**Libraries required**:

pandas

numpy

matplotlib.pyplot

matplotlib

seaborn

pprint

Datetime (and datetime from datetime)

**API Keys required:**

An API key is required for the FBI Crime database. To obtain one please go to → <https://api.data.gov/signup/>

**Getting data from the GitHub repository:**

The /raw\_data folder in this repo contains recent exports from our personal databases. You are welcome to build applications that draw directly from this repository.

* The CSV file /raw\_data/[fraudTest.csv](https://github.com/Data-2-CFG/group-1-project/blob/main/raw_data/fraudTest.csv) provides the initial simulated test dataset for the credit card transactions. This is not updated as we used this as our base dataset.
* The CSV file /raw\_data/[fraudTrain.csv](https://github.com/Data-2-CFG/group-1-project/blob/main/raw_data/fraudTrain.csv) provides the initial simulated train dataset for the credit card transactions. This is not updated as we used this as our base dataset.
* The CSV file /raw\_data/COVID19API\_2019-01-01\_2020-12-31 provides the 2020 stringency data for the entire world: This will be useful for the COVID visualisations
* The CSV file /raw\_data/COVID19API\_2021-01-01\_2021-12-31 provides the 2020 stringency data for the entire world: This will be useful for the COVID visualisations

**Order in which to run the Notebooks:**

There is an order to run the notebooks to reach the visualisations as they encourage you to save datasets as a csv file onto your local server to run codes for future use.

1. The notebook [COVIDdata\_API.ipynb](https://github.com/Data-2-CFG/group-1-project/blob/main/API%20-%20data_retrieval/COVIDdata_API.ipynb) , this is to retrieve COVID stringency data. For the COVID visualisations the date ranges input were 2019-01-01 to 2020-12-31 and 2021-01-01 to 2021-12-31. However, these data sets have also been added to the data repository.
2. The notebook [CRIMEdata\_API.ipynb](https://github.com/Data-2-CFG/group-1-project/blob/main/API%20-%20data_retrieval/CRIMEdata_API.ipynb) , this is to retrieve CRIME values of USA data. For the CRIME visualisations, please follow Niki’s instructions. Please ensure you have an API key for the code to be able to retrieve from the database. There is the option to save all 4 API datasets into csv files but for the purposes of our project we only used API\_1.
3. The notebook Transaction dataset.ipynb (Niki shall be uploading this), this is to merge the fraudTest and fraudTrain datasets to obtain a more complete dataset to work with for the visualisations.
4. The notebook CovidVisualisations\_Joey.ipynb, this is to run the data cleansing and preparation aspect before finally coming to creating the visualisations. All the datasets here are either retrievable from the previous steps or will be created into dataframes which can then be exported into csv files. Comments and descriptions are provided throughout the notebook to guide you through the code. Feel free to use the graphs for your own perusal.
5. The notebook data visualisation\_Nicol\_def.ipynb, (Niki shall be uploading this), this is to run the data cleansing and preparation aspect before finally coming to creating the visualisations.