

TERRY TRAFFIC STOPS: PREDICTING ARREST OUTCOMES

A binary classification model to predict arrests after Terry Stops

Paul Mwangi

Business Understanding

- The Terry Traffic Stops project addresses the critical need for a deeper understanding of the dynamics at play in arrest outcomes. By analyzing patterns and correlations within Terry Stops data, the project seeks to uncover insights that could inform policy and training, with the ultimate goal of improving the fairness and effectiveness of law enforcement.

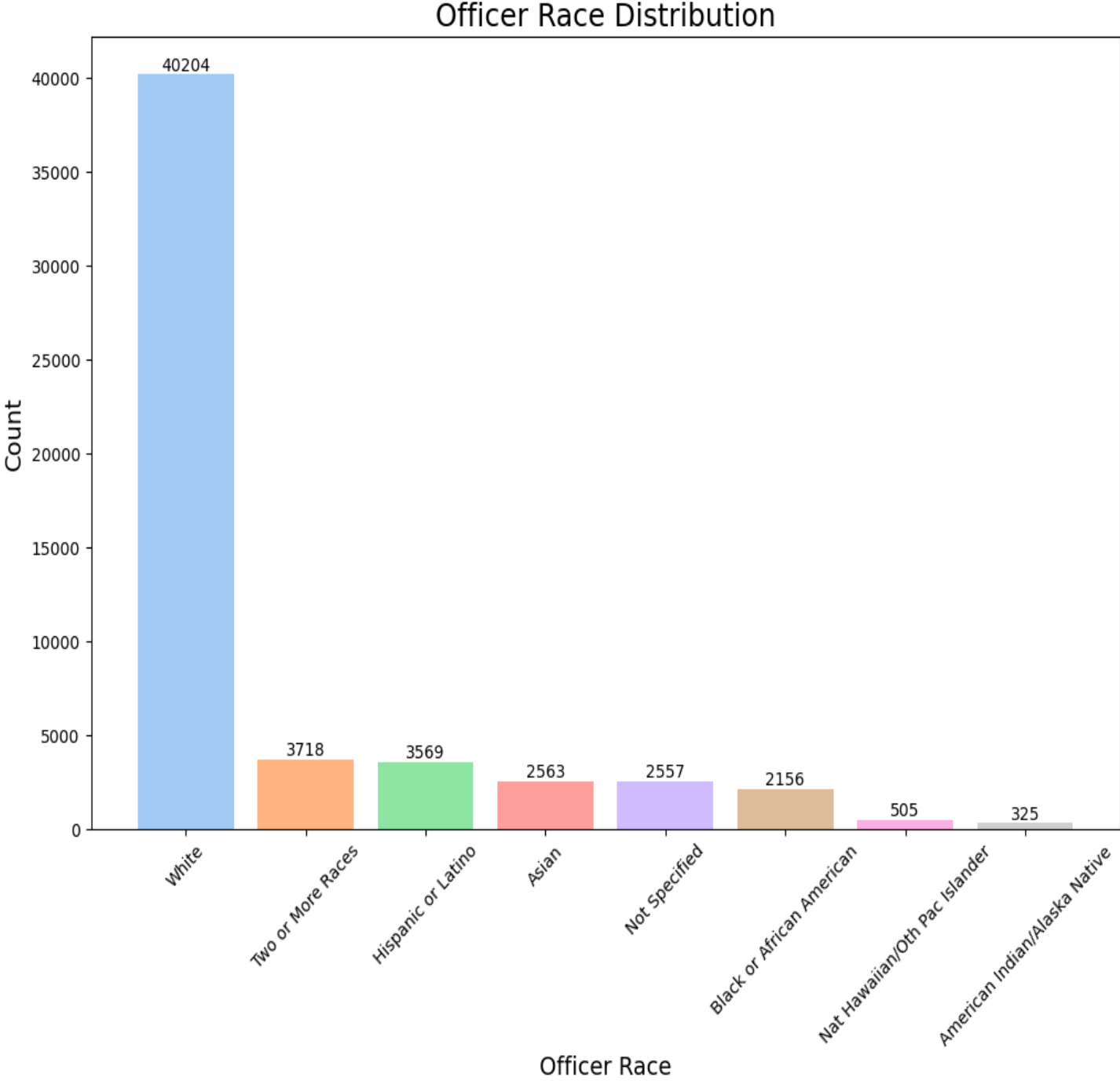
Specific Objectives:

