****

**Project**

**Visualizing the Impact of COVID-19: A Data-Driven Approach**

**Name: Amit Kumar Teacher: Ms. Partima**

**Roll No: 23MCA20150 Subject: Business Analytics**

**Section: 23MCA-3/A**

**Title**

1. Aim
2. Requirement
3. Methodologies
   1. Data cleaning
   2. Data analyzing
   3. Chart creation
4. Data Visualization
5. Result and key insights

**Aim**

* The aim of this project is to analyze and visualize COVID-19 data using various charts in Excel to provide clear insights into the spread, impact, and trends of the pandemic.
* By visualizing key metrics such as case numbers, deaths, recoveries, and geographical distributions, the project seeks to make complex data more accessible and understandable to a general audience. This way, we can also visualize which countries suffer from the most casualties.

**Requirement**

* **Data set** -- We can import our dataset from various site
* **MS-Excel** – We will do our visualization in MS Excel. MS Excel provides us the advanced charting tools which will be helpful in data visualization.

**Data Collection**

* For this project, the data on COVID-19 was collected from WHO site which is publicly available, reputable sources to ensure accuracy and reliability. ([COVID-19 data | WHO COVID-19 dashboard](https://data.who.int/dashboards/covid19/data)).
* This dataset includes detailed global records of COVID-19 cases, deaths, and recoveries.
* The data covers the time period from the early stages of the pandemic in 2020 to the most recent available update, allowing for a comprehensive view of how the pandemic evolved over time.

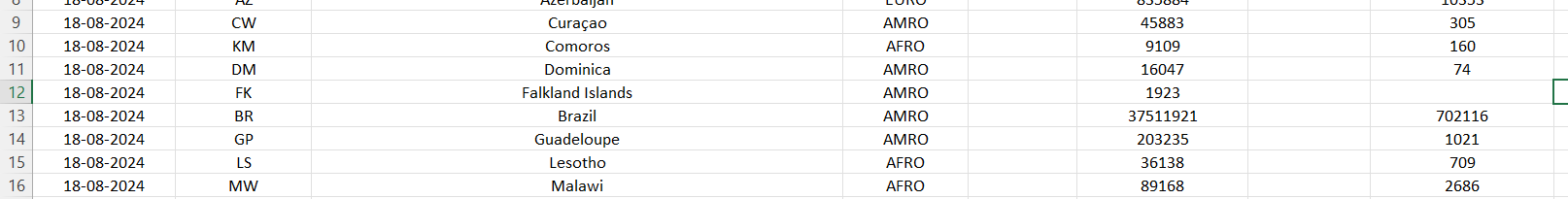
**Methodologies**

* It the step-by-step process for how you will use Excel to visualize the data. This may include:
* **Data cleaning**: How you handle missing values, organize columns, or prepare the data for analysis.
* **Analysis Techniques**: Mention any basic statistics or techniques like calculating daily case averages, growth rates, etc.
* **Chart Creation**: Specify the types of charts you’ll use for different kinds of data (e.g., line charts for trends, pie charts for proportions).

Implementing the above thing one by one

1. **Data cleaning**

* In data cleaning we do the things like handling missing data, normalize the data formats, removing duplicates etc.
* Now the data we are using is mostly cleaned but we still find some missing data so we fill those data by taking the average.

