Investigating the effect of coronavirus lockdowns on Crime Rates within the UK

Elliot Linsey  
210764484Prof. Norman Fenton  
MSc Data Science and AI

*Abstract*—The coronavirus pandemic of 2020 – 2021 resulted in a restriction of public mobility within the UK. A combination of classical and Bayesian statistical methods has been implemented to investigate the effect of lockdowns on crime rates within different crime types, regions, as well as model the true crime rates in contrast to reported crime rates.

Keywords—Coronavirus; ARIMA; Bayesian network causal model; Chi-Squared; Crime rates; UK.

# Introduction

The implementation of lockdown protocols by the United Kingdom Government on 26th March 2020 created a unique environment by restricting movement among non-essential workers and limiting social contact (Johnson, 2020). The primary reason was to reduce the spread of the disease and ease the burden on the National Health Service. However, the effects of this lockdown spread to all sectors, including hospitality, construction, education, travel, and the judicial system (ONS, 2020). From March 26th 2020 lockdown measures legally came into force and the UK police were given extra enforcement powers to reduce the spread of coronavirus, including the ability to instruct members of the public to return home or leave an area (UK Gov, 2020). The lockdown required non-essential shops such as pubs and retail stores to close and for non-essential workers to work from home if possible. Those that could not work from home were furloughed, with one in four people that had been employees having been on furlough at some point between March 2020 and June 2021 (ONS, 2021). Table 1 contains a full timeline of lockdown events from the beginning of lockdown on March 26th 2020 to the end of the third and final lockdown on July 19th 2021. These restrictions resulted in changes to the mobility of the general public to the point where even seismic noise was reduced by 50% for months at a time, with the period being described as “the great seismic quiet period” (Lecocq et al, 2020).

This report investigates the effects of the national lockdowns on crime rates and also arrest outcomes over the full coronavirus period of March 2020 to August 2021. The areas selected for study are the North, East Midlands, West Midlands, South West and South East police constabularies of England and Wales. Specific forms of crime have been selected; these are violence and sexual offences, theft offences, drug offences, public order offences & arson and criminal damage offences.

Data from 2017 to 2022 is used so that a pre-covid comparison can be established. The data used is from <https://data.police.uk/> which contains official UK police records of every crime reported within the 43 geographic police forces in England and Wales, the British Transport Police, the Police Service of Northern Ireland and the Ministry of Justice.

# Literature Review

Approximately 90 countries followed the example set by Italy in March 2020 by implementing their own forms of national lockdown and curfews during the coronavirus pandemic (Euronews, 2020). Whilst the lengths of time and requisite powers to enforce lockdowns among the public varied by country, the vast majority of these lockdowns followed similar rules. Non-essential workers were to stay home, gatherings were restricted and schools and retail stores were shut. Studies around the world have been conducted to investigate the relationship between the coronavirus and crime. They have found substantial variations in which types of crime the lockdown protocols impacted as well as how these changes were distributed over different cities and countries (Nivette, 2021).

Crime takes many forms and is strongly associated with both a target and a victim’s mobility (Farrell et al, 2020) and the situational opportunity for a crime to take place (Clarke, 2012). As the lockdowns restricted the ability to leave the house and interact with the public, it is simple to theorise those certain types of crime that rely on mobility as well as opportunity would decline. Crimes such as robbery and assault could be expected to decline due to less members of the public being present in opportunistic settings late at night, as well as the closure of venues decreasing the amount of alcohol being consumed in city centres.

Crimes such as burglary could be expected to decrease due to the increased guardianship of the property due to the stay-at-home order, restricting the potential burglar’s access and opportunity. However, other forms of crime such as domestic violence and child abuse could be expected to rise due to the closer proximity and greater time spent together of families within the home (Sri et al, 2021). These domestic situations may be exacerbated by the greater increase in alcohol being consumed at home, leading to increases in mental health issues, suicide, alcoholism, as well as previously mentioned domestic violence (Ramalho, 2020).

In tandem with the rise of alcohol use is the rise of substance abuse, with increases of overdoses being witnessed in the USA, particularly in relation to opioids (Abramson, 2021). Within the UK, respondents to anonymous surveys have reported an increase in drug consumption and frequency due to factors such as boredom, more free time, and stress (CREW, 2020). Combined with issues such as poorer drug quality (such as lower quality heroin and the greater availability of fentanyl), this has led to the highest number of deaths by drug misuse in the UK in 2020 since records began (OHID, 2021).

The lack of physical opportunity for some crimes may have led to an increase in cybercrimes. With the inability to generate an income through either physical crime or having been made redundant due to the pandemic, online fraud as well as email and social media hacking were seen to increase during the immediate months following the lockdown announcement in the UK (Buil-Gil, 2021).

Domestic violence within the home against women and girls has been dubbed the “shadow pandemic” (Mlambo-Ngcuka, 2020). Unfortunately, gender-based violence is known to rise during emergency periods (Unicef, 2022). The UK failed to take into account what effect the national lockdown would have on domestic violence rates, leading to 16 girls being murdered within the first month which was triple the rate at the same period in 2019 (Taub et al, 2020). The scale of isolation experienced by victims is a result of support networks such as shelters and help-seeking avenues being shut down during the pandemic, which forces victims to move back in with abusers (Women’s Aid, 2020). What makes coronavirus unique is the fact that “the accessibility of services and the ability of women to access these [aid] services will decrease” (Hersh, 2020). Due to the close proximity to their abuser, many women have felt unable to call help services and therefore text messages and emails to providers had sharply increased whereas calls had decreased (Graham-Harrison et al, 2020).

A number of research papers have been published focusing on the crime aspect of the pandemic. They use statistical methods such as wavelet analysis and spatial point pattern test for retrospective analysis of data. They also use machine learning such as support vector regression and linear regression to identify relationships between covid levels and crime rates (Ma et al, 2022). Others have used time series analyses as well as predictive models such as Auto-Regressive Integrated Moving Average (ARIMA) to try and predict what crime rates would have been during 2020 if the pandemic had not occurred, and therefore see if there are significant differences between what was predicted and what the true rates were (Payne et al, 2020).

# Method

## Data

The data used is openly sourced from https://data.police.uk/ which contains information from 43 British police forces from 2017 – 2022. All police forces are used except for British Transport Police, Northern Ireland and Greater Manchester. Due to Greater Manchester switching over its system during 2020, some of their data reporting is inconsistent; therefore, they have been removed entirely. The website describes the data acquisition process as: “Every month each police force generates a Crime and ASB file and a Police Outcomes file in a set format. The forces upload these to a private server managed by the Single Online Home National Digital Team in the Government network, where the files undergo quality assurance. Copies of the data from police forces is then sent to the Ministry of Justice (MoJ), where they try to match the crimes with any court results contained in their own records. The MoJ send any matching court results back to the Single Online Home National Digital Team, where they are integrated with the existing data. All data is then anonymised before being published.” (Data.police.uk, 2022). From the data we are able to extract the following features: Which force reported the crime, the crime types, the outcome types, region, month and year.

Region is created by assigning each row a value from [South West, South East, West Midlands, East Midlands, North] depending on which police force reported the crime. There also appears to be a lag at collecting data towards the end of 2021 as all crime levels dramatically decrease. This has been attributed to missing data that has not been cleaned and included within the dataset at the point of analysis. Regardless, the period of interest of March 2020 to August 2021 has not been affected.

Five crime types are being investigated: Violence and sexual offences, theft offences, criminal damage and arson, public order offences, drug offences.

Theft offences has been created by combining offences such as 'Other theft', 'Burglary', 'Shoplifting', 'Theft from the person', 'Robbery' and 'Bicycle theft'. The other main categories have not been joined with any other subcategories.

Four outcome types are being investigated: No suspect identified, unable to prosecute suspect, suspect charged, other outcome.

‘Other outcome’ is created by combining subcategories such as ‘Offender fined,’ ‘Offender given a caution’ but also includes ‘Defendant found not guilty’ and ‘Offender sent to prison’. Due to the large number of outcomes possible (24 in total), it made sense to keep the main variables of interest untouched and combine the rest into a single variable.

The inclusion of data from 2017 is to be used as a control, in this way the data contains three years of non-covid crime statistics (2017 – 2019) and two years of covid statistics (2020 – 2021).

## Seasonal ARIMA:

The Auto Regressive Integrated Moving Average model is used within time series analysis to predict future values. It contains three terms, ‘p’, ‘d’ and ‘q’ that specify the level of autocorrelation, differencing, and partial autocorrelation. Put simply, this method finds identifiable patterns from the previous timesteps to predict future timesteps. The first step is determining whether the timeseries is stationary in order to set the level of differencing. This can be determined by looking at the plotted line graphs and identifying trends. From looking at the data, it appears that the data is non-stationary. A more conclusive method is to use an Augmented Dickey-Fuller test. The null hypothesis (p=0.05) is that the timeseries is non-stationary, therefore by failing to reject the null hypothesis indicates the non-stationarity of the time series, which is true in all crime type investigations.

Once non-stationarity has been determined, the auto regression and moving average levels must be set. This can be established using autocorrelation function (ACF) and partial autocorrelation function (PACF) as well as the Akaike Information Criterion (AIC) to establish goodness of fit. Using the ‘auto\_arima’ function from statsmodels within Python 3, different variants of ‘p’, ‘d’, and ‘q’ within the ARIMA model can be examined with the time series data and the optimal parameters can be chosen by minimising the resulting AIC. Lastly, the data is being analysed by year (with monthly subdivisions), so a seasonal attribute of 12 is included to represent the monthly repetition.

Once the optimal parameters have been found, the data for each crime type from periods 2017 – 2019 are used to predict the crime rates of 2020 and 2021. Included within this are the 95% confidence intervals of the predictions. These are then compared to the actual crime rates observed within 2020 and 2021 and whether there are any significant differences between them.

