

COURSERA CAPSTONE

IBM Applied Data Science Capstone

CLUSTERING BOSTON CITY WITH CRIME DATA

By Manasa Devi Chakka

INTRODUCTION

Boston is a global city and placed among the top 30 most economically powerful cities in the world. Boston's colleges and universities exert a significant impact on the regional economy. Boston attracts more than 350,000 college students from around the world, who contribute more than US\$4.8 billion annually to the city's economy. The area's schools are major employers and attract industries to the city and surrounding region. The city is home to a number of technology companies and is a hub for biotechnology, with the Milken University rating Boston as the top life sciences cluster in the country. As with all the major cities crime rate is also high in Boston city which might have been attracted by its rich economy.

BUISINESS PROBLEM

Can a solution be found that best describes the area along with nearby venues where it is safe for new immigrant can find accommodation?

TARGET AUDIENCE

As mentioned above with all the resources like colleges, universities, schools and technology companies established it is relevant that all the employers , students and immigrants who have to relocate might want to choose a safe neighbourhood for accommodation. Our goal here is to provide such people with info about areas and respective crime data.

DATA

To solve our business problem we need the following data

- Boston crime data with category of crime, crime name and location of occurrence
- Venues data around the location of crime

SOURCES TO FIND DATA

Boston crime data can be found from the Kaggle link (<https://www.kaggle.com/ankkur13/boston-crime-data>). This is a dataset containing records from the new crime incident report system, which includes a reduced set of fields focused on capturing the type of incident as well as when and where it occurred. This dataset has more than 2 lakh rows and 17 columns.

The venue data around required location can be found using foursquare API. Using explore end point we can achieve this. Foursquare is one of the most used API to obtain various kinds of details of the specific location. Explore endpoint is a regular call and a we can make enough calls with free personal developer account.