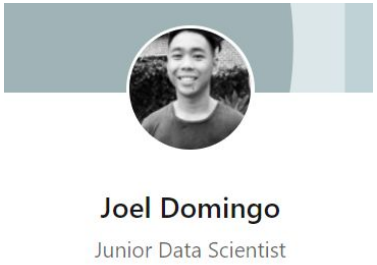
# IOD Capstone Project:
## East London Police Use of Force Dataset

Joel Domingo

# About me and Agenda

- Joel Domingo
- 24 year year old Graduate
- Studied IT and Accounting at University of Technology Sydney (2020)




UNIVERSITY OF TECHNOLOGY SYDNEY

## Agenda:

- Introduction
- About the dataset
- Preliminary Insights
- Modelling
- Conclusions

# Introduction

*"Studies show law enforcement with some college education use force much less often than those with no higher education..."*

*"... only 1% of police forces within the United States have education requirements for those looking to join their forces."*

Rydberg, Jason; Terrill, William (2010). "The Effect of Higher Education on Police Behavior". *Police Quarterly*. **13** (1): 92–120.

**Main Stakeholders:**

- Metropolitan Police Department (London)
- Police Agencies (worldwide)
- Civilians (their safety)
- Officers
- Government (funding changes)

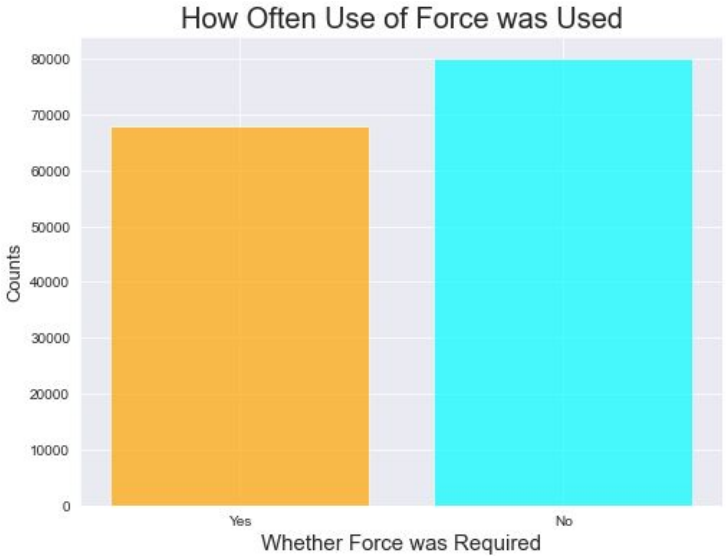

## Goal:

- To explore the **factors** which cause Police to use force, and **aim to reduce** that **internally** through resource allocation, and training and modelling.
- Create a model which can **predict whether use of force** may be **required**, and give **insights and prepare** officers responding to incidents.

# About the Dataset

- Dataset is a collection of features gathered from incident reports from the London Metropolitan Police Service, hosted by the London Government (source: https://data.london.gov.uk/dataset/use-of-force)
- Contains data ranging from 1st April 2020 to 31 January 2021.
- Contains 147,895 records of officer accounts, each with 271 features.

