**Analysis of how the pandemic affected the levels of different crimes across the UK**

Emma Davis, Erin Ludwig, Tia-Louise Way

[**Defining the question**](#_de02g76zp029) **2**

[**Data collecting**](#_79wesivv1znv) **3**

[**Data audit**](#_he05czd99rd1) **4**

[**Data cleansing**](#_3kfzzeres687) **5**

[**Data Analysis**](#_r6ooaalfaj1i) **5**

Was there a difference between the lockdown months compared to the rest of the pandemic? 5

Comparing the lockdowns to the whole data period (2019-2021) 9

Comparing different areas over the 3 years 11

[**Conclusion**](#_am2h8uefl4f) **and lessons learnt 19**

**Defining the question**

Using data from crimes recorded in England and Wales from January 2019 to December 2021, this project will seek to answer the following question:

**How did the level of crime in the UK change during the Covid-19 pandemic?**

We had an initial idea of what we thought the results would show, namely that crime went down during the lockdown periods, but not enough of an idea to create a hypothesis, which is why we went with a broad question. This will also allow us to investigate further based on our initial findings and further questions they might raise.

# **Data collecting**

There are several different options for accessing datasets of UK crime statistics, including the Office of National Statistics, and this project will use a 3-year data collection from data.police.uk.

This site provides open data for crime and policing in England, Wales, and Northern Ireland.

The specific dataset being used is "December 2021 (1.5 GB) - Contains data from Jan 2019 to Dec 2021- 98c67474d3607171b772826f52ad1703". To compare data during the pandemic to before the pandemic this is a suitable dataset to use because it contains 2019 and Jan-Mar 2020, which is pre-pandemic, included in the same set of files as Apr 2020 onwards.

**Data audit**

This dataset downloads with folders for each month in the three-year period, with CSVs for each police force covering reported crimes, stop and search data, and outcomes.

There are over 4600 files in the dataset with a total size of over 5GB.

The files are all in a standard format however, so combining them shouldn’t be an issue as far as column headers etc are concerned.

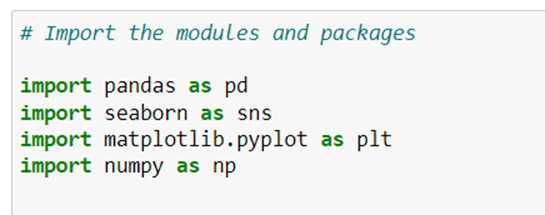
We considered expanding the timeframe to include more pre-pandemic years, but given the size of the data we have already downloaded we did not feel that this was feasible.

**Data cleansing**

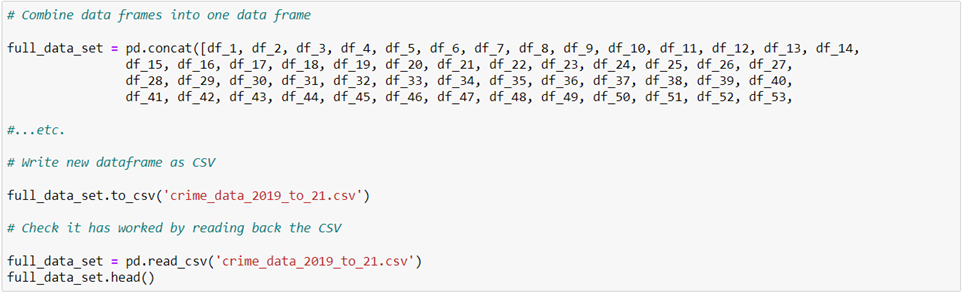
The project only requires the street CSVs, which contain details on crimes reported, the location, the month, and the latest known outcome. All other files can be discarded. This leaves 1582 files.

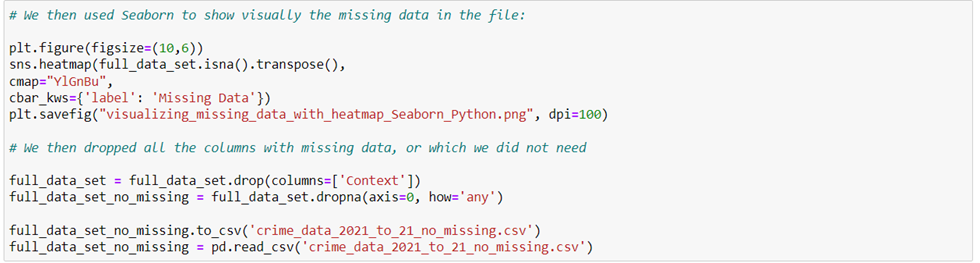
All of the CSVs for the individual months and individual police forces need to be compiled into one file/dataframe.