## Percentage Changes:

This investigated the percentage changes for the total rate of each crime type over the 2020 - 2021 period, as well as drilling down into the regional differences for each crime type. Firstly, the median value is selected for each crime type from 2019, this provides a baseline value that can be compared against. Each month of 2020 – 2021 is then compared against this baseline value and differences are noted for both the total crime over England, as well as the regional data. This allows us to see whether over time the crime rates start to fall back to the pre-covid rates. An issue is that this method does not take into account seasonal trends as every point is compared to one median value from 2019.

Secondly, a percentage change analysis is conducted with a 12-month time lag. Therefore, the change between March 2020 and March 2019 is determined and so on for the time period. What this monthly difference provides is a step-by-step comparison between a year with no lockdowns or restrictions, and a year that has lockdowns. What is interesting is the comparison between years 2021 and 2020 as both years contained lockdowns. From these two years we can determine whether a potential lockdown effect reduces over time.

## Chi-Squared test:

Using crosstabulation, we can put the data into the correct format for a chi-squared test to be performed. In this situation, the chi-squared test is being used to test the independence of the two features, year (lockdown or non-lockdown) and crime type. The null hypothesis (p=0.05) is that the number of crimes is independent of the level of lockdown that the crimes were committed under. The test will generate the expected frequency of our variables. If the expected frequency differs significantly from the observed frequency, we can reject the null hypothesis. However, this does not necessarily mean that they are directly related, due to the potential for a confounding variable to exist that could make both variables appear to be related. It is also important to understand how a p-value is interpreted. A p-value is the probability that you would receive results of this extremity (in this case the difference between the expected and observed results) if the experiment was repeated multiple times and the null hypothesis was true. This means that even with a low and statistically significant p-value of p=0.03, you could expect your value to be this extreme 3% of the time if the null hypothesis was correct.

## Probability of an outcome:

To account for the different numbers of crimes occurring each year, we can group the data by year, crime type or region, then have outcome type as our output variable. We can then normalise these figures between a range of [0,1] to generate probabilities for each specific outcome type. For example, we can find the probability that the outcome “No suspect identified” would occur for a crime type of “Criminal damage and arson” in the year 2019. This can also be done for regions, for example we can find the probability that the outcome “Suspect charged” occurred in the East Midlands in 2020. To make things easier to read, these figures can be converted to percentages as well as plotted in stacked bar charts, stratified by year. From this information we can make observations as to how the probability of an outcome type changed from year to year, answering queries such as whether a suspect was more likely to be identified for certain crimes or in certain regions during covid or non-covid times?

## Bayesian Networks:

An issue with using the police records is that we do not have the full number of crimes committed, only the number of crimes reported. This uncertainty is excellent to model with a Bayesian network, which can create causal links between features as well as model testing accuracy. Using AgenaRisk, a network is constructed that contains a branch for each crime type being investigated. The data we have is:

• Reported Number of Crimes

• Police density per 10,000 citizens

• Number of covid months

• Police level of focus on each specific crime type (integrating prior knowledge)

From these features, we are estimating the accuracy of the reported crime figures as well as estimating the true number of crimes for each crime type.

Each branch contains two halves, one estimates the true number of crimes and the other estimates the accuracy of measurement. On branch one (estimating the true number of crimes), we use data from 2019 and 2020 and a ‘covid’ probability table is created. During this time period of 24 months, 8 months contained covid lockdowns, therefore the NPT for covid is 0.33 for true and 0.66 for false. This then links to the next node of ‘Actual number of crime type’. This is an estimated value, for example if we know that approximately 50% of violent crimes are reported to the police, we can infer that the true number of violent crimes is double the reported number of violent crimes. The true number of crimes is constructed using triangle distributions to avoid separated peaks and to allow for greater overlap. On the other half of the branch is the data used to create the accuracy NPT, this is a standard measurement idiom for Bayesian networks (Fenton et al, 2019). Firstly, police density which only contains two outcomes, that of greater than 20 police officers per 10,000 citizens or less than 20. This leads into the Accuracy node and a higher proportion of police officers is expected to have a higher accuracy in reporting crimes. The other input to Accuracy is the Police policy on each crime type. A high police focus is expected to lead to a higher accuracy overall. Accuracy is defined explicitly as the level of underestimation of the true crime rates. A minimal underestimation (or highest accuracy) would expect the reported crime rate to be very close to the true crime rate. A maximal underestimation is defined as the reported crime rate being 50% of the true crime rate.

To estimate the true crime rates, external information (prior knowledge) is used. The rates are estimated for non-covid periods. For drug offences, the Crime Survey for England and Wales (CSEW) estimates 120,000 drug users every month (ONS, 2020). The UK also has a high focus on drug crime, with the government announcing a 10-year plan to help tackle drug abuse and reduce crime (Johnson, 2022)

For violent crime, the CSEW estimated that 49% of violent incidents were reported to the police (ONS, 2020). The UK has a high focus on violent crime, using a “Whole-System” approach with Violence Reduction Units being setup in 2019 in 18 areas of England and Wales (Home Office, 2022).

For criminal damage and arson, the CSEW estimates 82,000 incidents a month. (ONS, 2020). The Metropolitan police is considering criminal damage as a “lower level” crime that it may not pursue in efforts to save £400m (The Independent, 2017).

For theft, the CSEW estimates approximately 126,000 thefts each month. Again, the police force treats theft as a low-level crime that it may not pursue as with criminal damage above. It may see a level of reporting close to the actual value, due to insurers usually requiring a police report to confirm a theft has taken place (abi.org.uk, 2022).

For public order, I’ve based it on incidents such as racial hatred and workplace violence and threats where both are claimed to be highly underreported. (Institute of Race Relations, 2020 and Arnetz et al, 2015).

# Results

From initially plotting the data from 2019 to 2022 (Figure 1, all figures are available in the Appendix), there is a clear drop at the start of lockdown (first vertical line) in violence and sexual offences, theft offences and a small drop in criminal damage and arson. Both drug offences and public order offences do not appear to be negatively impacted at this point and instead they start to rise. Immediately following the initial lockdown, all crime types begin to rise with all of them reaching their pre-lockdown levels apart from violence and sexual offences which exceeds previous levels. This rise could be attributed to the gradual reopening of services that occurred during the summer of 2020 such as the rule of six and the eat out to help out scheme. The following decline could be the result of more strict lockdown measures coming into place again during the autumn and winter of 2020. From January 2021 to March 2021 the UK entered a third national lockdown, this could explain the bump that occurs in all crime types (except drug offences) around March and April 2021 as the lockdown is lifted, with schools and non-essential shops opening within that time.

At the end of the total lockdown period (second vertical line), the crime rates appear to have stabilised. However, due to a lag in the reporting of these statistics by each police force and the time that this data was analysed, a large drop appears for all crime types towards the end of 2021. This should be ignored as it is not representative of the true data.

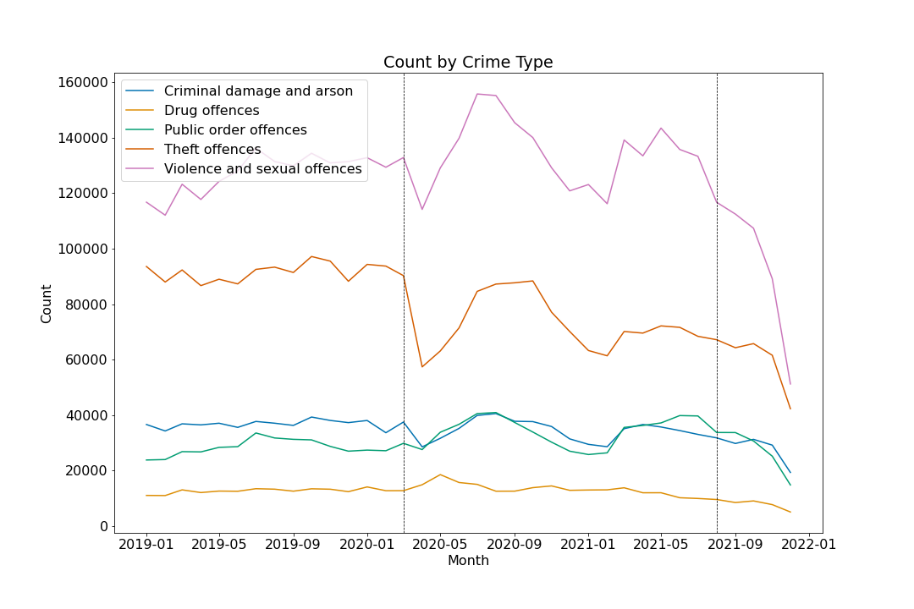


Figure 1 - Count by Crime Type

## ARIMA:

Using data from 2017 to 2019, we can predict the expected crime rates of 2020 and 2021 (Fig 2). Using the historical data, is it easier to see consistent patterns such as in public order offences with the peaks and troughs from 2017 up to 2020. This makes it easier to see the effect of lockdown as we can not only identify anomalous patterns that occur from 2020 onwards but compare them to the predicted crime rates.

Starting with total crime, the initial drop of lockdown does fall outside of the 95% confidence interval of the prediction, stating a significant difference in what occurred and what was expected. It then rises up and stays within the predicted boundaries for the majority, apart from exiting it again briefly at the beginning of 2021. Both these significantly different points have occurred when lockdowns were announced, indicating a strong correlation between lockdowns and the initial reduction of crime rates.

When drilling down further into each respective crime type, we see more noticeable deviations. Criminal damage and arson follow a similar pattern to total crime, with two significant dips happening at the first lockdown and at the January 2021 lockdown. The prediction on this crime type appears slightly off however, with the 2017 to 2019 data containing a downward slope but the prediction rising. The ARIMA model may not have captured the trend so effectively on this crime type.