# Insights: Use of Force Frequency



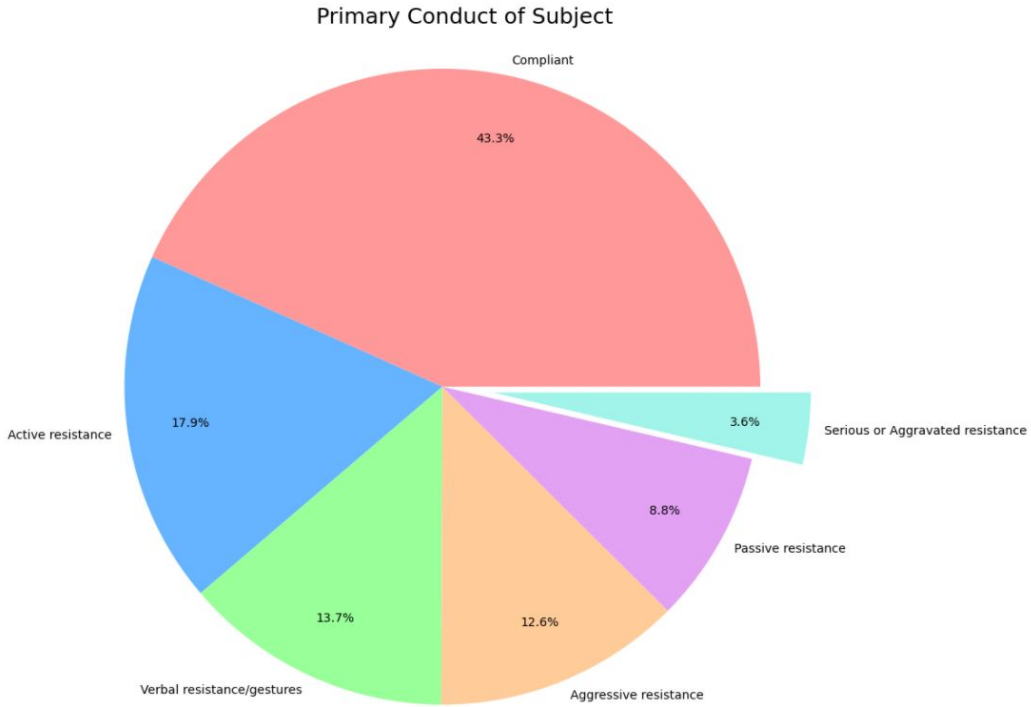

**Q: What counts as 'Use of Force'?**

- Anything beyond **compliant handcuffing** and **tactical communications.**

| | **Totals** |
|---|---|
| **Required Force** | 67,950 |
| **Did not require Force** | 79,945 |
| **Total** | **147,895** |

# Insights: Subjects Analysis

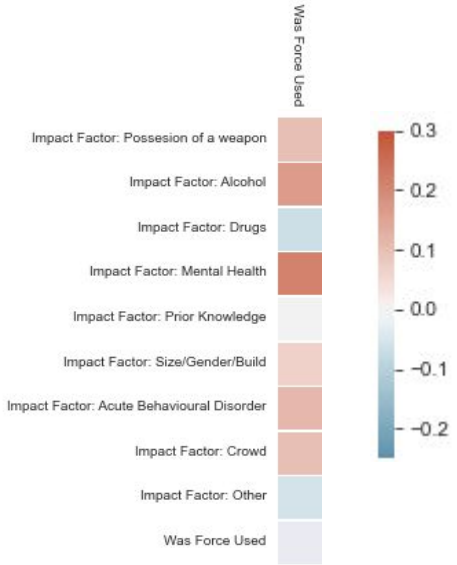
| Conduct | Counts |
|---|---|
| Compliant | 64,058 |
| Active Resistance | 26,525 |
| Verbal Resistance | 20,243 |
| Aggressive Resistance | 18,655 |
| Passive Resistance | 13,017 |
| Serious of Aggravated Resistance | 5397 |



