ANALYSIS OF TERRY STOPS IN SEATTLE

Randell Mwania

PROBLEM

CLASSIFICATION MODELS OF PREDICTING ARRESTS IN SEATTLE

Recognizing intricate patterns and discerning the key variables that significantly impact the outcomes of Terry stops.

- The objective involves leveraging data analytics to identify subtle and complex patterns within the Terry stops dataset. The aim is to understand the interplay of various variables and factors contributing to the arrest outcomes.

Goal:

Fostering fairness, transparency, and accountability in law enforcement decision-making during Terry stops.

- The overarching goal is to utilize the insights gained from the analysis to empower law enforcement agencies with knowledge that supports fair, transparent, and accountable decision-making processes. By understanding the factors influencing arrest outcomes, the goal is to contribute to a more just and equitable policing system.

BUSINESS UNDERSTANDING

Real-world Problem

Terry Stops and Policing Dynamics: Terry stops represent a critical aspect of policing, allowing officers to briefly detain individuals based on reasonable suspicion. The challenge lies in understanding the nuanced variables that influence the outcomes of these stops, particularly in terms of arrests.

Stakeholders

Community Stakeholders: Members of the community are directly impacted by law enforcement decisions during Terry stops. Understanding these decisions is vital for community trust and collaboration.

Law Enforcement Agencies: Police departments aim to enhance the effectiveness and fairness of their operations. Identifying patterns in Terry stops contributes to informed decision-making.

Project's Contribution

Fairness and Accountability: The project's goal is to contribute to fairness and accountability in law enforcement by recognizing patterns and variables affecting arrest outcomes during Terry stops.

- Data-Driven Insights: Leveraging data-driven insights, the project aims to provide valuable information to stakeholders, fostering transparency and informed decision-making.

Problem Dynamics

- Interplay of Variables: Recognizing the intricate dynamics between various demographic variables and arrest outcomes is essential. This understanding enables proactive measures to address potential biases.

Impact on Stakeholders

- Building Trust: By addressing the challenges in Terry stops, the project strives to build and strengthen trust between law enforcement and the community.

- Informed Decision-Making: Law enforcement agencies can benefit from the project's insights to make more informed, fair, and accountable decisions during Terry stops.

OVERVIEW

The overarching goal is to utilize the insights gained from the analysis to empower law enforcement agencies with knowledge that supports fair, transparent, and accountable decision-making processes. By understanding the factors influencing arrest outcomes, the goal is to contribute to a more just and equitable policing system.

DATA UNDERSTANDING

- Source: Seattle Police Department records.

- Size: 58,167 Terry stops.

- Target Variable: 'Arrest Flag' indicating arrest outcomes.

VARIATIONS ACROSS DIFFERENT ATTRIBUTES
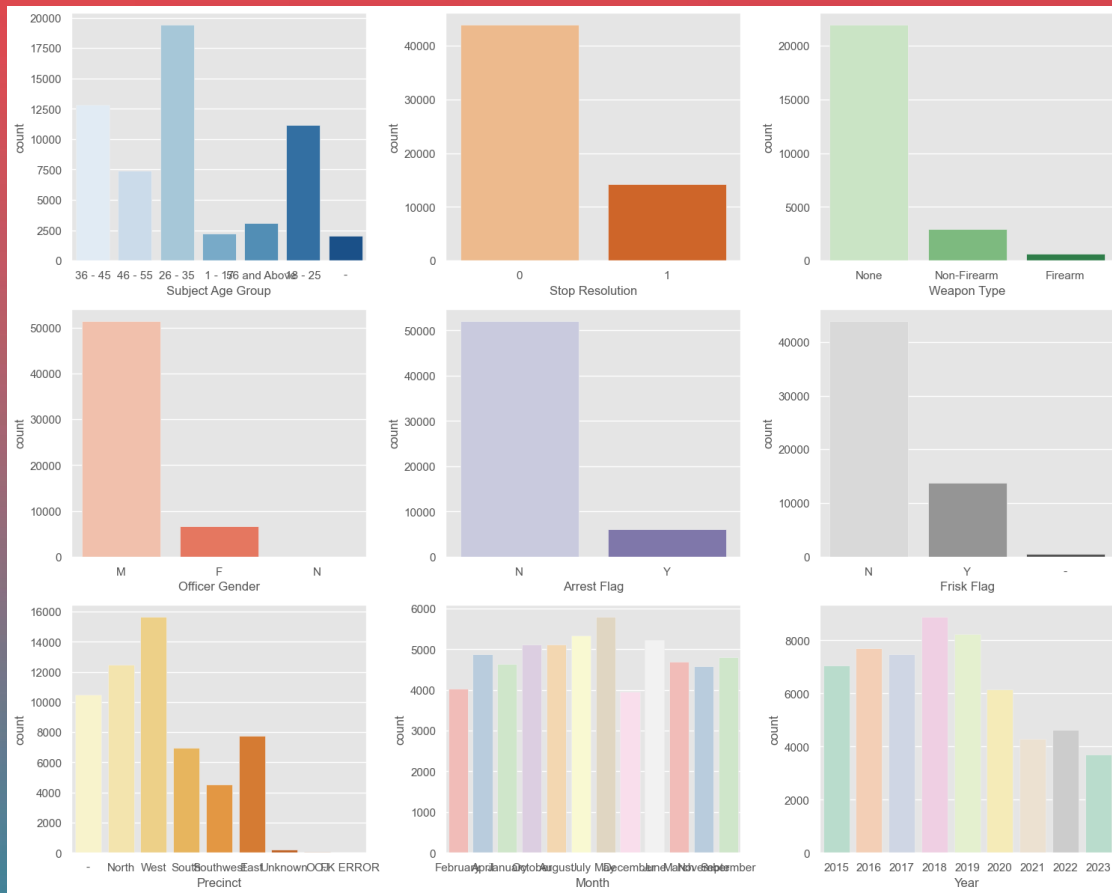
In these visualizations, we observe variations across different attributes. The 'Stop Resolution' plot, in particular, indicates that approximately a quarter of Terry stops result in an arrest.