Public order offences and violence and sexual offences appear to have been impacted the least. Although violence does drop significantly during the initial lockdown, both crime types stay within the predicted range. This could be due to confounding variables, such as different types of violent crime or public order dropping but other types increasing to counteract the difference overall.

Both theft and drug offences have the greatest difference to their predictions. Theft sees the greatest drop at the beginning of lockdown putting it strongly out of the predicted range, whilst drug offences see a massive rise at the same period. Theft and drug offences do revert quickly to more stable levels during the summer of 2020, however all crime rates drop during the lockdown of early 2021, except for drugs which rises again before dropping steeply.

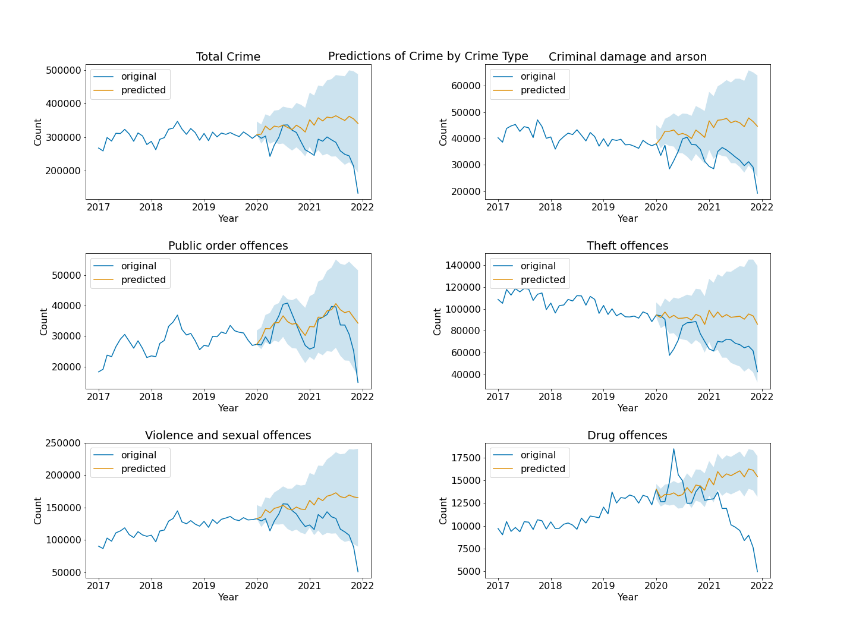


Figure 2 - Predictions of Crime 2020 to 2021

## Regional Changes:

The changes in crime rate can also be examined regionally (Fig 3, Fig 4 & Fig 5). By normalising by 10,000 citizens, trends can be identified and the regions can be directly compared. All appear to follow a similar pattern during the coronavirus lockdowns, with an initial dip followed by a rise during the summer, then another large dip and another rise as lockdowns start to end during mid-2021. A notable difference between the regions that the West Midlands dips earlier and sharper at the end of 2020 in comparison to the rest of the regions. This could be related to the local lockdowns that some regions and cities (such as Leicester) implemented separately from the rest of the country. What is also interesting to note is the fact that population size does not appear to have an impact on crime. The population numbers used are in Table 2 (in appendix), the South East has more than double the population of the East Midlands yet only has 2/3 the crime per 10,000 citizens. Similarly, the North has approximately 1/6 the population of the South East but records higher crime per 10,000 citizens.

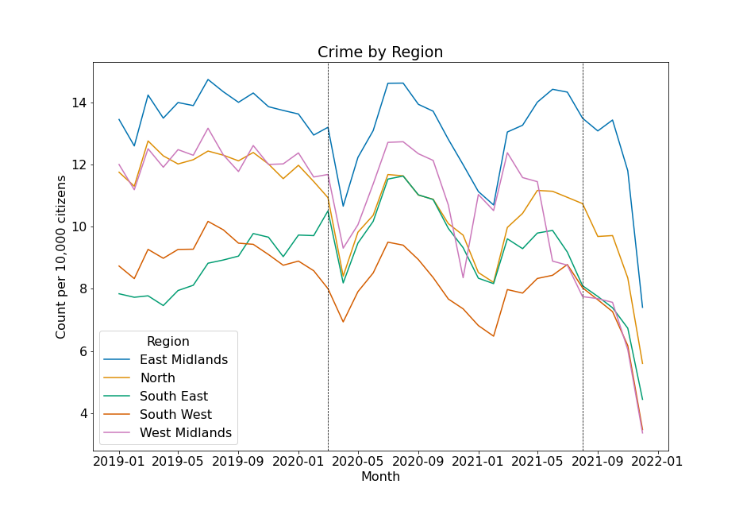


Figure 3 - Regional Crime per 10,000 Citizens

Figure 4 calculates the percentage change between each month and the same month of the year prior (for example, April 2020 compared to April 2019) for each crime type and region. All regions appear to follow the same pattern for each crime type, the noticeable difference being the South East following the same shape but shifted higher during the 2020 period. As this is percentage change, this indicates that the South East had a lower crime rate during 2019 compared to 2020 for all crime types. All regions and crime types follow the initial dip of April 2020, some like theft offences stay far lower than the previous year records whilst others such as public order offences & violence and sexual offences increase by up to 40% compared to the previous year. Drug offences sees the largest spike up to a 70% increase in the South East. What is interesting to look at are the values at April 2021, as this is demonstrating the change between a year exactly after the lockdown. Every crime type and region spike upwards at this point, emphasising the sudden crime rate drop that occurred a year before.

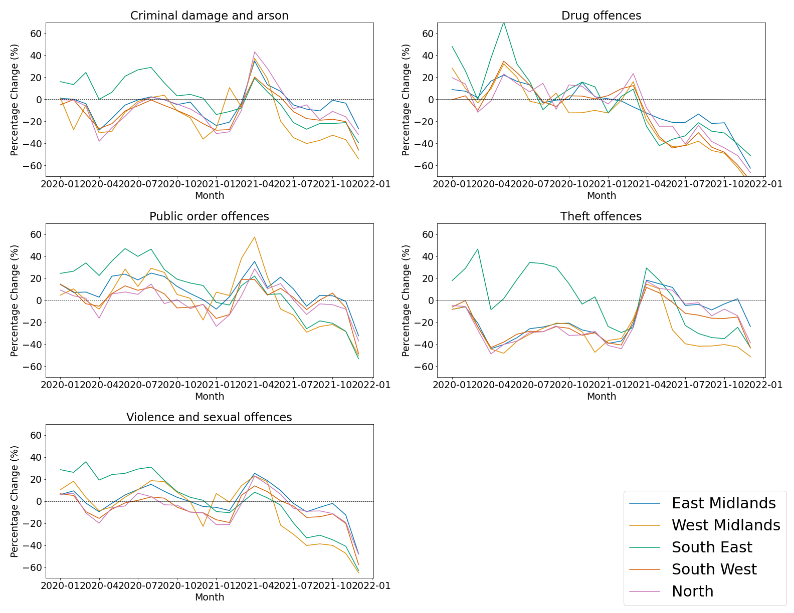


Figure 4 - Percentage Change to Previous Year by Region

Figure 5 repeats the above but instead calculates the percentage change to a median 2019 value for each crime type and region. This allows a comparison between both covid years of 2020 and 2021 to a non-covid year of 2019. Similar patterns are observed as in Fig 4 but the difference is that we can directly compare each month to an overall non-covid period. To demonstrate, after the initial bump for drug offences, we find that all regions drop down to approximate 2019 levels until March 2021, before dropping even further than previous non-covid levels. Again, the South East appears to have the greatest increase of crime compared to the other regions, especially within theft offences.

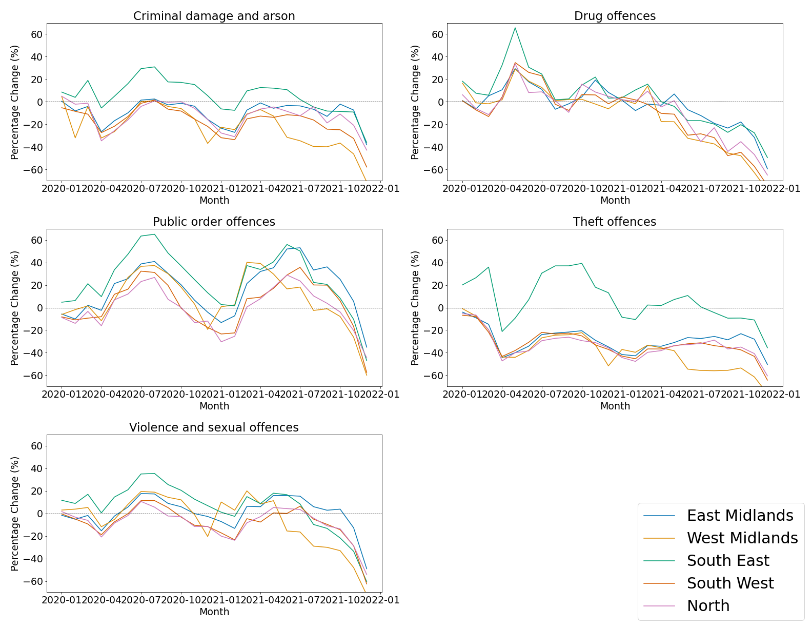


Figure 5 - Percentage Change to 2019 by Region

It is important to roll up at points when visualising this type of data. Fig 6 shows the actual counts of the crime types through the period. Whilst drug offences have a large spike in Fig 4 & Fig 5, the actual increase is only 20,000. It is clear to see that criminal damage, drug offences & public order offences don’t have as large raw changes as theft offences & violence and sexual offences, with those differing in the hundreds of thousands.

