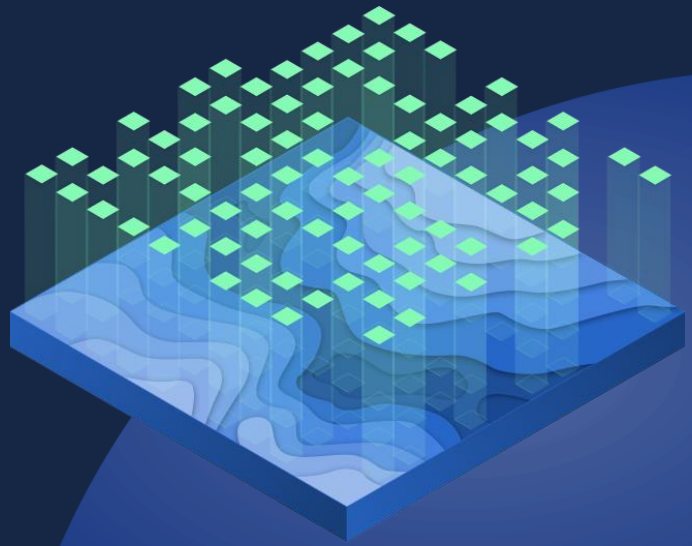
CSI:

CRIMINO-SPATIAL INVESTIGATORS



CASA0007: Quantitative Methods

Our mission:

Debunk crime myths with data

Fact or Myth?

