# Seattle Terry Stops Analysis

Spencer Hadel

# Summary

- Preliminary investigation into the Seattle Terry Stops
- Connection between the identified races of the subjects and officers in question, and outcome of Terry Stops
- Primary findings show that meaningful connections exist, but that deeper analysis is necessary for clearer results

# Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

# The Problem

Terry Stops (also known as 'stop-and-frisks'), when a police officer uses theur right to legally temporarily detain a person based on 'reasonable suspsicion' that the person may be involved in criminal activity. The officer has the right to physically 'frisk' the subject, and take whatever action they feel is necessary properly handle the situation.

This analysis will investigate the recorded Terry Stop data of Seattle.gov in an attempt to identify racial disparity in the outcomes of theses stops.

# Data

This analysis utilizes Terry_Stops.csv, a dataset of over 52,000 entries of Terry Stops in Seattle from 2015 to the present (2022).

Null values, extraneous data, outliers and probable errors were removed from the data to increase model performance.

Notable and important features include Subject Perceived Race, Officer Race, Arrest Flags, Frisk Flags, and Stop Resolution.

# Methods

- Iterative approach to model selection and tuning

- Main Scores Observed: **Accuracy, Recall, Precision, and F1 scores**

- Visualizations: **Confusion Matrices** for each model, to observe relationships between **True Positives, False Positives, True Negatives, and False Negatives**

# Methods

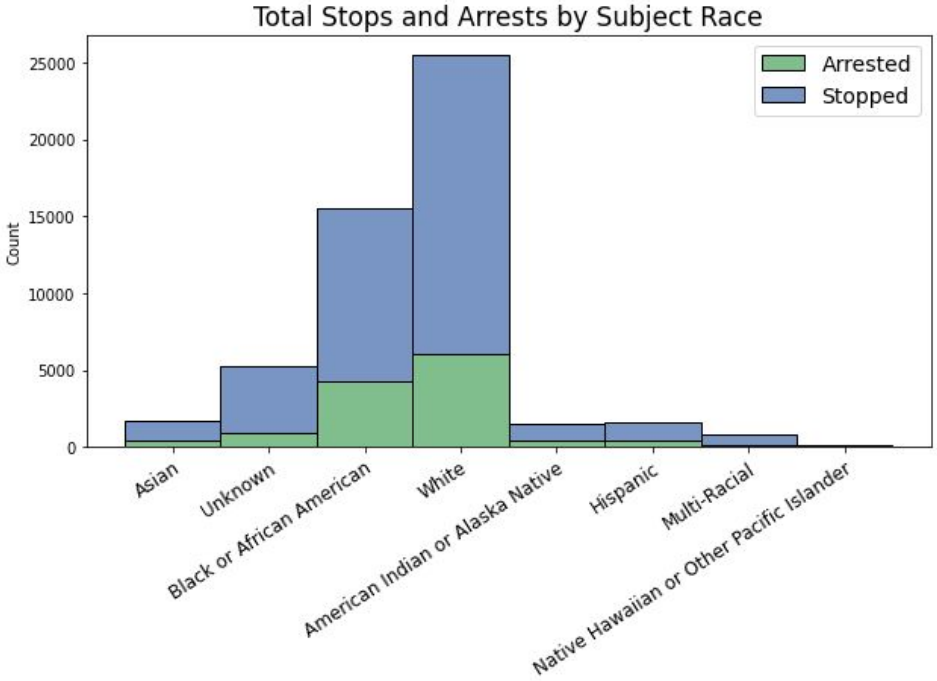
Classification Algorithms Used:

- **Logistic Regression**
- **K Nearest Neighbors**
- **Decision Trees**
- **Random Forest**
- **XG Boost**

