DATASETS

1. **US Police Shootings from 2015- Sep 2022**

This datasets contains a lists of people killed by law enforcement in the United States, both on duty and off duty. Also includes the circumstances under which these victims were shot. The data contains the follwoing 17 columns:

Name, id, date, mannerofdeath, armed, age, gender, race, city, state, signsofmentalhealth, threat\_level, flee, body cam, longitude, latitude, isgeocording exact.

DISCUSSION CONCLUSION: we concluded that the dataset is large enough to lead to an interesting analysis using machine learning tools and it also contains variable number of data types. The main subject and issue captured by this dataset have been the bone of contention in the current political climate. Hence, we decided the study of this dataset will lead to interesting discoveries.

2. **House Rent Prices of Metropolitan Cities in India**

This is a dataset containing Rental Prices of Approx 2L houses of Metropolitan cities in India to understand their key factors behind their cost. It is made up of 9 columns. The dataset from kaggle appears to have already been cleaned so that made it less preferable to the dataset about police shootings.

3. **Credit Card Approval Prediction**

The datasets contains information about bank customers’ credit card approvals. Despite it containing 18 columns, the data types contained in the dataset is mostly Boolean. The group decided the dataset doesn’t offer much data type variety to warrant being chosen as our main dataset for the project.