**Contents**

[Crime Data Analysis Report for Real Estate Decision Making](#_nrt713epcvwv)

[Methodology](#_yxaoeggsxvu)

[1. Executive Summary](#_sdfrc0leadud)

[2. Data Quality Assessment](#_2fxpmkklc956)

[2.1 Data Collection & Preprocessing](#_5p95rg7hcnec)

[Data Processing Workflow](#_40oqmyf8i2fd)

[2.2 Data Cleaning & Missing Values](#_pt0yagc4qhqv)

[3. Exploratory Data Analysis (EDA)](#_ncrgy1wqaypt)

[Top 5 Areas by Crime Volume (excluding 'Unknown'):](#_txxyjdydmtib)

[Figure 2](#_8f9hpx89bxi8)

[3.1 Crime Distribution by Police Force](#_ay21cl76rcqh)

[3.2 Crime Type Analysis](#_imwvs1a2o5ce)

[3.3 Violent Crime Hotspots](#_5qwml653x1mf)

[3.4 Temporal Trends](#_3akahdxtu3v3)

[3.5 Theft Crime Insights](#_7m5rp4p1h5b8)

[3.6 Property Crime Insights](#_pf3udgp1705o)

[3.6 Crime outcome rates](#_n6iqhqc6j6h7)

[4. Recommendations for Real Estate Strategy](#_mke2q8rc6i9v)

[4.1 Top Recommendations](#_rvcpl3m5gt55)

[4.3 Further Research Needed](#_d19564nvlwhd)

[5. Conclusion](#_jsyn2g80ohy8)

[Appendix: Data Sources & Tools](#_4wq51fficml9)

# **Crime Data Analysis Report for Real Estate Decision Making**

**Prepared for:** Nadine Green, Head of Sales  
**Prepared by:** EO   
**Date:** 10.06.2024

**Data Source**

* **Source:** [UK Police Street Crime Data](https://data.police.uk/data/)
* **Time Period:** March 2022 – March 2024
* **Police Forces Included:**
  + Hertfordshire
  + Kent
  + Lincolnshire
  + Northamptonshire

**Column Descriptions**

* **Reported by**: The police force that recorded and submitted the crime data.
* **Falls within:** Currently the same police force that reported the crime. This setup is under review and may change soon.
* **Longitude and Latitude:** The crime's anonymised location coordinates. For more details, refer to the [*Location Anonymisation* section](https://data.police.uk/about/#location-anonymisation).
* **LSOA code and LSOA name:** Identifiers for the Lower Layer Super Output Area (LSOA) where the anonymised crime location falls, based on [boundaries](http://data.gov.uk/dataset/lower_layer_super_output_area_lsoa_boundaries) set by the Office for National Statistics.
* **Crime type:** Categorisation of the crime, as defined in the [Police.UK FAQ](https://www.police.uk/pu/about-police.uk-crime-data/) list of crime types.
* **Last outcome category:** Indicates the most recent outcome recorded for the crime.
* **Context:** An optional field for police forces to add readable notes or extra details about the crime. As of now, it is left blank in newly uploaded CSV files.

### **Methodology**

The analysis was conducted using Python, with the following key steps:

1. **Data Collection:** Downloaded street-level crime datasets from UK Police Data Portal.
2. **Data Cleaning & Preprocessing:** Handled missing values, standardised LSOA codes, and removed irrelevant columns.
3. **Exploratory Data Analysis (EDA):** Analysed crime distribution, types, temporal trends, and regional comparisons.
4. **Visualisation:** Generated charts using Matplotlib, Seaborn, and Plotly to highlight key insights.

**Tools Used:**

* Python (Pandas, NumPy)
* Jupyter Notebook
* Visualisation: Matplotlib, Seaborn, Plotly, Folium

### **1. Executive Summary**

### **Total Crimes Analysed:** 917,007

### **Crime Distribution by Police Force:**

### Kent: 44% (403,483 incidents)

### Northamptonshire: 24%

### Hertfordshire: 16%

### Lincolnshire: 16% (lowest crime volume)

### **Violent Crime Rates:**

### Lowest: Hertfordshire (29.7%)

### Highest: Northamptonshire (41.1%)

### **Top Crime Types:**

### Violence & sexual offences (37.2%)

### Anti-social behaviour (22.5%)

### Criminal damage & arson (12.1%)

**Objective:** To identify safest UK regions for real estate investment using 24 months (2022-2024) of street-level crime data from four police forces.