Q1 **Nowhere is safe in London**

Q2 **Westminster has seen crime increasing over the years**

Q3 **Wealthier neighbourhoods see less crime**

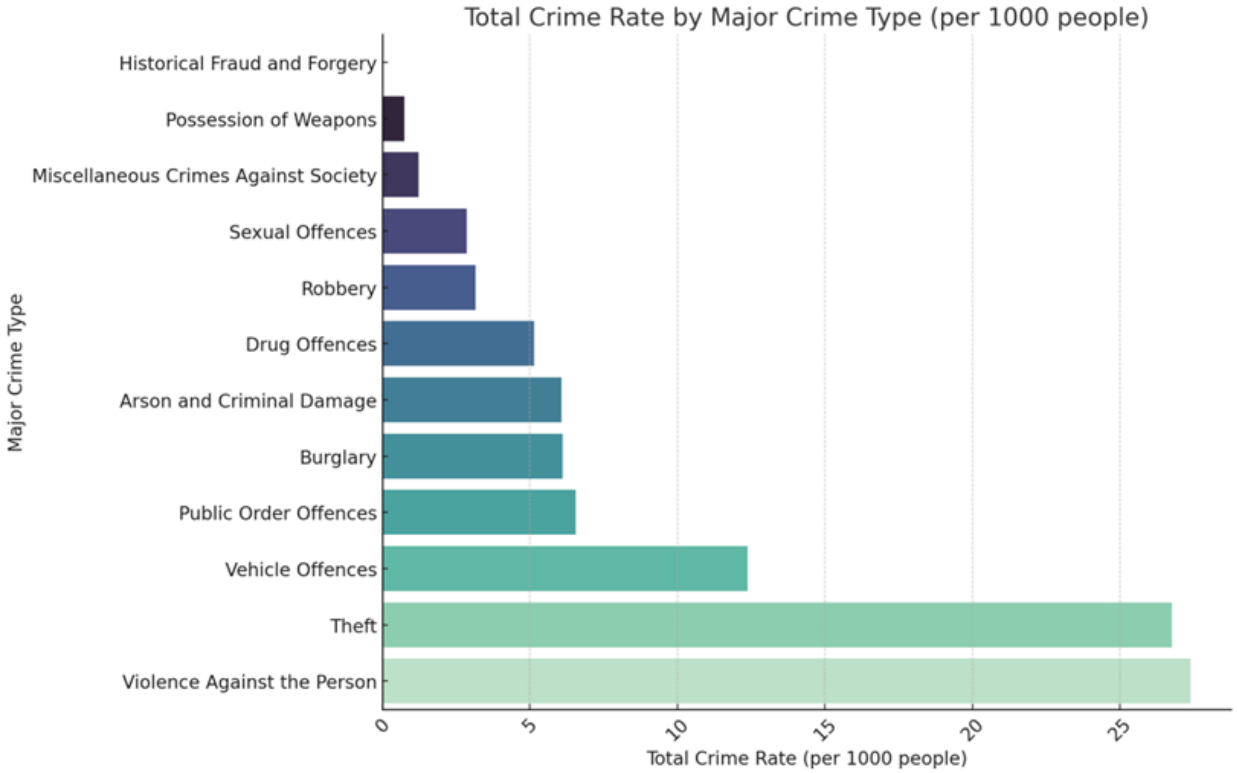
Datasets used in this research

Dataset	Source	Contains	Counting method	quantity
Crime in London 2010-2023 (Crime Numbers)	**Metropolitan Police Service**	**Date**	**Monthly**	**2010.4 - 2023.10**
		Crime Type	**Major & Minor**	**12 Major types**
Population Density	Office for National Statistics	Date	Yearly	2012-2022
		Density	Every 5 years	2011/2016/2021
Workless Households (Proportions)	Office for National Statistics	Date	Yearly	2012-2022