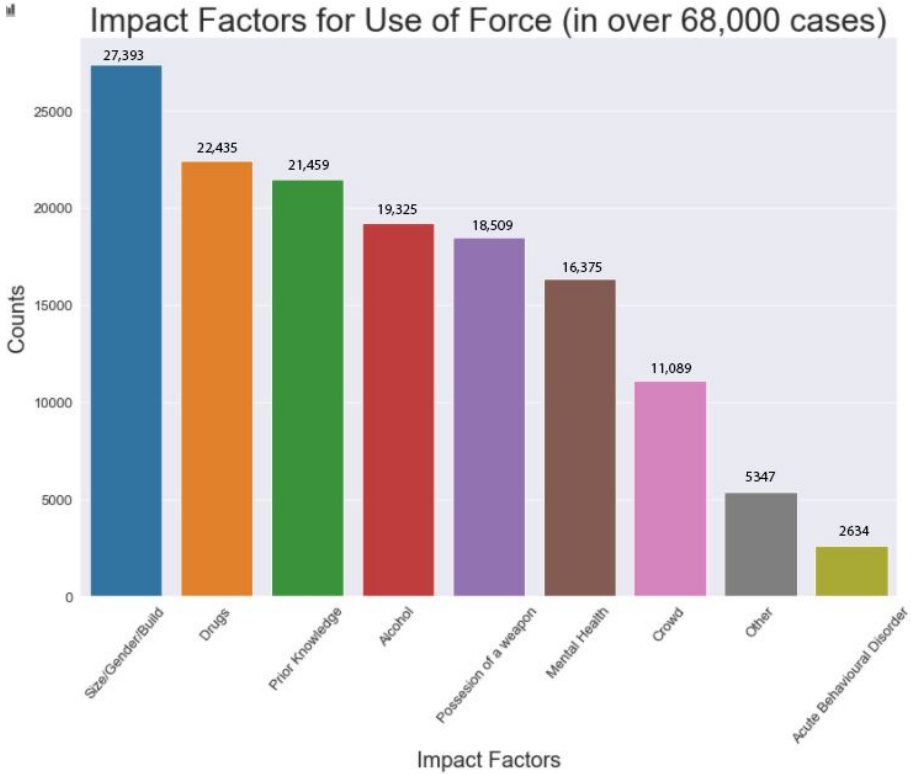

Further subject demographic analysis can be found in the provided report.

# Impact Factors for Use of Force

Correlation:

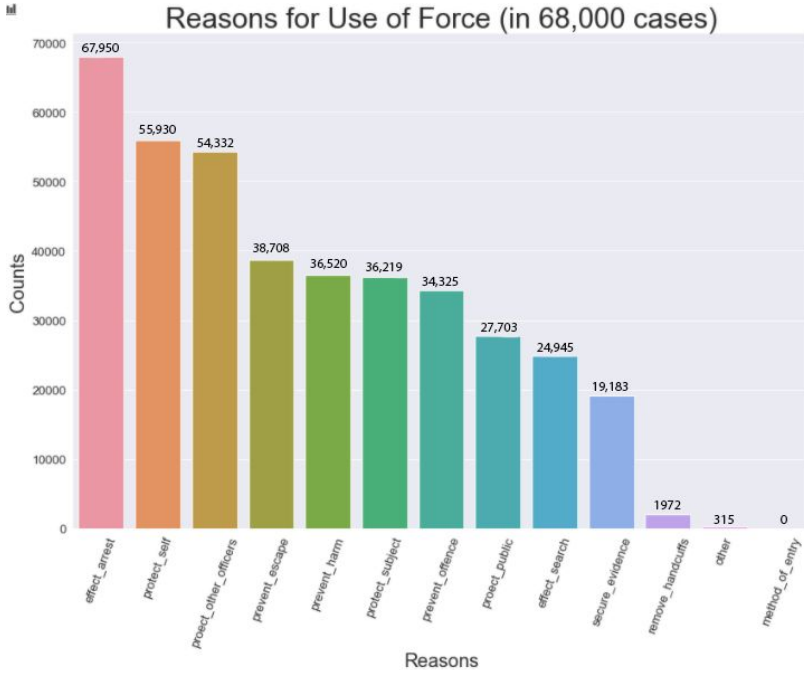


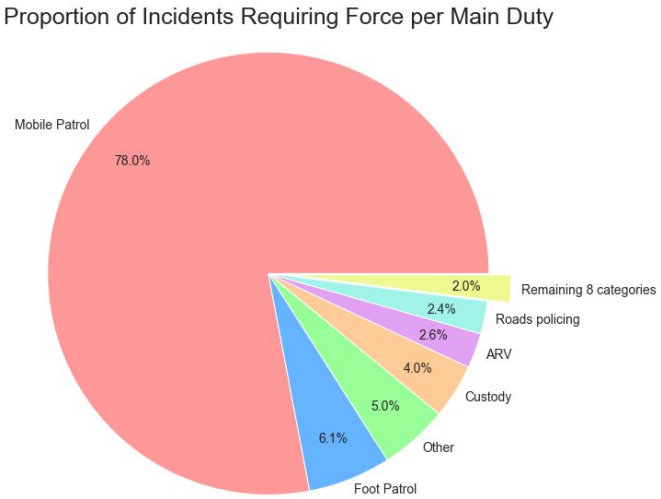

| Impact Factors | Counts |
|---|---|
| **Size/Gender/Build** | 27,393 |
| **Drugs** | 22,435 |
| **Prior Knowledge** | 21,459 |
| **Alcohol** | 19,235 |
| **Possession of Weapon** | 18,509 |