Key findings include:

* **Safest Police Force Areas:** Lincolnshire (lowest overall crime) & Hertfordshire (lowest proportion of violent crime)
* **Highest Crime Area:** Northamptonshire (highest proportion of violent crime) & Kent (highest percentage of crime)
* **Top Crime Types:** Anti-social behaviour (most frequent), Violence & sexual offences.
* **Temporal Trends:** Crime peaked in May 2022 but declined in May 2023.

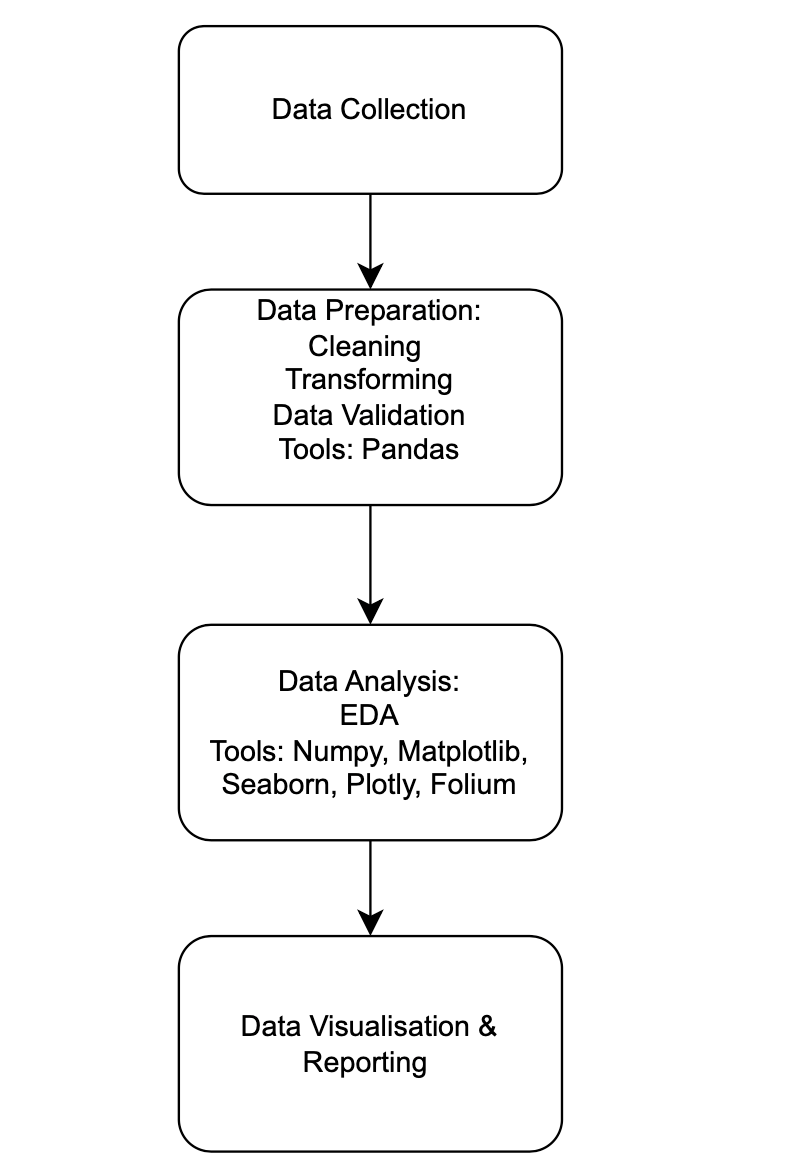
**Recommendations:**

1. Prioritise developments in Lincolnshire & Hertfordshire
2. Conduct micro-location analysis using LSOA-level data, also considering population numbers
3. Avoid high risk areas e.g in Northamptonshire and Kent

