CAPSTONE PROJECT : SAFETY-RISK ANALYSIS

Created by : Vito Anthoni Stanpo / 14th February 2020

1.Introduction

1.1.Background

Most of the times in our daily lives,we should take precaution for something that's unexpected. As an old saying said, we should expect the unexpected. In a metropolitan city, whether you want to open a new business or running your current one, precaution for safety should be considered. You can optimize the budget you should be spending. Using data science we can get the most vulnerable neighbourhood to crime compared to the other location. We can choose place far from the crowd to gain the maximum safety compared to other places. On the other hand, some business especially like restaurants should not get too far from the crowd in order not to be forgotten. This project hopefully helps to find the optimum points by finding the least vulnerable to crimes neighbourhood yet do not stayed too far from the crowds. For this project we are focusing representing crowd as restaurants competitor as we conciously assumed we want to open restaurant business

1.2.Target Audience

The main objective of this project is analysing which neighbourhood is the most vulnerable to the crime rate in San Francisco. With data science methods and tools such as clustering we can determine how much precaution should be taken in terms of security investments as relativity so that business owner do not overcommit for security investments.

2.Methodology

2.1.Data Requirement

To solve this problem, we will need below data:

- List of neighborhoods in San Francisco
- Latitude and Longitude of these neighborhoods
- Venue data related to Crime In Toronto,Canada

2.2.Data Usage

- The scrapping of Toronto neighborhoods via Wikipedia
    - Used for getting the dataframe for showing the first components for the projects and as the base for the maps
- Getting Latitude and Longitude data of these neighborhoods via Geocoder package
    - Used together for data scrapping from Wikipedia to draw the boundary in the maps
- Using Foursquare API to get venue data related to these neighborhoods
    - Getting the list of venue and that adding it into choroleth data for crimes in Toronto maps
- Using CSV files for crime rate
    - Used for creating chroropleth maps for crimes in Toronto