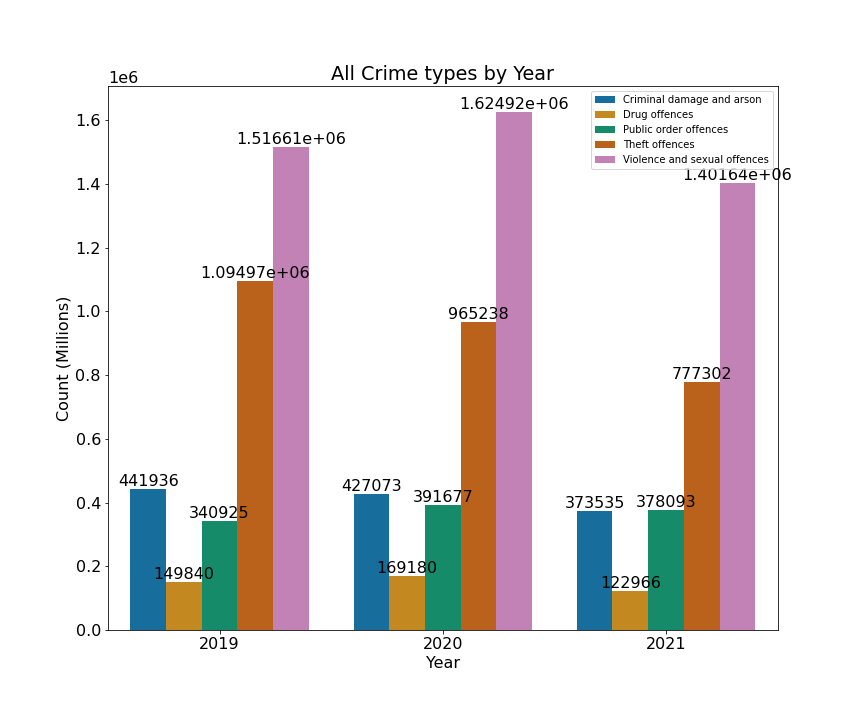


Figure 6 - Crime Counts by Year

## Outcomes:

The recorded levels of crime appear to have changed over the lockdown periods, however the outcomes regarding these arrests does not appear to have moved significantly. The highest change is within theft offences between 2019 and 2020, with an increase of 5% that there will be no suspect identified. This could be a result of less thefts occurring during covid (fig 6) as well as police resources being stretched thin to enforce lockdown, therefore theft crimes that may have usually been investigated would have not received the normal amount of attention.

Another change is noticed in drug offences, with ‘other outcomes’ rising by around 5% each year from 56% in 2019 to 65% in 2021. This difference comes directly opposite ‘suspect charged’ which decreases 5% each year. This is may be evidence of the police and the judicial system attempting to reduce drug crime using other methods such as rehabilitation programmes, rather than charging drug users which may not provide them with the help to kick their addiction. Drug crimes also include drug dealers which would be adequately charged and it is unlikely they would be given another outcome such as a caution or fine.

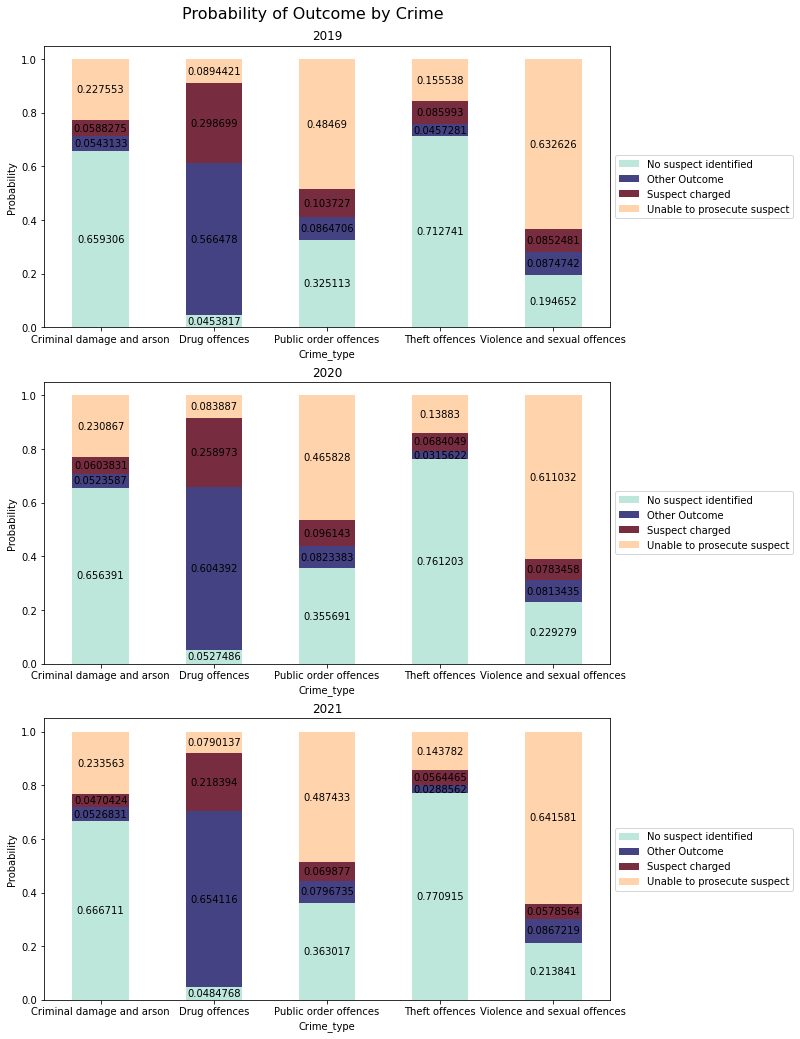


Figure 7 - Probability of Outcome by Crime

The outcomes can also be stratified by region. Not all regions have the same reaction to lockdowns being introduced, with the South East increasing the ‘no suspect identified’ from 47% in 2019 to 56% in 2020 whilst all other regions decrease in the same outcome. The difference for South East compared to the other regions could be related to the disproportionate increase of certain crimes such as theft offences as seen in Figures 4 & 5. As seen in Figure 7, it is known that theft offences have a high probability of no suspect being identified, therefore a high increase in theft offences for a given region may increase the overall probability of no suspect being identified.

The probability for a suspect to be charged also decreases for all regions from 2019 to 2021. Like in Fig 7, both ‘other outcome’ and ‘suspect charged’ are the minority probabilities, with only small changes occurring over the period, however each region has very similar ranges (approximately 5% to 10%) for both of these outcomes. This differs to the ranges of ‘no suspect identified’ and ‘unable to prosecute suspect’, with all regions having very different levels. For example, in 2020, the South East has a 56% probability of ‘no suspect identified’ whilst the South West has half this value at 28%. What this appears to show is that some regions are better overall at identifying suspects, but may lack the ability to actually prosecute them. This is evidenced by the fact that all regions have similar ‘suspect charged’ probabilities yet wildly varying ‘unable to prosecute suspect’ ranges. A number of factors could influence this such as the size of the region, the number of police officers as well as the public attitude towards police, each regions level of focus on different crime types and the prevalence of certain types of crime in each region.

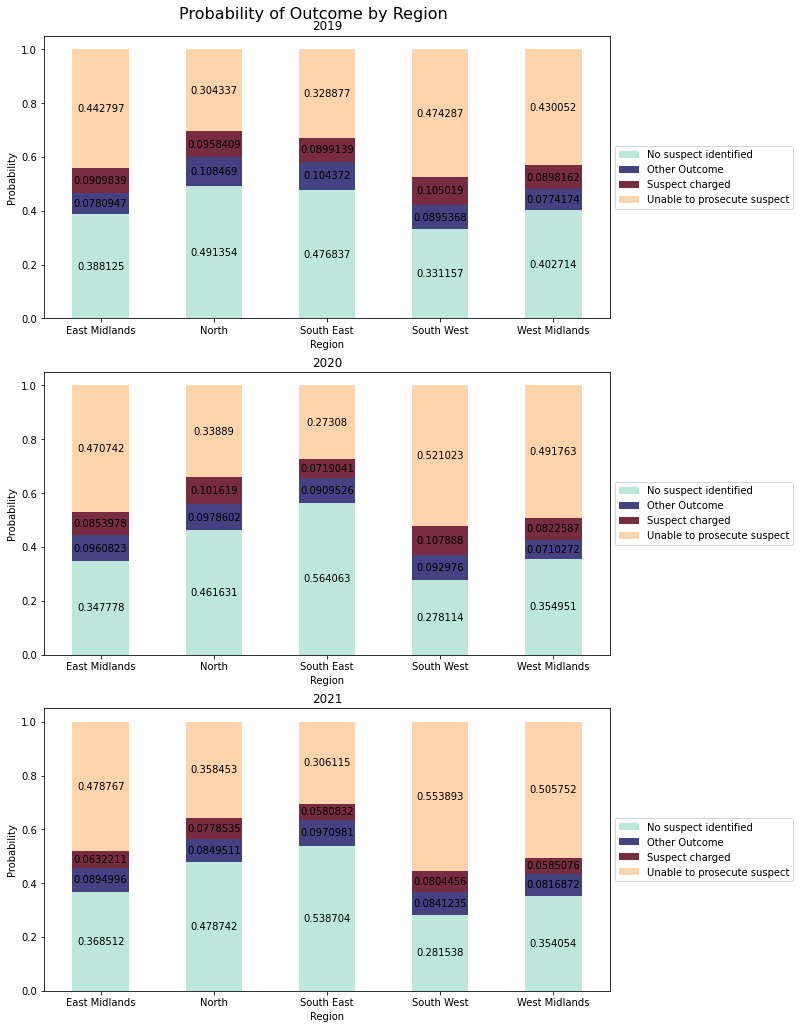
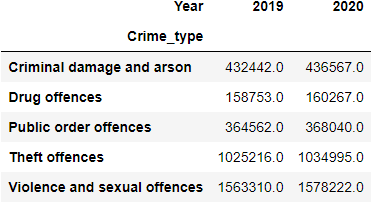
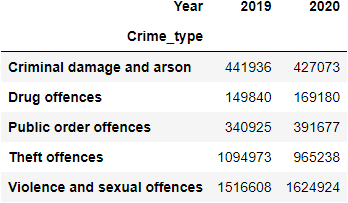


Figure 8 - Probability of Outcome by Region

## Chi-Squared Test:

The Chi-Squared test highlighted an issue with statistical tests of this nature and relying on p-values (Lin et al, 2013).