## **2. Data Quality Assessment**

### **2.1 Data Collection & Preprocessing**

### **Data Processing Workflow**



**Figure 1**

### **2.2 Data Cleaning & Missing Values**

|  |  |  |
| --- | --- | --- |
| **Column** | **Missing Values** | **Action Taken** |
| Crime ID | 159,240 | Dropped (irrelevant for analysis) |
| Month | 0 | Dropped as Month/ year columns were added |
| Reported by | 0 | Dropped ( same data as falls within) |
| Falls within | 0 | None |
| Location | 0 | 57413 - marked as “On or near” which was changed to unknown  7309 put down as having “No Location” |
| Longitude/Latitude | 7,309 | Retained (not critical for regional analysis)  set to 0 for unrecorded values |
| LSOA Code/Name | 7,310 | Imputed with "Unknown" where missing.  Changed all code names to 2011 naming standard to ensure uniformity as some codes had both the 2011 and 2021 naming standard under a single code. |
| Context | 100% missing | Dropped (no useful data) |
| Last Outcome | 159,240 | Marked "Unknown" if blank |

**Key Data Decisions:**

* Removed irrelevant columns (Crime ID, Context, LSOA code, Reported by).
* Kept records with missing coordinates since analysis focuses on broader regions.
* Added year and month column in order to analyse data by year and month
* Changed all code names to 2011 naming standard to ensure uniformity as some codes had both the 2011 and 2021 naming standard under a single code. Attempted to use the 2021 LSOA codes and names but this left more blank data as boundaries had changes so some code names and numbers that existed in the original dataset, were no longer valid for the 2021 LSOA codes/ names

## **3. Exploratory Data Analysis (EDA)**

### **Top 5 Areas by Crime Volume (excluding 'Unknown'):**

### 

|  |  |
| --- | --- |
| **LSOA Name** | **Crime Count** |
| Northampton 021F | 3836 |
| Dartford 006D | 3686 |
| Canterbury 020F | 3367 |
| Swale 010E | 3042 |
| Dartford 003C | 2857 |

### **Figure 2**

* Four areas in Kent appear in the top 4 ( Dartford, Canterbury, Swale, Darftford) However, the area with the highest Crime count is in Northampton.

### 

### **3.1 Crime Distribution by Police Force**

### **Figure 1**

**Figure 3**

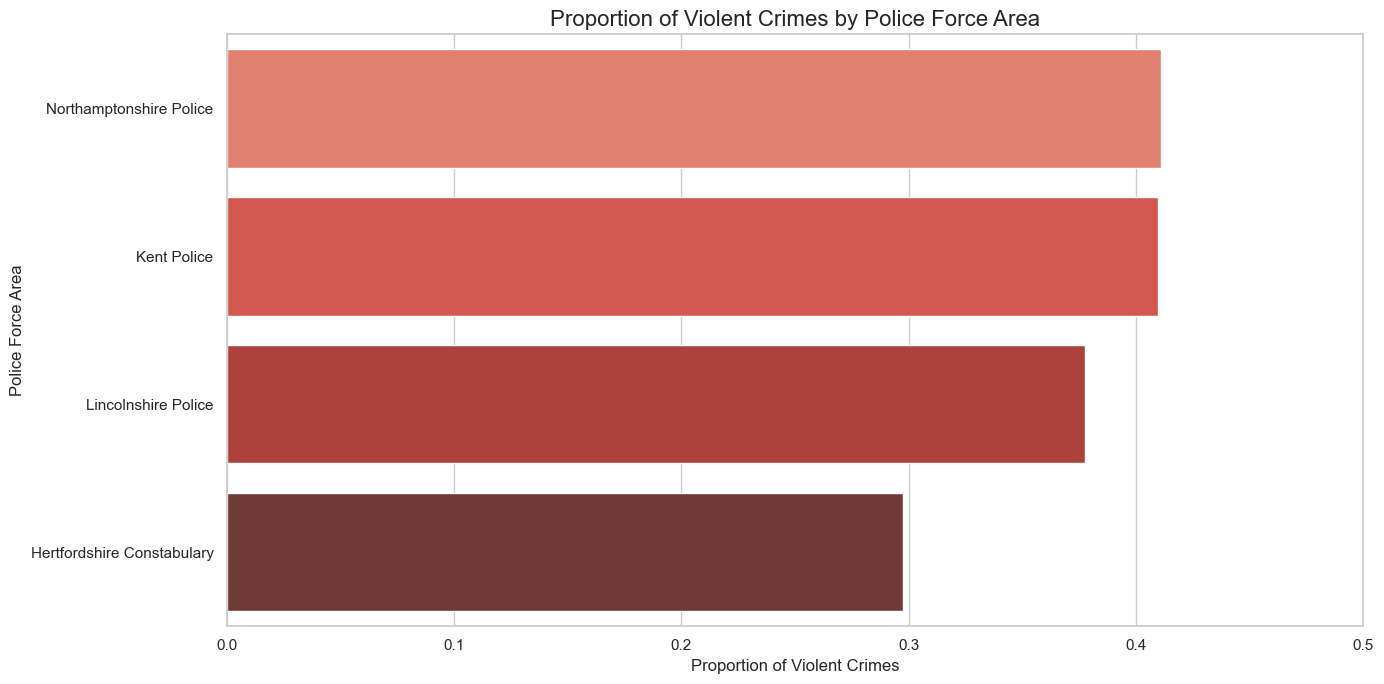
* Kent accounts for 44% of total crimes (highest volume).
* Lincolnshire has the lowest amount of crime (16%) making it a safer investment option.