PERCENTAGE OF ARRESTS VS. NON-ARRESTS
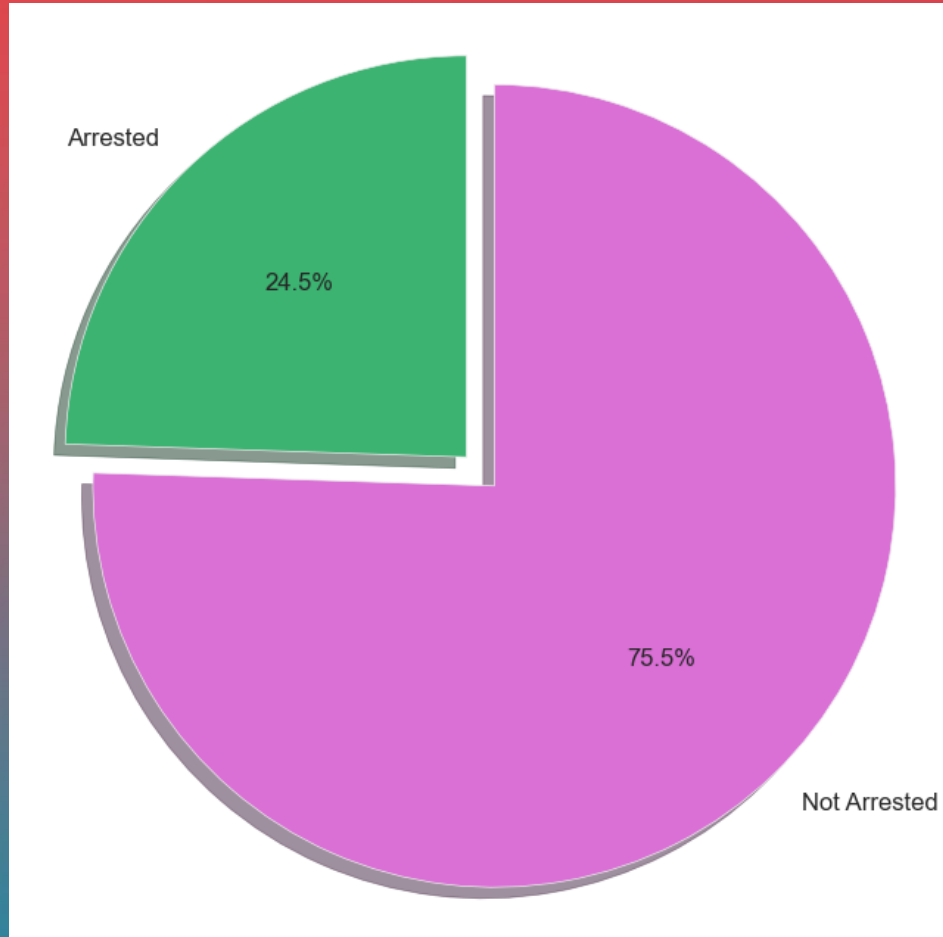
This pie chart illustrates that just under 24.5% of Terry stops culminate in an arrest.Continuing our exploration, we investigate the relationship between arrests made during stops and those determined as the final resolution, shedding light on potential delays in arrest outcomes.