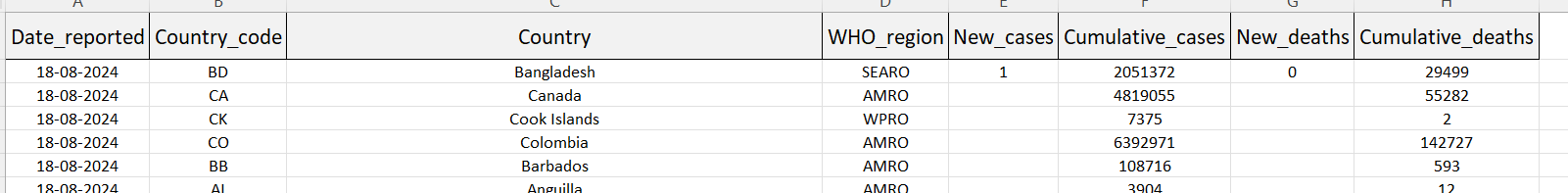
****



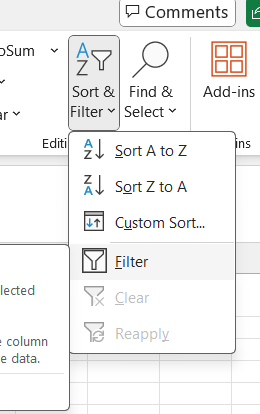
* After that we can remove the duplicate values and format the data. Here all the data is clean with no duplicate values and the format is also correct.

1. **Analysis Techniques**

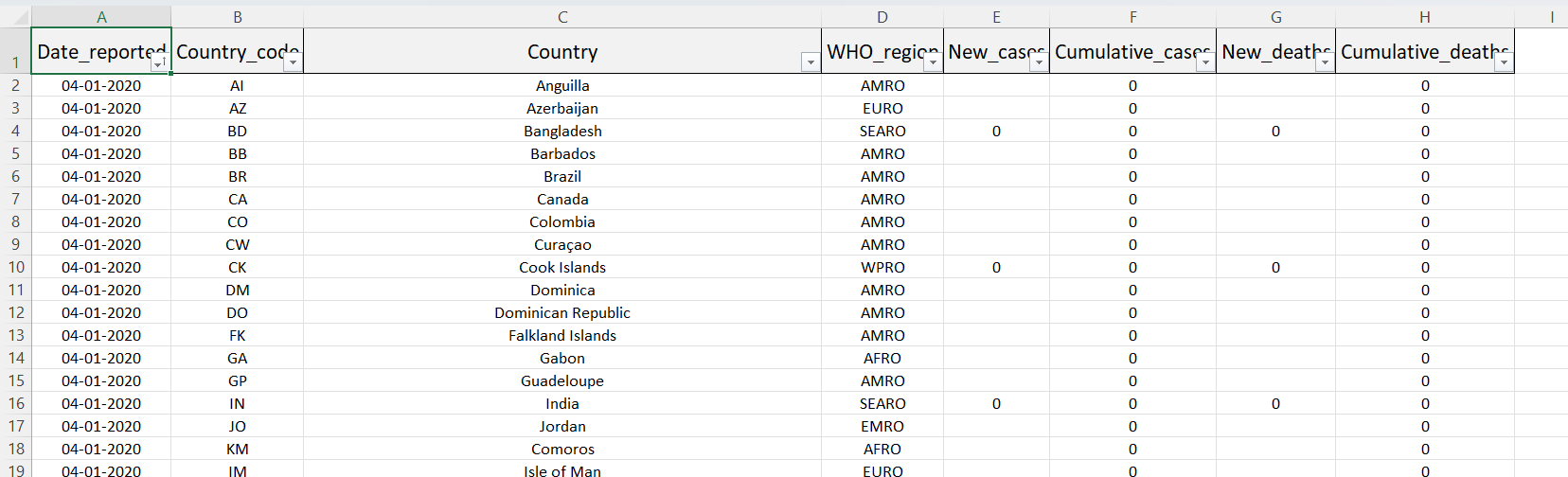
* Now I start to analyze the dataset by adding the sorting and filtering options. Before that I change the font and background of the heading to make it more readable