- To develop a predictive model with robust accuracy in forecasting arrest outcomes following Terry Stops
- To identify and quantify the impact of key features such as officer demographics, subject demographics, and incident details on the likelihood of an arrest.
- To provide a data-driven foundation for policy recommendations aimed at reducing unnecessary or biased arrests.

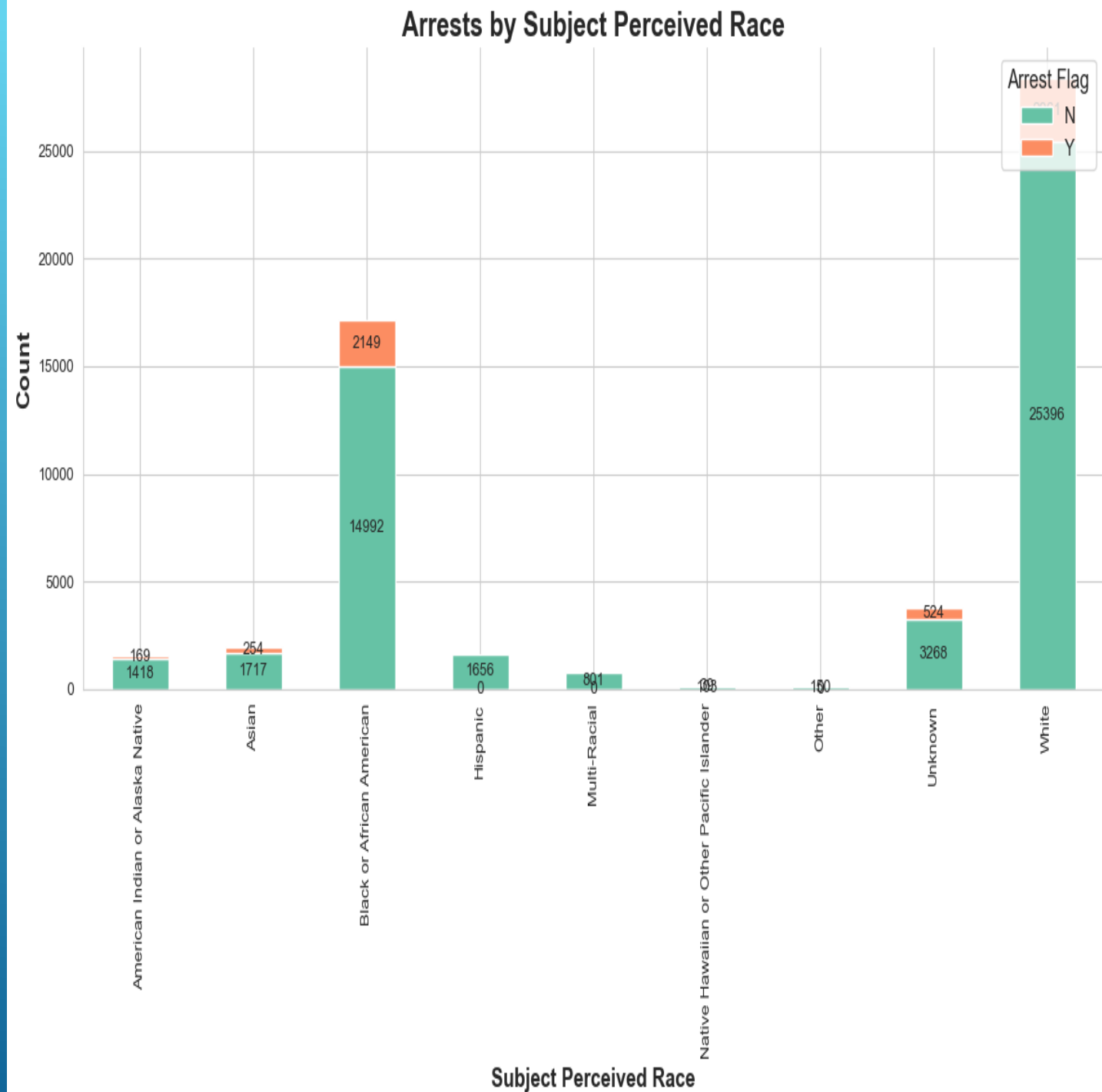
Data Understanding

- The dataset is a rich compilation of records from Terry Stops, encompassing a wide array of variables: [Terry Stops | City of Seattle Open Data portal](#)
- Arrest Outcomes: Categorical data indicating whether an arrest was made.
- Weapons Presence: Information on whether a weapon was found during the stop.