Tables 3 & 4 - Actual Crime Rates and Expected Crime Rates 2019 - 2020



Tables 3 and 4 represent the reported crime counts and the expected crime counts. Due to the large sample size, the resulting chi-squared statistic is extremely high at 16686, which results in a p-value of 0. Initially, this may be seen as a positive and statistically significant result. However, when the same test is performed on the data from 2017 and 2018 (tables 5 and 6 in appendix A), two non-covid years. A similarly high test statistic of 20915 is recorded with a p-value of 0. These p-values would allow us to reject the null hypothesis that both year and crime rate are independent of one another but due to the fact that 2017 and 2018 are also significantly different, it does not appear to be useful.

The chi-squared test is not taking into account any external factors such as coronavirus or historical data in generating its expected values, it is simply using the original values to proportion differently. However, removing the reliance on p-values and test statistics, it is still interesting to see how the true values differ from the expected values. In all cases in the reported values, the difference between 2019 and 2020 is greater than the difference in the expected values. Also, the difference between 2019 and 2020 for all crime types is approximately the same or less than the difference for all crime types from 2017 to 2018. For example, violent offences increased by approximately 110,000 from 2019 to 2020 whilst they increased 210,000 from 2017 to 2018. This highlights another issue with relying on p-values, that although these are technically statistically significant results, the magnitudes are not adequately displayed. Once over a certain sample size, a difference of 100,000 and 200,000 will both result in very low p-values.

## Bayesian Network:

Firstly, a small network was learned on the data to investigate the relationships between the variables of region, lockdown level, crime type and outcome type (Fig 9). This provides a more interactive version of the data produced in figures 7 & 8. To keep the proportions the same, each year has been given its own level of lockdown, with 2019 being ‘No Lockdown’, 2020 being ‘High’ and 2021 being ‘Medium’. It is found that there is a higher probability of a public order offence occurring during 2020 and 2021, this could be due to the enforcement of lockdown restrictions as well as mask mandates leading to more public order disturbances. This method also allows us to investigate some of the minority probabilities (suspect charged and other outcome) that were seen in figures 7 & 8. When selecting these in ‘Outcome type’, we find that if a suspect is charged, it was most likely within 2020 (35.5%), in the South East (37.44% and for a violent or sexual offence (42.16%). Similarly, if ‘other outcome’ is selected, it is most likely to have occurred during 2019 (40.57%) and again within the South East (32.48%) for violent or sexual offences (40.7%). The model is available at: <https://crimeduringcovid.staging.agenarisk.app/> Please email elliot.linsey@gmail.com to request access.



Figure 9 - Small Bayesian Network

The larger Bayesian model (Fig. 10 in appendix) is used to model the uncertainty regarding the true number of crimes committed versus the reported number of crimes. Using this model in combination with the prior knowledge allows us to estimate the true number of crimes within different covid periods. The Bayesian network is a use of ‘smart data’ rather than ‘big data’ which is what the previous statistical methods have been using. The Bayesian network is able to answer questions of an interventional and counterfactual nature, whereas classical statistics are limited to finding associations between variables and lack the causal structure behind them. Using the integration of prior knowledge, the Bayesian network is able to answer questions such as: “How many drug offences could we expect to be reported every month during a lockdown period, if the police have a high focus on this crime type and there are less than 20 officers per 10,000 citizens?” A classical machine learning model may only learn the association between drug offences and these variables, whereas a Bayesian network is not limited by what is explicitly shown in the data and can make inferences that big data lacks.

Table 3 records the rates (rounded to the nearest 1000) of actual and reported crimes during both lockdown and non-lockdown periods. The reported figures here are when the accuracy value is unknown, if we were to set the accuracy level for violence and sexual assaults to ‘maximal underestimate’ it would report 50% of the actual number of violent crimes, so approximately 132,000 would be reported.

Table 7 - Actual and Reported Crime estimation during Lockdown and non-Lockdown periods

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Lockdown** | | **Non-Lockdown** | |
| **Crime Type** | **Actual** | **Reported** | **Actual** | **Reported** |
| Drugs | 168,000 | 117,000 | 120,000 | 83,000 |
| Violence | 264,000 | 190,000 | 220,000 | 158,000 |
| Criminal Damage | 72,000 | 47,000 | 82,000 | 54,000 |
| Theft | 95,000 | 59,000 | 126,000 | 79,000 |
| Public Order | 65,000 | 43,000 | 52,000 | 34,000 |
| **Total** | 664,000 | 456,000 | 600,000 | 408,000 |

Overall, the table and model are reporting the fact that both actual and reported crime increased in total over the coronavirus period, however some forms decreased such as theft & criminal damage and arson. The other three crime types of drugs, violence and sexual offences & public order offences increased over the lockdown period. This model is available to use at <https://actualvsreportedcovidcrimes.staging.agenarisk.app/> please email elliot.linsey@gmail.com for access.

A highlight of using these types of Bayesian networks is the accessibility of the models to the general population. This has only been a recent development, but a high level of mathematical understanding is no longer required to use the models presented. The simple process of selecting the choices (as well as being able to leave unknown information blank) means these models are easy to use and the results can be personalised to each situation. There is no formal training required in their usage and any knowledge of Bayesian methods needed is minimal.

# Conclusion:

In this study, a number of methods were used to investigate the relationship between coronavirus lockdowns and reported crime rates, outcomes & actual crime rates over five crime types. A Bayesian network was used to model the uncertainty regarding true crime rates, due to the fact that all the data was reliant on police reported figures. The coronavirus pandemic was unique in how it significantly altered the way of life of the UK (as well as global) population. It was found that the effect of the lockdown did result in significant changes for most crime types when compared to their predicted values. However, the effect did not appear to be long-lasting and instead is characterised as a sudden drop, then a quick rise to stable levels.

This effect was seen in all regions over the UK, with extremely similar patterns being observed. However, some were more affected than others with greater increases and decreases. The outcomes were not greatly affected, instead the differences between the regions and crime types were already present before coronavirus occurred.

There was a stark difference in the use of classical statistical methods such as the chi-squared test which relies on the use of p-values compared to using a Bayesian network. Due to the sample size, the chi-squared test became difficult to interpret with no ability to discern the magnitude of the yearly changes. The use of time-series based data may not have been ideal for the chi-squared test, due to the difficulty in proving that each year is inherently independent from one another. The Bayesian network on the other hand was able to model the effect that lockdown had on crime rates and produce information that was not solely based on association.

# Further Research:

This paper has opened up many more avenues of investigation into the coronavirus and crime topic. The subdivision of each crime type could be more deeply researched. For example, within violent and sexual offences there is the indication that crimes such as domestic violence would have increased whilst others such as drunken fighting may have decreased. Drilling down into these subdivisions may show evidence that coronavirus lockdowns had different levels of effect within each crime type and could lead to the identification of areas that need support in a post-covid world. As the data is subdivided, it could also raise occurrences of paradoxes such as Simpson’s paradox, these would be interesting and cause lead to greater cause for investigation into whichever variable contains the paradox.

Similarly, all the regions were concatenated into 5 groups from 42 police forces. Whilst the trends showed that all moved in very similar patterns, there is definite scope for looking at each individual police force and identifying those that had anomalous results. From these forces, further research into their specific policing methods could be conducted so as to identify causes or variables that may be useful in combating crime around the UK.

The Bayesian network has room for expansion, with the identification of more causes and variables to produce more accurate modelling of the lockdown effect, as well as to model crime post-covid. Some factors that could be investigated would be the mobility of the population. Google analytics produced community mobility reports which showed how populations changed their movement trends during the pandemic. Another factor could be how the economic state of the UK and its citizens affected crime rates. Were those citizens that received furlough less likely to commit crimes than those that did not receive or received less furlough?