****

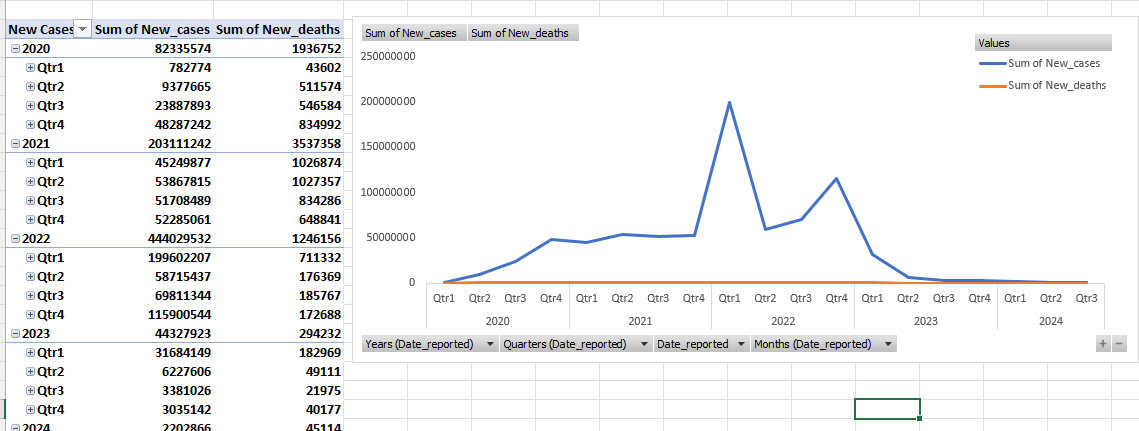
* After that I will go the editing tab and choose the **filter option**. Then the drop-down arrow will appear which you can use to filter.

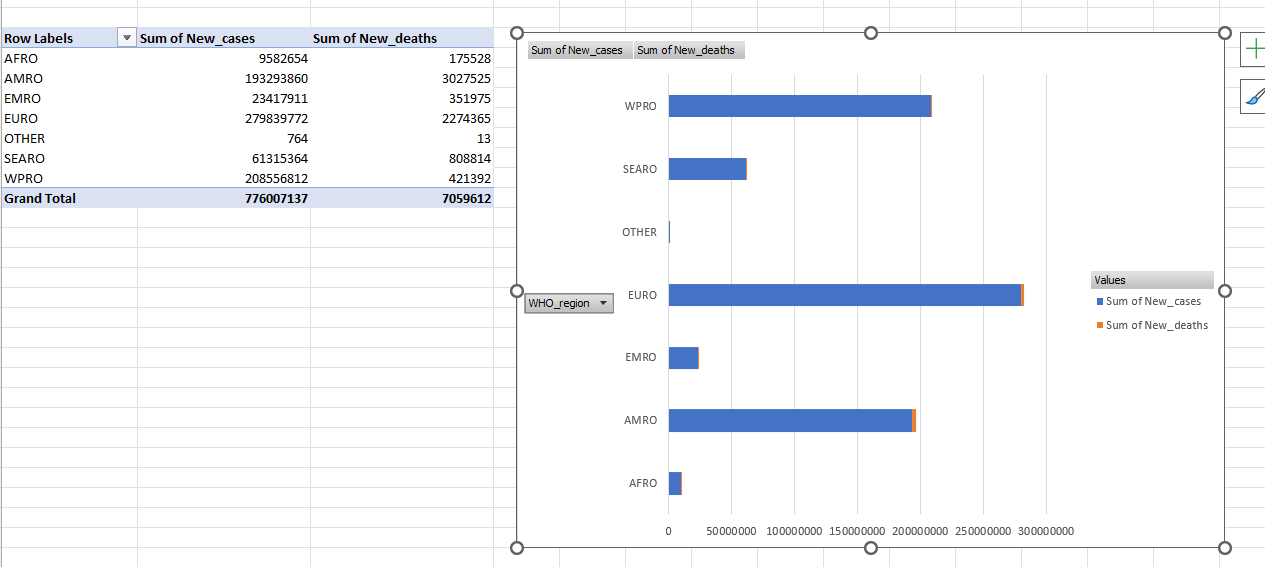


* We can also sort the data from ascending and in descending order. This way we can visualize data in ascending and descending order.

