# Defining the Question (The Problem Statement).

Covid-19 had a huge impact in the world. In this exploratory project, we attempt to find out the countries that were the most affected by it. Further, the study tries to find out the following:

* Which country (having more than a million residents) had the highest percentage of citizens infected by covid?
* Highest death by country and continent.
* Countries with the highest death rate
* Countries with the highest vaccination %

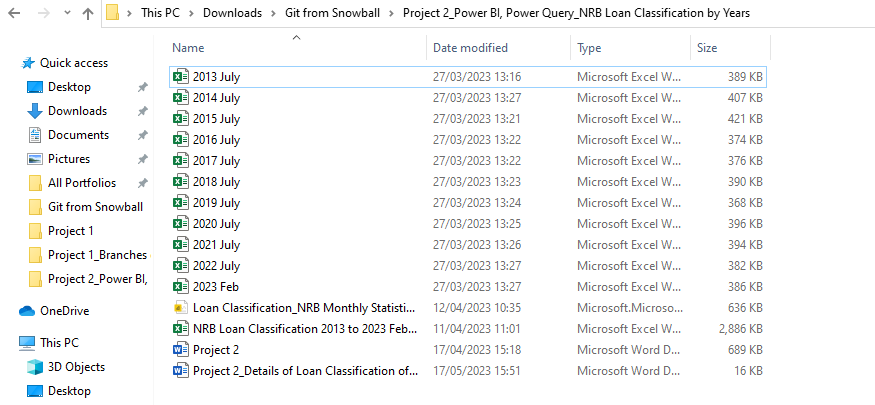
# Collecting the Data.

Data was collected from <https://ourworldindata.org/covid-deaths>

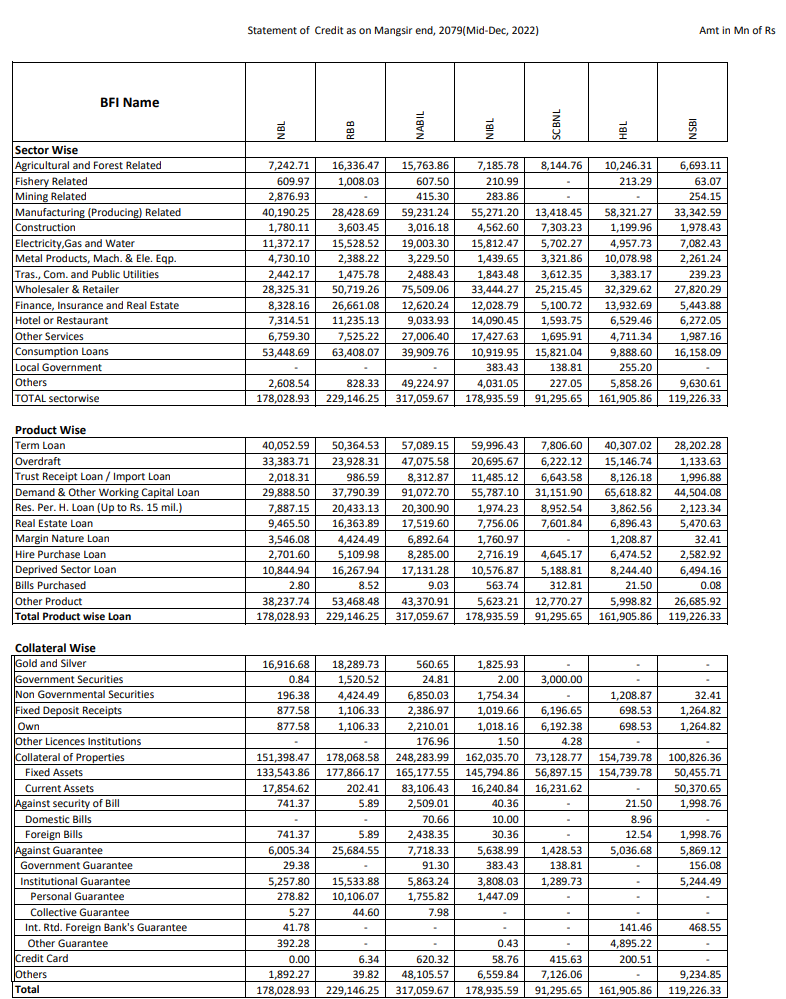
**Time period of the study:** 3rd Jan 2020 to 31st May 2023

**Background**: NRB publishes provisional Banking and Financial Statistics each month. Each publication is around 60 pages long and contains details of financial statements of loans and deposits of individual and industry-wise data of Commercial Banks, Development Banks, Finance Companies, and only industry-wise data of Microfinance Institutions.

Before July 2021, only pdf files were uploaded. We have converted the pdf files to Excel files using [www.ilovepdf.com](http://www.ilovepdf.com), and kept in a folder.



Following is a snapshot of the data from the central bank’s website.



# Cleaning the Data.

Each of the 11 Excel files had at least 40 tables. Only 10-15 tables were required for our study.

