Introduction:

This report presents an exploratory data analysis (EDA) of crime data in South Wales, UK. The report was obtained from December 2019 to November 2022. The data was obtained from an official government source and includes information on the types of crimes, location of crimes, and dates of crimes. The goal of this EDA is to gain a better understanding of the crime patterns in South Wales and to identify any potential areas for further investigation. The analysis is performed with the intent of testing some hypotheses about the data, such as:

1. The most common type of crime in South Wales
2. The relation between crime type and location
3. Crime rate fluctuates within the time of the year(s)

The EDA is structured to first provide a general overview of the data, then to investigate specific relationships and patterns in the data. The analysis will include both univariate, bivariate, and multivariate techniques to understand the underlying trends in the data.

Data Description:

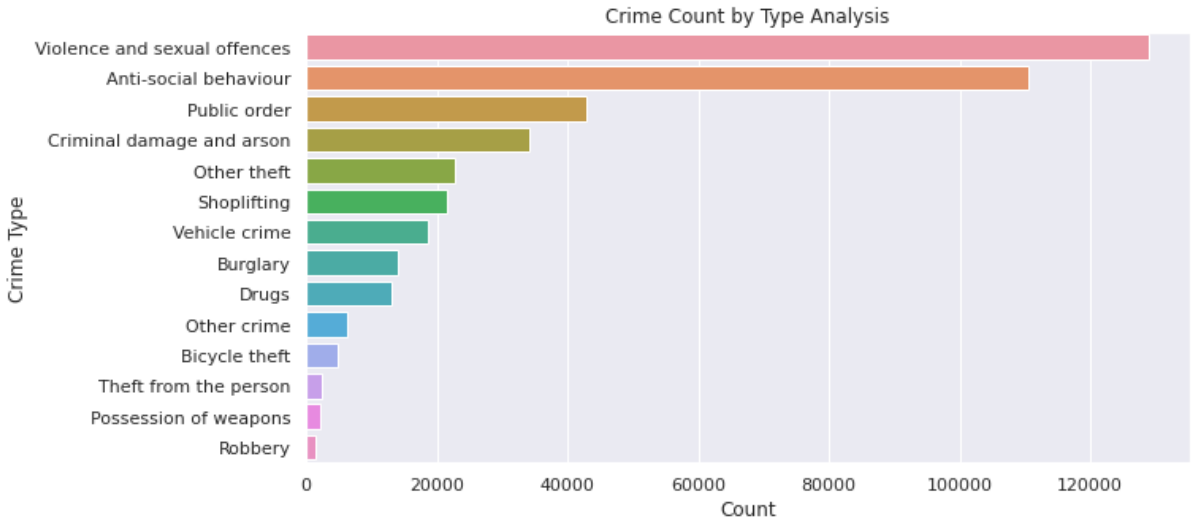
The crime data used for this EDA includes a total of 423944 observations and 10 variables. The variables in the dataset include crime type, location, date, and other demographic information (e.g., police force area). Some variables have missing values, which were handled by using the fill python command, and the rest were considered as outliers hence getting dropped.

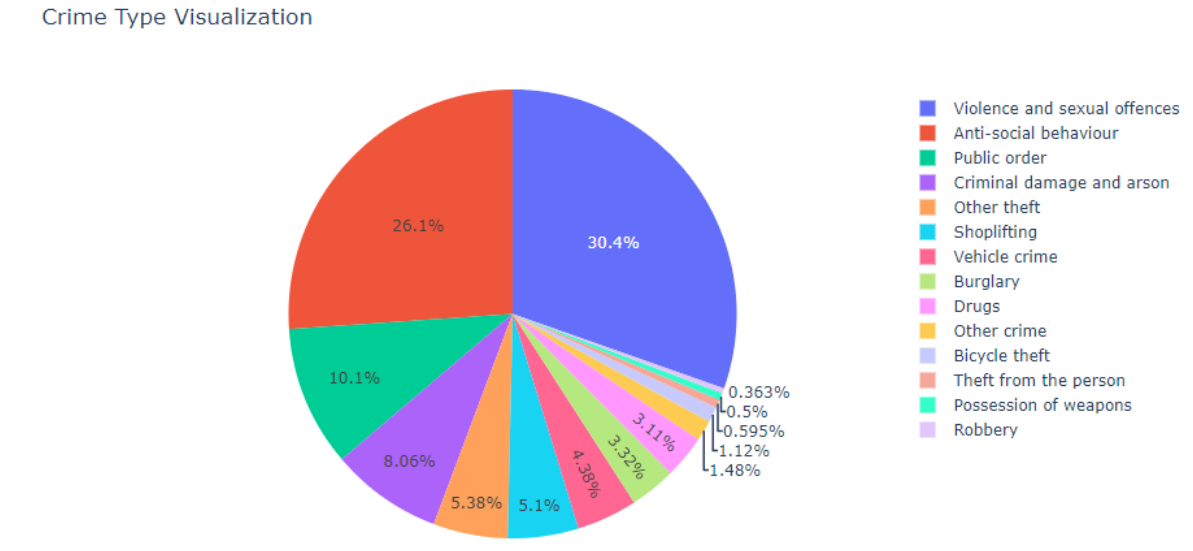
Univariate Analysis:

The most common type of crime reported in South Wales was Violence and sexual offences. This crime type accounted for 30.4% of all reported crimes. Other common crimes included:

Violence and sexual offences 128819
Anti-social behaviour 110550
Public order 42962
Criminal damage and arson 34191
Other theft 22798
Shoplifting 21616
Vehicle crime 18548
Burglary 14071
Drugs 13167
Other crime 6293
Bicycle theft 4745
Theft from the person 2523
Possession of weapons 2120
Robbery 1541

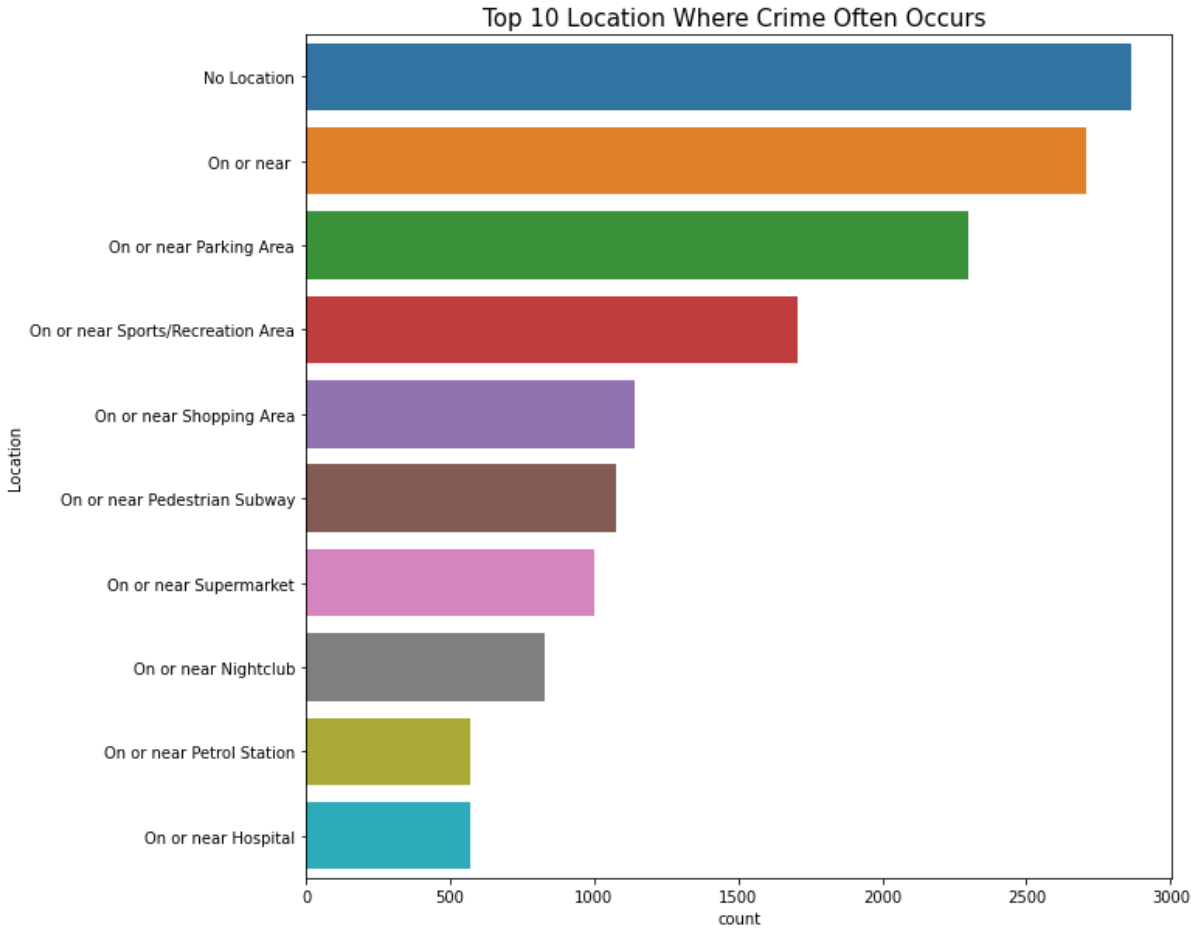
Below are the visualizations to the crime type