### **3.2 Crime Type Analysis**

### **Figure 3**

* **Most Frequent Crimes:**
  + Violence & sexual offences (37.2%)
  + Anti-social behaviour (22.5%)
  + Criminal damage & arson (12.1%)
* **Rarest Crimes:** Theft from person (0.6%) and robbery (0.7%).

### **3.3 Violent Crime Hotspots**

****

**Figure 4**

**Violent Crime Summary**

. Violent Crime Statistics:

- Mean: 37.4%

- Median: 39.4%

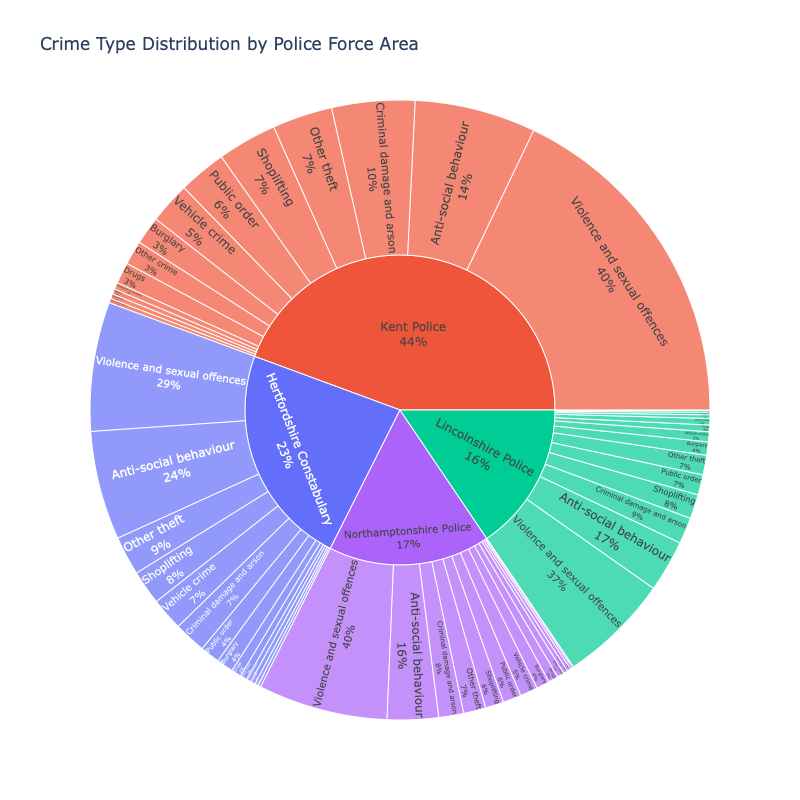
- Standard Deviation: 5.3%

Violent Crime Range by Area:

- Lowest: 29.7% (Hertfordshire Constabulary)

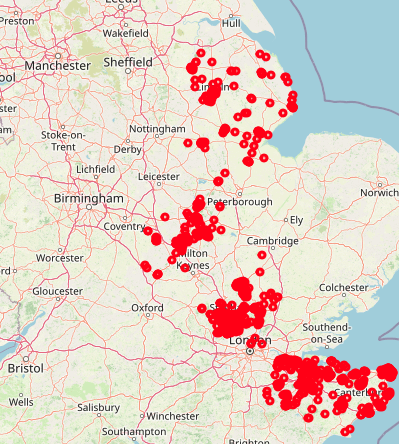
- Highest: 41.1% (Northamptonshire Police)

* **Northamptonshire** has the **highest violent crime rate (41.1%)**, exceeding the mean (37.4%).
* **Hertfordshire** has the **lowest (29.7%)**, making it a safer choice.