# References:

Home Office. (2022). ‘Whole-system’ approach to tackling violent crime is working. <https://www.gov.uk/government/news/whole-system-approach-to-tackling-violent-crime-is-working> accessed [22/07/22]

The Independent. (2017). Met Police to stop investigating some 'low-level' crimes in response to £400m funding cuts. Link: <https://www.independent.co.uk/news/uk/crime/met-police-spending-cuts-400-million-funding-london-crimes-not-investigated-burglary-assault-a8002746.html> accessed [22/07/22]  
  
Institute for Government (IfG). (2022). Timeline of UK government coronavirus lockdowns and restrictions. <https://www.instituteforgovernment.org.uk/charts/uk-government-coronavirus-lockdowns> accessed [29/06/22]

Institute for Race Relations (IRR). (2020). Racial violence statistics. Link: <https://irr.org.uk/research/statistics/racial-violence/> accessed [22/07/22]   
  
Johnson, B. (2020). Prime Minister’s statement on coronavirus 25th March. <https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-25-march-2020> accessed [29/06/22]

Johnson, B. (2022). From harm to hope: A 10-year drugs plan to cut crime and save lives. Link: <https://www.gov.uk/government/publications/from-harm-to-hope-a-10-year-drugs-plan-to-cut-crime-and-save-lives/from-harm-to-hope-a-10-year-drugs-plan-to-cut-crime-and-save-lives> accessed [22/07/22]  
  
Langton, S. Dixon, A. Farrell, G. (2021). Six months in: Pandemic Crime trends in England and Wales. Doi: <https://doi.org/10.1186/s40163-021-00142-z> accessed [29/06/22]  
  
Lecocq, T. et al. (2020). Global quieting of high-frequency seismic noise due to COVID-19 pandemic lockdown measures. <https://www.science.org/doi/10.1126/science.abd2438> accessed [30/06/22]

Ma, L. Wu, C. (2022) The Impact of COVID-19 on Crime: A Study from the Spatial-temporal Perspective in the Montgomery County, AL. <https://pesquisa.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/resource/pt/covidwho-1798809> accessed [01/07/22]

Lin, Mingfeng. Lucas, Henry Jr. Shmueli, Galit. (2013) Research Commentary—Too Big to Fail: Large Samples and the *p*-Value Problem. Information Systems Research 24(4):906-917. <https://doi.org/10.1287/isre.2013.0480> accessed [26/07/22]

Mlambo-Ngcuka, P. (2020). Violence against women and girls: The shadow pandemic. <https://www.unwomen.org/en/news/stories/2020/4/statement-ed-phumzile-violence-against-women-during-pandemic> accessed [01/07/22]  
  
Nivette, A.E., Zahnow, R., Aguilar, R. et al. (2021). A global analysis of the impact of COVID-19 stay-at-home restrictions on crime. Nat Hum Behav 5, 868–877 <https://doi.org/10.1038/s41562-021-01139-z>

Office for Health Improvement & Disparities. (2021). Adult substance misuse treatment statistics 2020 to 2021: report. <https://www.gov.uk/government/statistics/substance-misuse-treatment-for-adults-statistics-2020-to-2021/adult-substance-misuse-treatment-statistics-2020-to-2021-report> accessed [01/07/22]  
  
Office for National Statistics (ONS), (2021). An overview of workers who were furloughed in the UK. Link: <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/anoverviewofworkerswhowerefurloughedintheuk/october2021> accessed [29/06/22]  
  
Office for National Statistics (ONS), (2020). The impact of the coronavirus so far: the industries that struggled or recovered. Link: <https://www.ons.gov.uk/economy/economicoutputandproductivity/output/articles/theimpactofthecoronavirussofartheindustriesthatstruggledorrecovered/2020-12-09> accessed [29/06/22]

Office for National Statistics (ONS), (2020). Drug Misuse in England and Wales – Appendix table. Link: <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/datasets/drugmisuseinenglandandwalesappendixtable> accessed [22/07/22]

Office for National Statistics (ONS), (2020). The nature of violent crime in England and Wales. Link: <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/articles/thenatureofviolentcrimeinenglandandwales/yearendingmarch2020#reporting-violent-crime-to-the-police> accessed [22/07/22]

Office for National Statistics (ONS), (2020). Crime appendix tables. Link: <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/datasets/crimeinenglandandwalesappendixtables> accessed [22/07/22]

Payne, J. Morgan, A. (2020) COVID-19 AND VIOLENT CRIME: A COMPARISON OF RECORDED OFFENCE RATES AND DYNAMIC FORECASTS (ARIMA) FOR MARCH 2020 IN QUEENSLAND, AUSTRALIA. <https://osf.io/preprints/socarxiv/g4kh7/> accessed [01/07/22]

Ramalho, R. (2020). Alcohol consumption and alcohol-related problems during the COVID-19 pandemic: a narrative review. Australasian Psychiatry. 2020;28(5):524-526. doi:10.1177/1039856220943024

Sri AS, Das P, Gnanapragasam S, Persaud A. (2021). COVID-19 and the violence against women and girls: ‘The shadow pandemic.’ International Journal of Social Psychiatry.;67(8):971-973. doi:10.1177/0020764021995556

Taub, A., Bradley, J. (2020). As domestic abuse rises, UK failings leave victims in peril. <https://www.nytimes.com/interactive/2020/07/02/world/europe/uk-coronavirus-domestic-abuse.html> accessed [01/07/22]

UK Gov. (2020). Police given new powers and support to respond to coronavirus. <https://www.gov.uk/government/news/police-given-new-powers-and-support-to-respond-to-coronavirus> accessed [29/06/22]

Unicef. (2022). Gender-based violence in emergencies. <https://www.unicef.org/protection/gender-based-violence-in-emergencies> accessed [01/07/22]

Women’s Aid. (2021). SHADOW PANDEMIC – SHINING A LIGHT ON DOMESTIC ABUSE DURING COVID. <https://www.womensaid.org.uk/wp-content/uploads/2021/11/Shadow_Pandemic_Report_FINAL.pdf> accessed [01/07/22]

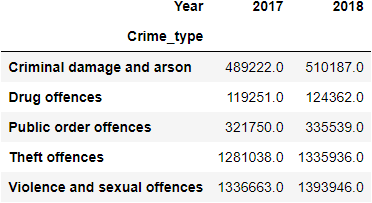
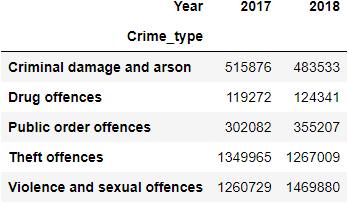
## Appendix A Tables :

Table - Timeline of Lockdown events

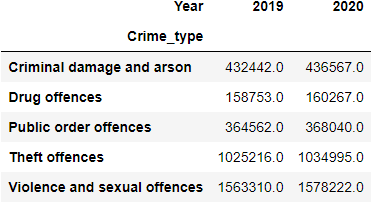
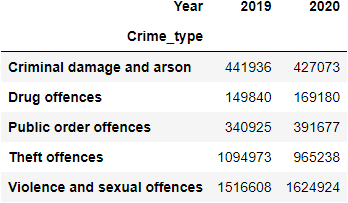
|  |  |
| --- | --- |
| March 26 (2020) | Lockdown measures legally come into force; police given powers to enforce restrictions |
| Apr 3 | Figures from transport office demonstrate greatly reduced usage of motor vehicles, rails and buses (Department for Transport, 2020) |
| april 16 | Lockdown extended for “at least” 3 weeks |
| April 30 | PM says “we are past the peak” of the pandemic |
| may 10 | People that cannot work from home should return but avoid public transport |
| june 1 | Phased re-opening of schools in England. The “Rule of six” is implemented, allowing six people from different households to meet outdoors. |
| june 15 | Non-essential shops reopen in England |
| june 23 | Relaxation of restrictions and 2m social distancing rule |
| june 29 | First local lockdown is announced in Leicester |
| july 4 | More restrictions are eased, with pubs and restaurants being allowed to operate |
| july 17 | Public transport can be used for non-essential purposes |
| july 18 | Authorities gain additional powers to enforce social distancing, such as restrictions on those not wearing face coverings (DHSC, 2020) |
| august 1 | Shielding guidelines for the vulnerable cease. Permitting 2 million people to leave home and return to work |
| August 3 | “Eat out to help out” scheme with discounts on meals at restaurants |
| august 14 | Further relaxation of lockdown, theatres and bowling alleys open. |
|  |  |
| september 14 | “Rule of six” gatherings are banned |
| september 22 | Return to working from home and a 10pm curfew for hospitality sector |
| october 14 | New three-tier system of restrictions is implemented |
| november 5 | Second lockdown comes into force |
| november 24 | Announced that 3 households can meet during the Christmas period |
| december 2 | Second lockdown ends |
| december 15 | Rules are to be relaxed over Christmas, but advised to keep celebrations small |
| december 19 | Tier 4 “Stay at Home” restriction is introduced. |
| december 21 | London and South East England put under tier 4 restrictions |
| december 26 | More areas of England enter tier 4 restrictions |
| january 6 (2021) | Enter third national lockdown |
| february 15 | Hotel quarantines for travellers from high-risk countries |
| march 8 | Schools open for primary and secondary school students. Outdoor recreation allowed between two people. “Stay at home” order remains in place |
| march 29 | “Rule of six” is allowed again, for outdoors. “Stay at home” order ends |
| April 12 | Non-essential retail shops reopen. Outdoor venues including pubs and restaurants reopen |
| may 17 | “Rule of six” is now allowed for indoors, 30 people allowed to mix outdoors. Indoor venues of pubs and restaurants reopen. |
| july 19 | Most legal limits on social contact removed, final sectors reopen such as nightclubs |

Table - Region Populations

|  |  |
| --- | --- |
| **Region** | **Population** |
| South East | 22,145,000 |
| South West | 9,937,900 |
| North | 3,977,308 |
| East Midlands | 10,475,600 |
| West Midlands | 9,584,002 |



Tables 5 & 6 - Actual Crime Rates and Expected Crime Rates 2017 - 2018



Tables 3 & 4 - Actual Crime Rates and Expected Crime Rates 2019 - 2020

Table 7 - Actual and Reported Crime estimation during Lockdown and non-Lockdown periods

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Lockdown** | | **Non-Lockdown** | |
| **Crime Type** | **Actual** | **Reported** | **Actual** | **Reported** |
| Drugs | 168,000 | 117,000 | 120,000 | 83,000 |
| Violence | 264,000 | 190,000 | 220,000 | 158,000 |
| Criminal Damage | 72,000 | 47,000 | 82,000 | 54,000 |
| Theft | 95,000 | 59,000 | 126,000 | 79,000 |
| Public Order | 65,000 | 43,000 | 52,000 | 34,000 |
| **Total** | 664,000 | 456,000 | 600,000 | 408,000 |