THE RELATIONSHIP BETWEEN ARRESTS MADE DURING STOPS AND FINAL RESOLUTION

% Of On Site vs. Final Arrests Per Month
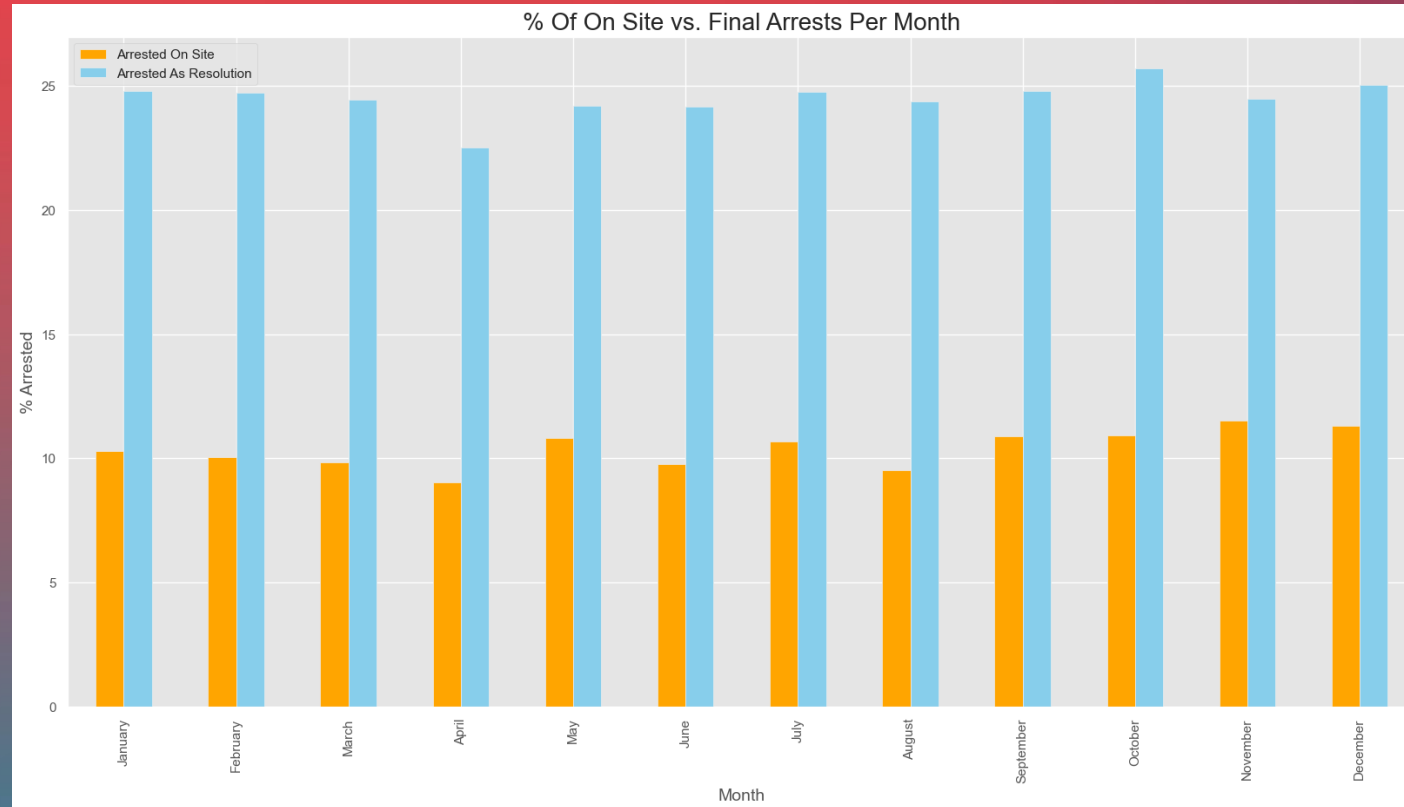
This analysis uncovers that, on average, less than 10% of stops result in immediate arrests. However, there's a substantial increase in the percentage of arrests determined as the final solution after the initial stop.