		Percentage	Households (Working/Mixed/ Workless)	2012-2022
House Price (£)	Land Registry	Date	Monthly	1995.1 - 2023.9
GCSE results (Proportions)	Department for Education	Date	Yearly	2012-2022
		Pupils by gender	Number/Percentage (English and maths)	Number of people
Residence Weekly Earning (£)	NOMIS	Date	Yearly	2012-2022
		Earnings	Median	2012-2022

***All data are divided by 32 boroughs.**

Q1:

"Nowhere is safe in London"



Why 2022?

- Most Up-to-Date Information
- Post-Pandemic Analysis
- Comparative Baseline

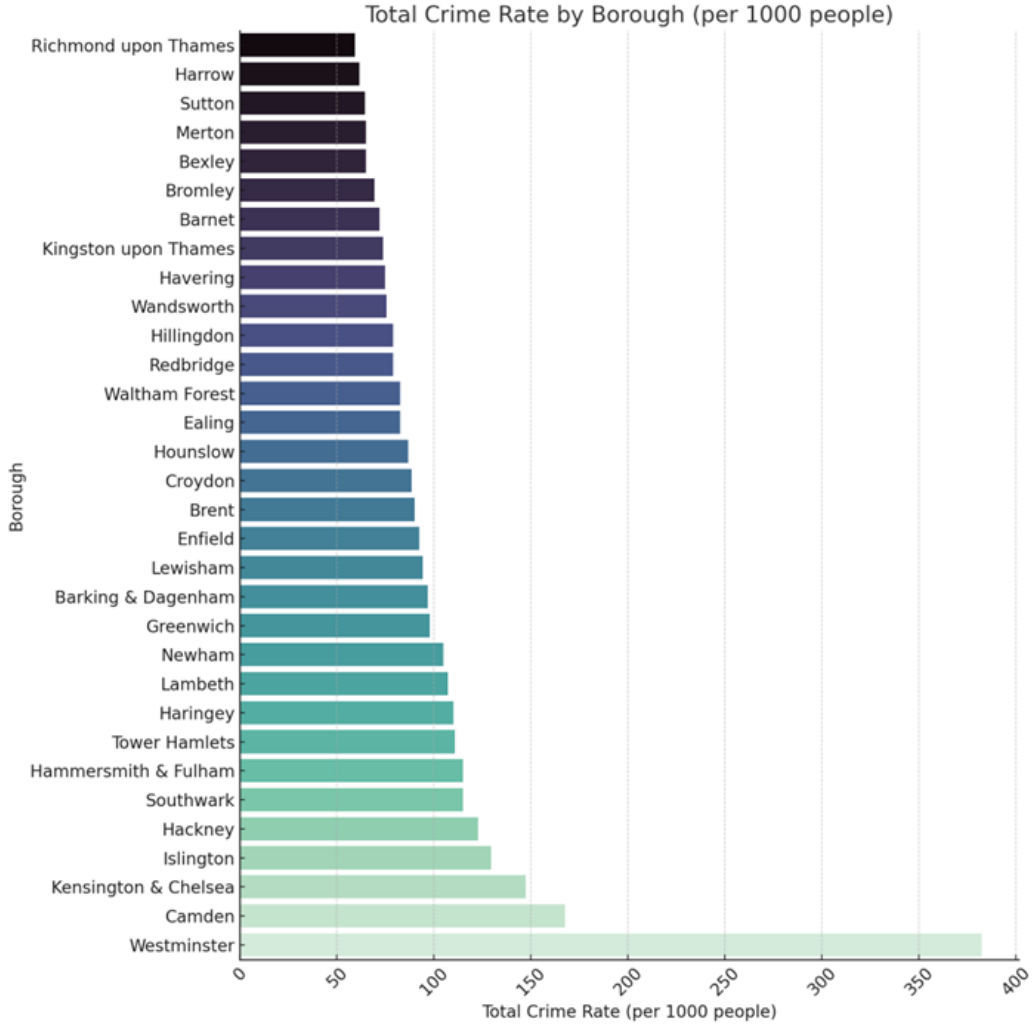
The fact is: In 2022, there were a total of 864,422 recorded incidents of crime in London.

