**Modules used:**

Python:

* Pandas
* Numpy
* Matplotlib
* API
* GMAPS

**Files in this project:**

* “Crime.ipynb”
* “README.md”
* “Presentation.pptx”
* “2020-08-west-midlands-street.csv”
* “2021-08-west-midlands-street.csv”

**Introduction:**

Let’s talk about crime in pandemic…

Anyone asked about what he remember from 2020-2021 definitely will mention COVID-19 and how that impacted everyone’s life. Change was noticeable on all levels of existence from work through personal life, global economics, spending habits… This list can go on and on. In our project we decided to check how Covid affected crime rate in West Midlands area in two selected months (August) in 2020 and 2021.

To be specific we were about to check out how crime rate changed year after hard lockdown and beginning of the pandemic when this was new to everyone and no one knew how to live with restrictions and new reality.

**Data Cleaning:**

We obtained data from <https://data.police.uk/docs/> - API documentation.

2 CSV files with data from August 2020 and August 2021 has been downloaded and put in data frame using Pandas module. Files store quite significant amount of columns and rows (28452 rows, 12 columns for 2020 file and 32096 rows, 12 columns for 2021 file). While merged that gave us total of 60548 rows and 12 columns.

We decided to merge both data frames on all column headers which will reduce further cleansing. In result we ended up with one data frame with 12 columns and 60548 rows.

After further analysis and discussion we decided to reduce size by dropping spare columns and empty cells which will affect analysis and outcome. As result we ended up with one final data frame which had all columns that we will be using (7 of them), additionally by dropping empty cells we downsize it to 53090 rows.

That was solid base for next activities that we planned.

**Difference in Crime numbers:**

We were hoping that first visualisation should gave us basic information about change in number of crimes between two years. To get a total number of crimes for each individual year we decided to simply lock frame for 2020 and 2021 count values by “Crime ID” which is unique value for each crime. Results has been placed in new data frame “comparison”. To visualise numbers we decided to use simple bar chart.

**Fully recorded crimes in August 2020: 23644**

**Fully recorded crimes in August 2021: 29446**

Chart

Description automatically generated

**Difference in Crime numbers:**

As total number of crimes for each month was significant, 5 top crimes were selected for further comparison and analysis as we were interested if these will be the same crimes in both months.

In order to pick top 5 we had to assemble new data frame “type\_of\_crime\_combined” which consist two columns with count of crimes for each year (lock by “Month” and groupby by “Crime type”). Next step was to select top5 by nlargest option for each month.

Results has been added to new data frame for simplicity in building new visualisations.

Bar chart with comparison of top 5 for each month.

Chart, bar chart

Description automatically generated

And two pie charts for each August:

Chart, pie chart

Description automatically generated

Chart, pie chart

Description automatically generated

4 out of 5 selected crimes had increase in 2021.

One which alternative tendency was “Anti social behaviour”

A picture containing chart

Description automatically generated

In 2020 this has been committed 4807 times while in 2021 we noticed reduction to only 2649. This might be caused frustration that boredom that most of people felt in 2020 during lockdown but we are not social experts so we can only guess.

**TOP CRIME OUTCOMES:**

After crime part we decided to investigate law enforcement work and check how are they doing in catching bad guys. No surprises there as in August 2020 roughly 88.897% of reported crimes ended up either as “Unable to prosecute suspect” or “Investigation complete; no suspect identified”. For 2021 that was: 90.355% - actual increase.

Graphical user interface, table

Description automatically generated

Results has been presented in bar chart for comparison:

Chart

Description automatically generated

Also 2 pie charts has been created for each month.

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with medium confidence