**Capstone Project**

**IBM Data Science Professional Certificate**

**By: Ajeo Mathew Cherukathara Mathew**

**INTRODUCTION**

Crimes happen ubiquitously no matter how large or educated the people are. Crimes committed across the world are increasing every day and it is difficult to tackle this problem. Traditional approaches like being cautious and improving the safety can’t guarantee us in reducing the crime rates. It is very difficult to understand the patterns of crimes but with the help of predictive analysis in machine learning, we can take necessary measures in order to prevent the crimes.

In a recent report by the Los Angeles Times, we could see that the L.A. homicides are down again and police credit thousands of extra patrol hours. Assuming that police can more often patrol a district assigned to them or the area around their police stations, the business question that I’m trying to answer is “Does the presence of constant police patrolling impact the number of homicides within the police district?”. In this analytics project, we visualize homicides occurred in the city of Chicago, Illinois along with the police districts and the police stations in the city and try to answer the business problem. This will help the police department in the city to identify areas of lesser presence and increase the patrolling in the city in reducing crimes.

**DATA SCOURCE**

The data for this analysis project is taken from multiple sources.

* Chicago Crimes\_2019 Dataset accessed from the Chicago Data Portal. This dataset is an open dataset which has records of crimes which happened in the year 2019 within Chicago. The dataset is extracted from the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) system. We will be more interested in the Latitude, Longitude, and District fields of the dataset where the Primary Type field is Homicide.
* For studying the different police districts along with their geographic boundaries, we use the Boundaries - Police Districts (current) dataset from the Chicago Data Portal. The data will be a ‘.geojson’ file type.
* Finally to access information about locations of the different police stations in the city of Chicago, we use the FourSquare API.

**REFERENCES**

* <https://www.latimes.com/california/story/2020-01-16/l-a-homicides-are-down-again-police-credit-thousands-of-extra-patrol-hours> , report on the homicide rate in Los Angeles
* <https://data.cityofchicago.org/Public-Safety/Boundaries-Police-Districts-current-/fthy-xz3r> , for the police district location data
* <https://data.cityofchicago.org/Public-Safety/Crimes-2019/w98m-zvie> , for the crimes 2019 dataset