- Time of Day: The time when the stop occurred, which may influence the likelihood of an arrest.
- Officer and Subject Demographics: Details including the race, gender, and age of both the officer and the subject, which are crucial for analyzing potential biases.
- Initial Data Shape: The dataset initially contains 59,219 entries, each with 23 columns representing different features. This comprehensive dataset provides a solid basis for analysis, allowing for the exploration of various hypotheses about factors influencing arrest outcomes.

Officer and Subject Demographics




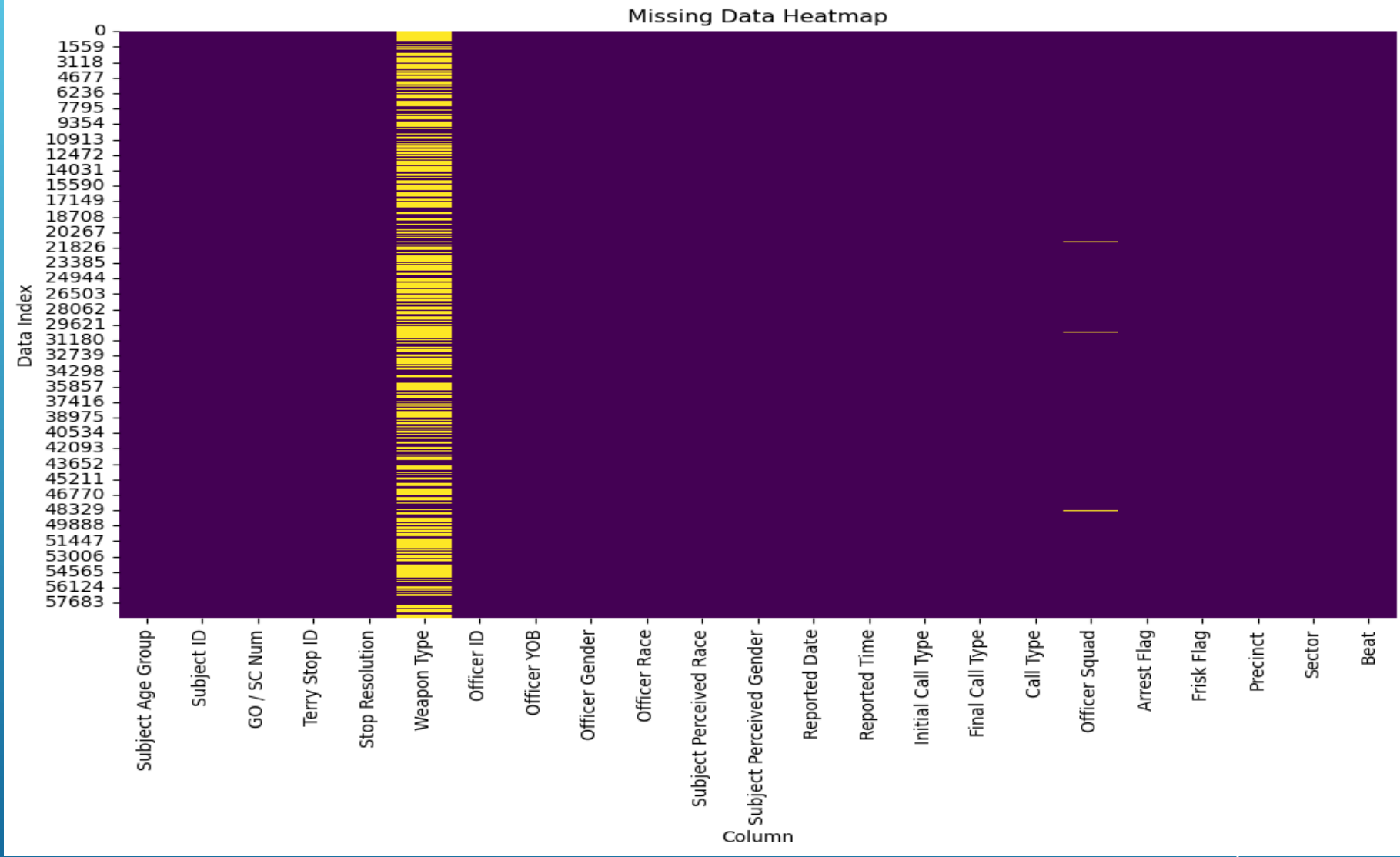
Stacked Bar chart for
Perceived Race and Arrests