**Figure 5**

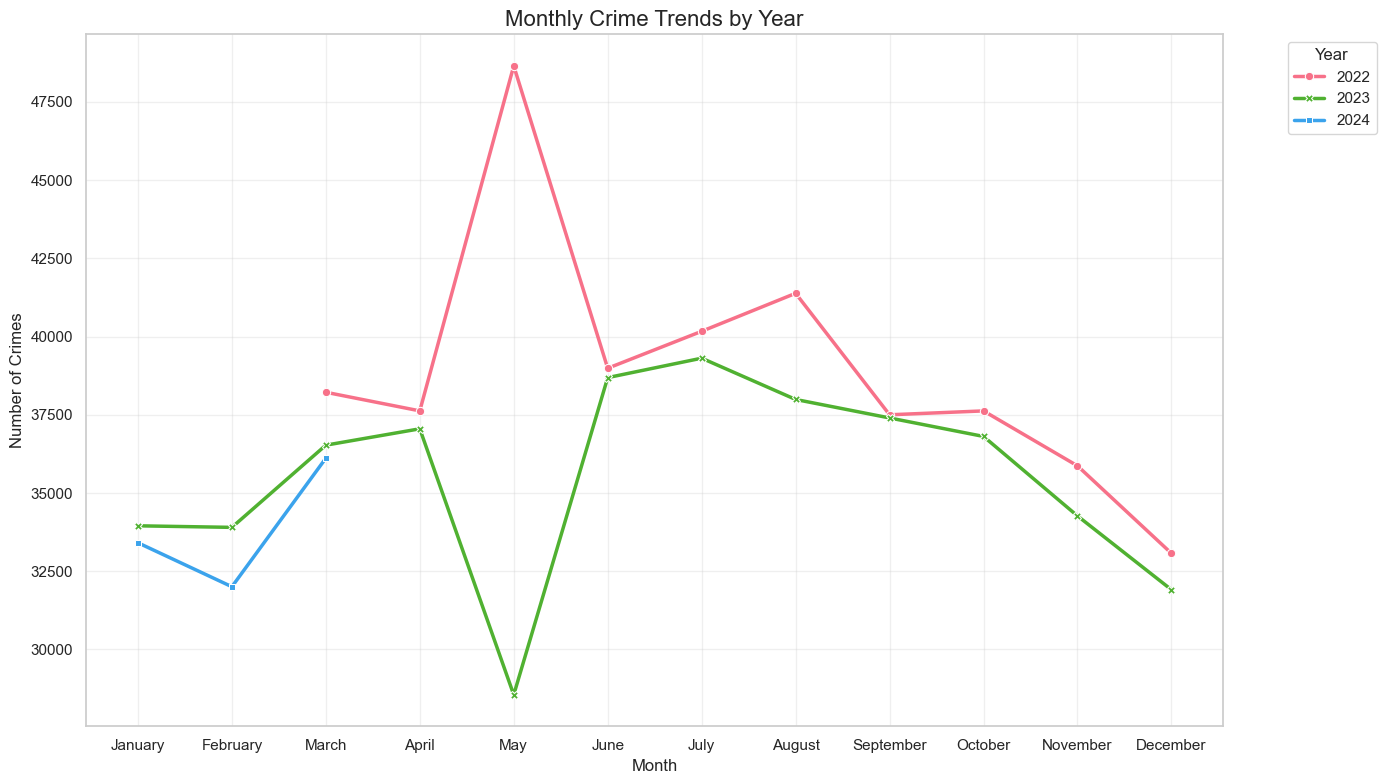
* All Police forces most frequent crime was violence and sexual offences, followed by anti social behaviour
* The third most frequent crimes for each forces are as follows :
  + Kent, Northamptonshire and Lincolnshire- criminal damage and arson,
  + hertfordshire - other theft,



