

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

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

For this project we need the following data

- Boston crime data with location of occurrence
- Venue data of respective location

Boston crime data can be obtained from Kaggle link as given below.

<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 2,60,760 rows and 17 columns.

Venue data can be find using Foursquare API with explore endpoint.