Terry STOP Prediction Project

By Alex Muturi

Table of contents

Business Understanding

What is a Terry Stop?

It’s a legal investigative ‘stop and frisk’ that a police officer is allowed to make if they have reasonable suspicion about a suspect.

Why make a Terry Stop?

A successful stop leads to the rightful arrest of a suspect, helping an officer clear and prevent crimes.

Business Understanding

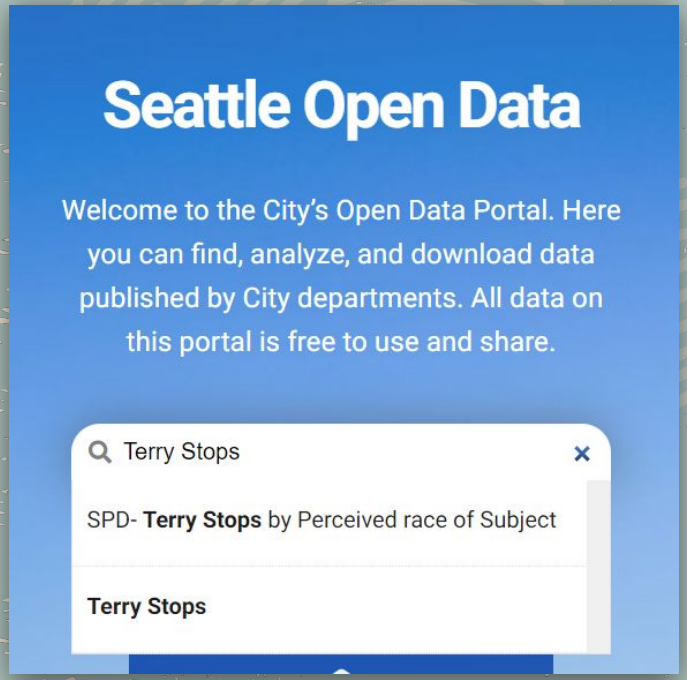
What is reasonable suspicion?

Based on prevailing factors such as the environment, time of day and subject particulars, an officer can use their reasoning to decide to stop a subject.

What does this project have to do with it?

The project uses machine learning algorithms to predict whether an arrest should be made based on the prevailing factors.

Data Understanding



The data was obtained from Seattle Open Data portal.

It represents records of police reported stops under Terry v. Ohio, 392 U.S. 1 (1968).

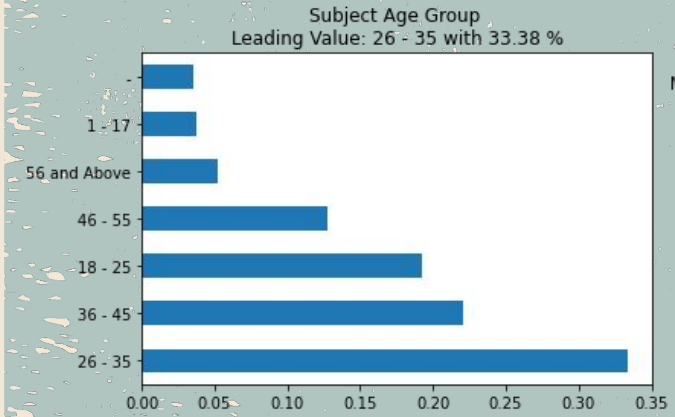
Each row represents a unique stop.