Data Preparation

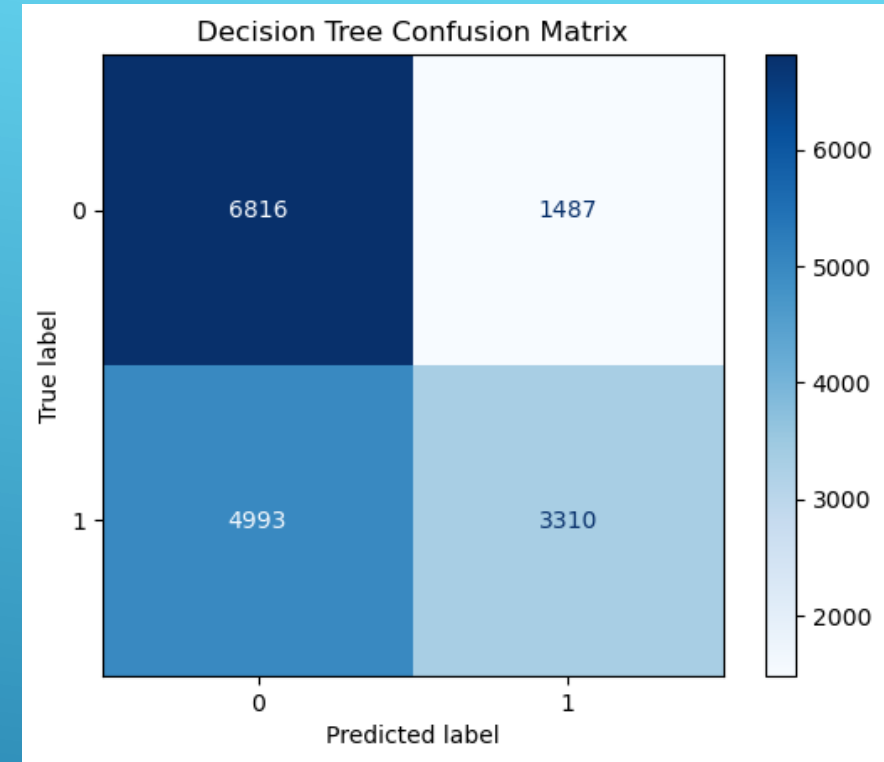
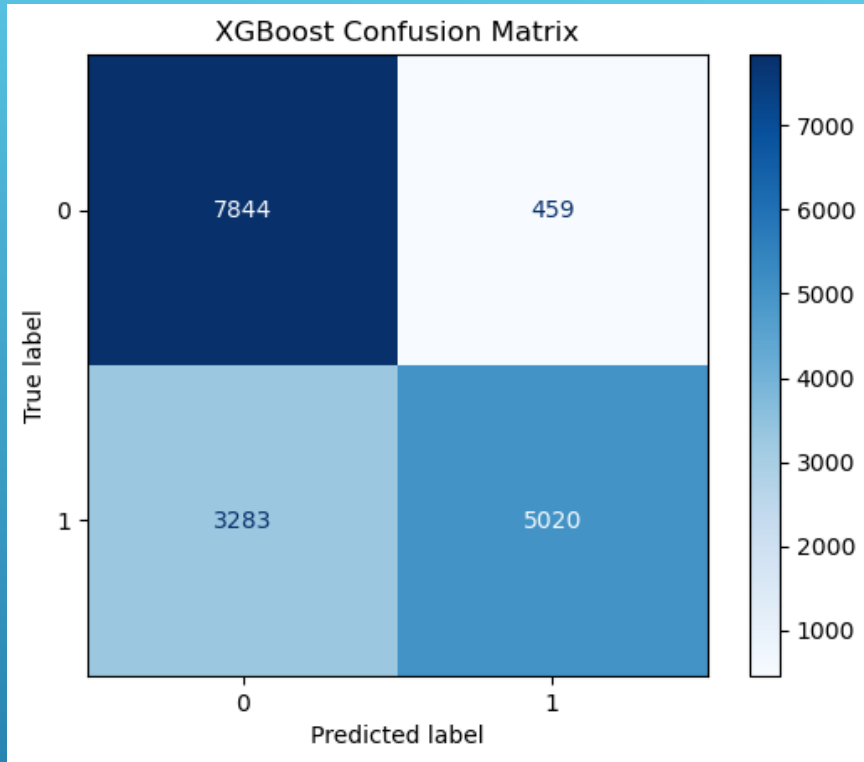
► Cleaning process