**Figure 6**

* Lincolnshire crimes more evenly distributed and there is a significantly lower density of crimes compared to other areas

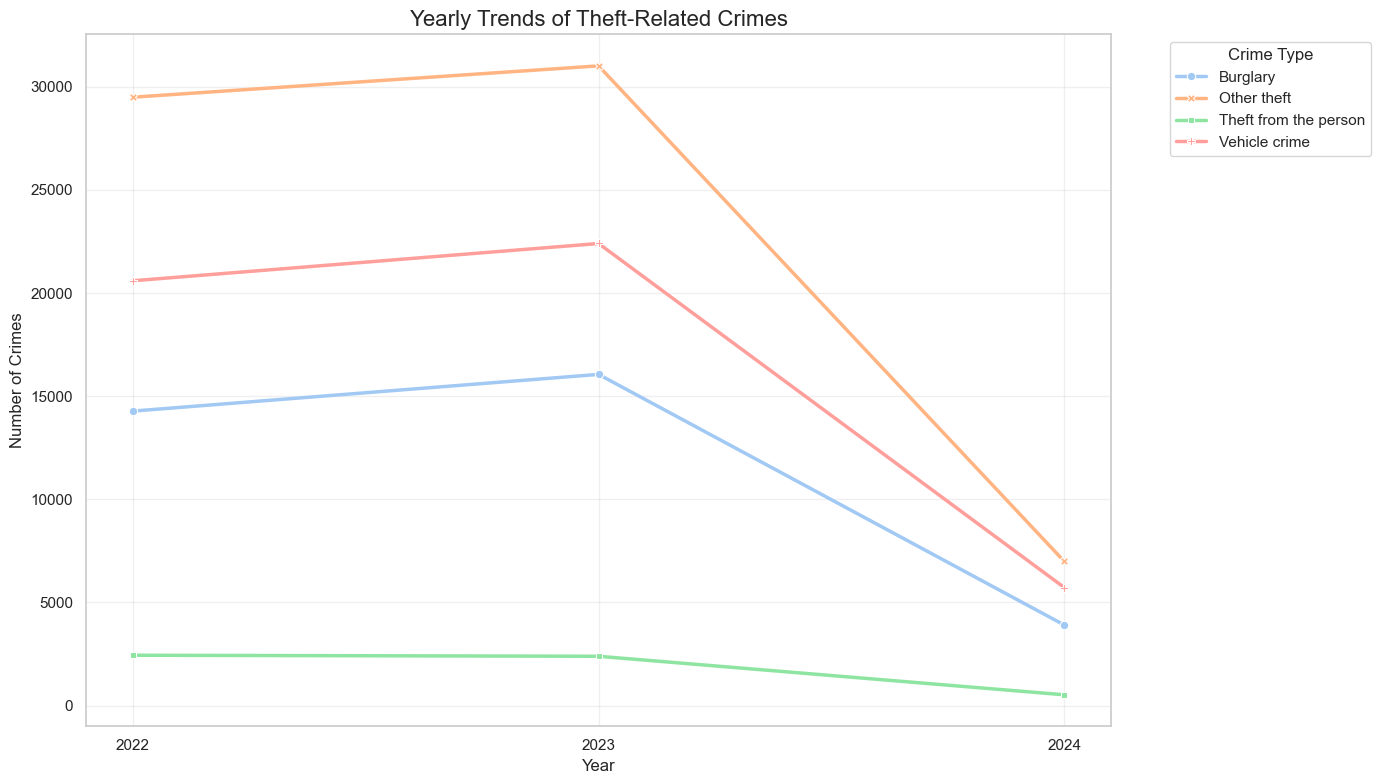
### **3.4 Temporal Trends**

****

**Figure 7**

* In 2022 crime peaked in May, on the contrary in 2023, crime declined the most in May
* In 2023, crime started to plateau from september. In 2022 it started to plateau from October
* March 2024 had fewer crimes than March 2022.

