Fouad Sfarijlani

fouad.sfarijani@student.hu.nl

Fundamentals of data science

Extended Abstract



MASTER DATA DRIVEN DESIGN 2020-2021

**Crime Rate During the Pandemic**

**Introduction:**

Over the past decade, datafication has grown to become an accepted new paradigm for understanding sociality and social behavior (Van Dijck, 2014, p. 199). Kitchen (2014, pp. 1–3) argued that gaining access to datasets can help answer particular questions, and open data movements seeks to answer these questions. He continue his argument by stating that open data provides easy-to-use research tools. Through transparency, sharing and collaboration, the value of these datasets for society can be realized. We are witnessing exceptional circumstances with a global pandemic that started in early 2020. The COVID-19 pandemic of 2020 is unquestionably one of the most significant world-wide events in recent history, impacting culture, government operations, crime, economics, politics and social interactions for the foreseeable future (Stickle & Felson, 2020a, p. 531). Data serve for illustrating developments and assessing policies during the pandemic, while media coverage relies heavily on statistics that provide quickly accessible overviews (Nguyen, 2020). Because of this extraordinary situation, I chose to examine the crime and incidents rates before and during the lockdown caused by the pandemic. Countries that are going to be examined are United Kingdom, The Netherlands and The United States of America.

**Theoretical Framework:**

Stickle and Felson (2020b, p. 532) argued that entire countries, ceased to operate or significantly reduced day-to-day travel. Doing so, commutes from home to work has been eliminated, as well as leisure activities, social gatherings, shopping and more. And therefor, the scope of nature and crime during the pandemic changed with the lockdown as the main factor. We are going to examine this theory by dissecting data of crime incidents and safety indices in the following countries: United States of America, United Kingdom and the Netherlands.

Crime Incidents

Pandemic Lockdown

Cities Safety index

**Problem definition:**

The pandemic and the lockdown associated with it has an impact on every person’s life. Crime and incidents are expected to change during the lockdown as their movement has become limited. And generally, citizens are occupying their households. It is important to examine the types of incidents that are taking place during the pandemic. Addressing this issue will give more insights on which categories of incidents and crime should governments focus.

**Methodology:**

Multiple data sources where taken when developing the dashboard. They are as follows:

* **COVID-19 Data:**
  + Source: Narrativa – COVID-19 Project Tracker.
  + Method: API Endpoint
  + Link: ﻿<https://api.covid19tracking.narrativa.com/api/>
* **Crime and Incidents Data:**
  + United Kingdom Data:
    - Source: UK Police Data.
    - Method: API Endpoint
    - Link: <https://data.police.uk/docs/>
  + Netherlands Data:
    - Source: Central Bureau of Statistics – Netherlands.
    - Method: Web scraping.
    - Link: ﻿<https://www.cbs.nl/nl-nl/visualisaties/welvaart-in-coronatijd/veiligheid/>
  + United States:
    - DC:
      * Source: DC Open Data
      * Link: <https://opendata.dc.gov/search?q=crime&sort=-created>
      * Method: CSV
    - Chicago:
      * Source: Chicago City Open Data
      * Link: <https://data.lacity.org/A-Safe-City/Crime-Data-from-2010-to-2019/63jg-8b9z>
      * Method: API Endpoint
    - Houston:
      * Source: Houston City Open Data
      * Link:<http://www.houstontx.gov/police/cs/Monthly_Crime_Data_by_Street_and_Police_Beat.htm>
      * Method: Excel files
    - Philadelphia:
      * Source: Philadelphia Open Data
      * Link: <https://www.opendataphilly.org/dataset/crime-incidents>
      * Method: Excel files.
    - Dallas:
      * Source: Dallas City Open Data
      * Link: h[ttps://www.dallasopendata.com/Public-Safety/Police-Incidents/qv6i-rri7](https://www.dallasopendata.com/Public-Safety/Police-Incidents/qv6i-rri7)
      * Method: API Endpoint
    - San Francisco:
      * Source: San Francisco Open Data
      * Link: <https://data.sfgov.org/Public-Safety/Police-Department-Incident-Reports-2018-to-Present/wg3w-h783>
      * Method: API Endpoint
  + Safety Index Data:
    - Source: NUMBEO
    - Link: <https://www.numbeo.com/crime/rankings.jsp?title=2018>
    - Method: Web scraping

It is important to mention that data from the United States was collected based on 6 cities due to time limitation and availability of the data.

