Capstone Project - The Battle of Neighborhoods (Week 2)

Introduction: Business Problem

The main objective of this project is to determine the safest place in terms of the criminal situation for opening a business in Denver, USA.

First, needed to choose the safest borough with a sufficient list of neighborhoods for business development by analysing crime data. We will use the studied course material and such standard libraries as pandas, numpy, matplotlib, etc. We will focus on finding the safest borough and its neighbourhoods, where few grocery stores.

Data

Based on the problem, we must: Find the safest area with a enough list of neighborhoods Identify its most common places

We will use the geographical coordinates of the city of Denver and the coordinates of the crimes recorded from 1st January 2015 to 17 July 2020.

Following data sources will be needed to extract/generate the required information:

https://www.kaggle.com/paultimothymooney/denver-crime-data dataset include type, date and location of crime in Denver.

https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Denver There we have a neighbourhoods and which borough they include.

Also, we will needed to create a new dataset of the Neighborhoods, boroughs, and the most common venues and the respective Neighbourhood along with coordinates. This data will be fetched using Foursquare API to explore the

neighbourhood venues and to apply machine learning algorithm to cluster the neighbourhoods and present the findings by plotting it on maps using Folium.

Methodology

In this project we will direct our efforts on detecting areas of Denver that have low criminal level, particularly those with low number of grocery stores.

In first step we have collected the required data. We have identify boroughs with highest/lowest crime cases and neighbourhoods.

Second, look for a type of crime and choose that have directly related with commercial/business sector, then choose boroughs where crimes against business is less common.

In third and final step we will focus on most promising areas and within those create clusters of locations, that are most fit for opening business there.

We will present map of all such locations but also create clusters (using k-means clustering) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

Analysis

Let's perform some basic exploratory data analysis and derive some additional info from our raw data. First let's count the number of all types of crime:

Results and Discussion

Our analysis shows that Northwest a good borough that include good neighbourhoods with different venue. It's a very good borough with low crime level and different types and various leisure options. Also that places very close to center of Denver, and have well space for commercial business.

Conclusion

Purpose of this project was to identify Denver areas close to center with low crime level for opening there business in order to aid stakeholders in narrowing down the search for optimal location. By calculating most common venue in Northwest from Foursquare data we have first identified types of commercial places, and then generated extensive collection of locations which satisfy some basic requirements. Clustering of those locations was then performed in order to create major zones of interest and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.