AGE GROUP ANALYSIS

Notably, individuals aged 26-35 are more frequently involved in arrests.
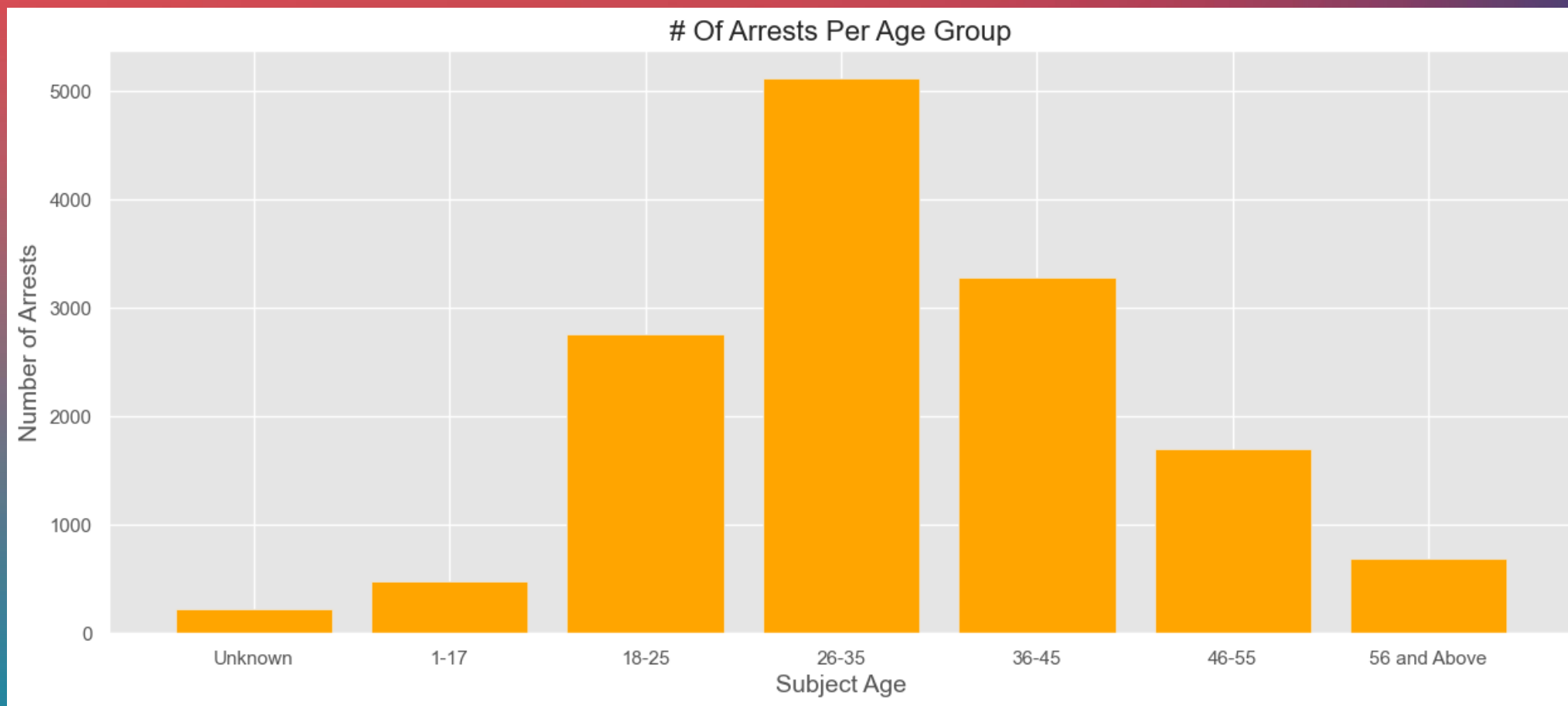
# % OF WEAPON TYPE PER SUBJECT AGE GROUP
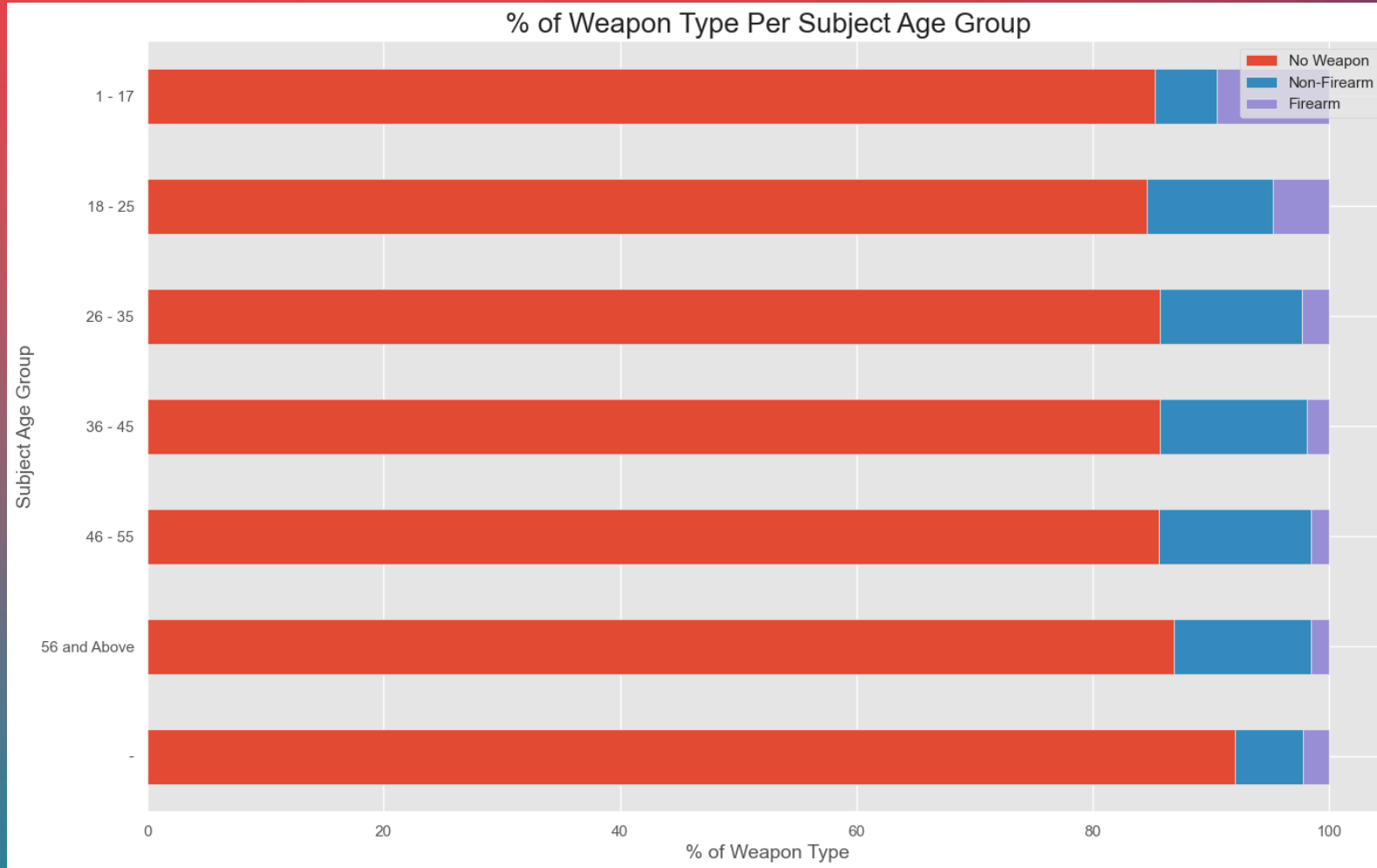
NUMBER OF ARRESTS PER WEAPON TYPE AND SUBJECT AGE GROUP