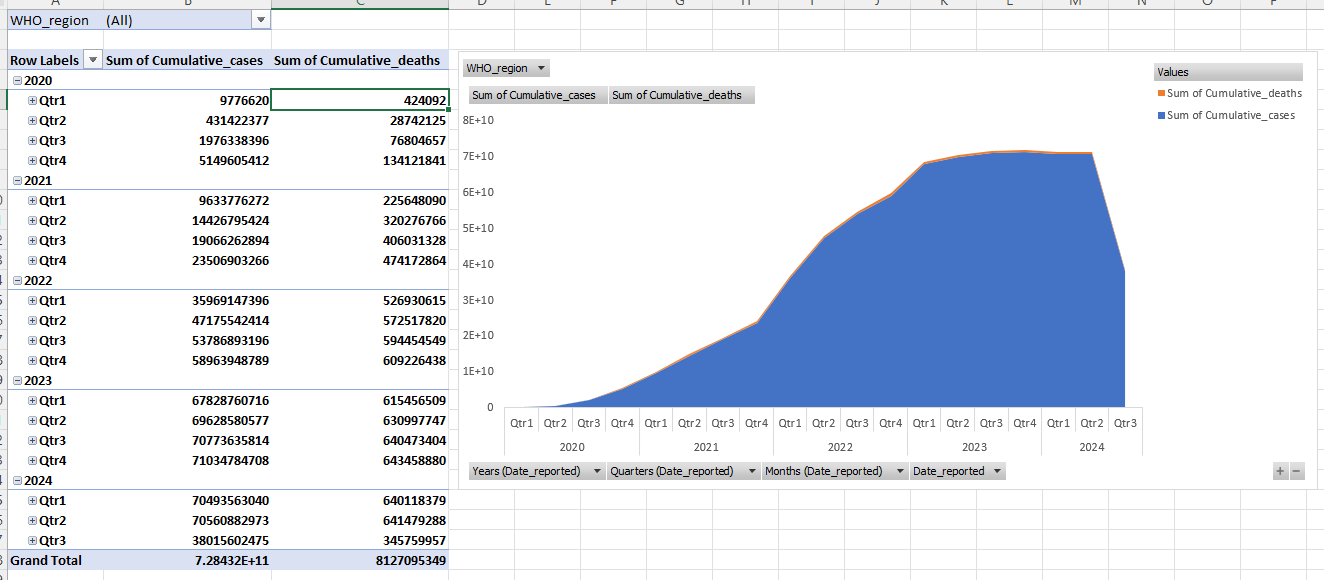


1. **Chart creation**

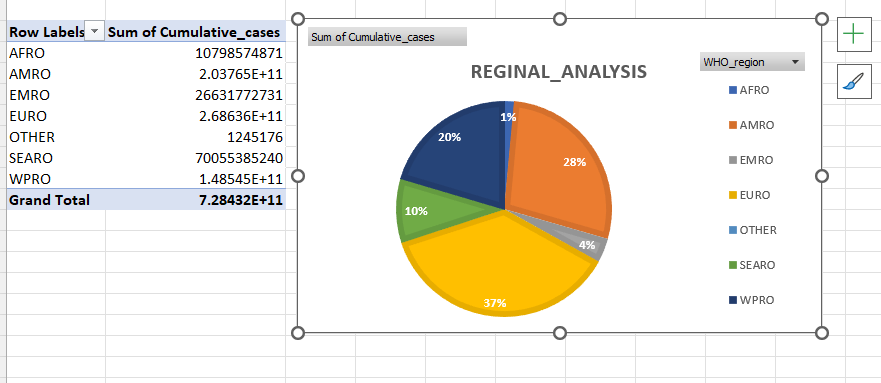
* **Time Series Analysis:** We plot line chart where we show the number of new cases and new deaths over time (date\_reported). This helps visualize trends and patterns.
* Here what we show is we have time on y axis and sum of new cases and sum of new deaths on x axis. This way we can see how many cases on that day and how many new deaths are there. Chart used here is line chart.
* **Comparative Analysis:** Comparing the total number of new cases and new deaths across different countries or regions. Here we show data regions wise.
* Here on Y axis we take the different regions and on x axis we have the total number of new cases and sum of deaths. We use bar chart here



* **Cumulative trends:** Show the cumulative cases and cumulative deaths over time. It shows how the cumulative deaths and cases are increased over time. And then decreased at one point.
* Here on x axis, we have the time period and on y axis we have cumulative cases and cumulative deaths. And for filter, I have added the region, so that we can get data basis on the region. We use the area chart for this.