The code used for data cleansing can be found in the file 'CFG Project - ED EL TW.ipynb', with snippets below:

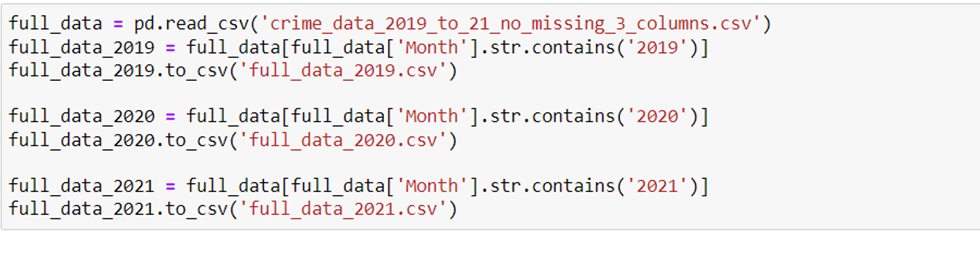








Having completed the above, we found that the file was still over 1GB (it had started at over 4GB), so we chose to save each of the three years separately in their own files.



# 

# **Data analysis**

## Was there a difference between the lockdown months compared to the rest of the pandemic?

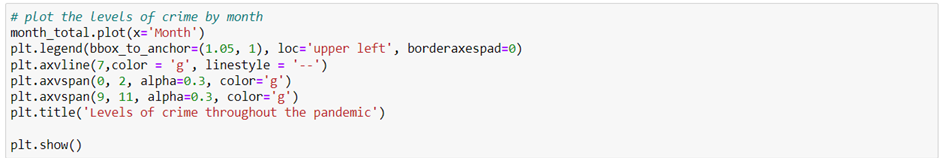
We determined that the months we should use for comparing lockdown statistics are:

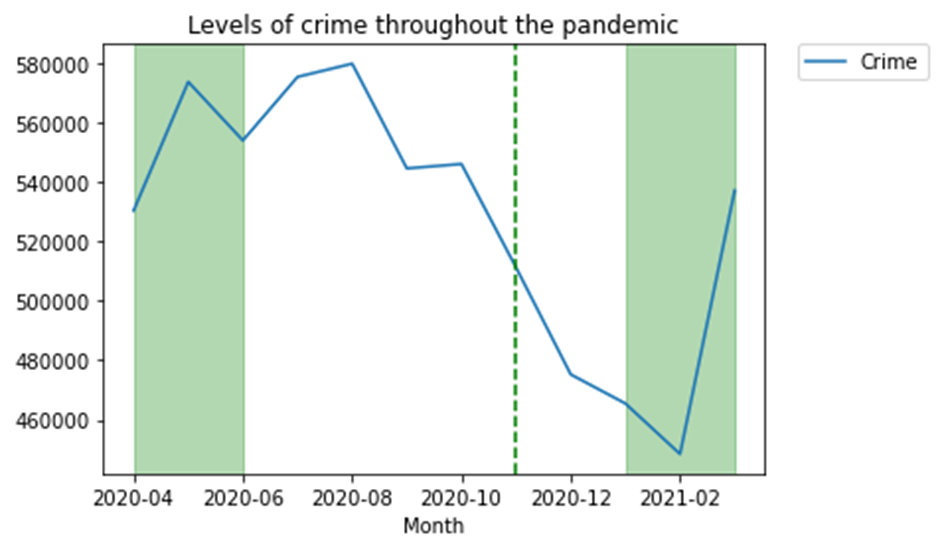
* Lockdown 1 - April 2020 (lockdown came into force 26th March), May 2020, June 2020 (restrictions relaxed on 23rd June)
* Lockdown 2 - November 2020 (Lockdown started 5th November, ended 2nd December)
* Lockdown 3 - January 2021 (lockdown starts 6th January), February 2021, March 2021 (lockdown ends 29th March)

To determine whether there was a difference in the crime levels between the lockdown periods and the rest of the pandemic, we will look at April 2020 to March 2021. This gives us the start of the first lockdown until the end of the third lockdown, and all the non-lockdown months in between.