# Pre-Modeling:

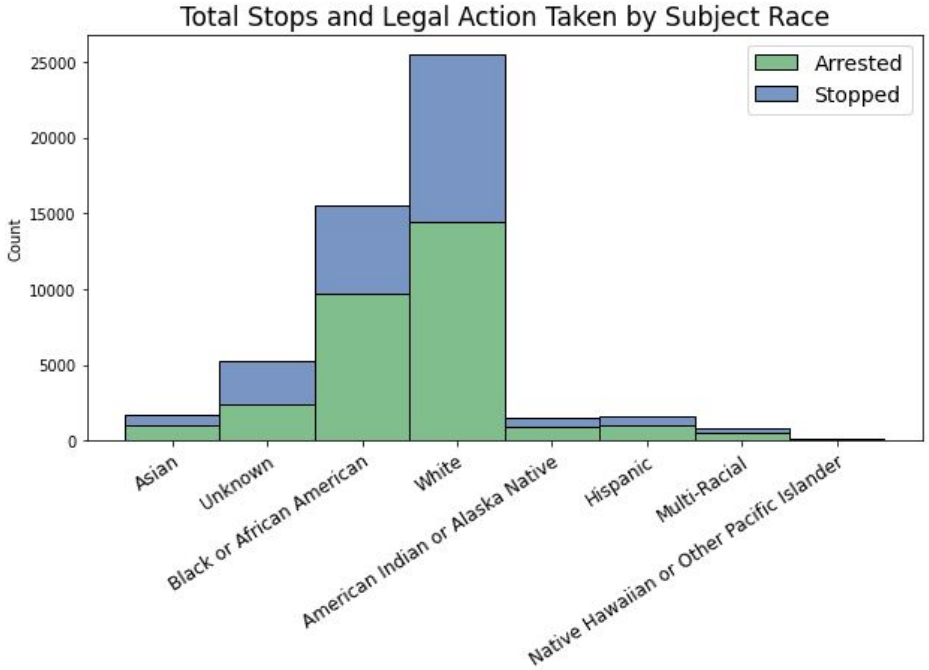
# Data Visualization

# Results - Data Visualization



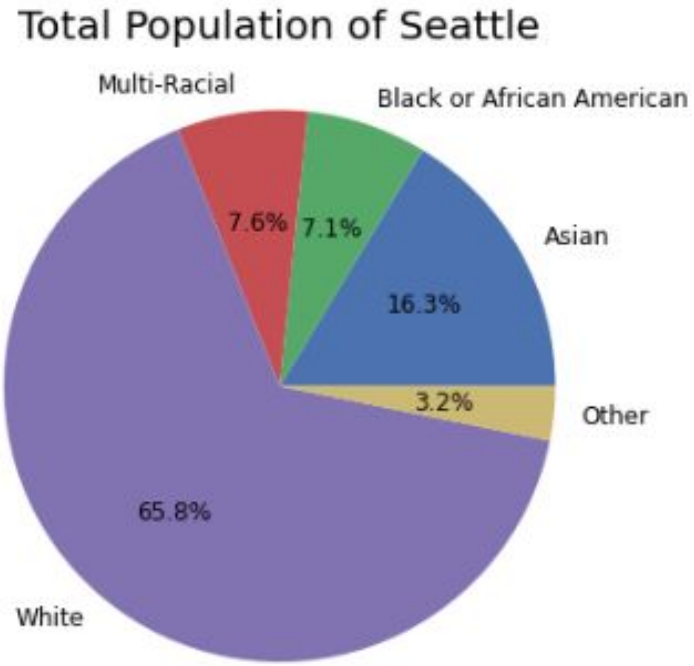

- Initial data analysis showed some basic findings about arrest data at Traffic Stops
- This shows how many Subjects of each race are stopped vs Arrested at a stop