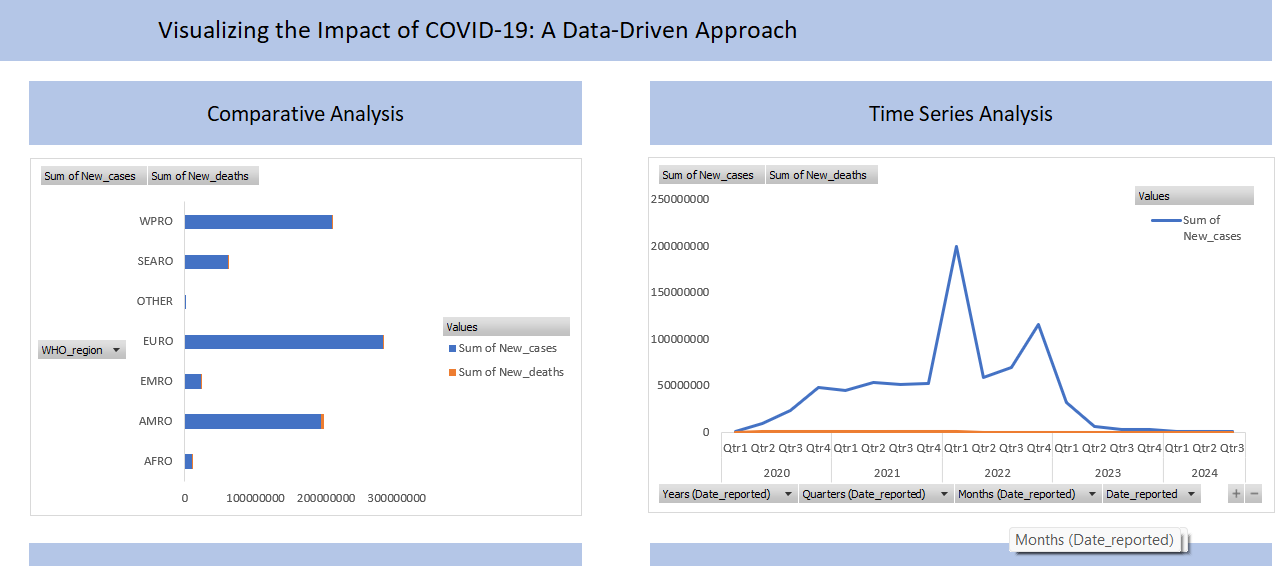


* **Regional Analysis:** Display the proportion of new cases or cumulative cases by WHO region. This way we can recognise how much the cumulative cases increase with region bise.
* Here we use the pie chart.

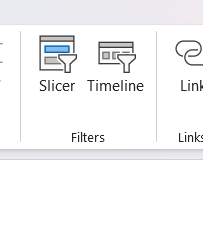


**Data Visualization:**

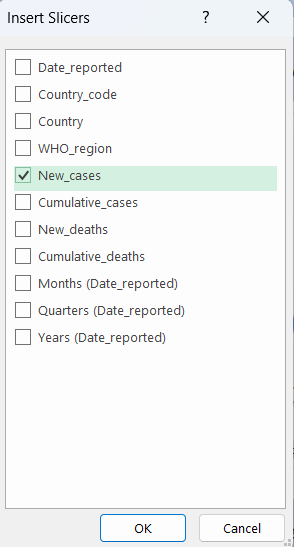
* Now we need to visualize the data all in the same place. We will copy all the charts in the same places for the better visualization.



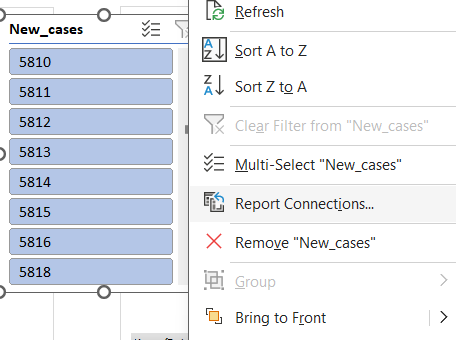
* Then we will add the slicer. For that select a table and then navigate to INSERT and there you will select the slicer.