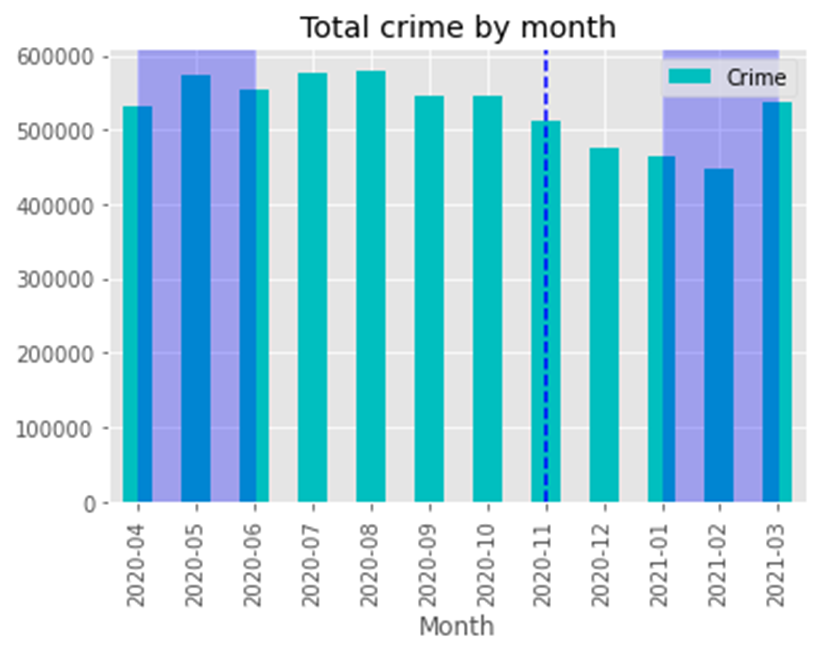






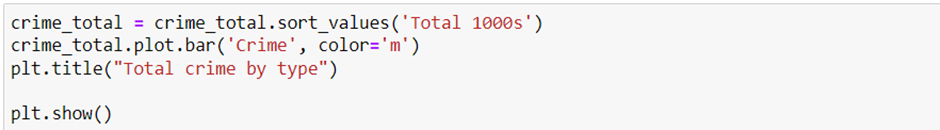


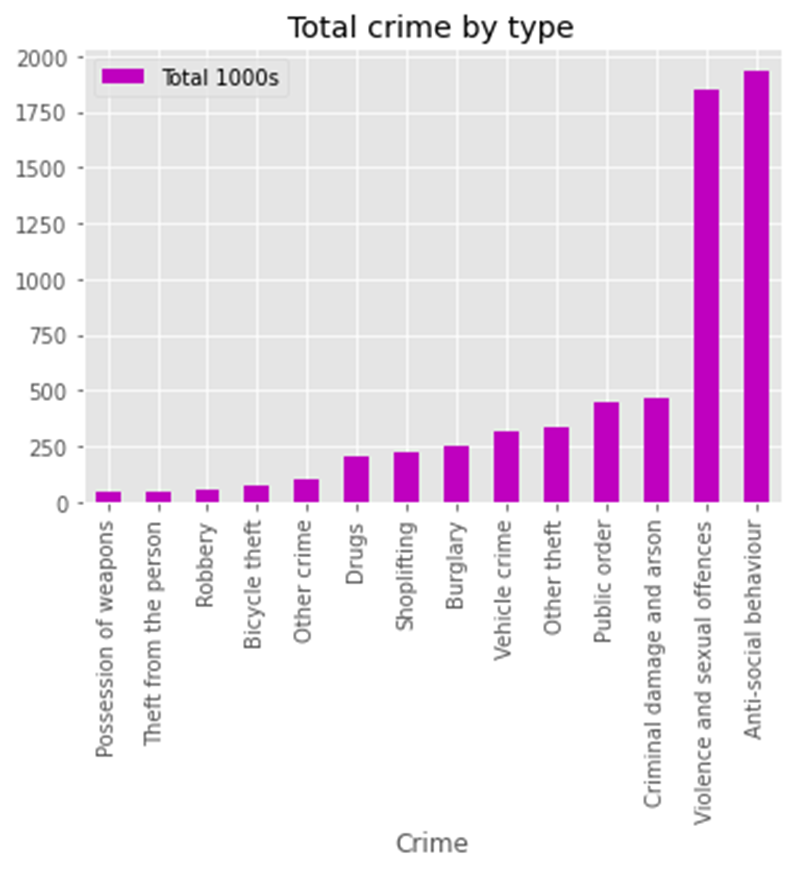
This would look better as a bar graph and with the y axis starting at 0 so you get a true reflection of the trend.



