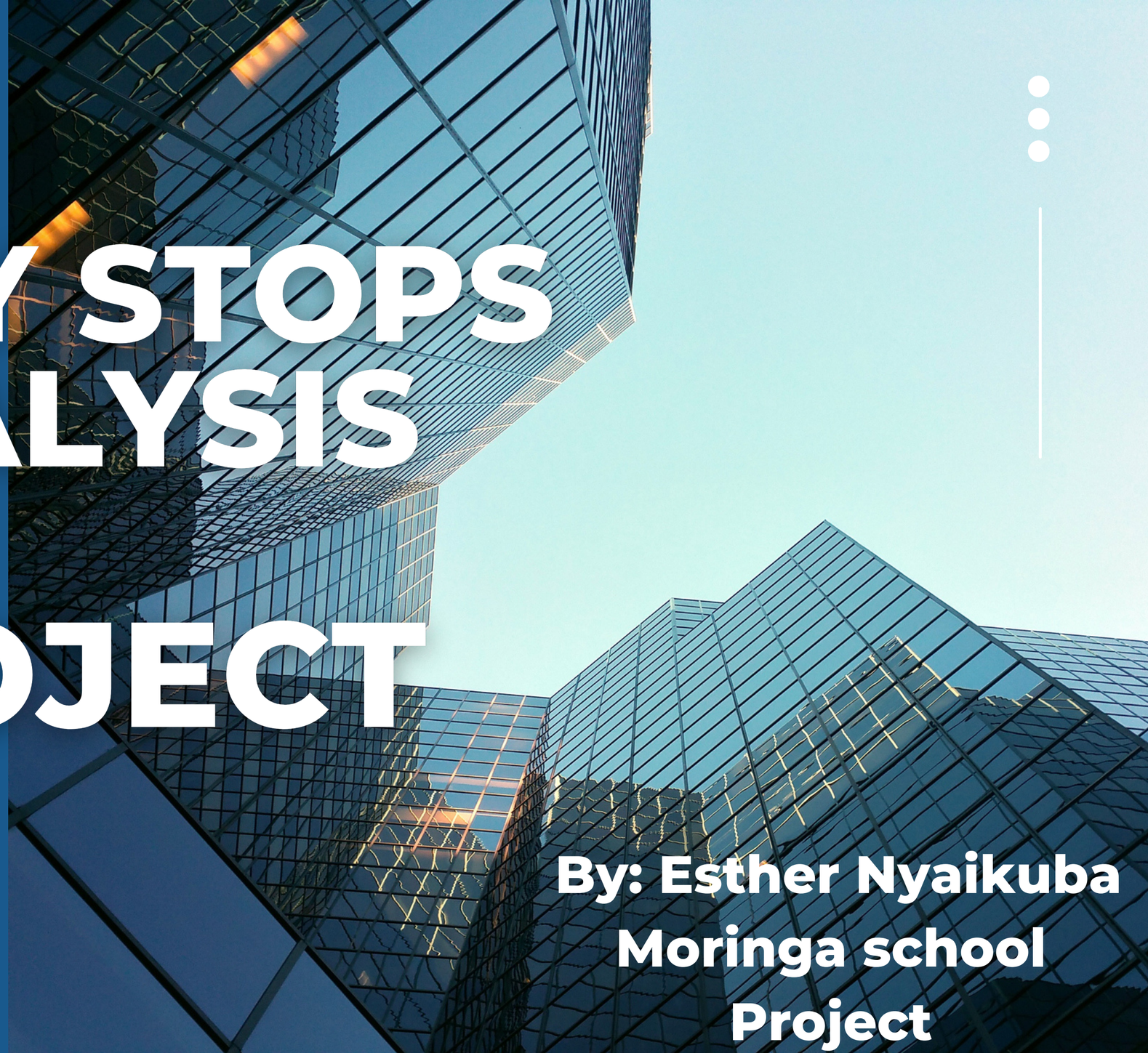