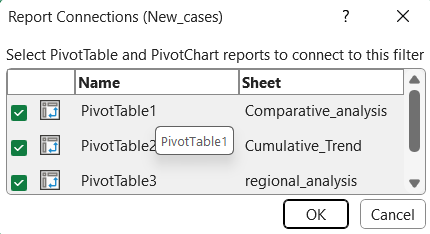
* Then we will select what are the fields we require for the slice. Here I choose the new cases.



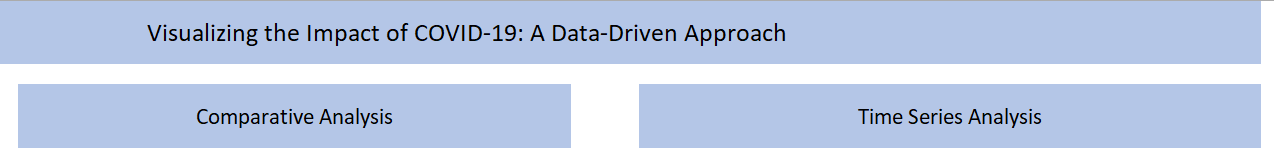
* Once the slicer will appear we will connect it with the tables. For that right click on the tab and then select the report connections. After selecting it we will

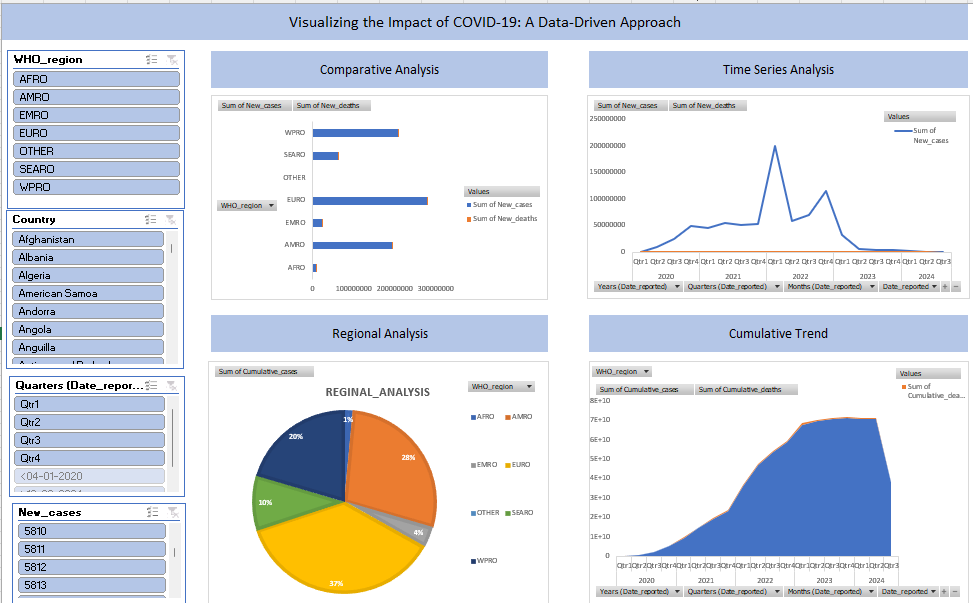


* Then we will select the tables where we want to connect our slicer with.



* Finally we will add the headings about what chart is showing which fields.





**Result and Key insights**

* By visualizing the data set we are able to get the detail information of the data and able to take out the useful information from it.
* Means by doing the time series analysis how many new corona cases and deaths occur in that quarter of time.
* By doing the comparative analysis, we able sum of new cases and sum of new deaths in the all regions. This way we can analyze that which region affect with maximum number of cases and how many deaths occur in that period.
* By doing the regional analysis, we can analyze which region has maximum number of cumulative growth rate of the cases.
* And finally by doing the cumulative analysis we can analyze what is growth rate of the number of cumulative cases.