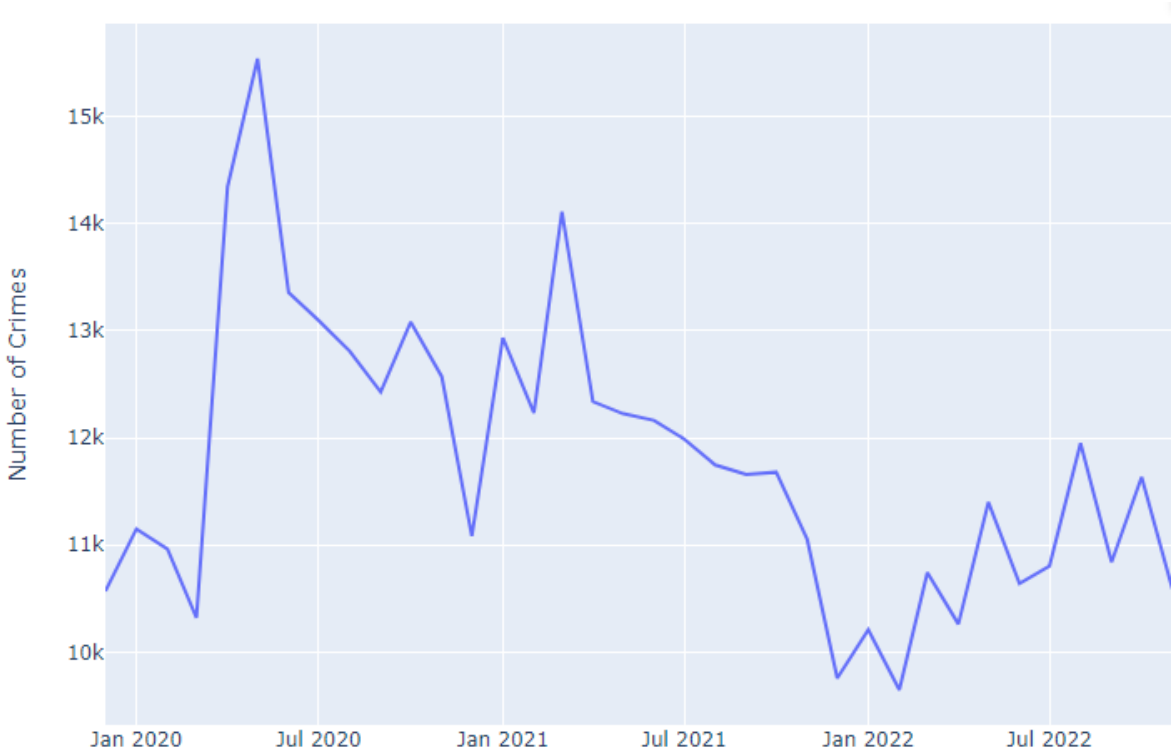
Bivariate Analysis:

A bivariate analysis was conducted to investigate the relationship between crime type and location. The analysis revealed that the highest incidence occurred in a No location zone, while the lowest incidence of this crime type was observed on or near Hospital. Below is a bar char chart visualization portraying the above hypothesis.