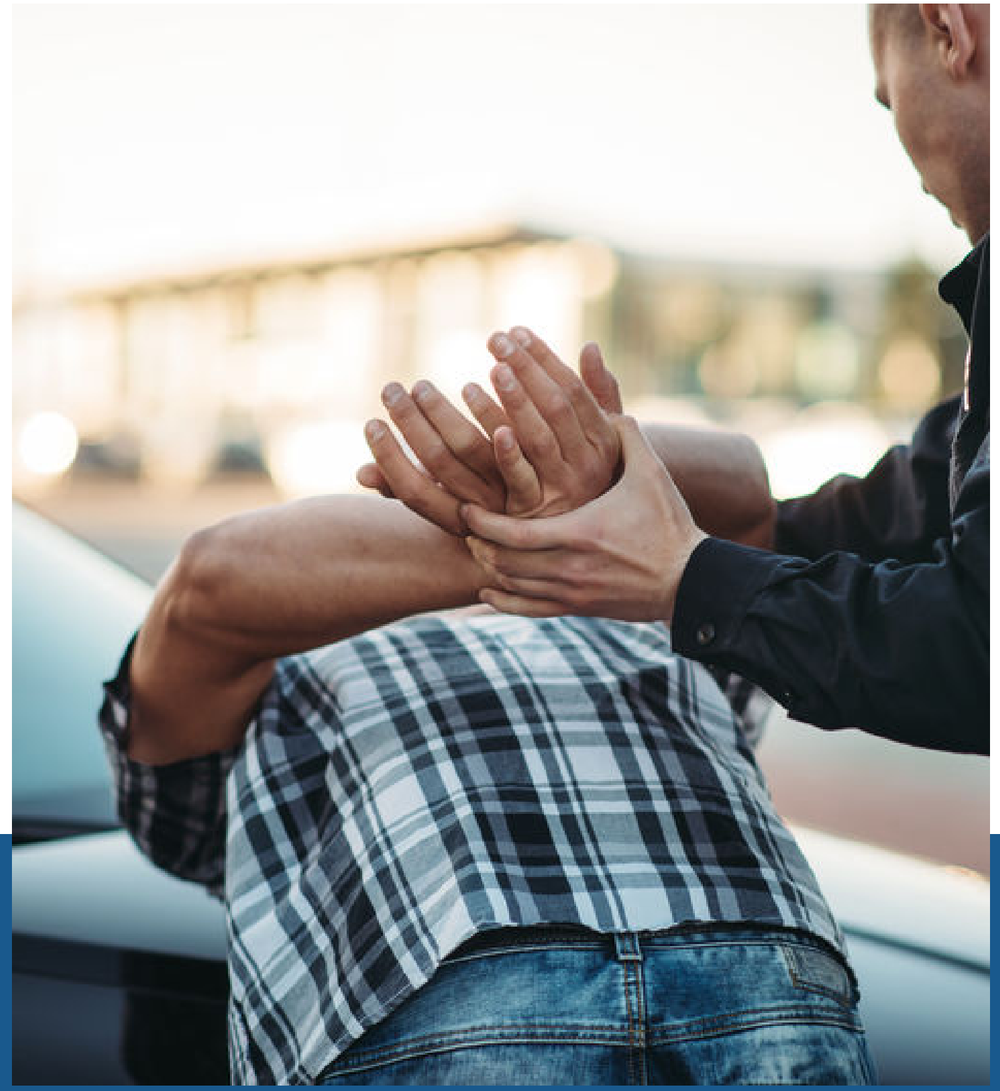
TERRY STOPS ANALYSIS