# Results - Data Visualization



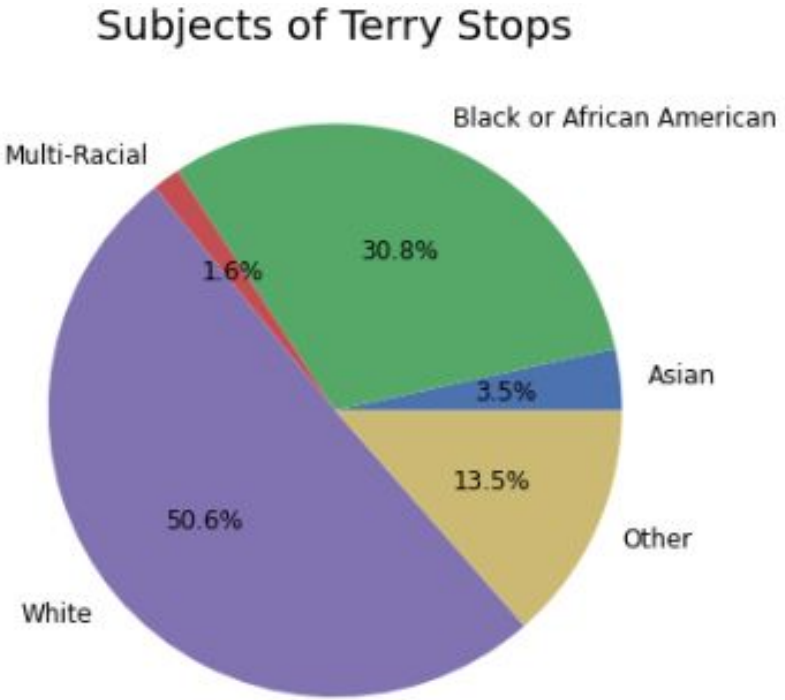

- Initial data analysis showed some basic findings about arrest data at Traffic Stops
- This shows how many Subjects of each race face legal action when stopped

# Results - Data Visualization




- The prior results seem relatively unbiased until we learn that Seattle's population is **65% White**, showing that our Terry Stop Data has been drastically affected by racial bias.

# Results - Data Visualization



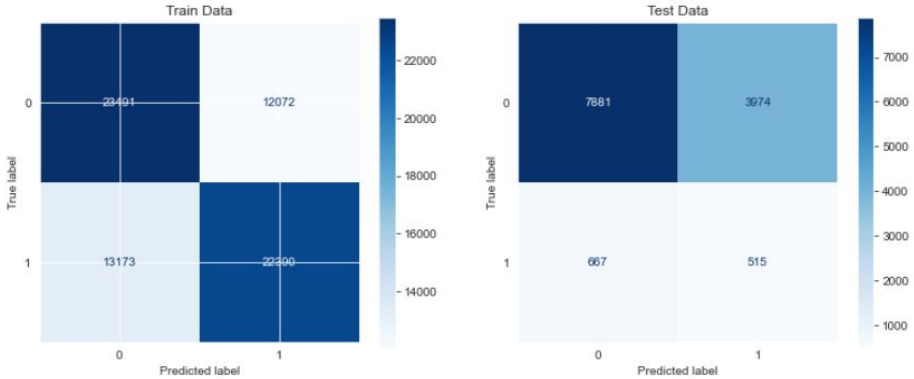

- For example, if Black or African American subjects only make up **7.1%** of the population, why do they make up **nearly 31%** of the Terry Stop Data?

