Analyzing the Correlation between NYPD Complaints Location and Nearby Venues

Daniel Ramos
September 29, 2019

NYPD Criminal Complaints

The city of New York is one of the most popular cities and one iconic cultural and financial center in the United States. It is one of the densest city in the world with a population of around 20 millions of habitants.

Bigger cities have bigger problems, security is one of them. The New York Police Department (NYPD) have around 95,883 complaints reported in 2018.

Data acquisition and cleaning

The main dataset used in this project is made available publicly by the New York Police Department.

https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Current-Year-To-Date-/5uac-w243/data

In this dataset are all the complaints reported to the Police Department in the current year and only in the borough of Manhattan.

The API service provided by Foursquare is also used to get info about the venues around.

Time of day and location of the complaints

00:00 to 07:59

08:00 to 15:59

16:00 to 23:59

Most popular venues categories nearby the complaints

From 100 most popular venues around the location of most of the complaints i obtained 58 unique venues types.

This is a high variation in the venues types. This difficult make conclusions about the influence of certain venues types to the occurrence of the event

	Venue Category	Venue
32	Italian Restaurant	5
28	Gym	5
53	Theater	5
43	Pizza Place	5
14	Coffee Shop	5
17	Cuban Restaurant	3
18	Dance Studio	3
5	Bakery	3
6	Bar	3
50	Taco Place	3

Clustering complaints by location

This is the clusters obtained using KMeans with the Latitude and Longitude data points.

This helps segmenting the borough in 10 unique groups to further analysis
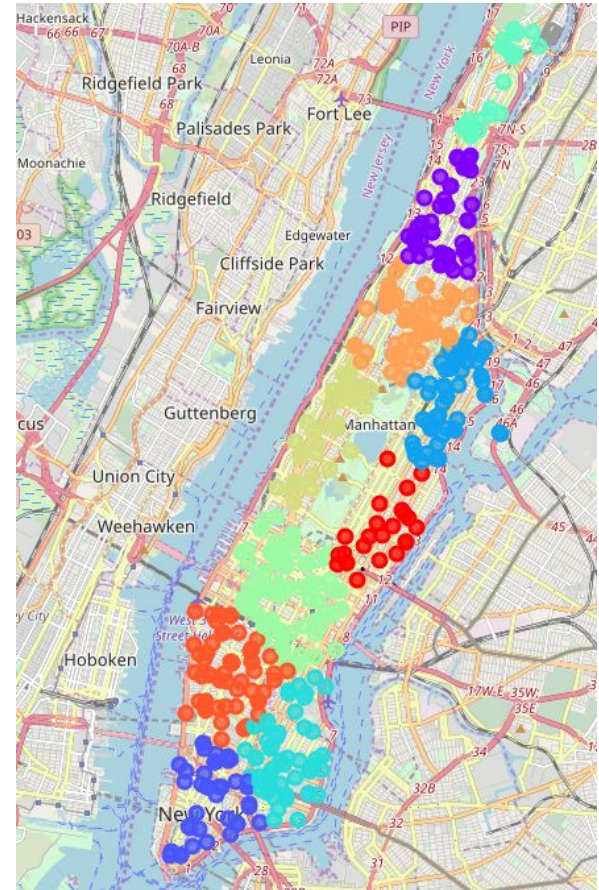
Rate of occurrence by cluster

With the clustered data point i can determine what zone is risky using the incident rate of each cluster and referencing the point using the centroid of the cluster latitude and longitude

	Incidents_Rate	Cluster	Longitude	Latitude
0	0.109279	9	-73.997819	40.739274
1	0.083239	1	-73.942302	40.829724
2	0.070389	7	-73.974556	40.786797
3	0.193768	6	-73.984566	40.756282
4	0.108597	8	-73.951042	40.810280
5	0.102456	4	-73.984331	40.725526
6	0.119627	3	-73.940346	40.795409
7	0.057426	5	-73.929363	40.856558
8	0.091199	2	-74.004067	40.714472
9	0.064021	0	-73.958713	40.770263

Thanks for Watching