But!

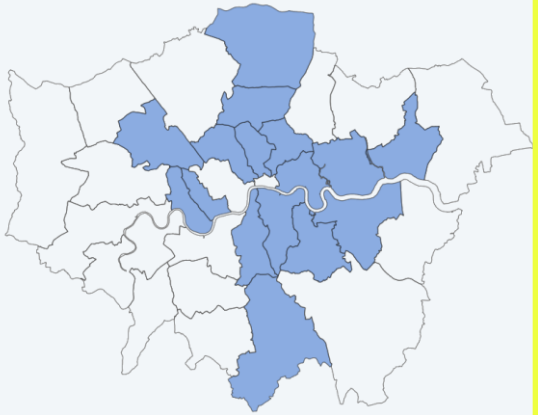
ANOVA Results
Statistic: 26.51
P-value: $1.51e^{-40}$

→ Statistically significant differences in crime rates for **different types of crimes across the boroughs of London.**

So, let's do some clustering!



K-means Clustering of London Borough Crime Rates in 2022



"Urban Dynamo"

- High dynamism and urbanization
- Vibrancy and busyness
- Higher crime rates

"Central Spotlight"

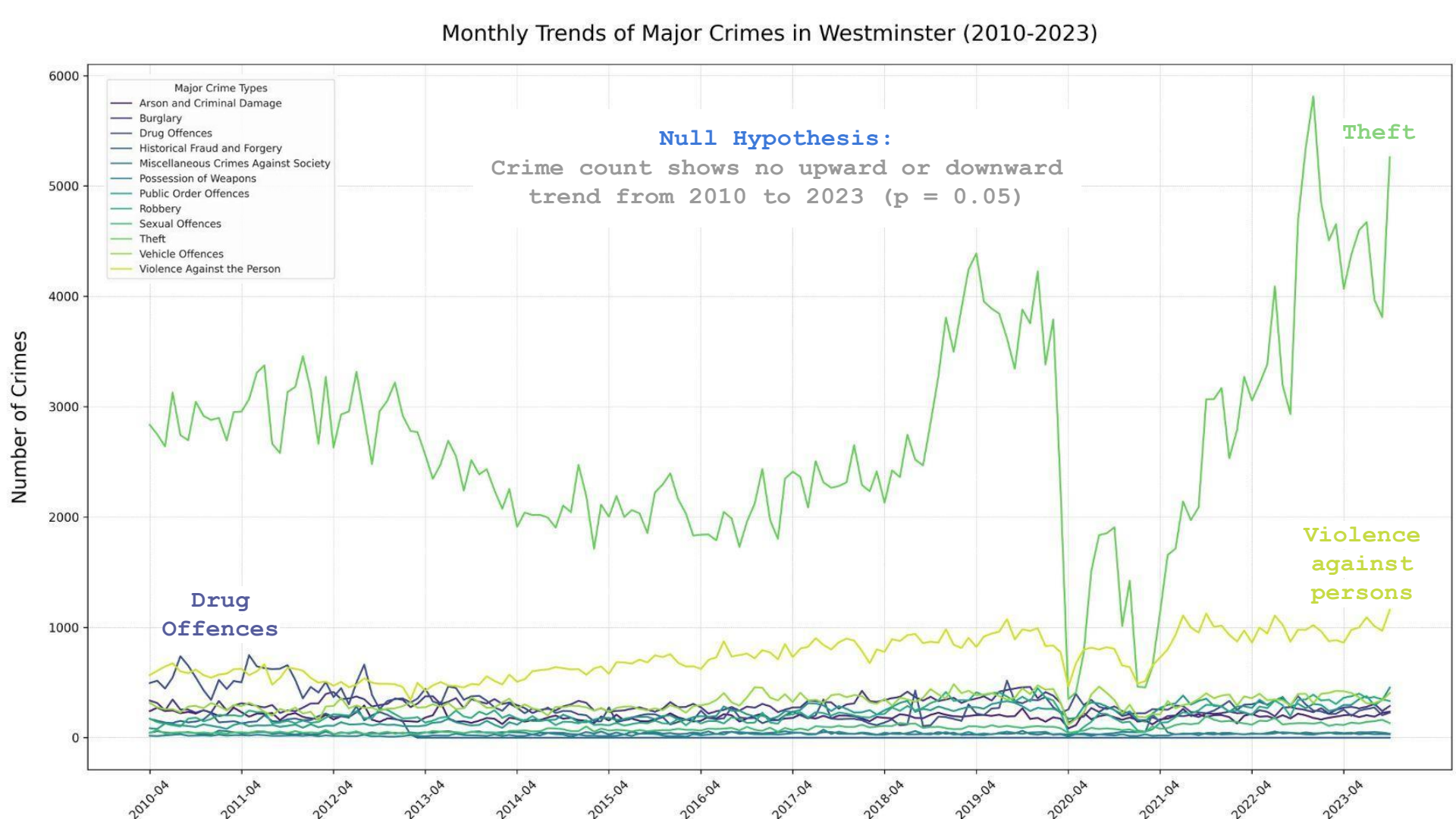
- Only Westminster!
- Heart of London
- Extremely high crime rates

"Suburban Oasis"

- Quiet and peaceful
- Suburb area
- Lower crime rates

Q2:

"Westminster has become more and more unsafe over the years"


Monthly Trends of Major Crimes in Westminster (2010-2023)
Major Crime Types
Arson and Criminal Damage
Burglary
Drug Offences
Historical Fraud and Forgery
Miscellaneous Crimes Against Society
Possession of Weapons
Public Order Offences
Robbery
Sexual Offences
Theft
Vehicle Offences
Violence Against the Person
Null Hypothesis:
Crime count shows no upward or downward trend from 2010 to 2023 (p = 0.05)
Theft
Violence against persons
Drug Offences
Number of Crimes
6000
5000
4000
3000
2000
1000
0
2010-04
2011-04
2012-04
2013-04
2014-04
2015-04
2016-04
2017-04
2018-04
2019-04
2020-04
2021-04
2022-04
2023-04