PROJECT



By: Esther Nyaikuba
Moringa school
Project



Project Overview

A Terry Stop in the United States allows a police officer to detain and question an individual who they have reasonable suspicion has committed or is about to commit a crime. If the officer has reasonable suspicion the person is armed, they may stop and frisk to search for a weapon.

Business Understanding

The aim of this project is to develop a binary classification model that predicts whether an arrest will occur after a Terry Stop. This analysis aims to provide insights into the factors influencing stop outcomes and explore the potential impact of features such as the presence of weapons, time of day, officer characteristics (gender, race), and subject characteristics (perceived race, perceived gender)

Data Understanding

- This data represents records of police reported stops under Terry v. Ohio(1968)

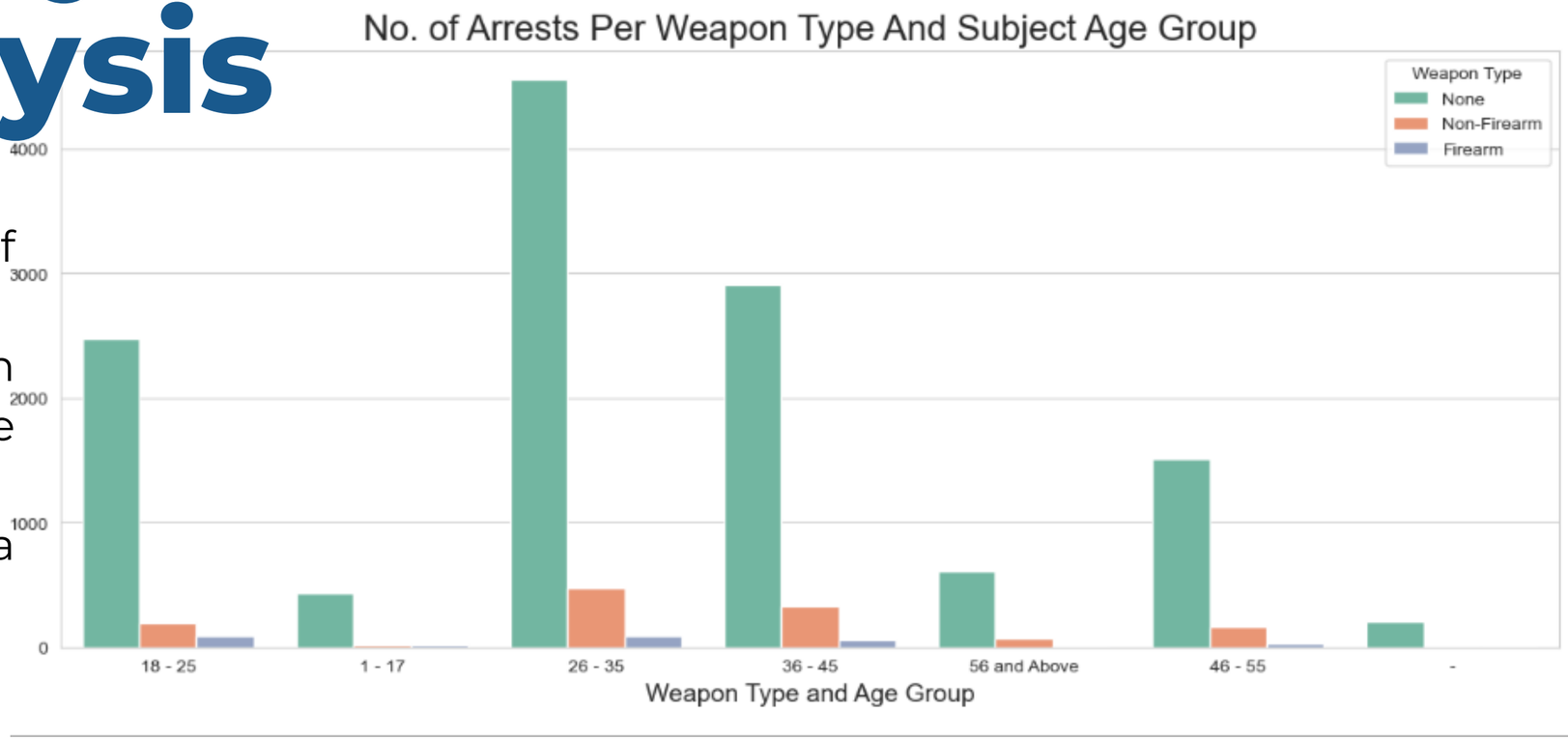
- Each record contains perceived demographics of the subject, as reported by the officer making the stop and officer demographics as reported to the Seattle Police Department.