# Why are police required to use force?

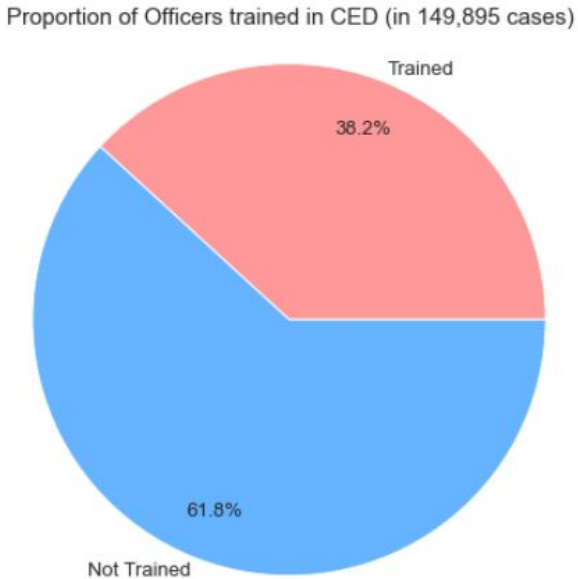
| Reasons | Counts |
|---|---|
| Effect arrest | 67,950 |
| Protect self | 55,930 |
| Protect Other officers | 54,332 |
| Prevent Escape | 38,708 |
| Prevent harm | 36,520 |
| Protect subject | 36,219 |
| Prevent offence | 34,325 |
| Protect public | 27,703 |
| Effect search | 24,945 |
| Secure evidence | 19,183 |
| Other | 315 |
| Method of entry | 0 |

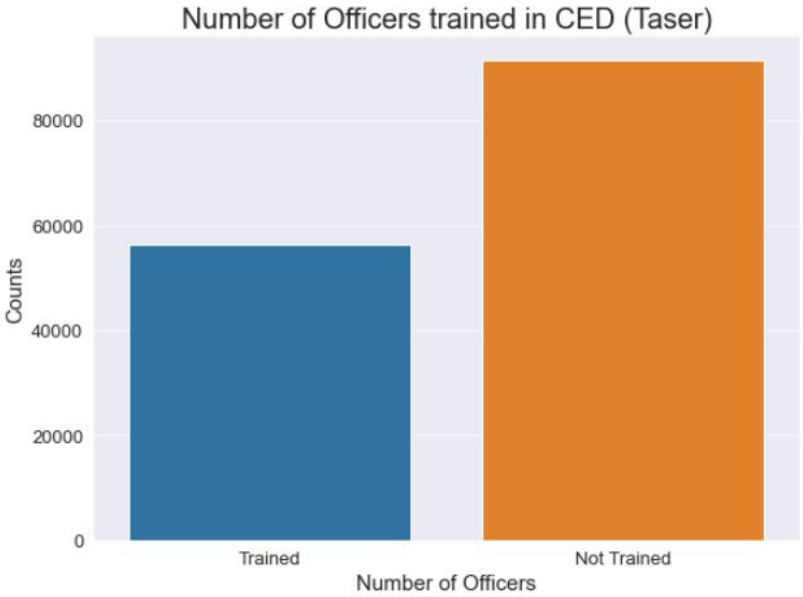


# Taser Use

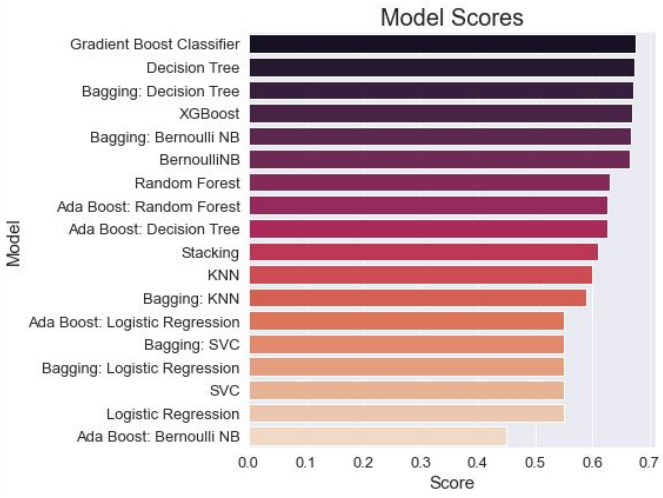
| Trained in Taser? | Counts |
|---|---|
| **Trained** | 56,450 |
| **Not Trained** | 91,445 |