Mann-Kendall test for Trend Detection

assumes no serial autocorrelation, therefore we must check ACF plots

No auto-correlation

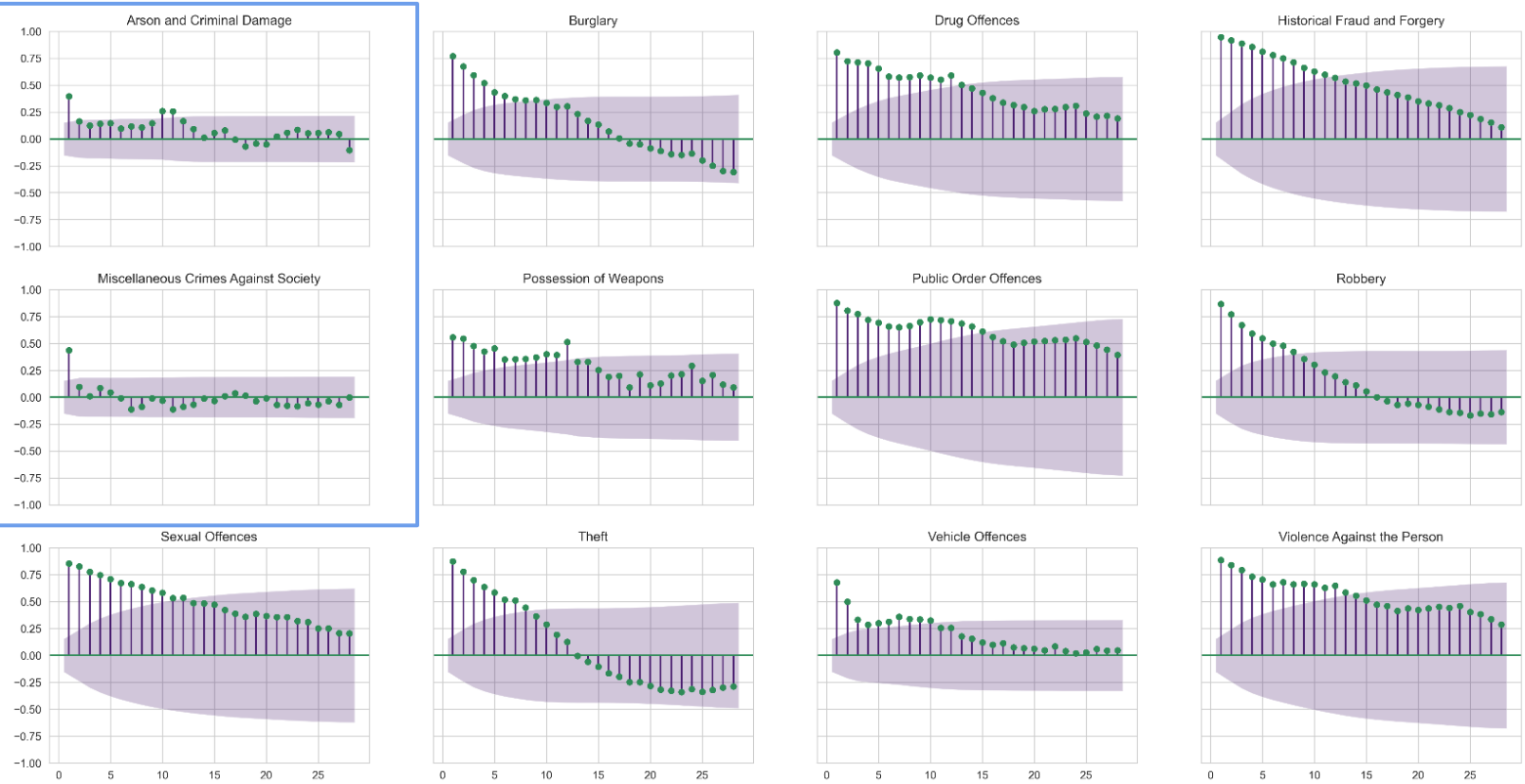
↓

Mann-Kendall test

Have auto-correlation?

↓

Modified Mann-Kendall tests



Reference: https://abhinaya-sridhar-rajaram.medium.com/mann-kendall-test-in-python-for-trend-detection-in-time-series-bfca5b55b

Crime Type \ MK Test Type	Mann Kendall Test	Modified MK Test Hamed Rao Approach	Modified MK Test Yue Wang Approach	Seasonal MK Test
Theft		no trend	no trend	no trend
Arson and Criminal Damage	no trend			no trend
Historical Fraud and Forgery		no trend	decreasing	decreasing
Burglary		no trend	decreasing	decreasing
Drug Offences		no trend	decreasing	decreasing
Miscellaneous Crimes Against Society	decreasing			decreasing
Possession of Weapons		increasing	increasing	increasing
Public Order Offences		increasing	increasing	increasing
Robbbery		increasing	increasing	increasing
Sexual Offences		increasing	increasing	increasing
Vehicle Offences		increasing	increasing	increasing
Violence Against the Person		increasing	increasing	increasing

Q: So, has Westminster become more unsafe over the years?