- The dataset has the following structural features;
- Number of rows = 58,370
- Number of columns = 23

- Classification/Target variable is the Stop resolution column.

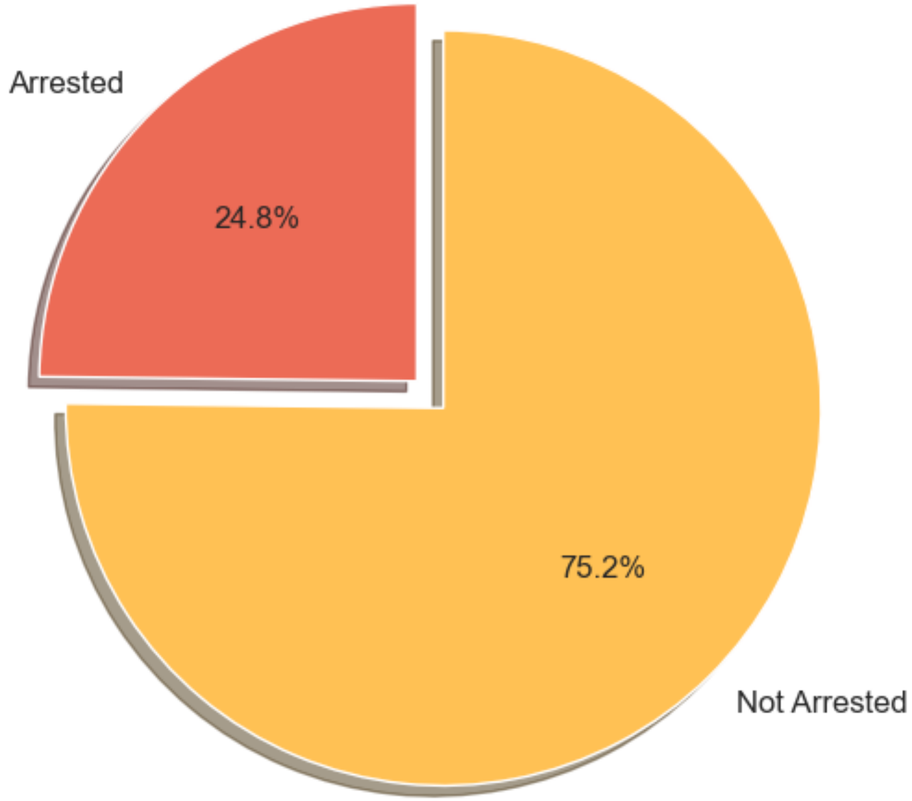
Exploratory Data Analysis

Most arrests show no involvement of weapons.
Notable spikes in weapons (both firearm and non-firearm) occur in the 18-25 and 36-45 age ranges.
The 36-45 age range also exhibits a spike in non-firearm weapons.