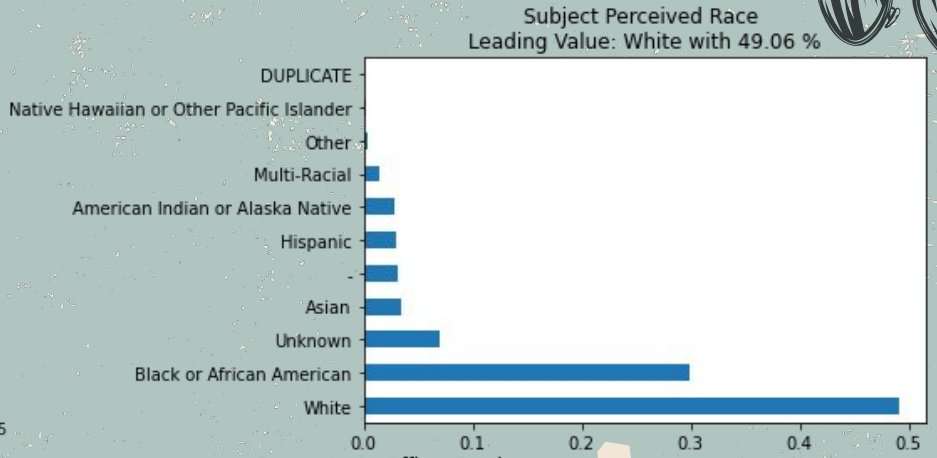
Selected Columns

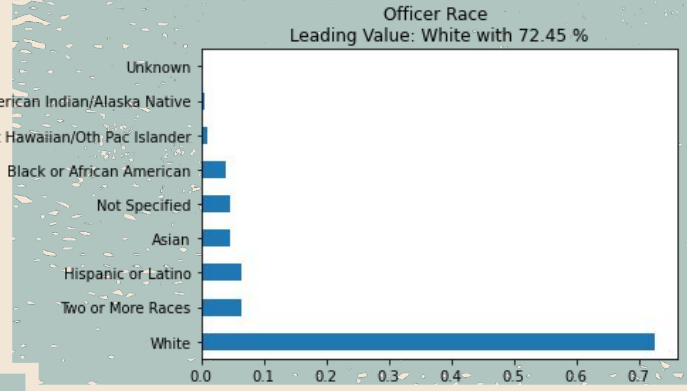
- Subject Age Group
- Subject Perceived Race
- Subject Perceived Gender
- Reported Date
- Reported Time
- Weapon Type
- Officer Gender
- Officer Race
- Officer YOB
- **Arrest Flag**