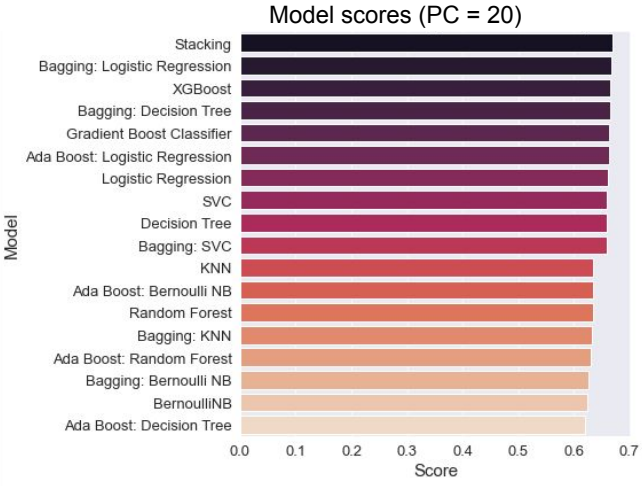



# Modeling:

| | Scores |
|---|---|
| Gradient Boost Classifier | 0.676 |
| Decision Tree | 0.674 |
| Bagging: Decision Tree | 0.672 |
| XGBoost | 0.670 |
| Bagging: Bernoulli NB | 0.668 |
| BernoulliNB | 0.666 |
| Random Forest | 0.630 |
| Ada Boost: Random Forest | 0.627 |
| Ada Boost: Decision Tree | 0.626 |
| Stacking | 0.609 |
| KNN | 0.600 |
| Bagging: KNN | 0.590 |
| Ada Boost: Logistic Regression | 0.550 |
| Bagging: SVC | 0.550 |
| Bagging: Logistic Regression | 0.550 |
| SVC | 0.550 |
| Logistic Regression | 0.550 |
| Ada Boost: Bernoulli NB | 0.450 |



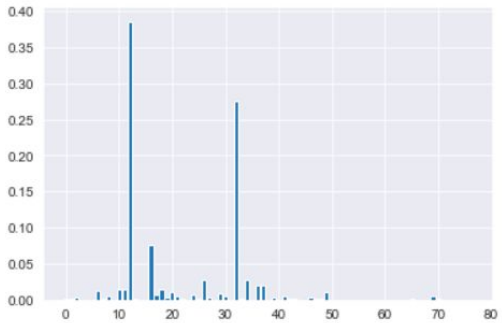

| | Scores |
|---|---|
| Stacking | 0.669 |
| Bagging: Logistic Regression | 0.667 |
| XGBoost | 0.666 |
| Bagging: Decision Tree | 0.666 |
| Gradient Boost Classifier | 0.663 |
| Ada Boost: Logistic Regression | 0.662 |
| Logistic Regression | 0.660 |
| SVC | 0.659 |
| Decision Tree | 0.658 |
| Bagging: SVC | 0.658 |
| KNN | 0.635 |
| Ada Boost: Bernoulli NB | 0.635 |
| Random Forest | 0.634 |
| Bagging: KNN | 0.632 |
| Ada Boost: Random Forest | 0.630 |
| Bagging: Bernoulli NB | 0.626 |
| BernoulliNB | 0.624 |
| Ada Boost: Decision Tree | 0.620 |




- Conducting dimensionality reduction reduced processing time slightly.
- It improved the average result of all models, with more homogeneous results.
- However, the top models were unable to exceed the results without dimensionality reduction.

# Final Model and Use Case - Decision Tree or Gradient Booster

| | Scores |
|---|---|
| Decision Tree | 0.662 |
| Gradient Booster | 0.662 |




**Hypothetical Scenario:** Officer responding to call with the following features. Using the models, predict whether the officer should expect/be prepared to use force on the subject.

- At Sports Stadium
- Officer called is mobile patrol
- They are Single Crewed
- subject age 24
- No disability
- Male
- Asian
- In Lewisham

| | Outcome | Chance that Officer required to use force |
|---|---|---|
| Decision Tree | 0 | 0.28 |
| Gradient Booster | 0 | 0.40 |

# Conclusions and Some recommendations:

- Analysis revealed observations significant enough to make recommendations to police agencies
- Large frequency of use of force is largely factored towards **unpreparedness**
- The goal is to find ways to reduce risk to both officers and public regarding safety, and we can do this by increasing preparedness
- This model and analysis varies on an international scale, but the fundamental principles and features can be applied anywhere in the world.

**Key Recommendations:**

- Implement further training in tactical communications
- Implement education requirements prior to successful recruitment
- Implement further training on dealing with mentally-ill individuals
- Implement suggested model and distribute to all London Police departments.