## Appendix B Figures

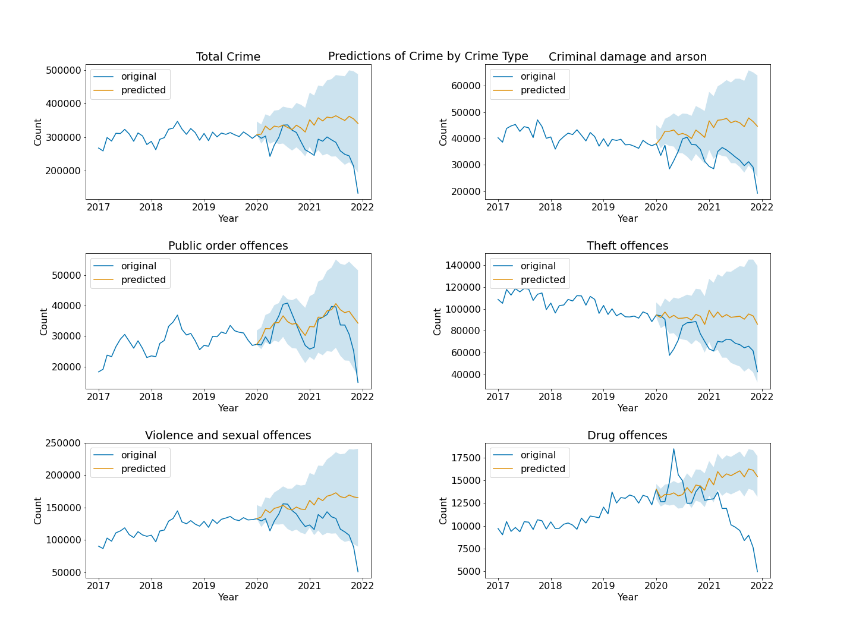


Figure 2 - Predictions of Crime by Crime Type

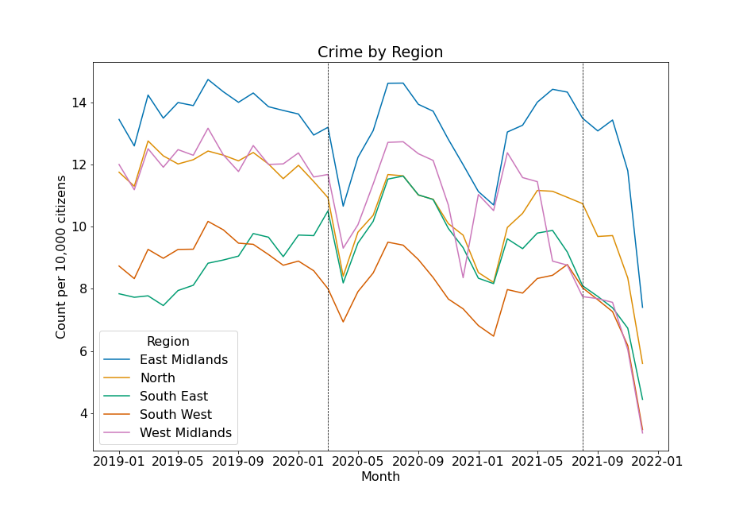


Figure 3 - Regional Crime per 10,000 Citizens

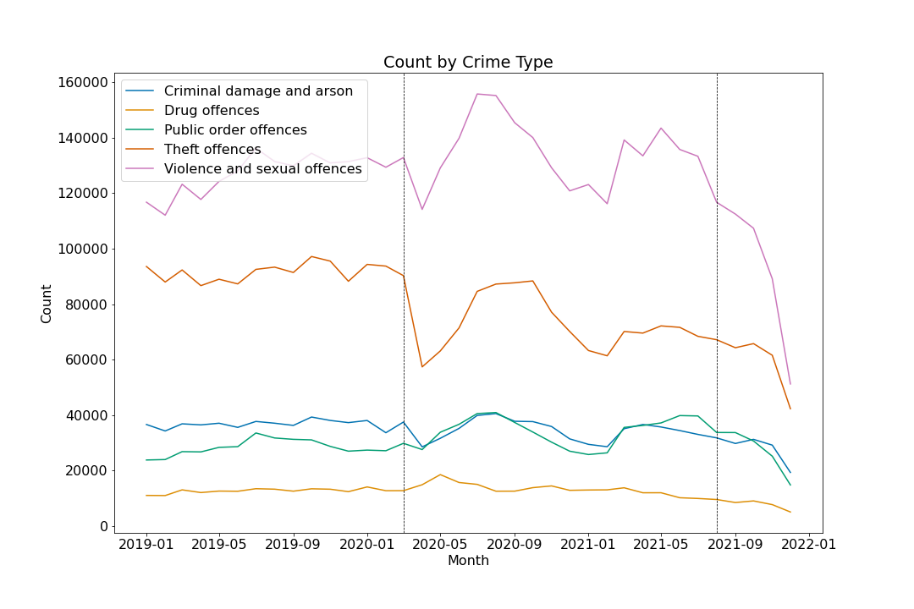


Figure 1 - Count by Crime Type

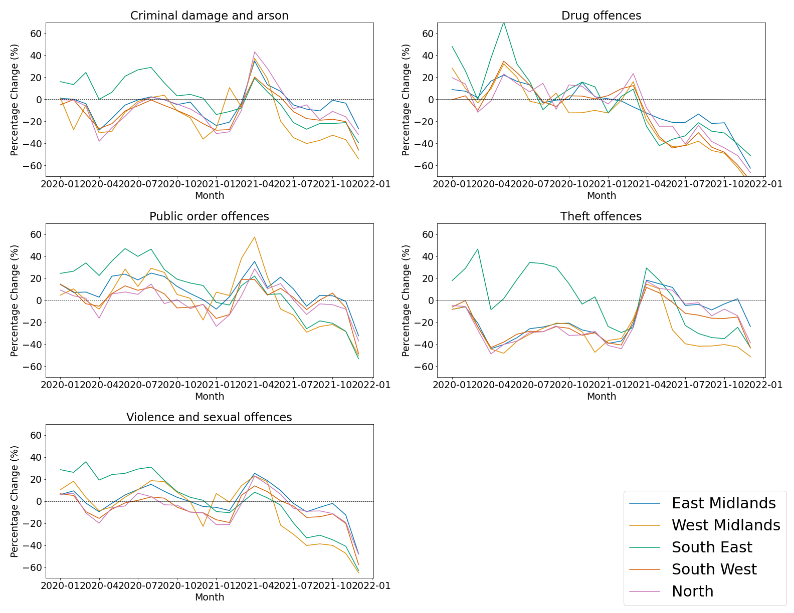


Figure 4 - Percentage Change to Previous Year by Region

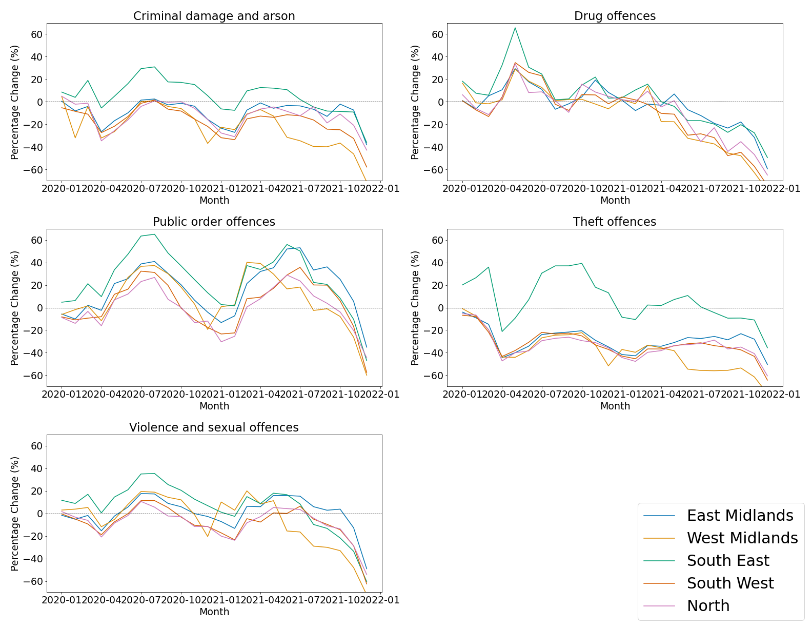