# Data Modeling

Target: ‘Physical Arrest’
Ideal Metrics: F1, Precision

# Results - Baseline Model

## Baseline Logistic Regression Model (with Synthetic Data)



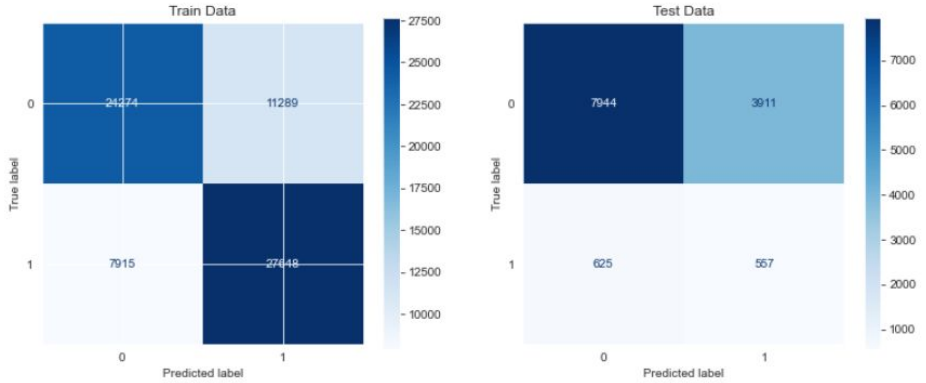

Train Data:
Accuracy: 0.6450665017012063
Recall: 0.6295869302364817
Precision: 0.6497011200742847
F1: 0.6394858978936094

Test Data:
Accuracy: 0.644013193219299
Recall: 0.4357021996615905
Precision: 0.11472488304744932
F1: 0.18162581555281254

- Baseline scores without synthetic data were too heavily affected by data imbalance
- None of these scores are ideal. Next we check scores with other classifiers

# Results - Testing Classifiers

## XG Boost Model Performance (with Synthetic Data)



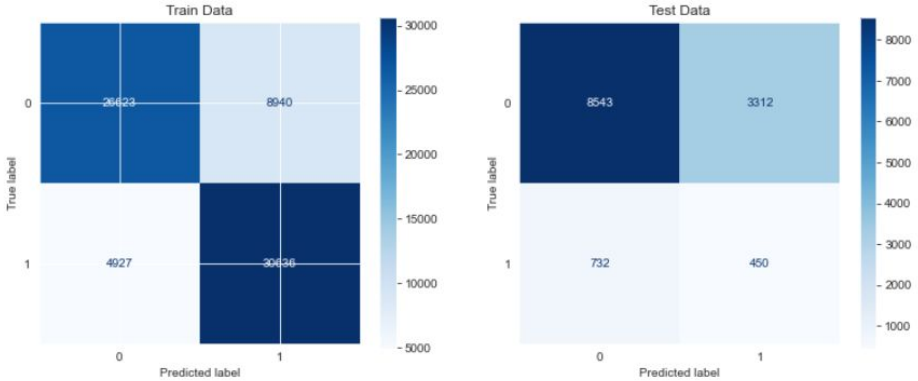

Train Data:
Accuracy: 0.7300002811911256
Recall: 0.777437224081208
Precision: 0.7100701132598813
F1: 0.7422281879194631

Test Data:
Accuracy: 0.6520671933727085
Recall: 0.4712351945854484
Precision: 0.1246642793196061
F1: 0.19716814159292034

- Only a slight increase in scores with best model tested
- F1 Score still performing incredibly sub-optimally
- Tuning hyperparameters of models may improve performance

# Results - Hyperparameter Tuning

## Tuned XG Boost Performance




Train Data:
Accuracy: 0.8050361330596406
Recall: 0.861457132412901
Precision: 0.7741055184960582
F1: 0.8154487017394428

Test Data:
Accuracy: 0.6898059369486845
Recall: 0.38071065989847713
Precision: 0.11961722488038277
F1: 0.18203883495145629