Python Libraries:

Various methods where used in Python in order to collect and compile the data. Libraries learnt during the course where extensively used. Also, some additional libraries were used such as **Selenium** in order to interact with websites that uses javascript.

Libraries used in the scripts are summarized as follows:

requests, json, bs4, Selenium, TQDM, Pandas, numPy, dash\_copre\_components, dash\_html\_components, plotly.express, dash.dependencies.

Obstacles:

Many obstacles were faced during the compilation of these data sets. Finding data was first to mention. Cleaning and compiling the data to make a uniform set was also a problem. Some cities open data websites offered excel sheets only, and historic data were available in monthly basis. Manual combination on Excel was required(mainly DC, Philadelphia and Houston). Original datasets are available in the submission file. Also, language barrier was an obstacle when dealing with Dutch websites.

Categorization:

After collection, incident data was categorized in order to aggregate the various incident reports.

An excel sheet available inside the submission file named “**Incident Categorization.xlsx**” contains the categorization.

Scripts:

Six Python scripts are available inside the submission folder. Five of them are for data collection and cleaning, and one for aggregation and visualization.

They are named as follows:

* USA\_crime\_data.py
* Uk\_crime\_data.py
* Safety\_index\_data.py
* NL\_crime\_data.py
* Covid19\_data.py
* Analysis\_dashboard.py

It is recommended to run all the data collection files and leave **Analysis\_dashboard.py** till the end.

Visualization:

With Dash as a primary library, 4 graphs were created. Interactivity also was added by adding filter options. They are as follow:

|  |  |
| --- | --- |
| Filter Type | Effect |
| Country Filter | Affect the whole dashboard |
| Incident Filter | Incident Graph (fig. 2) |
| Year Filter | Distribution graph and Safety Index (fig. 3 & 4) |
| City Filter | City Safety Index (fig. 4) |