A: Well, depends…

Theft, though high count, shows **no significant trend***

Arson, **Fraud**, **Burglary**, **Drugs**, **etc.** show **no trend** or have **decreased**

Other crime types, **incl. Violence** sadly seem to have **increased**

(*) MK test statistics may be affected by fluctuations during and after COVID lockdown

Q3:

"Wealthier boroughs have less crime"

Multiple Linear Regression for Factors Detection

Description of Variables for Regression

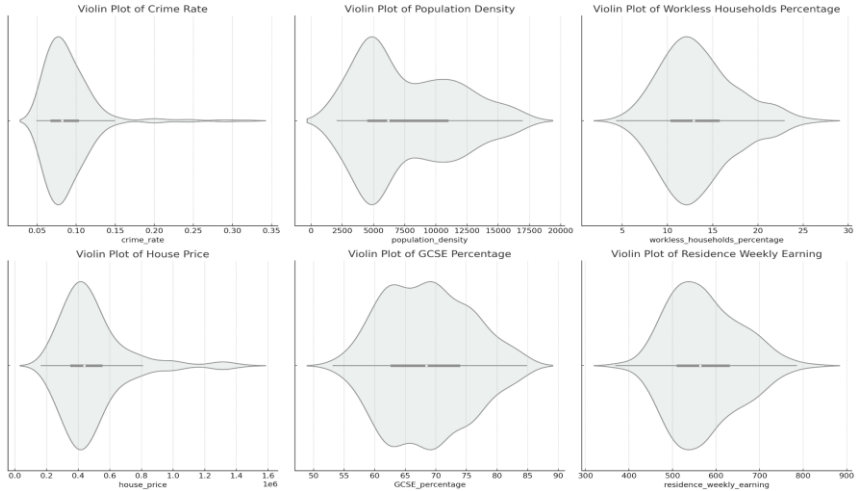
	Name	Type	Time span	Meaning
Dependent	**crime_rate**	continuous	2012-2022	Number of crimes/population in each borough
Independent	**population_density**	continuous	2012-2022	Population density in each borough
	workless_households_percentage			Proportions of households that are classified as workless in each borough
	house_price			Average house price in each borough (£)
	GCSE_percentage			Percentage of pupils who achieved a standard 9-4 pass (2016-2022); Percentage of pupils who achieved A*-C in English and maths GCSEs including equivalents (2012-2015)
	residence_weekly_earning			Resident based median weekly earnings (£)
	location	categorical		Part of London (Central, East, West, North, South) represented by 4 dummy variables

Slice of Datasets

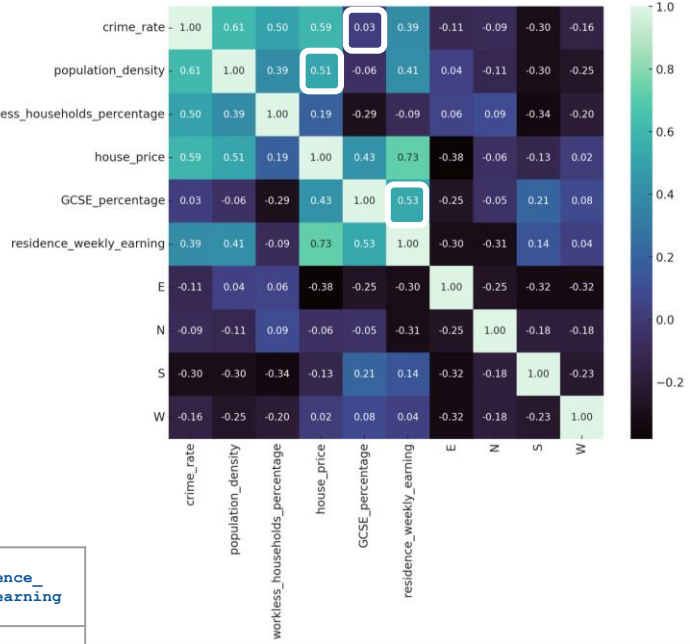
	crime_rate	population_density	workless_households_percentage	house_price	GCSE_percentage	residence_weekly_earning	position
0	0.149004	10324.381770	22.2	586165.9352	59.8	612.0	C
1	0.127191	10542.699980	21.6	667447.5587	61.9	629.5	C
2	0.118717	10777.999010	19.1	756486.4726	62.3	613.1	C
3	0.125663	11063.139030	21.4	787440.9329	57.9	588.1	C
4	0.122390	11298.208590	18.0	807118.9620	63.1	632.5	C