Now we will create a graph for the total number of crimes by crime type

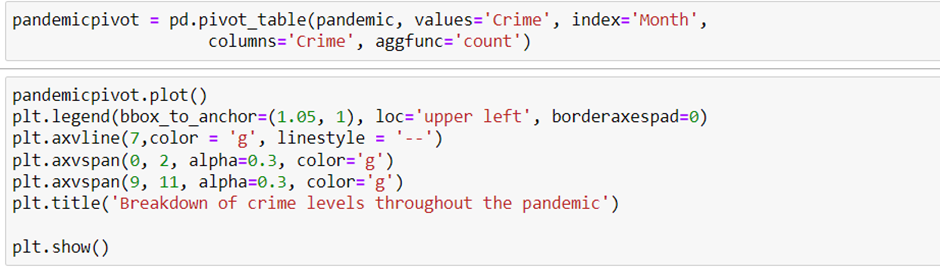


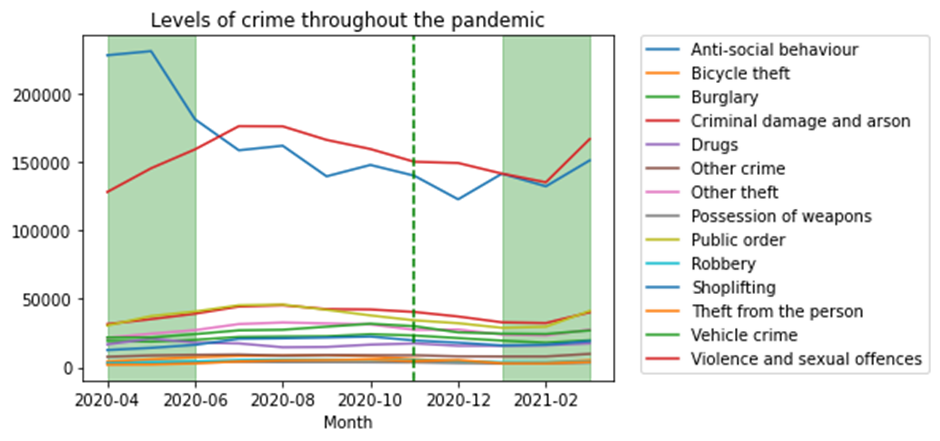




Although there is a slight increase in the total level of crimes in May-20, the number falls slightly in June and then increases again for the two non-lockdown months of July-20 and August-20. Then we can see that the total number of crimes began to decrease towards the end of 2020, even during the second lockdown in November. In March-21 the number rises again.

This would suggest that there isn't necessarily a difference in the level of crime during lockdowns and throughout the pandemic as a whole. The two most prevalent crime types are anti-social behaviour, and violence and sexual offences.





When we compare the green lockdown periods to the rest of the pandemic, we can see that anti-social behaviour starts very high, but towards the end of the first lockdown begins to drop and then drops again outside of the lockdown. It begins to rise slightly during the third lockdown.

The other noticeable observation is that violence and sexual offences increase during the first lockdown, continue to rise slightly afterwards, but then fall again until rising during the third lockdown.

All other crimes have a consistent level whether throughout lockdowns or non-lockdown periods.

## Comparing the lockdowns to the whole data period (2019-2021)

# 

# A wider look at the full datasets even further highlights the spike in anti-social behaviour as the first lockdown hit in April 2020. This month saw the highest levels of anti-social behaviour in the whole period from January 2019 until January 2021. It is hard to determine whether the lockdowns alone were the cause of this spike as the level of anti-social behaviour dropped during the second lockdown and then fluctuated during the third. There is no clear pattern to demonstrate the pandemic as a cause.

There is some small scale fluctuation shown in the graph as the lockdown period hits. Many crimes see a decrease in frequency as restrictions are enforced.

# Comparing the crime types year on year shows that the proportion of violence and sexual offences compared to other crimes have increased since 2019.