Multivariate Analysis:

A multivariate analysis was also conducted to investigate the relationship between crime type, location, and time of occurrence. The results indicate that May 2020 had the highest number of count of crime cases recorded. There was a slight gradual increase between January and July 2022 in the count of number of crime cases as shown in the below line graph visualization



Below is a count of cases in various months in descending order;

[ ]

```
#Count of offences each month in different years (in Descending order)
df['Month'].value_counts().sort_values(ascending=False)
```

2020-05-01 15541

2020-04-01 14346

2021-03-01 14112

2020-06-01 13357

2020-07-01 13098

2020-10-01 13082

2021-01-01 12935

2020-08-01 12812

2020-11-01 12572

2020-09-01 12429

2021-04-01 12340

2021-02-01 12235

2021-05-01 12227

2021-06-01 12164

2021-07-01 11991

2022-08-01 11953

2021-08-01 11749

2021-10-01 11680

2021-09-01 11661

2022-10-01 11636

2022-05-01 11401

2020-01-01 11150

2020-12-01 11084

2021-11-01 11052

2020-02-01 10962

2022-09-01 10840

2022-07-01 10804

2022-03-01 10743

2022-06-01 10641

2022-11-01 10581

2019-12-01 10569

2020-03-01 10322

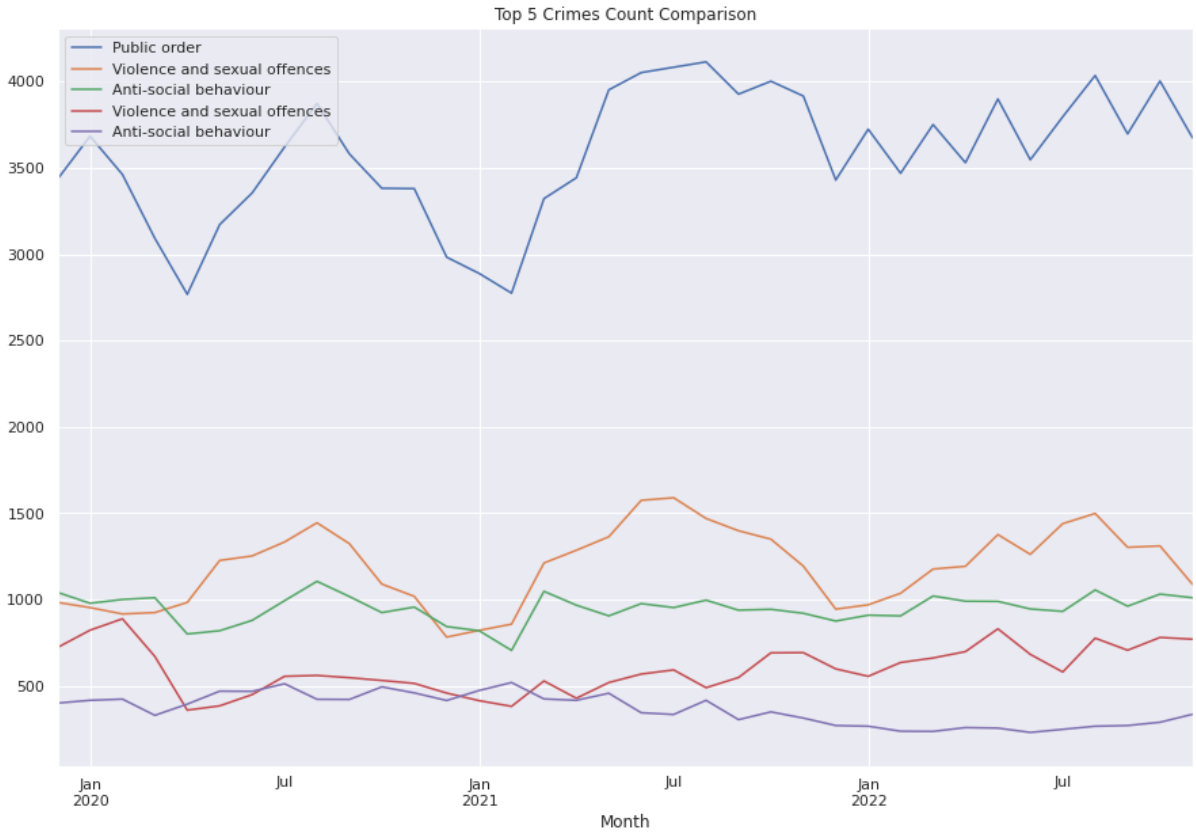
2022-04-01 10262

2022-01-01 10208

2021-12-01 9757

2022-02-01 9648

The below multi-line graph shows various variations of crime types across different month periods;



Conclusion:

Overall, this EDA has provided insight into the crime patterns in South Wales. The most common type of crime was violence and sexual offenses, and this crime type was more likely to occur in May 2020. The data analysis also supports or disproves the hypotheses posed in the beginning. Further investigation is needed to understand the reasons behind these crime patterns and to develop strategies for reducing crime in the area.