Live in Boston

'The battle of Neighborhoods'

IBM Apllied Data Science CAPSTONE

Claudia Cerasale

05/11/2020

1. Introduction:

How safe is Boston and which is the most liveable neighborhood? First let's start with some considerations.

From the polls, the 27% of surveyed as of end of 2018 said they considered Boston to fairly or very unsafe. According to recent surveys the city's low murder rate per capita (9 per 100.000) could likely be a contributing factor to this perception. The city also has the second-lowest incidence of vehicle theft at a rate of 258 per 100.000 and fares well on the other crimes rates.

Is it true to say that Boston was more violent than New York and Seattle, but less violent than Chicago and Las Vegas, according to numbers from the FBI, based on crimes committed back in 2015. In 12/21/18 nationally, Boston ranked 14 out of 50 according to US News. After digging up and analysing the data what we found out that in recent numbers from the Boston Police Department, or BPD, show that violent crime, as well as property crime, has continued to drop, and has been steadily dropping for past years.

1.1 Business Understanding:

Public safety is vital to public health and happiness and the safety of a state can be a crucial factor in deciding where to move with the family, or more simply for those who need to move to another city for study purposes or work, but more generally it can also be useful for a tourist to know which is the safest neighborhood where to stay. In fact, going around the various forums, it can easily be seen that the questions are often addressed to the safety of the neighborhood as well as which interesting places to visit. So the project aims to find the safest and most comfortable neighborhoods in Boston, both through the analysis of the crime rate in the various neighborhoods and through the evaluation of the best served areas for traveling, for example, or for buying food, college presence etc. .

Therefore we will select the safest district of Boston based on the total of crimes, we will explore the various neighborhoods to find the 10 most common venues in

each neighborhood and finally we will group the neighborhoods using k-mean clustering, identifying the most suitable neighborhood for the needs of the individual.

Crime statistics will provide an overview of this problem.

Having outlined our problem, the factors, therefore, that will influence our decision are:

1. The total number of crimes committed in each district in a year (2017)
2.The most common venues in each selected safest neighborhood.

To extract / generate the requested information the following data sources will be required:

- Preprocessing of a Boston Police Department (BPD) dataset showing Boston crimes from 2016 to 2018.
- Conversion of District Codes into the names of their neighborhoods, with web scraping (Beatiful Soup Libraries)
- Distinction between crimes and accidents, and consequent filtering of the types of crimes
- Creation of a new dataset of Boston's safest neighborhoods and generation of their coordinates: the neighborhood coordinates will be developed using the geocoding of the Google Maps API
- Foursquare to identify the most common venues
- Clustering (K-Means)