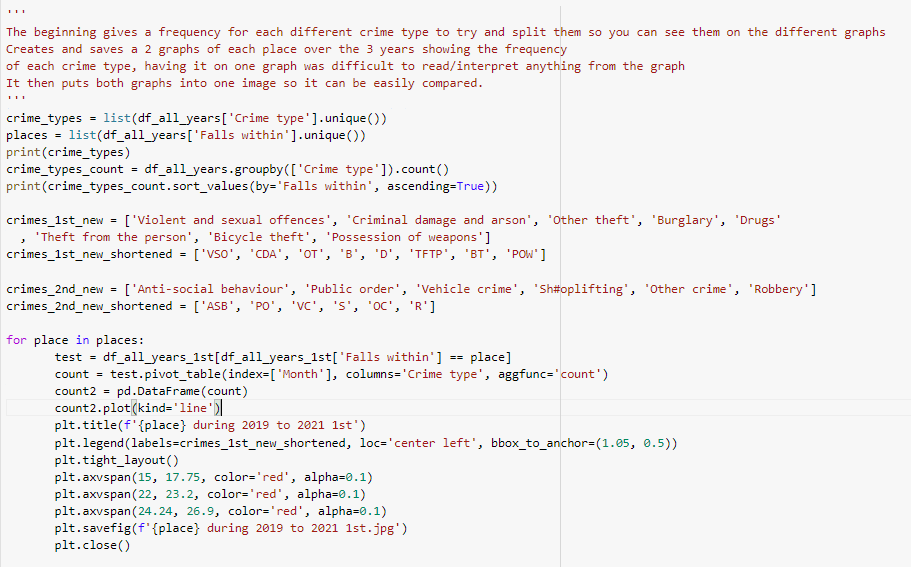
# 

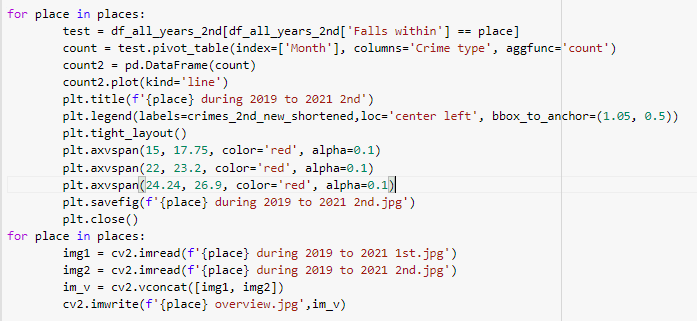
# 

# 

## Comparing different areas over the 3 years

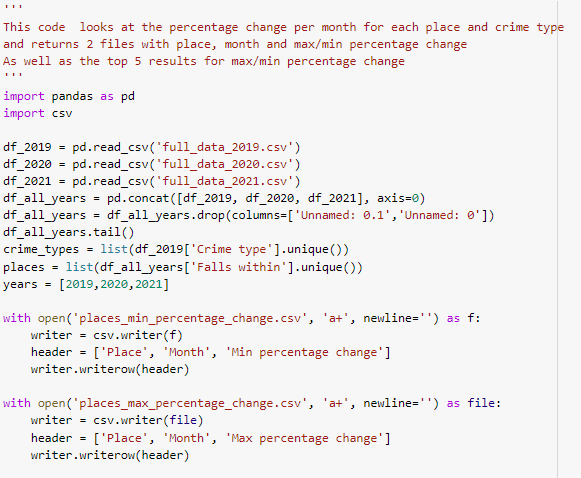
Originally we wanted to show all the crime types on one graph for each area.  
With 14 different crime types if was very difficult to see everything,



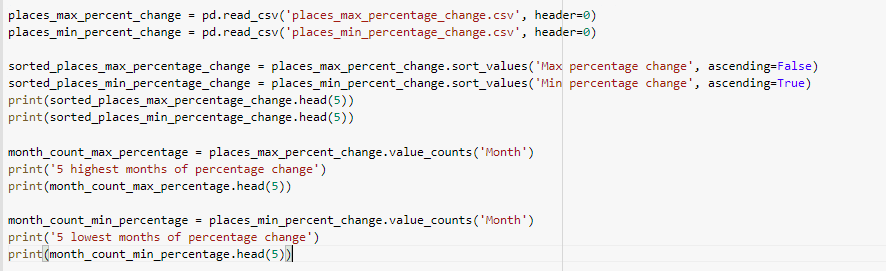


Looking at 45 different graphs and comparing them would be a very time consuming job.