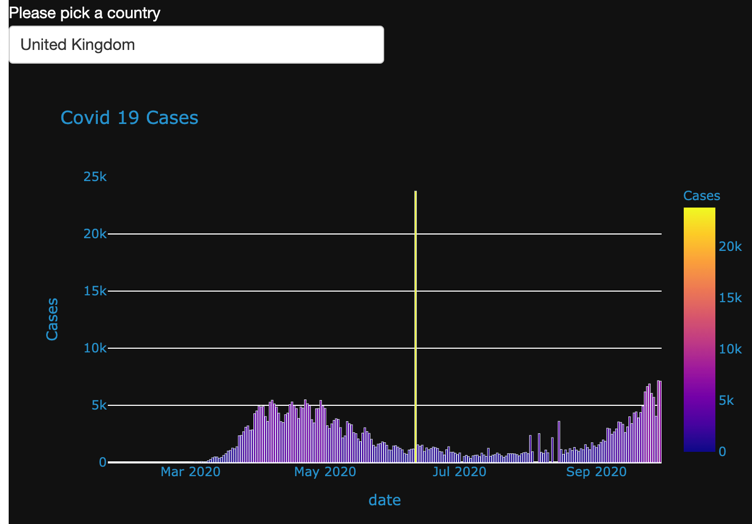


Figure 1- Sample of COVID-19 Graph with UK selected

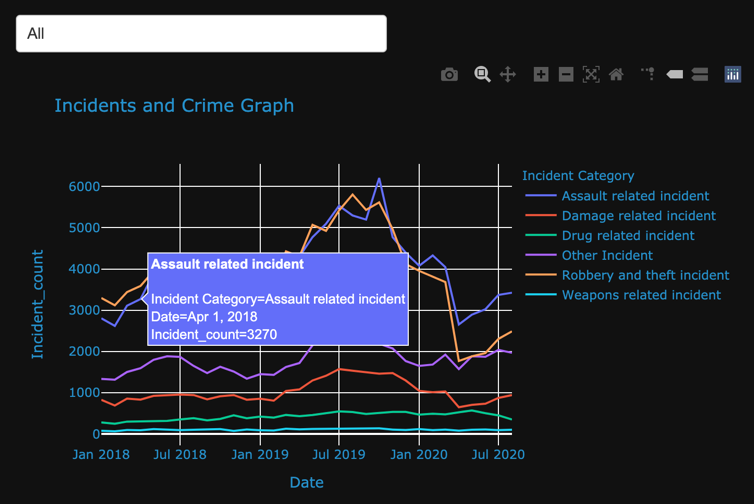


Figure 2 Sample of Crime incidents graph in the UK with “All” options selected

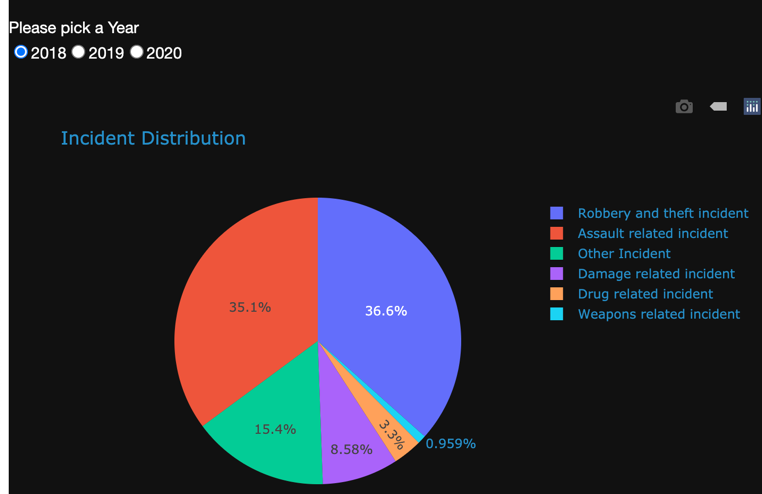


Figure 3 Sample of Incident distribution graph with UK selected and year 2018



Figure 4 Sample of Safety Index Graph in the UK, where year 2018 is selected and option “All” in the graph is selected

It is worth mentioning that lists that appears in drop down menu changes dynamically with each country selected. Because every country has different categorizations of incidents, cities and distributions.

**Findings:**

Overall, results are mixed and category dependent. Some crimes have dropped during the pandemic while others have increased. In the United States Assault related incidents have increased from 7810 incident in January 2020 to 9272 in June. Fraud cases have seen a steady decline from 1721 cases in January 2020 to 1473 in June. Alongside robbery with 18064 in January to 14799 in June. Sex offenses has dropped substantially from 573 recorded cases in January to 398 in June. A low record in April where the number was 273 recorded cases. On the other hand, reported domestic abuse incidents increased in comparison to 2018 and 2019. In the UK, crime overall has decreased during the pandemic. Except drug related crimes have remained fairly steady. In the Netherlands, Robbery crimes and damage related cases have decreased during the pandemic also. Except nuisance cases. The number have increased drastically during the whole period of the first wave of COVID-19 going from a respective peak of 35000 and 37000 in 2018 and 2019 to reach 50458 reported cases in April 2020. Safety indices for cities in the Netherlands have increased. While in the United Kingdom, London, Leeds and Birmingham experienced a slight decrease in their safety index of 5, 3 and 5 points respectively. The rest of the cities have increased. In the US, safety index of cities have increased generally. Cities such as Riverside has gone up from 56 in 2018 to 62 points in 2020.

**Conclusion:**

Crime pattern during the COVID-19 pandemic has changed overall. And cities safety index in countries examined also has gone up. Some crime incidents has decreased drastically in comparison to previous years. Yet at the same time, some incidents such as domestic abuse and nuisance reports has gone up. That being said, limitation of movements and curfews during the first wave of the pandemic has been a factor in reducing crime rates overall. At the same time, it gave rise to other instances where malicious behavior has moved indoor.

***References***

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197–208. <https://doi.org/10.24908/ss.v12i2.4776>

Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences* (1st ed.). SAGE Publications Ltd.

Nguyen, D. (2020). Mediatisation and datafication in the global COVID-19 pandemic: on the urgency of data literacy. *Media International Australia*, 1329878X2094756. <https://doi.org/10.1177/1329878x20947563>

Stickle, B., & Felson, M. (2020). Crime Rates in a Pandemic: the Largest Criminological Experiment in History. *American Journal of Criminal Justice*, *45*(4), 525–536. <https://doi.org/10.1007/s12103-020-09546-0>