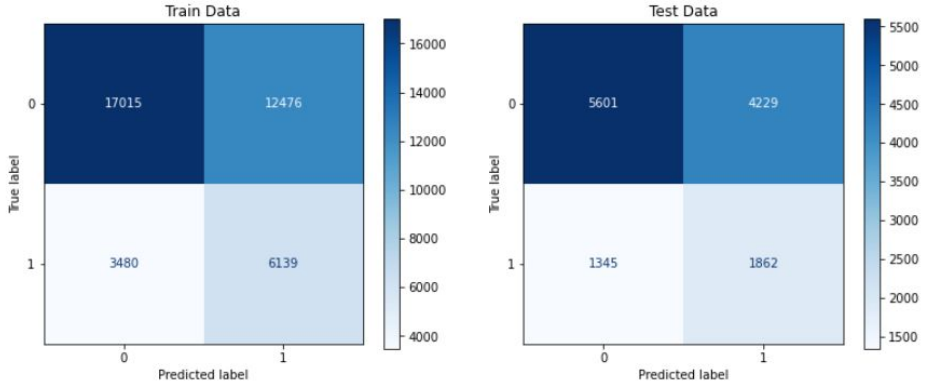
- Once again, only a slight improvement, this time in Accuracy
- Discrepancy between Trian and Test Data shows that with the current feature set, models are not successfully predicting outcomes

# Data Modeling

New Targets:
“Arrested” and “Legal Action Taken”

# Results - Target: ‘Arrested’

## XG Boost Model Performance (with Synthetic Data)



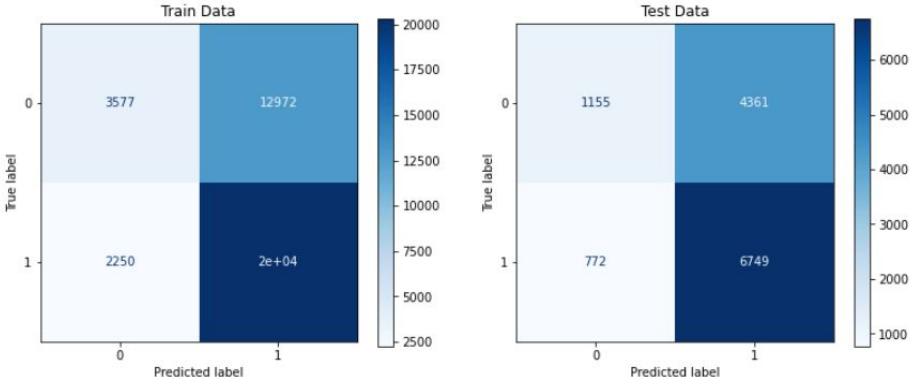

```
Train Data:
Accuracy:  0.5920225006392227
Recall:    0.6382160307724296
Precision: 0.32978780553317216
F1:        0.43486576468088123
```

```
Test Data:
Accuracy:  0.5724476489990028
Recall:    0.5806049267227938
Precision: 0.3056969298965687
F1:        0.40051624005162395
```

- Significant improvement in scores, though still not ideal
- Feature set with less class imbalance seems to be the main cause of increased scores.

# Results - Target: ‘Legal Action Taken’

XG Boost Model Performance (No Synthetic Data)




Train Data:
Accuracy: 0.6107900792636154
Recall: 0.9002703780860778
Precision: 0.610251479734399
F1: 0.727419239309505

Test Data:
Accuracy: 0.6062744496433229
Recall: 0.89735407525595
Precision: 0.6074707470747075
F1: 0.724491438999517

- Best scoring model so far
- High F1 (comparatively), but lower Accuracy and Precision than needed to claim to be a successful classification model.

# Final Observations

- The dataset as it existed in this analysis was sub-optimal for creating strong classification models.
- Data imbalance seemed to be the primary cause of low scores, and adding synthetic data did not increase scores as much as desired.
- Changes to the feature set and target variable were most effective, showing that future analyses could adjust these further to yield even stronger results.

# Conclusions

The analysis did yield the results it needed in creating classifiers to accurately predict Terry Stop Outcomes. This is could be caused by:

- Varied and inconsistent record keeping in Seattle.Gov's dataset make it hard to clean and interpret
- The need to investigate the dataset with even more feature sets and target variables

Next steps could include using different ***multi-target Classifiers*** to yield more meaningful results.

# Thank you

**Email:** shadel96@gmail.com
**GitHub:** @shadel96
**LinkedIn:** linkedin.com/in/spencer-hadel/