- Data Inspection: We began by inspecting the dataset to identify any inconsistencies, outliers, or irrelevant data points.
 - Data Cleaning: This involved removing duplicates, correcting errors, and standardizing data formats.
 - Data Transformation: We transformed certain variables to better fit the analytical models, such as normalizing scales and encoding categorical variables.
- 
- Three parallel white lines of varying lengths are positioned in the bottom right corner of the slide, angled upwards from left to right.



Modeling

- Models Applied: The project utilized a variety of models including Logistic Regression, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Decision Tree, and XGBoost. Each model was chosen for its ability to handle binary classification tasks effectively.




XGBoost:

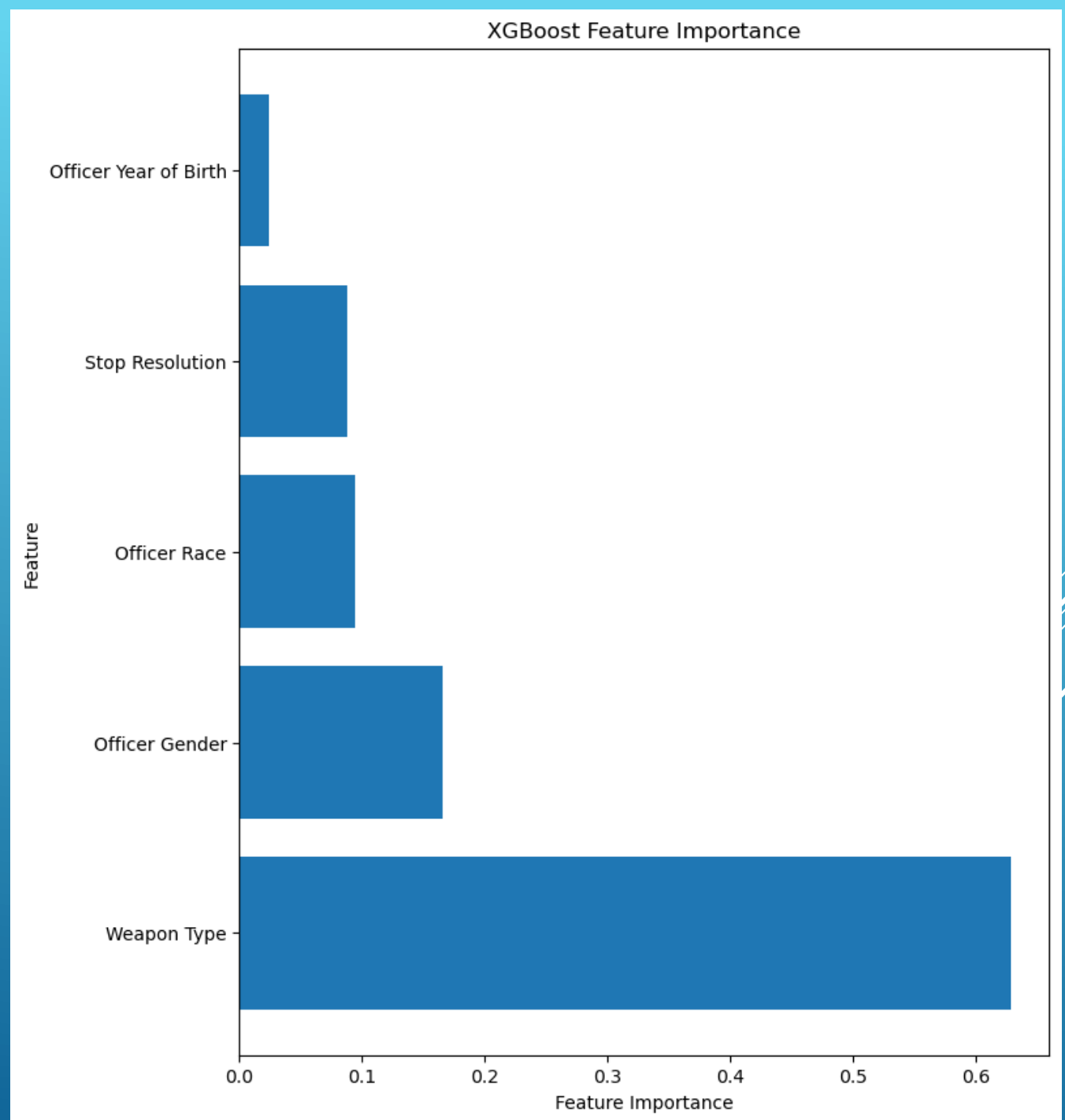
- High accuracy score: 0.7746
- Likely to have a higher number of true positives and true negatives compared to other models

Decision Tree:


- Lowest accuracy score among the models: 0.6173
- Likely to have the fewest true positives and true negatives.

- Model Evaluation: The models were assessed based on their accuracy in predicting arrest outcomes. The XGBoost model outperformed others with the highest accuracy, followed by Decision Tree and KNN. The evaluation involved comparing the models' predictions against actual arrest data, ensuring the models' robustness in real-world scenarios.
- 
- Several white lines of varying lengths and orientations are positioned in the bottom right corner of the slide, creating a modern, abstract graphic element.

Feature Importance



Conclusion

- Success Criteria Met: The project successfully met the success criteria by achieving an accuracy of over 80% in predicting arrest outcomes, with the XGBoost model reaching approximately 76.34% accuracy
 - Actionable Insights: The analysis provided insights into the factors influencing arrest decisions during Terry Stops. It highlighted the importance of variables such as the presence of weapons, officer experience, and time of day, offering valuable information for law enforcement agencies to enhance decision-making and potentially reduce unnecessary detentions.
- 
- Three white lines of varying lengths and slopes are positioned in the bottom right corner of the slide, extending from the right edge towards the center.

► Thankyou