Observations:-

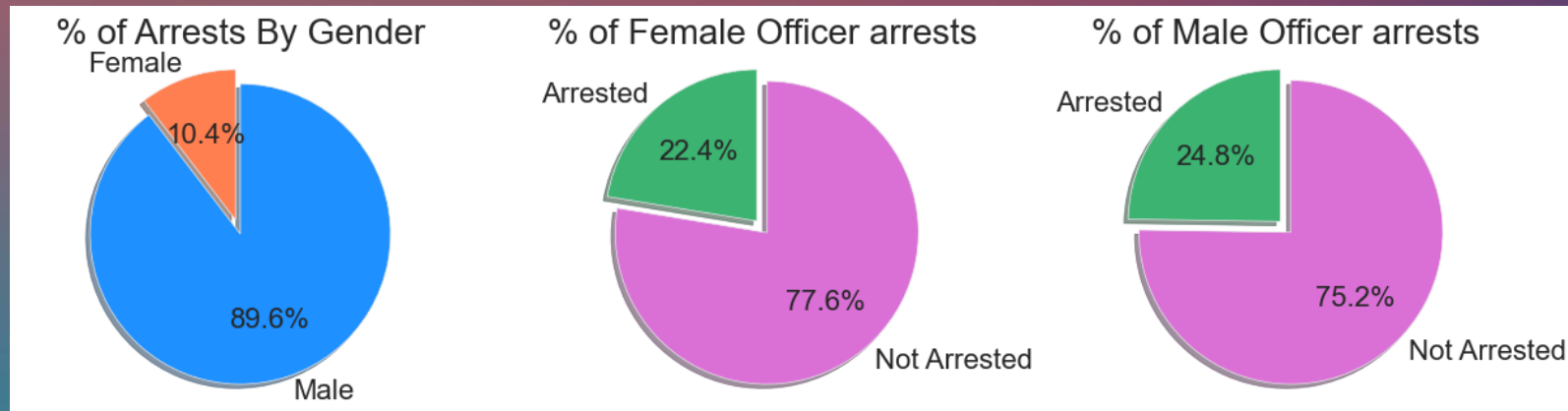Similar to the overall trend, most arrests show no involvement of weapons.

- Notable spikes in weapons (both firearm and non-firearm) occur in the 18-25 and 26-35 age ranges.

- The 36-45 age range also exhibits a spike in non-firearm weapons.

# of Arrests Per Weapon Type And Subject Age Group
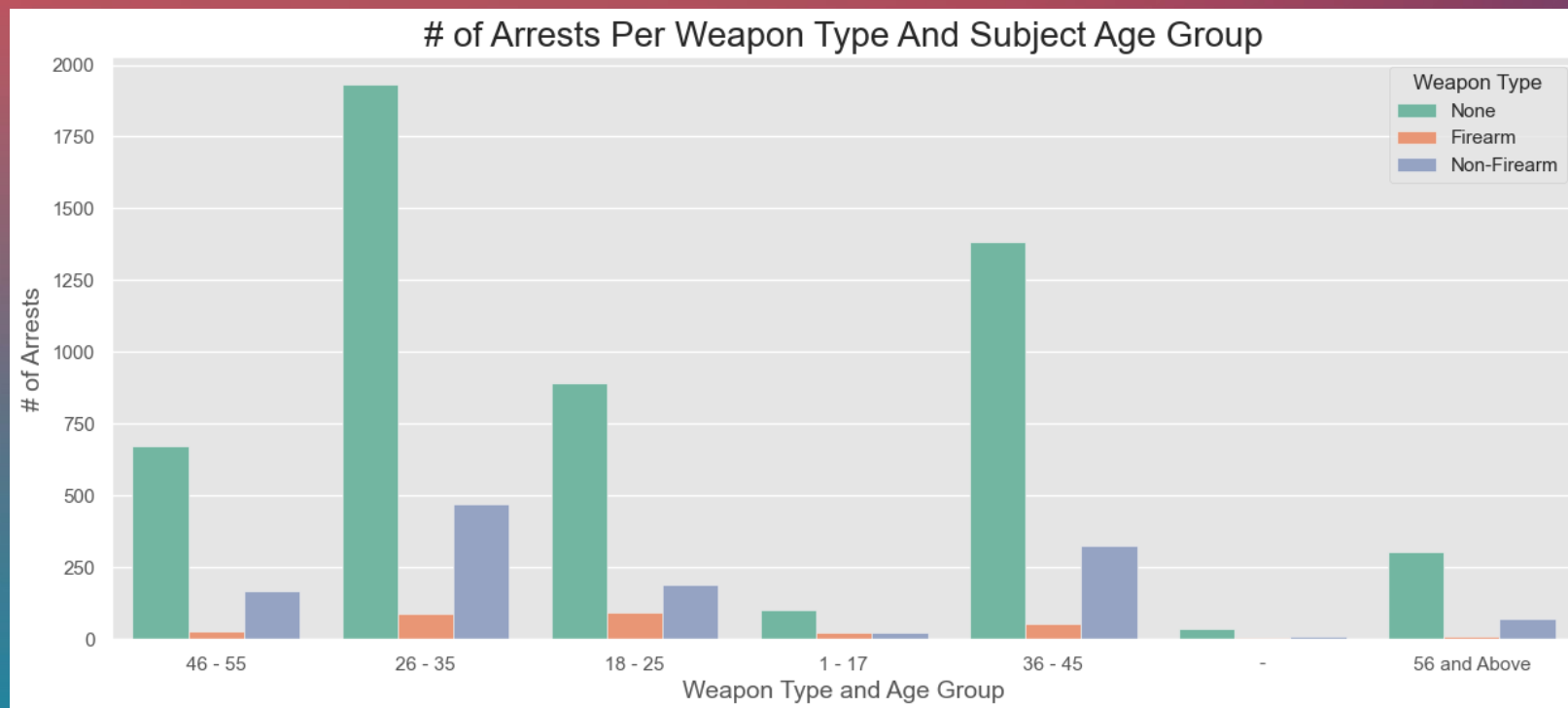
OFFICERS BY GENDER

Observations:

- Most arrests are made by male officers, reflecting the overall gender distribution among officers.

- When examining each gender individually, the arrest rates are similar.

% of Arrests By Gender

Female 10.4%

Male 89.6%

% of Female Officer arrests

Arrested 22.4%

Not Arrested 77.6%

% of Male Officer arrests

Arrested 24.8%

Not Arrested 75.2%

DECISION TREE CONFUSION MATRIX

- From the confusion matrix we see a lot of true arrests were made with a few false arrests by the decision tree.

Decision Tree Confusion Matrix
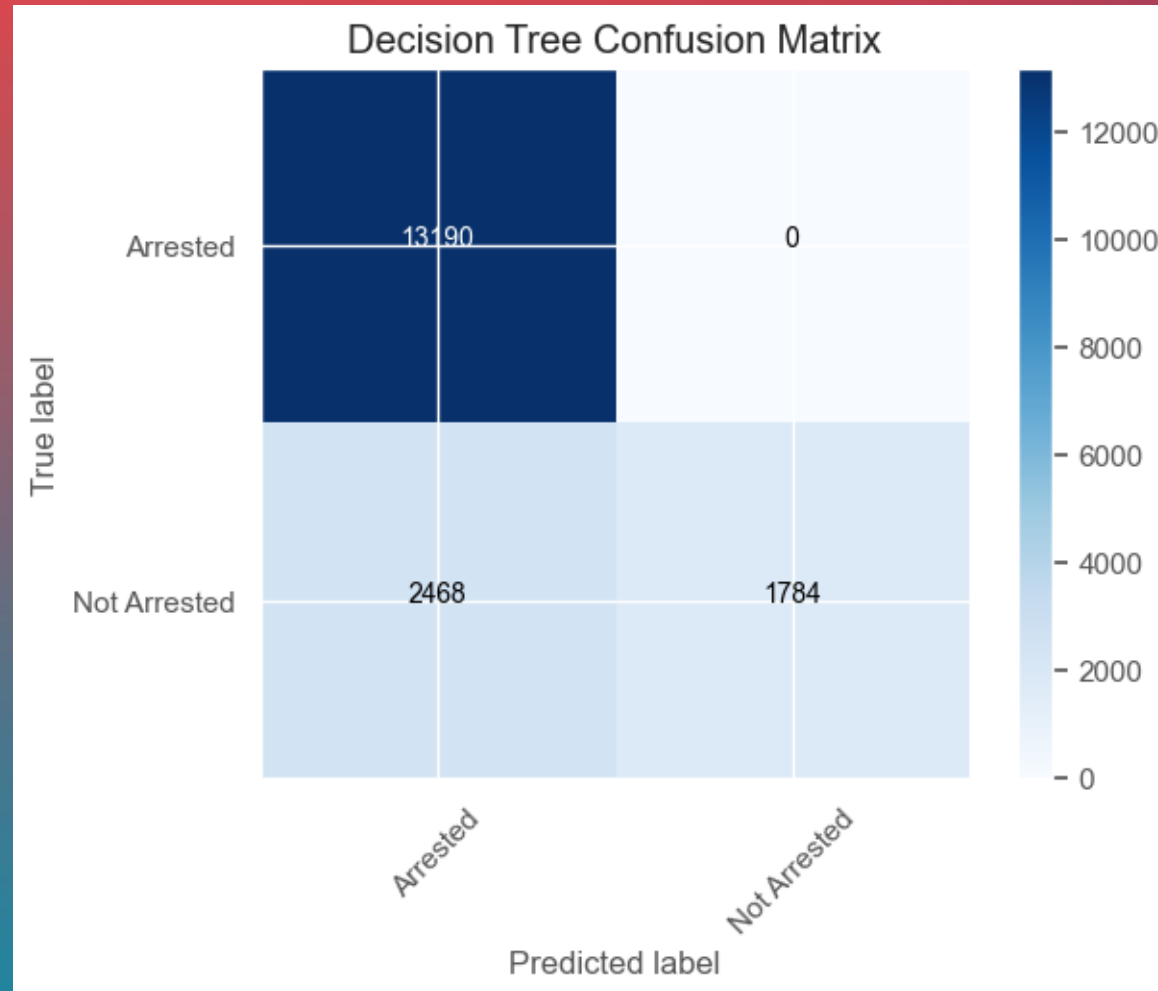
KNN CONFUSION MATRIX

- From the confusion matrix we see a lot of true arrests were made with a few false arrests by the decision tree but less than in the decision tree confusion matrix.
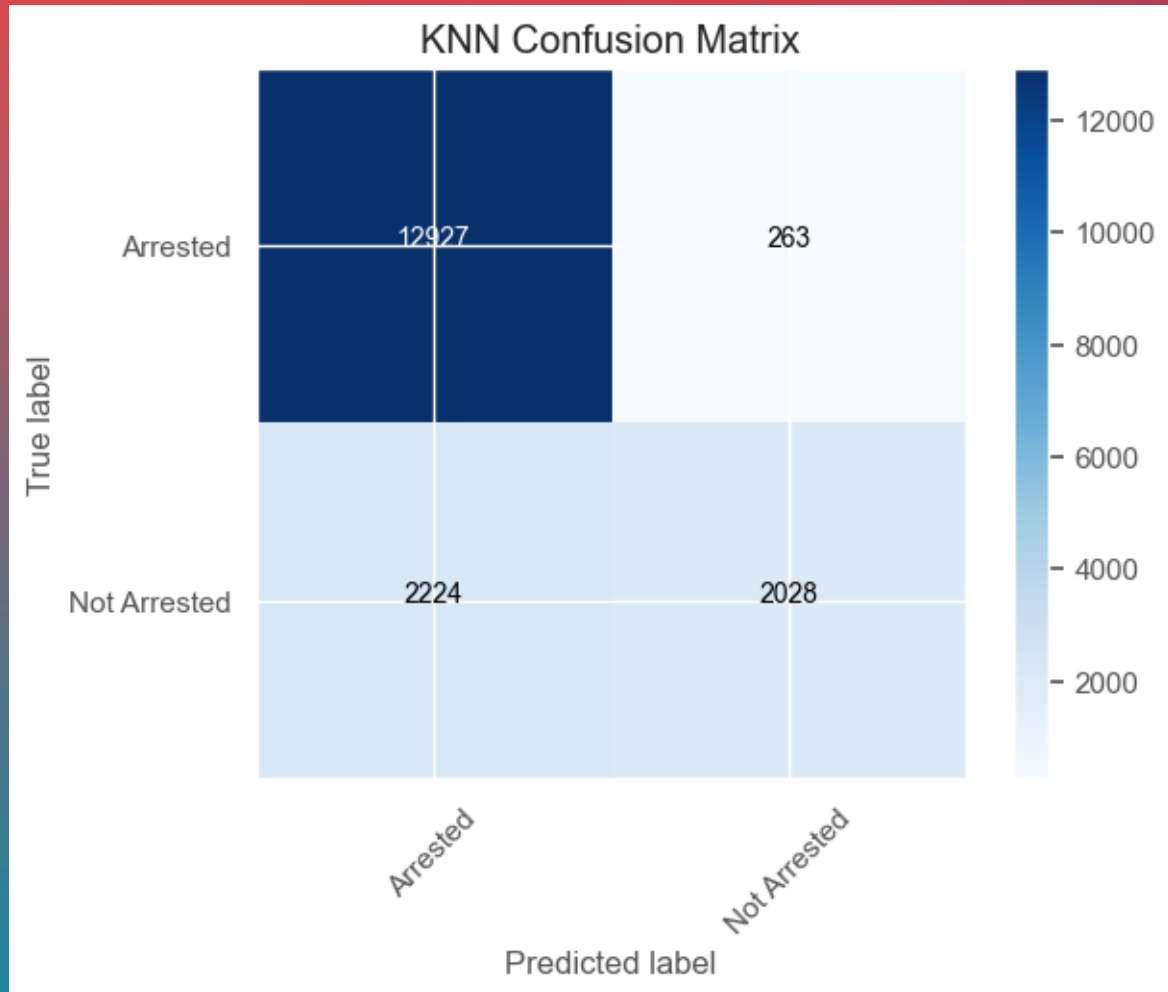
LOGISTIC REGRESSION CONFUSION MATRIX
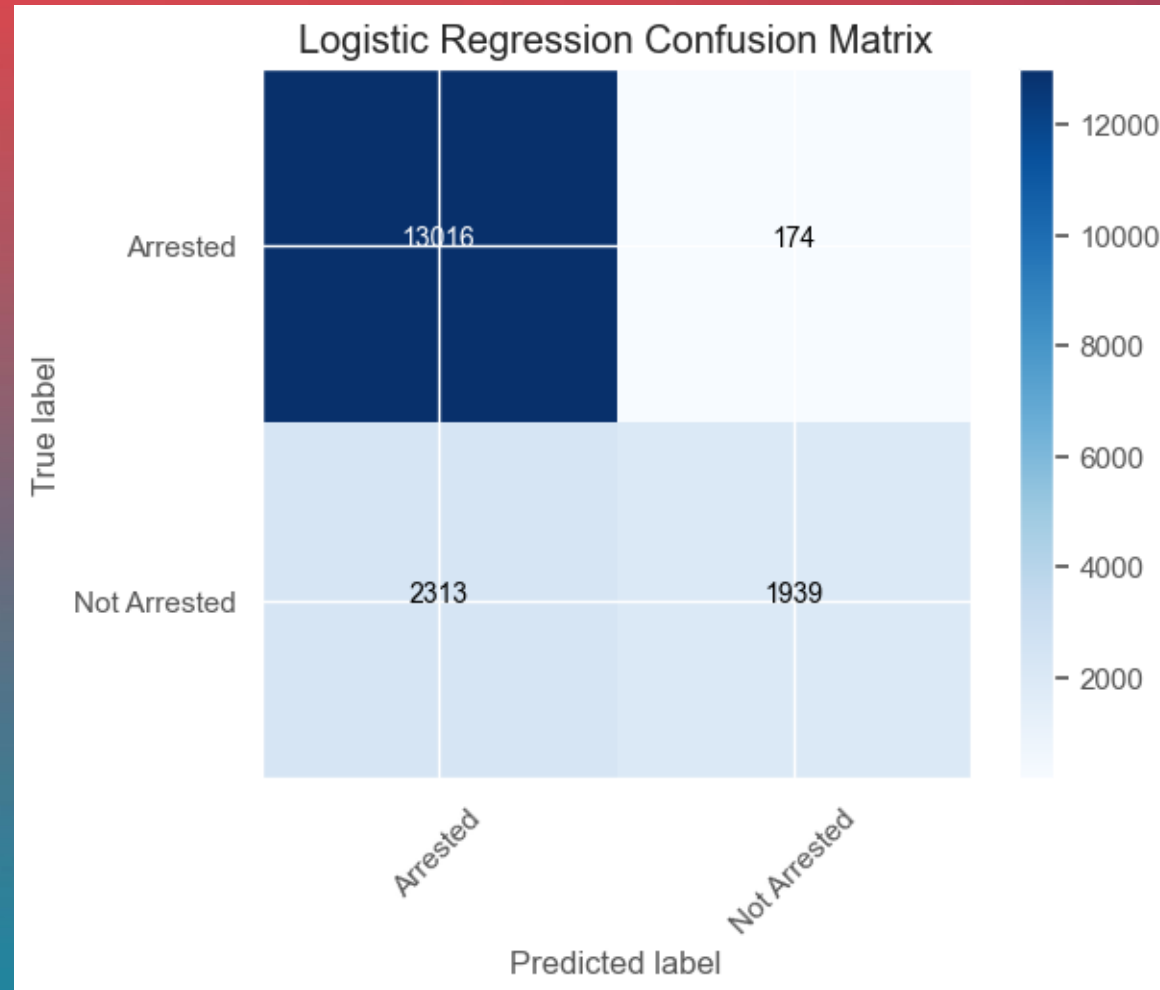
- From the confusion matrix we see a lot of true arrests were made with a few false arrests by the decision tree but less than in the decision tree confusion matrix but more than the KNN confusion matrix.