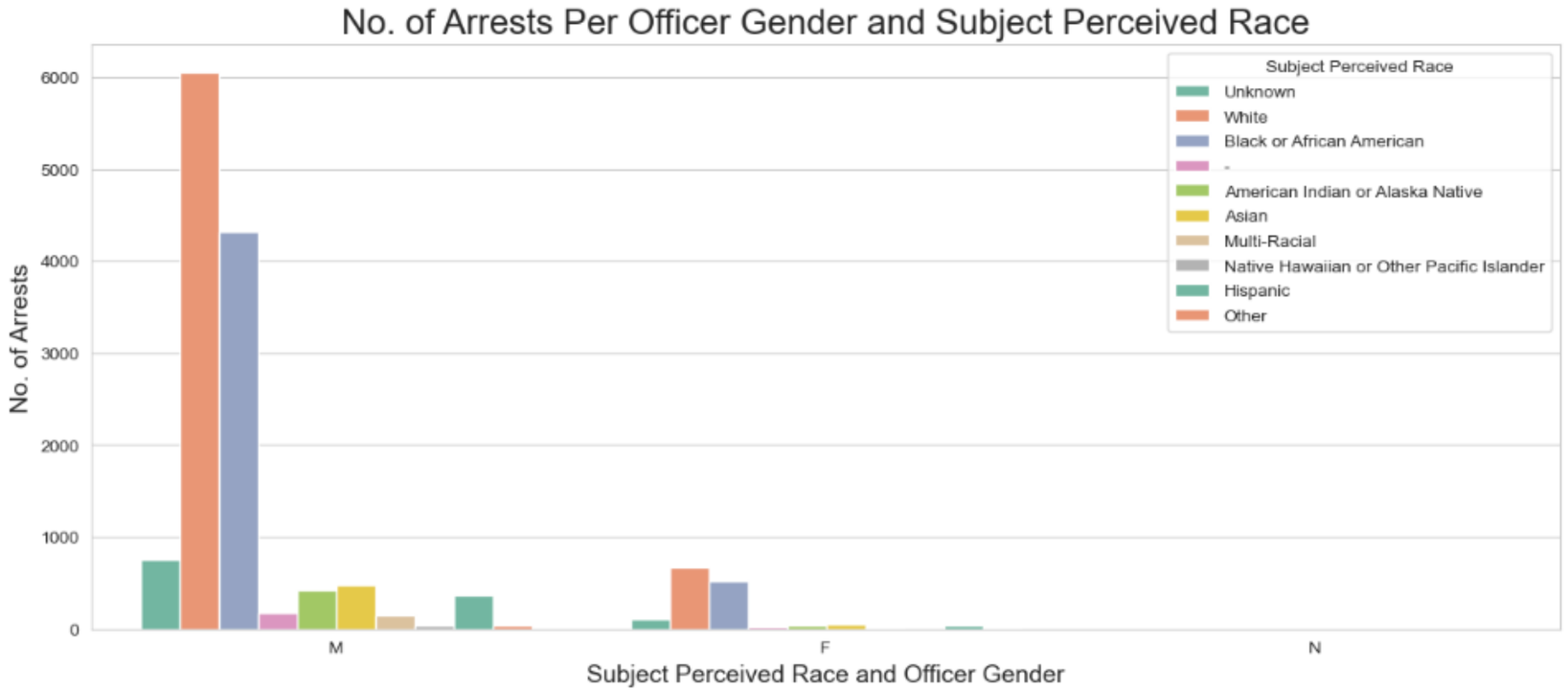
Arrest rate distribution

This pie chart shows the arrest rate distribution overall. The percentage of people arrested is 24.8% while those not arrested is 75.2%.