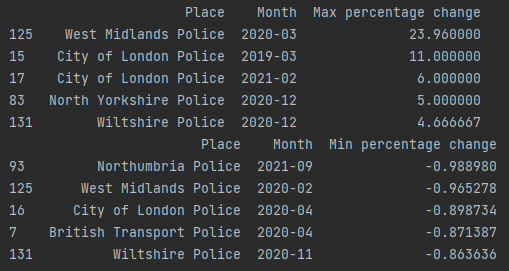
To try and only look at interesting graphs we next looked at percentage change over time for each area.







It gives us the following results.

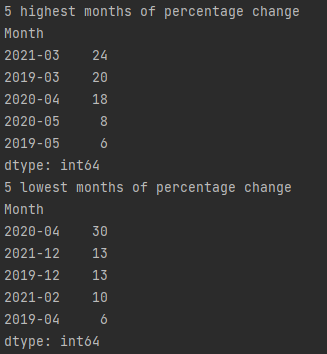


From this we can note the graphs which have either a large percentage change. I then looked at West Midlands Police Overview. You can very clearly see the change in gradient of the graphs.



And then looking at the City of London police overview.

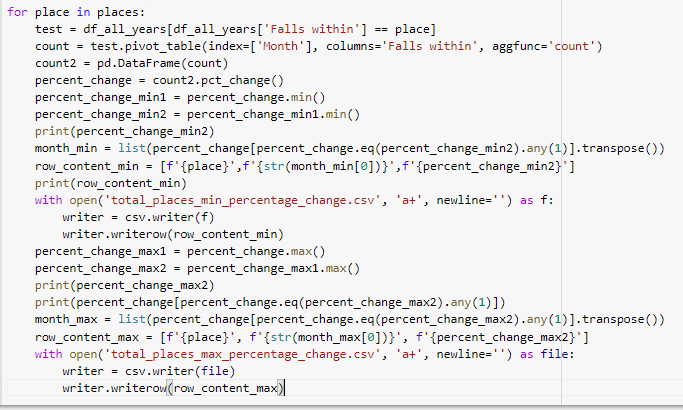


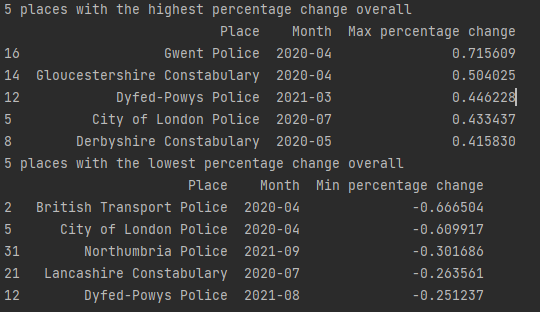


Interestingly enough when you count all the different months, you’ll notice April 2020 being in the top 3 of both lists.

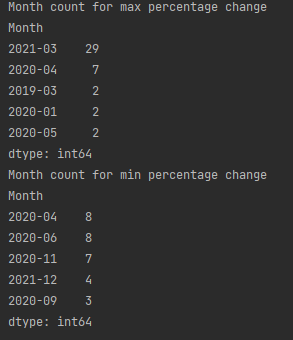
We then also looked at places and whether the overall crime level had any major increases or decreases.







Again you’ll notice April 2020 being top for both overall.

 Again lots of these top ones coincide with lockdown dates.

# 

# 

# 

# 

# 

# 

# 

# **Conclusion and lessons learnt**

The overall level of crime was lower outside the lockdown periods than during lockdowns, and violence and sexual offences increased most noticeably during the lockdowns.

It would be useful to have access to data from 2022 in order to compare the lockdown years (2021 and 2022) to a following year as well as the previous year (2019). This would enable us to determine whether crime rates will continue changing, or return to their original 2019 rates.

Dealing with such a large amount of data was problematic as run times were long

When dealing with such a specific period like a pandemic, while you can make observations of what happened, you cannot use your analysis to predict future situations. If another pandemic were to occur, we could assume that crime rates increased.

We looked at using the API however it didn’t fit the needs of gathering any new data in the same form as the data files.

Looking at particular areas it’s very interesting to look at the impact covid had on certain crime types.