The battle of Neighborhoods

1. Introduction & Business Problem

1.1. Background

The third-party delivery industry has been growing exponentially in recent years due to technology adoption. Nowadays, this service industry is based on mobile apps in which a client orders a product from the business listed and a delivery man picks up the items from a store and transports it to the client's address while receiving part of the fee charged for the service.

1.2. Problem

Orders are mostly accepted by the delivery man closest to the business selected by the client, because this means he has to invest the least amount of resources to get to the item, hence, he keeps a bigger profit from the fee. In order to have a substantial income from this activity, the delivery man has to carefully select the area in which he should be located to accept more orders without having to spend too many resources to get to the selected items. This project will propose which areas and times could be the most profitable for a delivery man based on the venue density and top visit hours, because it's inferred that these are also the top hours for take out orders, all this for venues in Manhattan.

2. Data

Only Manhattan will be analyzed in this project. We will analyze its neighborhoods' venue density by using the coordinates of each neighborhood, available here https://cocl.us/new_york_dataset, and use them as input to request the venues information available through Foursquare API. The data will be cleansed and wrangled in order to provide a dataframe that will be used for segmentation and clustering of the neighborhoods. The clusters will show which neighborhoods have more food business which will indicate a higher opportunity for the delivery man.

The neighborhood dataframe would look like this:

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

The dataframe merged with Foursquare API's info would look like this:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Dunkin'	40.877136	-73.906666	Donut Shop
4	Marble Hill	40.876551	-73.91066	Starbucks	40.877531	-73.905582	Coffee Shop

A map would show the clusters labeled by how much restaurant density the neighborhood has: