Data Analysis Report: Predicting Arrest Outcomes in Terry Stops

1. Introduction

This report summarizes the data analysis and modeling process conducted to predict whether a Terry Stop conducted by the Seattle Police Department (SPD) will result in an arrest. The analysis is based on the Terry Stops dataset provided by the Seattle Government, containing 64,699 records and 23 columns related to stop details, subject and officer demographics, and stop outcomes.

---

2. Business Understanding

Problem Statement

Terry Stops consume significant officer time and resources. The goal is to predict whether a stop will lead to an arrest (`Arrest Flag: Y/N`) to help SPD prioritize high-risk stops, allocate resources efficiently, and identify patterns for officer training to reduce low-yield stops.

Objectives

- Preprocess and explore the dataset to identify key features influencing arrest outcomes.
- Build and compare classification models (logistic regression and decision trees).
- Tune the selected model to balance precision and recall.
- Provide actionable insights for SPD leadership.

Success Metrics

- **Precision**: Minimize false positives (unnecessary stops).
- **Recall**: Ensure high-risk stops are not missed.
- **Baseline**: A dummy classifier predicting the majority class (no arrest) achieves ~85% accuracy due to class imbalance (15% arrests).

---

3. Data Understanding

Dataset Overview

- **Rows**: 64,699
- **Columns**: 23
- **Source**: Seattle Government

Key Columns

- `Subject Age Group`, `Subject Perceived Race`, `Subject Perceived Gender`

- `Weapon Type`, `Frisk Flag`, `Precinct`, `Call Type`
- `Arrest Flag` (target variable)

Initial Observations

- The dataset contains a mix of categorical and numeric features.
- Missing values are present in columns like `Weapon Type` (32,565 nulls) and `Officer Squad` (566 nulls).
- The target variable `Arrest Flag` is imbalanced (15% arrests).

---

4. Data Preparation

Data Cleaning

- **Missing Values**:
  - `Weapon Type`: Nulls and placeholder values (`-`) replaced with "Unknown".
  - `Officer Squad`: Rows with nulls dropped (566 rows).
  - `Frisk Flag` and `Subject Age Group`: Placeholder values (`-`) replaced with "Unknown" or dropped.
- **Duplicates**: No duplicates found.
- **Outliers**: Removed from `Officer Age` using IQR method.

Feature Engineering

- **`Weapon Type`**: Categories were consolidated (e.g., "Knife/Cutting/Stabbing Instrument" → "Knife/Cutting").
- **`Officer Age`**: Derived from `Officer YOB` and `Reported Date`.
- **`Arrest Flag`**: Mapped to binary values (`N` → 0, `Y` → 1).

Final Dataset

- **Rows**: 61,401
- **Columns**: 7 selected features + target variable:
  - `Subject Perceived Race`, `Subject Perceived Gender`, `Weapon Type`, `Frisk Flag`, `Precinct`, `Call Type`, `Arrest Flag`

---

5. Exploratory Data Analysis (EDA)

1. Arrests by Subject Perceived Race

- **Observation**: Disparities exist in arrest rates across racial groups.
- **Black or African American** subjects had the highest proportion of arrests (~13.7% of stops resulted in arrest).

- **White** subjects had a lower arrest rate (~11.4%).
- **Unknown** and **Other** races had varied arrest rates.

2. Arrests by Weapon Type

- **Observation**: Presence of weapons strongly predicts arrests. • Stops involving **Firearms** had the highest arrest rate (~28%).
- Stops with **No Weapon** had the lowest arrest rate (~2%).
- **Knife/Cutting** instruments also correlated with higher arrest rates.

---

6. Modeling Preparation

Feature Selection

The following features were selected for modeling:

- `Subject Perceived Race`
- `Subject Perceived Gender`
- `Weapon Type`
- `Frisk Flag`
- `Precinct`
- `Call Type`

Target Variable

- `Arrest Flag` (binary: 0 = No Arrest, 1 = Arrest)

Class Imbalance

- The target variable is imbalanced (15% arrests, 85% no arrests). Techniques like oversampling or class weighting may be required.

---

7. Next Steps

Modeling

- Train and compare logistic regression and decision tree models.
- Address class imbalance using SMOTE or class weights.
- Optimize for precision to minimize false positives.

Evaluation

- Evaluate models using precision, recall, F1-score, and ROC-AUC.
- Compare against a baseline dummy classifier.

Actionable Insights

- Identify key features driving arrest outcomes.
- Provide recommendations to SPD for resource allocation and bias mitigation.

---

8. Conclusion

The dataset has been cleaned, preprocessed, and explored to identify key predictors of arrest outcomes. The analysis reveals disparities in arrest rates across racial groups and confirms the significance of weapon presence in predicting arrests. The next phase will involve building and evaluating machine learning models to support SPD in optimizing stop outcomes.

Prepared by: Leslie Nungi

Date: 9-13-2025

Dataset Source: Seattle Government Terry Stops Dataset