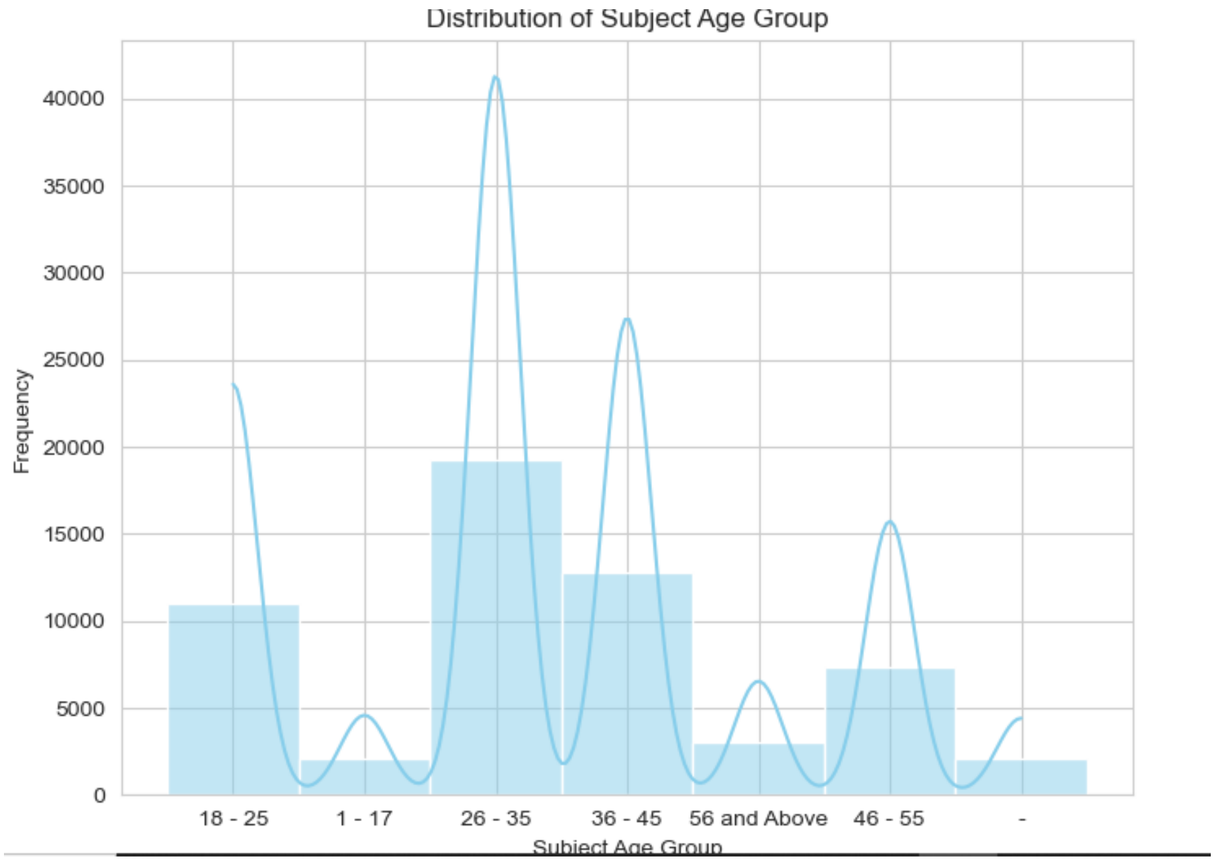
No. of arrests according to office gender

On average, both male and female officers stopped and arrested whites followed by black or african american then followed by asian.



Subject Age group

The age group "26 - 35" has the highest mean arrest rate among the groups, with approximately 26.34%. This suggests that, on average, individuals in the age range of 26 to 35 years old are more likely to experience arrest during Terry Stops. The mean arrest rates generally decrease as the age groups progress from younger to older individuals. This is evident in the descending order of mean arrest rates from "26 - 35" to "56 and Above". There is a category represented by "-" with a mean arrest rate of approximately 10.68%.



Modeling

Model	Accuracy Score (%)
Logistic Regression	74.90
Decision Tree	74.92
KNN	74.97

Evaluation

The Logistic Regression model achieved an accuracy score of 74.90%. This means that, on average, it correctly predicted the target variable for 74.90% of the instances in the dataset.

Decision Tree:

The Decision Tree model achieved a slightly higher accuracy score of 74.92% compared to Logistic Regression. It suggests that the Decision Tree model performed slightly better in terms of accuracy on the same dataset.

KNN (K-Nearest Neighbors):

The KNN model achieved the highest accuracy score among the three models, with a score of 74.97%. This indicates that the KNN model, using the chosen configuration, had the highest accuracy in predicting the target variable.

Recommendations

Conduct training programs for law enforcement officers to enhance their judgment on deciding when it is suitable to make an arrest during a Terry stop. Providing clear guidelines on differentiating situations requiring immediate action from those that can be addressed later can contribute significantly to reducing unnecessary arrests.

Emphasize the importance of recording the officer's precinct in all Terry stops. This additional data point can enhance the predictive capabilities of the model.

Implement training modules for officers to recognize optimal situations for conducting a 'frisk' during Terry stops. Understanding the appropriate circumstances for frisking individuals can serve as a crucial indicator in predicting arrests accurately.

Next steps

- Further analyze unknown or missing values.
- Experiment with SMOTE.
- Tune Support Vector Classification.

Thank You!