**DAB 322 CAPSTONE PROJECT 1- 911 Calls for Service 2021**

**Section-A**

**Prepared For**

**Prof. Savita Seharawat**

**Prepared By**

**Navdeep Singh (0817650)**

**Nitika (0820943)**

**Navjot Kaur (0818071)**

**Project Abstract**

"911 Calls for Service 2021" is the capstone project. To increase safety, the 911 Calls Capstone Project examined statistics on the frequency of emergency 911 calls. Python with NumPy, Panda's data analysis libraries, together with Matplotlib and seaborn data visualization tools are used to construct this project on a Jupiter notebook. Advanced methods, including machine learning, are used to analyze the gathered data once it has been recorded in a central database. In this research, exploratory data analysis and data visualization are used to determine the frequency of 911 emergency calls throughout the year as well as the most frequent causes and occurrences that lead to 911 calls and when they typically occur.

**Problem Understanding:**

* **Priority Based Approach**: There has been a significant number of priority-based calls in the given dataset, which helps us to understand the need for help.
* **Incident Location**: With a population of 576498 the high population density in Baltimore contributes to increased incidents, increasing the emergency factors.
* **Timing patterns:** 911 Calls for Service 2021 pose a significant detailed concern for incident timings in Baltimore, leading to location, and other different factors.

**Dataset:**

<https://catalog.data.gov/dataset/911-calls-for-service-2021-f0a4e>

The dataset used for the capstone project "911 Calls for Service 2021" is obtained from a data catalogue that specifically focuses on emergency calls data in the Baltimore City (USA). The dataset encompasses information about incidents occurring on several locations. The data is sourced from various law enforcement agencies such as Baltimore City police department,

The dataset provides detailed statistics about each 911 emergency call, including information about the location and incident type. It serves as a valuable resource for analyzing and understanding the characteristics of accidents and priority patterns.

The dataset was originally published in the year 2021, and has been periodically updated, with the latest modification made on June 30, 2022. It contains a total of 1048576 samples or instances and comprises 20 different features or variables.

Given the large number of observations, data cleaning and preprocessing steps will be necessary to handle missing values and remove irrelevant columns. The focus will then shift towards selecting specific factors that are likely to have a significant impact on incidents. These factors may include Incident type Priority, Community Statistical Areas, Location.

**Questions**

1. How many incidents can occur in future years?
2. What are the major reasons for emergency calls?
3. Are there any significant patterns for these emergency calls?
4. At what time do most of the incidents occur?
5. Predicting the severity of incidents based on different factors?

**Keywords**

Classification.

**Techniques & Theme**

As our data is categorical, we must use a classification model in order to analyze the intensity of road rashes. We are going to implement these two models:

* K-nearest neighbor
* Support Vector Machine (SVM)

Following are the tools which we will use to perform these tasks:

* Python
* For visualization Tableau or Power Bi or MS Excel

**References**

* <https://www.studypool.com/documents/15561890/911-calls-capstone-project->
* <https://data.baltimorecity.gov/datasets/baltimore::911-calls-for-service-2021-/about>