## **3.5 Theft Crime Insights**



**Figure 8**

* Highest n.o crimes is other theft
* Lowest is theft from the person
* For all crime types, there appears to be a decline from 2023 - 2024

### **3.6 Property Crime Insights**

### 

### Burglary Trends: Average annual decline of 0.32%.

### Safest LSOAs for Property Crime:

### Milton Keynes 017F (Northamptonshire)

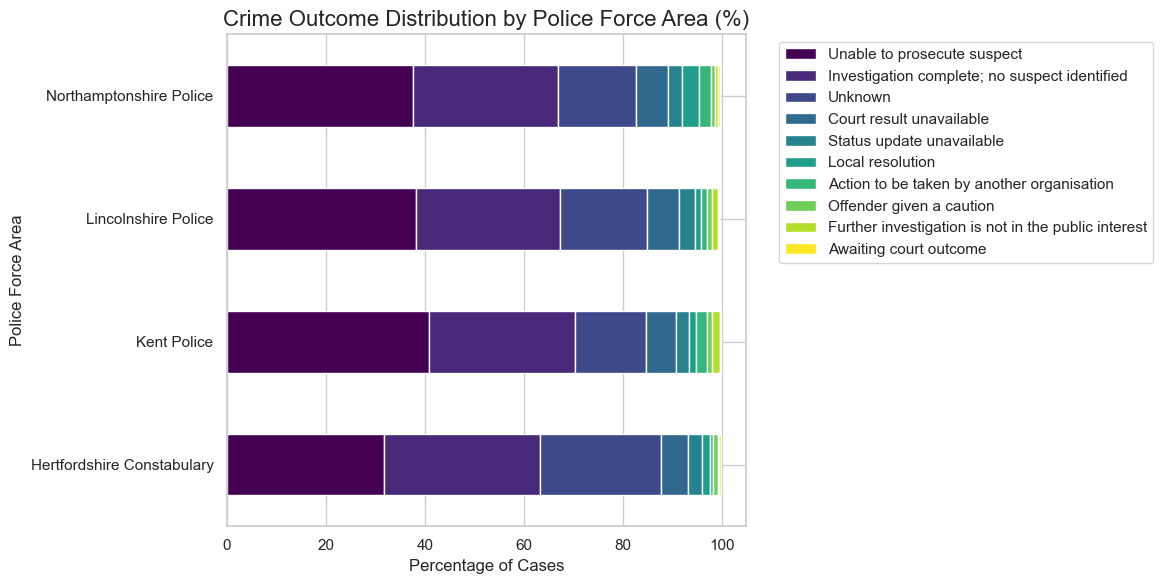
### Thurrock 020E (Kent)

### North East Lincolnshire 023F (Lincolnshire)

### North Lincolnshire 023D

### Wealden 005D (Kent)

### 3.6 Crime outcome rates



**Figure 9**

* For all police forces most crimes resulted in the suspect not being prosecuted, or not being identified at all.

## **4. Recommendations for Real Estate Strategy**

## **4.1 Top Recommendations**

Tier 1

* **Lincolnshire:** Lowest crimes committed overall - particularly 'North East Lincolnshire 023F', 'North Lincolnshire 023D' ( appeared in overall top safest areas for property)
* **Hertfordshire:** Lowest proportion of violent crime

Tier 2 Conditional Investment (if Expansion needed)

Areas in Kent nd Northamptonshire where property crimes are the lowest: Milton Keynes 017F (Northamptonshire), Thurrock 020E (Kent), Wealden 005D (Kent).

### **4.3 Further Research Needed**

* **Demographic data**: population density, income levels.
* **Local infrastructure** police stations, CCTV coverage.

## **5. Conclusion**

This analysis identifies Lincolnshire and Hertfordshireas the most promising regions for real estate investment due to lower crime rates and improving trends. Northamptonshire requires caution due to high violent crime rates.

**Next Steps:**

1. Look deeper into LSOA-level crime data for micro-location insights, considering factors like population numbers.
2. Compare with property price trends to assess ROI in safe vs. high-risk areas.

### **Appendix: Data Sources & Tools**

* **Jupyter Notebook**
* **Data Sources:** [UK Police Data Portal](https://data.police.uk/about/#columns)
* **Visualisation Tools:** Matplotlib, Seaborn, Plotly, folium, Numpy

**Prepared by:** EO