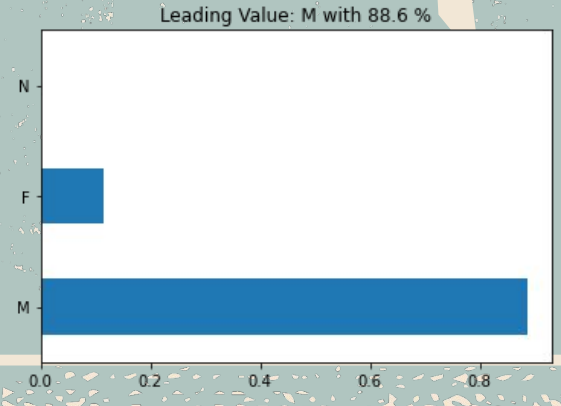
*This our target variable

Exploratory Data Analysis





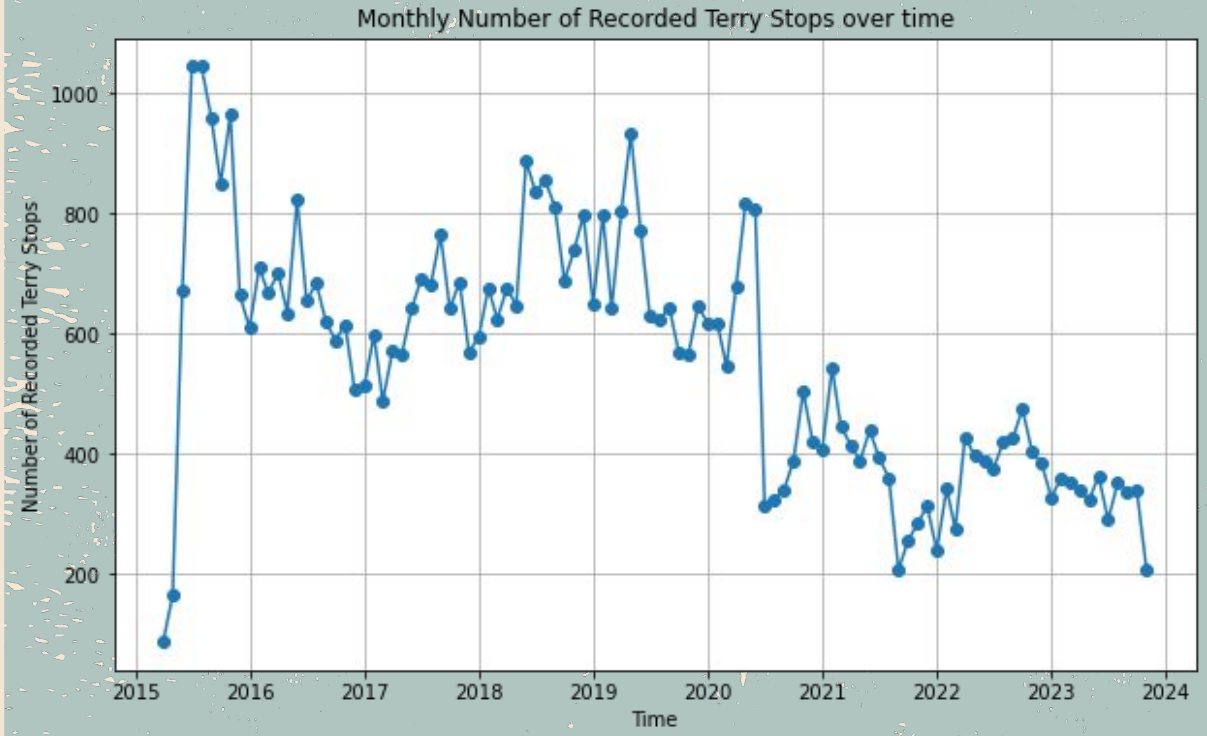


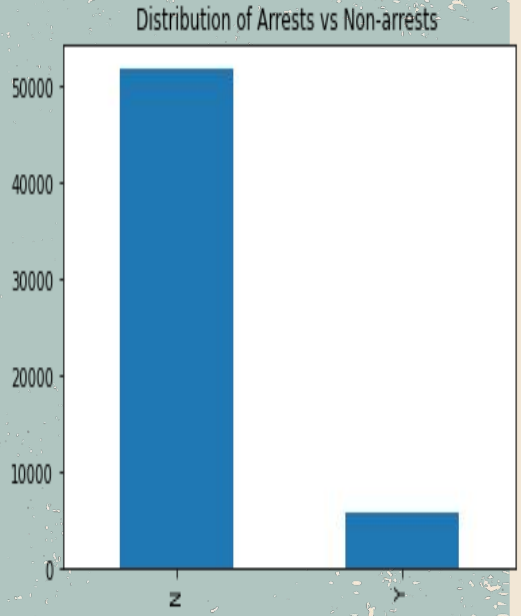


Quick Insights

1. Most Terry stops (90%) ended up with **no arrests**. This may suggest class imbalance
2. **Male subjects** were stopped most, at 79%
3. People with '**White**' as their perceived race were stopped most, at 49%
4. Officers with '**White**' as race on record made the most stops, at 49%
5. The most common subject age group was 26-35 years (33%)
6. More than 85% of subjects had **no weapons**