Figure 5 - Percentage Change to 2019 by Region

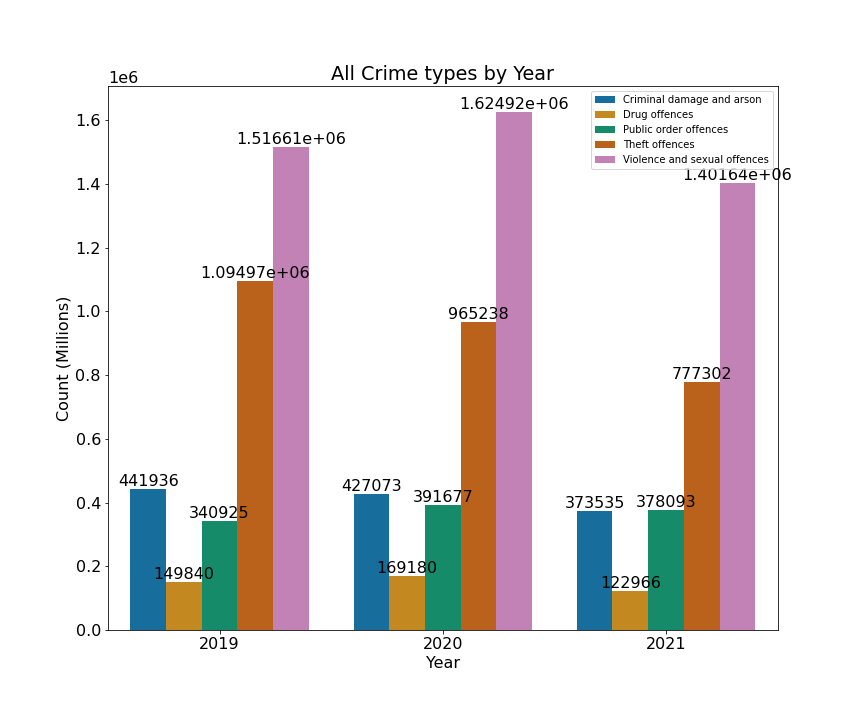


Figure 6 - Crime Counts by Year

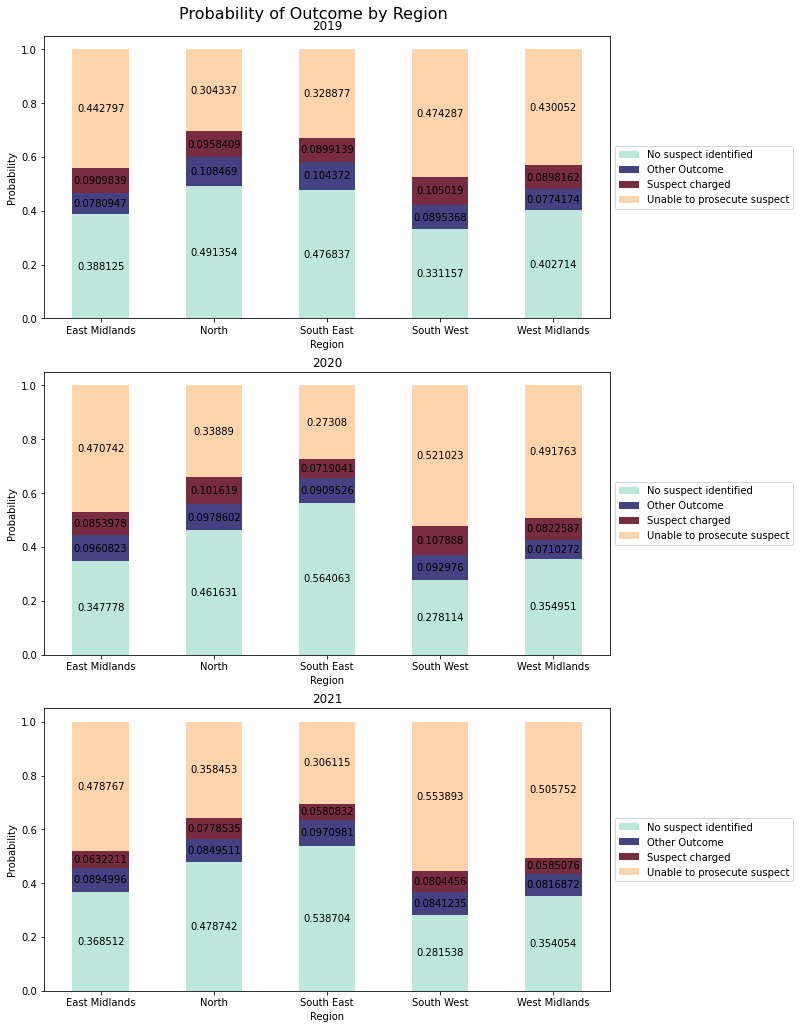


Figure 8 - Probability of Outcome by Region

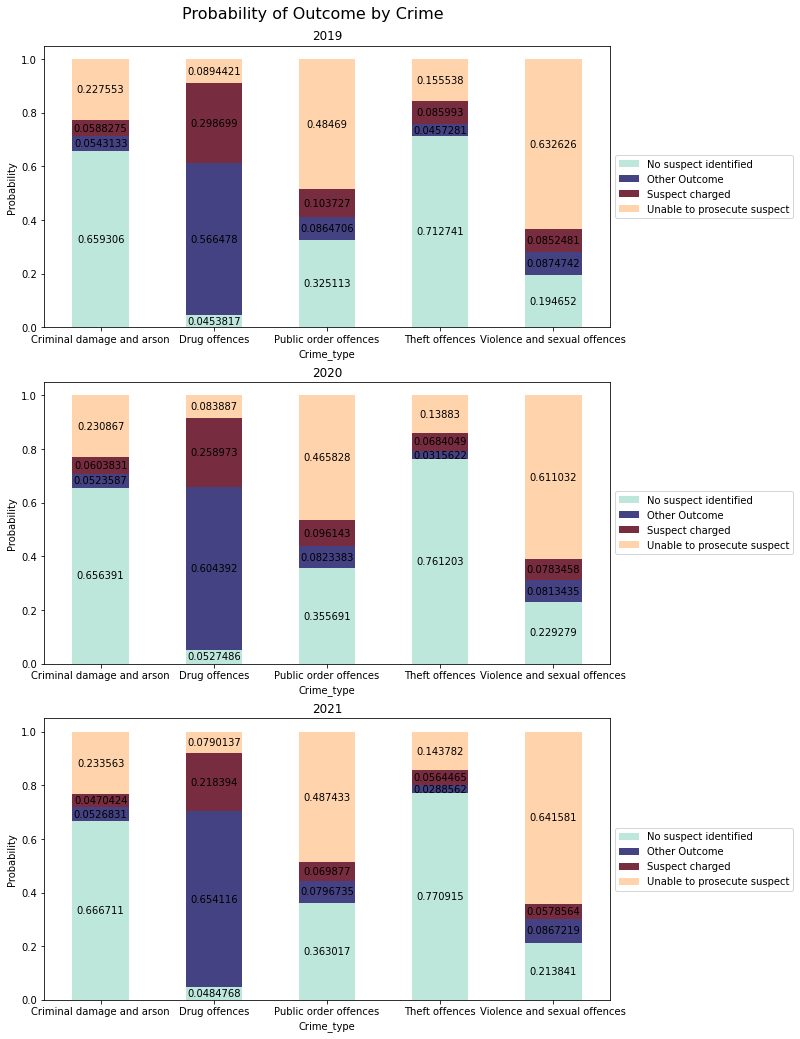


Figure 7 - Probability of Outcome by Crime



Figure 9 - Small Bayesian Network

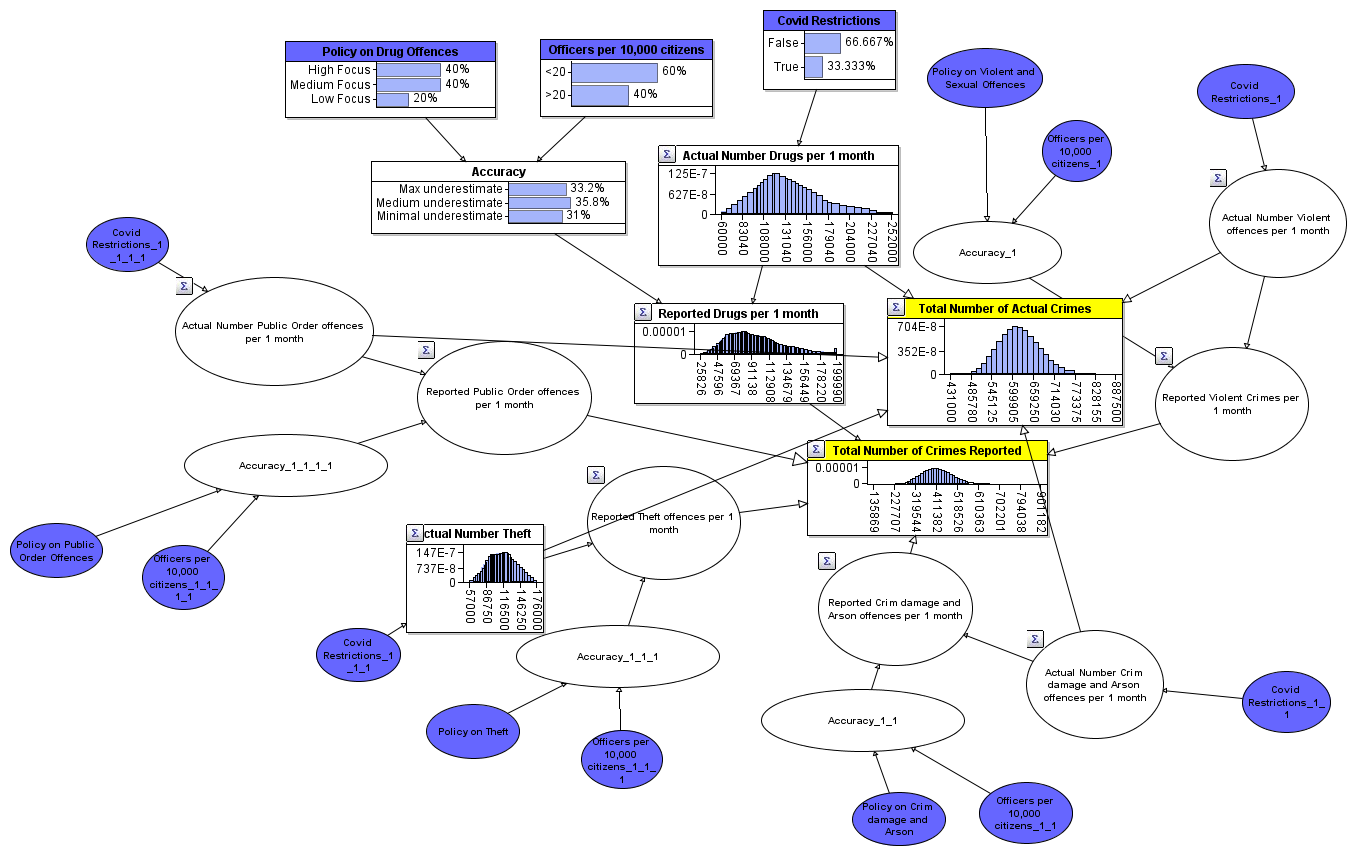


Figure - Estimating Actual Crime Bayesian Network