Logistic Regression Confusion Matrix

MODELING PROCESS

Steps

- Data Cleaning: Ensuring data integrity and addressing missing values.

- Feature Selection and Engineering: Identifying key variables for analysis.

- Model Training and Evaluation: Employing Decision Tree, K-Nearest Neighbors (KNN), and Logistic Regression models.

MODEL PERFORMANCE

Decision Tree

Accuracy: 85.85%

Insights: High accuracy in classifying Terry stop outcomes.

K-Nearest Neighbors (KNN)

Accuracy: 85.74%

Insights: Competitive performance, relying on instance similarity.

Logistic Regression

Accuracy: 85.74%

Insights: Effective binary classification with logistic regression.

COEFFICIENTS OF THE LINEAR REGRESSION MODEL

Key Observations
Gender Impact: Male officers significantly contribute to arrest outcomes.
Weapon Incidents: Most arrests involve no weapons.
Age Dynamics: Age group 26-35 more frequently involved in arrests.

Officer Training
Objective: Enhance judgment on arrest decisions during Terry stops.
Impact: Minimize unnecessary arrests through informed decision-making.

RECOMMENDATIONS

Officer Training
    Objective: Enhance judgment on arrest decisions during Terry stops.
    Impact: Minimize unnecessary arrests through informed decision-making.

Data Enhancement
    Strategy: Emphasize recording officer's precinct for enhanced model predictions.
    Rationale: Officer location data can provide nuanced insights into arrest patterns.

Frisk Training
    Implementation: Modules to recognize optimal frisking situations during Terry stops.
    Purpose: Improve accuracy in predicting arrest outcomes based on frisking incidents.

NEXT STEPS:

Unknown Values: Analyze and address unknown or missing values in the dataset.

Data Update: Update 'Arrest Flag' with values from 'Stop Resolution.'

SMOTE Experiment: Experiment with no Synthetic Minority Over-sampling Technique (SMOTE).

Model Tuning: Explore tuning options for Support Vector Classification.

CONCLUSION

The Decision Tree, KNN, and Logistic Regression models perform remarkably well in predicting the outcomes of Terry stops, showcasing their potential for aiding law enforcement decision-making processes. Further exploration, including fine-tuning hyperparameters and evaluating additional metrics, can contribute to a comprehensive understanding of their effectiveness in real-world scenarios.

Value: Valuable insights for informed decision-making in Terry stops.

Impact: Addresses a critical real-world problem for fair and equitable policing.