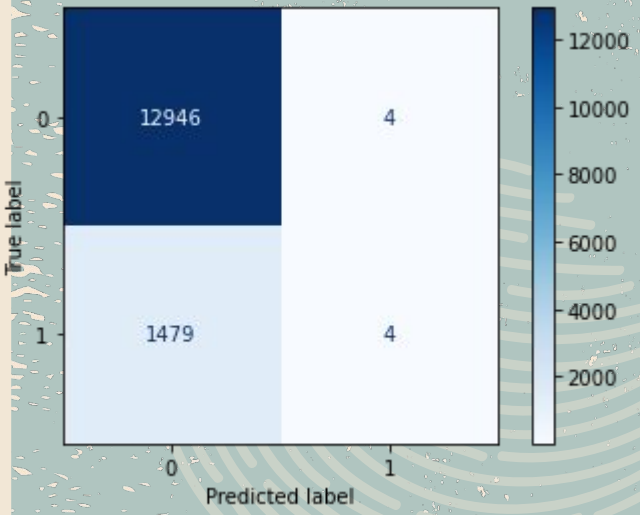
Monthly Terry Stops Trend, 2015 to Date





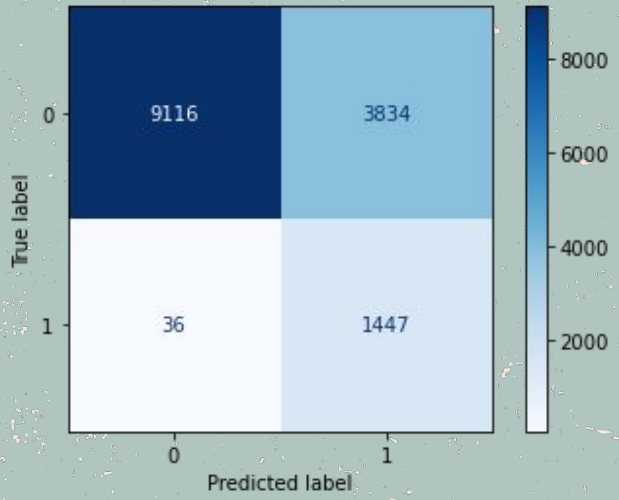
Modelling Results

Model I



Baseline Logistic Regression model

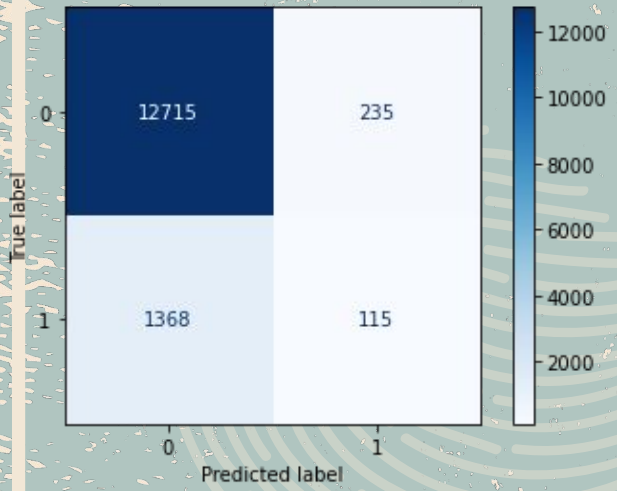
Model II



Weighted Logistic Regression

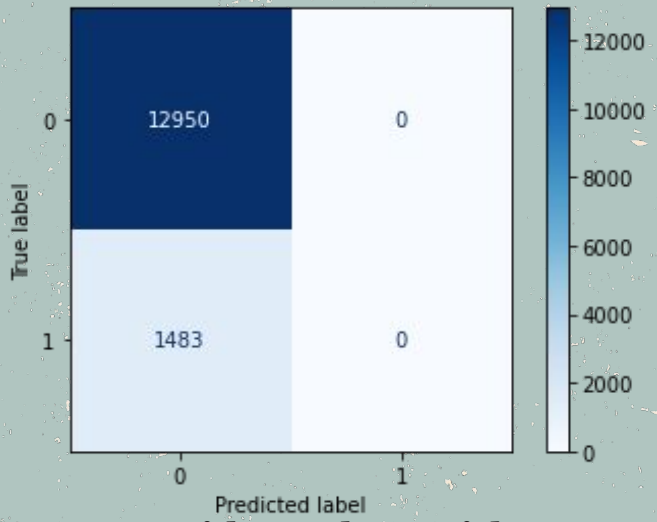
Modelling Results

Model III



Baseline Decision Tree

Model IV



GridsearchCV with Decision Trees

Evaluation

Model 3, logistic regression with inverted class weights emerged the best for use in our case.

While it had a 27% precision, it also had 98% recall which means we would be arresting almost everyone that needs to be. However, quite a number of innocent people would be inconvenienced which leads us to the recommendations

Conclusion & Recommendations

Predictions using the chosen model would help police officers reduce crime by stopping and arresting suspects.

<u>Recommendation</u>

The dataset could benefit from more information such as an area's crime index as well as more information about the suspects

Thanks!

Do you have any questions?

github.com/alexmuturi

Please keep this slide for attribution