Distribution Plot of Variables for Regression

Violin Plots of Variables for Regression



Correlation Matrix of Variables



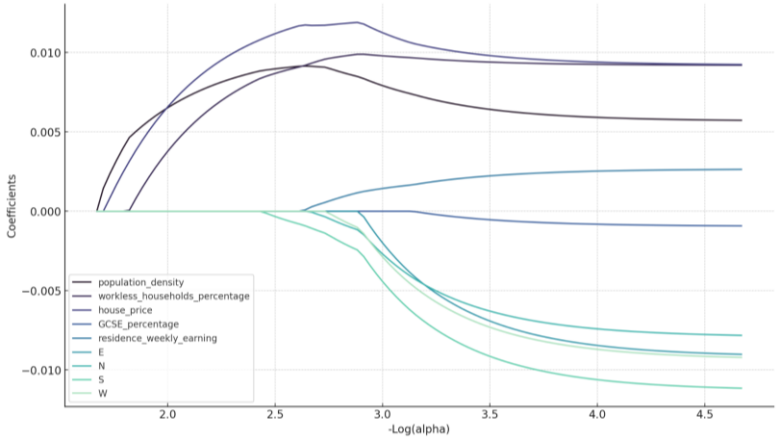
Summary Statistics of Key Fields

	crime_rate	population_density	workless_households_percentage	house_price	GCSE_percentage	Residence_weekly_earning
Mean	0.090	7630.31	13.42	487,220.60	68.53	571.94
Standard Deviation	0.0349	3914.25	4.068	214,905.50	6.88	81.10
Min/Max	0.0498/0.3207	2095.40/16926.82	4.4/26.5	165,863.90/ 1,450,443.00	53.2/84.8	369.80/833.10
Summary	slightly right-skewed	right-skewed	somewhat normal but with a slight right skew	right-skewed distribution	relatively normal	normal distribution

- Most variables are **correlated**, except **GCSE_percentage**
- **Autocorrelation** may exist

Lasso regression

Lasso Paths



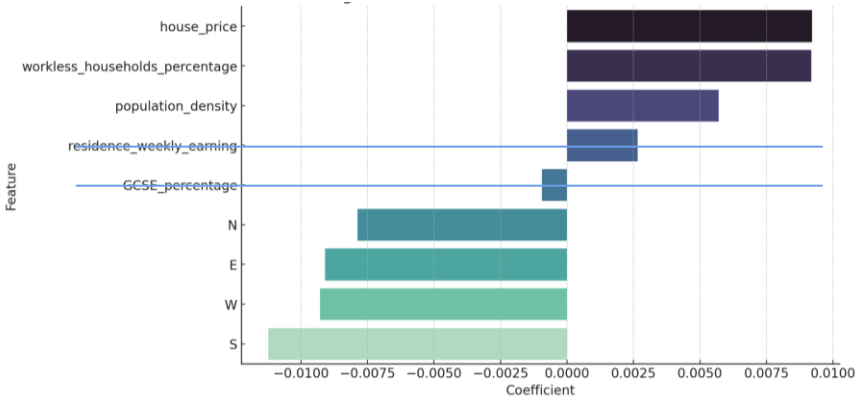
So, it seems that

- the **wealthier** you are, the **more dangerous** boroughs you tend to live in;
- you are generally safe unless you're in **Central London!** Sth London is the safest.

GCSE_percentage,
residence_weekly_earning
insignificant

house_price,
population_density,
workless_households_percentage
significant

Lasso Regression Non-zero Coefficients



Model Fit Evaluation Table

R-squared	0.6155
Adjusted R-squared	0.6051
The cross-validation R-squared value	0.5946
The standard deviation of the cross-validation R-squared value	0.1075
MSE	0.0004788

The model fits well!

Limitations and Further investigations suggested

- Biases from data sources
- Spatial dependence and autocorrelation of variables
- Positivism vs. Structuralism when discussing crime
- Deeper look into causes of crime and also consequences of crime on society

Verdict:

Q1 Nowhere is safe in London. **False!**

Q2 Westminster has seen crime increasing over the years. **Depends!**

Q3 Wealthier neighbourhoods see less crime. **False!**

Stay safe out there!