Macros were used initially to convert the tables into Excel recognizable tables and nomenclature was given to each table as “Table**N**”, where “N” stands for 1,2,3…., N.

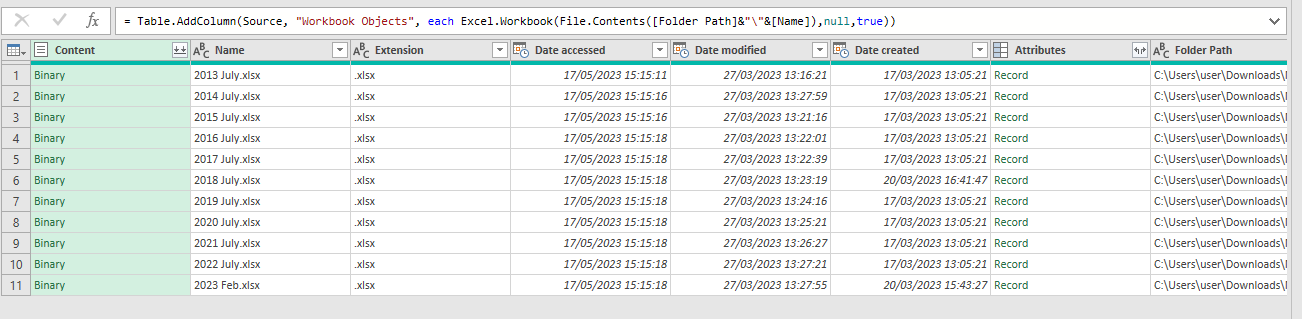
This nomenclature was required to distinguish the tables required for the study from other tables.

All 11 Excel files were loaded in a new Excel file using **Power Query**.

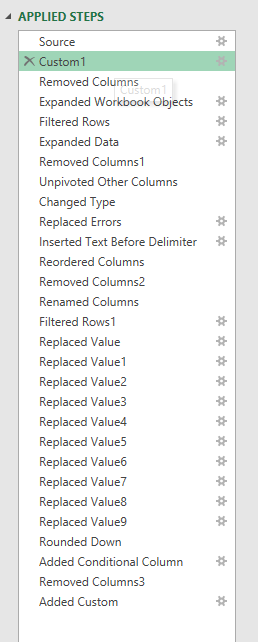
Only required tables were selected for the study using:





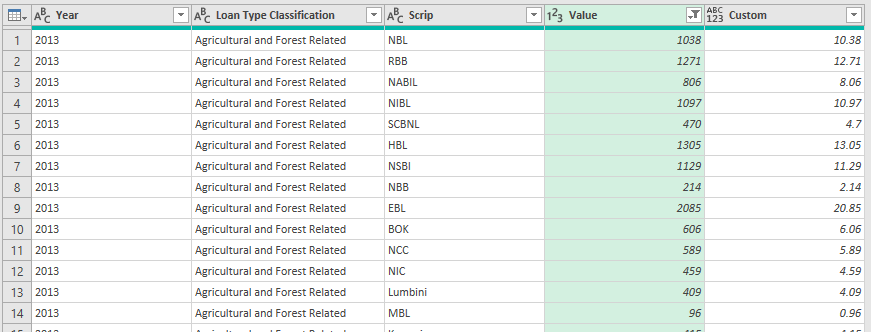


The following steps were used to clean the data:



Multiple “Replaced Value” was used to filter out the non-listed, merged, or acquired banks.

Finally, the data were transformed into five columns and 33,908 rows. It was then loaded to the Excel sheet and added to the data model.



Following relationships were created among the three tables based on a primary fields “Scrip” and “Loan Type Classification”.

A screenshot of a computer

Description automatically generated with medium confidence

# Analyzing the Data.

**Pivot Tables** were used to analyse various useful trends.

A screenshot of a graph

Description automatically generated with low confidence

A screenshot of a graph

Description automatically generated with low confidence

# Visualizing the Data.

A Dynamic visualization was possible using **Pivot Charts and Power BI** and the questions defined in the problem statement were answered.

* 80% of loans from Commercial banks, Development banks, and Finance companies were backed by Fixed assets-based collateral. This indicates a strong position of banks in Nepal to recover losses as they are mostly backed by fixed assets.

A screenshot of a computer

Description automatically generated with medium confidence

* We saw a change in NIC Asia Bank’s business model from wholesale to retail from 2017 onwards as fixed assets-backed loans started to shoot up while the current assets-backed loans continued to decline.

A screenshot of a graph

Description automatically generated with low confidence

* We figured out that the DSL, which was in increasing trend from 2013 to 2022, declined by 1,500 crores in Feb 2023. This might have negative implications for microfinance companies in the years to come.

A screenshot of a graph

Description automatically generated with medium confidence

**Dashboard in Excel:**

A screenshot of a computer

Description automatically generated with medium confidence

**Dashboard in Power BI:**

