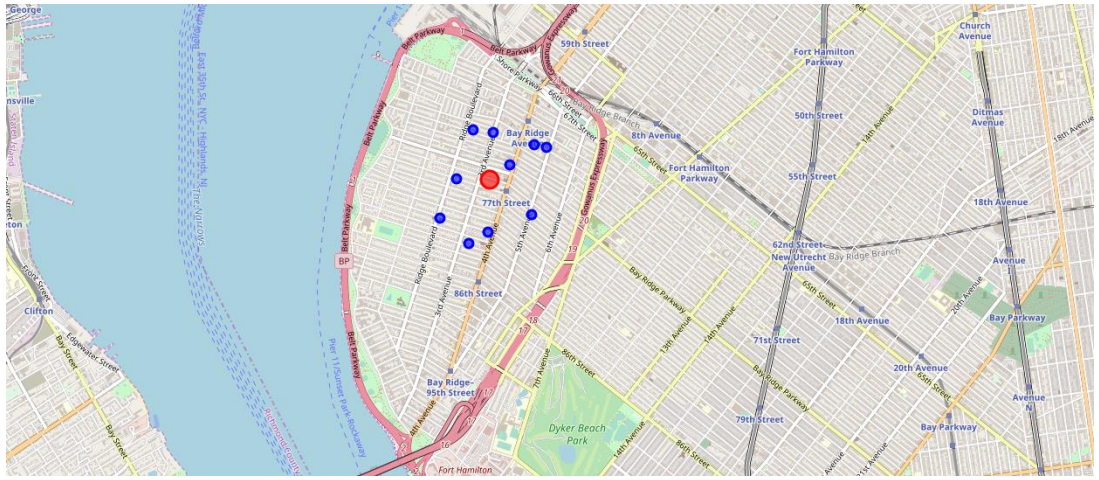
- First step included downloading the json data of New York.
- The data was converted to a pandas dataframe.
- The data had to be filtered to represent only the data for Manhattan as the locations are in Manhattan.
- Format and normalize the data.
- Geolocator-Nominatim was then used to find the geo-coordinates of the two locations.
- Foursquare was then used to obtain various data pertaining to schools in the vicinity of the two locations.
- Foursquare was then used to obtain various data pertaining to trending venues in the vicinity of the two locations.
- Represent the data visually with maps using Folium.

# Result 1: Schools nearby

**Location 2 and schools nearby**

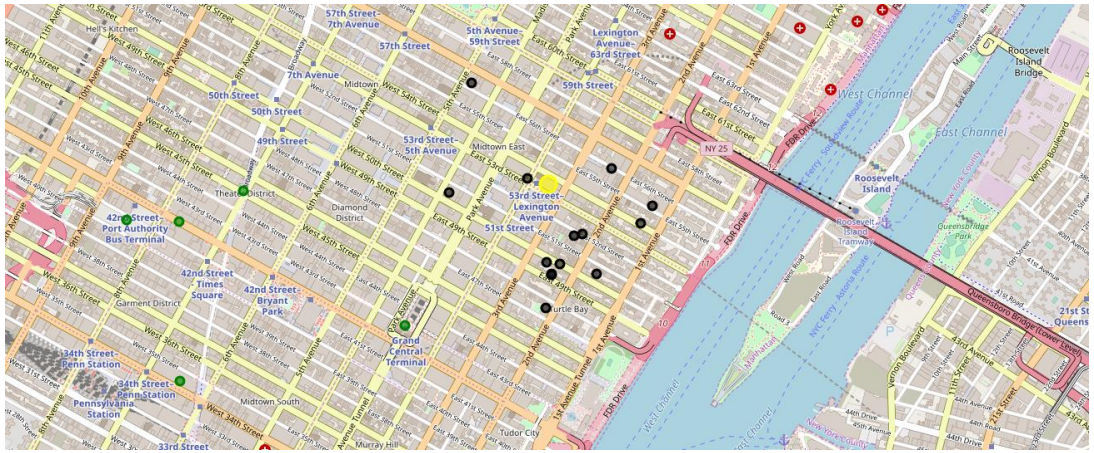



**Location 1 and schools nearby**

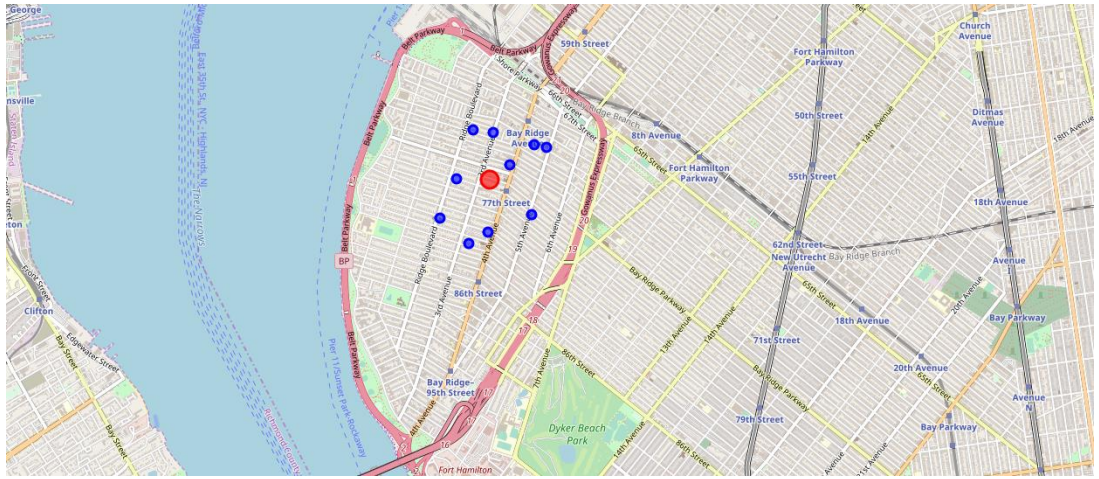


# Results 2: Schools and Trending venues nearby

**Location 2: Schools and Trending venues**




**Location 1: Schools and Trending venues nearby-no trending venues**



# QUESTIONS??

Thank you!!!!!!!!!!