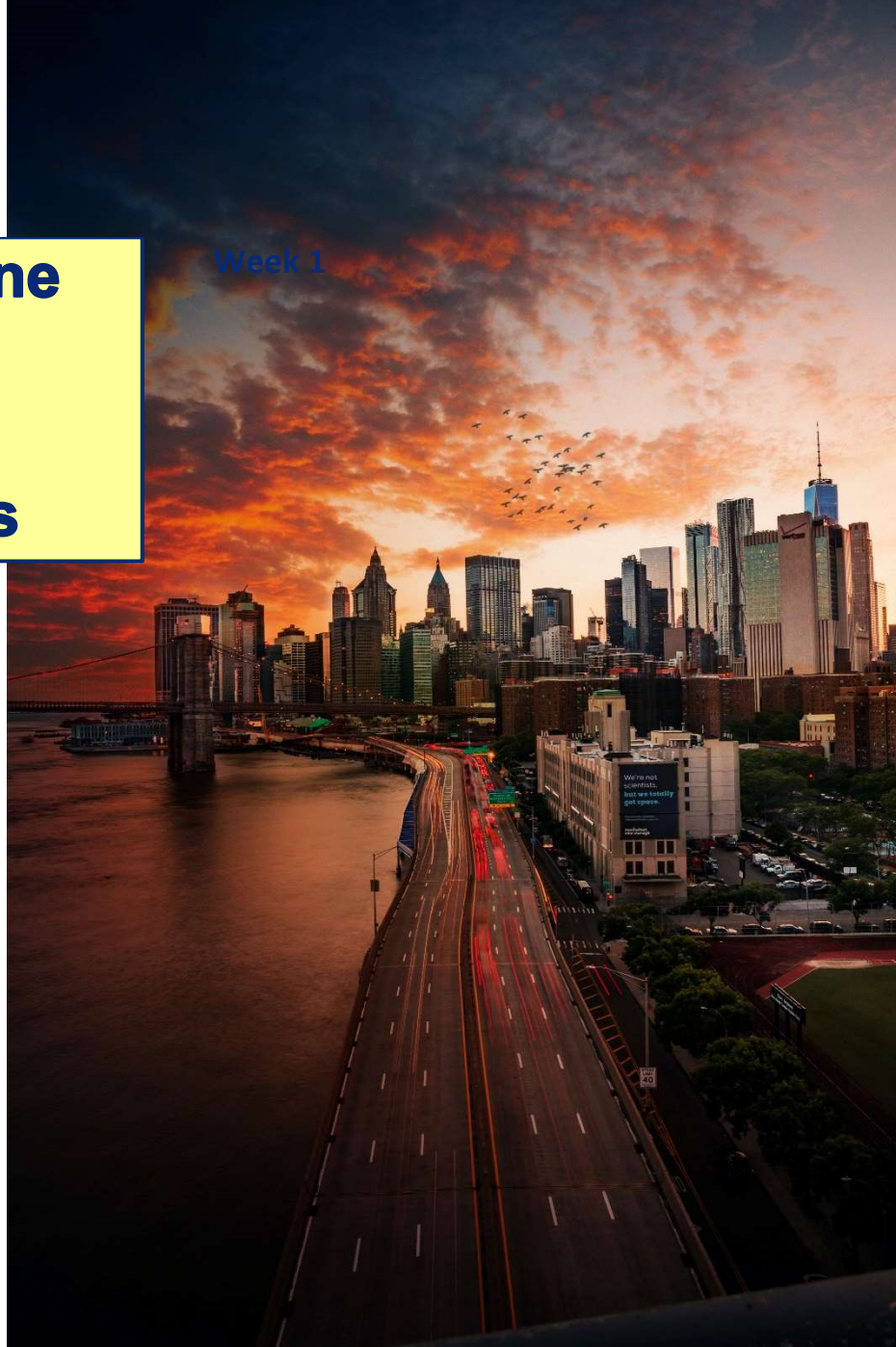


Coursera Capstone Project

NYC Restaurant Location Analysis

Week 1



JULY 4, 2020

Authored by: Michael Iannacci

Introduction/Business Problem

Background

A group of investors is looking to invest in a new restaurant concept and select a location in New York City to establish their first location. It is vital to the investment group that this location is successful and acts as a model location for future locations and potential franchises.

Given that New York City is the most populous city in the United States with an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles, it would be the ideal location to establish the first 10 locations over time before expanding throughout the region.

A member of the group has some general awareness of data science and knows that it is a powerful tool that can be used to help his group make the best decision possible for the location of the initial launch. He has convinced his team to hire me as a data scientist to run an analysis to guide their decision making process.

The type of restaurant that the group wants to create has a niche target of active people who are trying to eat healthy. Their concept combines both a grab-and-go section as well as a sit down section. Their primary focus will be health food and smoothies that will appeal to the active members of the community. They will also offer some traditional fare in order to have options to capture a wider customer base. The overall strategy is to strategically locate their restaurant around venues that their target consumer is most likely to frequent such as gyms, pools, yoga studios, etc.

After initial brainstorming session with members of the investment group, we developed a list of comprehensive venues to target in a given locale. A restaurant location strategically placed near as many of these venues as possible will maximize the foot traffic to the restaurant. Because the success of the restaurant depends on this

foot traffic, I recommended that we also take a look at the crime rate and take that into consideration when the ultimate location is selected.

Business Problem

After a second meeting, the requirements of the analysis were finalized:

- 1) Location will be in one of the 5 boroughs of New York City
- 2) Location will be centrally located near as many of the identified target athletic/fitness venues as possible
- 3) Location should minimize the proximity to existing competitors
- 4) Location should not be in a high crime area in order to promote foot traffic
 - After further discussion, it was agreed that only felonies in the latest calendar year of available data will be considered.

Data

The following data will be used for this analysis:

1. The city of New York will be analyzed for this study. New York City is comprised of 5 boroughs: Manhattan, Brooklyn, Queens, the Bronx and Staten Island. There are a total of 306 neighborhoods in these 5 boroughs. In order to analyze these neighborhoods, the free dataset at https://geo.nyu.edu/catalog/nyu_2451_34572 will be used. This dataset contains information, including latitude and longitude coordinates, which will allow for the boroughs and neighborhoods to be segmented and explored.

-
2. New York City geographical coordinates data will be used as input for Foursquare API. This will allow for venues within each neighborhood to be explored and classified.
 3. New York City crime data will be used from Kaggle at <https://www.kaggle.com/adamschroeder/crimes-new-york-city> . Only data for felony crimes over the full 2015 calendar year was extracted and will be used to give an overall picture of high and low crime areas.
 4. Wikipedia will be used for miscellaneous information and data https://en.wikipedia.org/wiki/New_York_City