**PROJECT TITLE : COVID-19 CASE ANALYSIS**

**COVID-19 CASE ANALYSIS :**

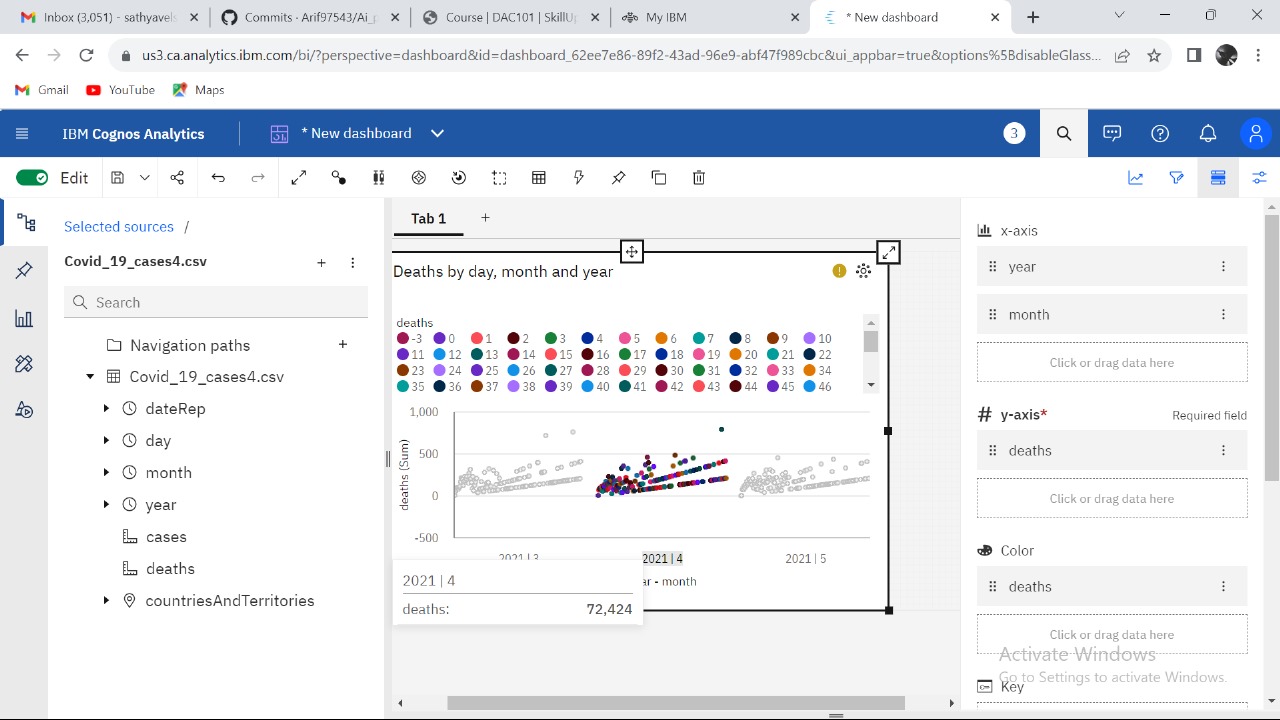
Building phase of the project :

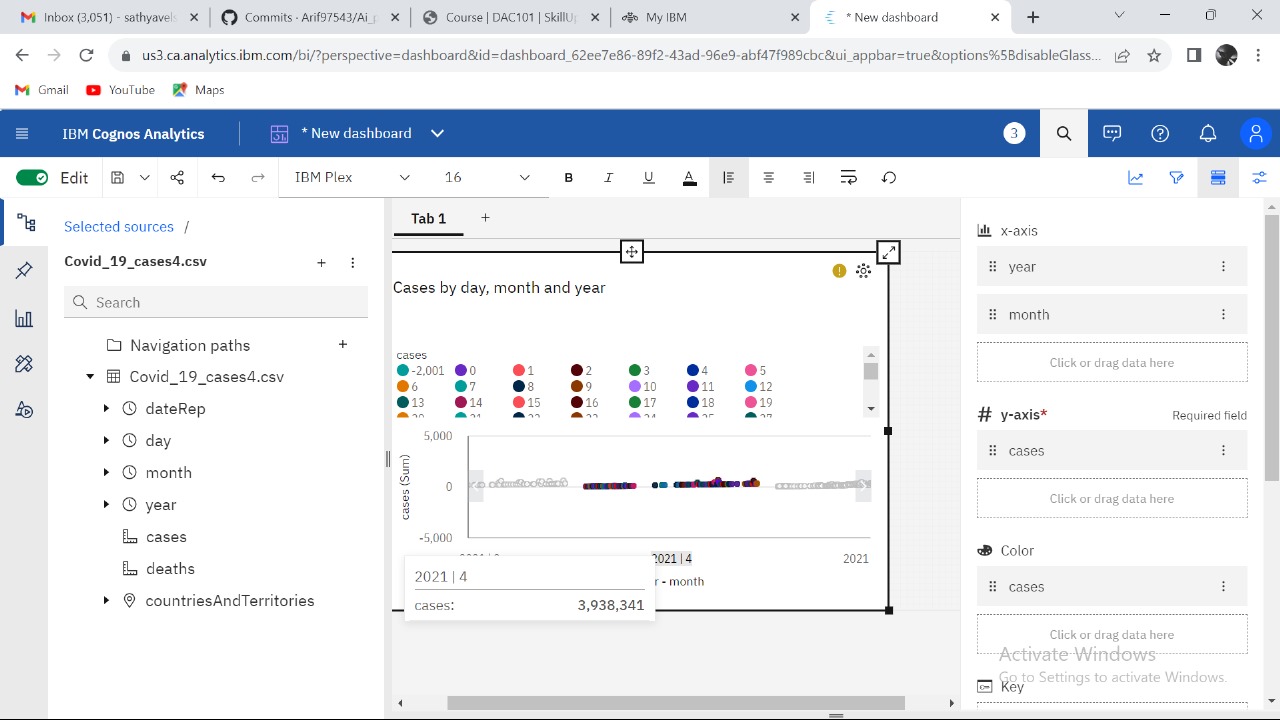
COVID-19 paved the way towards subsequent nationwide lockdowns, resulting in a rise of loans or credit applications from financial institutions as the prime source of project financing. However, financing institutions are becoming cautious as current, short-term, and long-term economic growth is still uncertain. Therefore, financial institutions opt to reduce the approval rate for financing, including for construction projects, by implementing additional evaluation processes.

Correlation analysis and scatter plot were employed for the quantitative data. We used Spearman’s correlation analysis to determine relationship strength between cases and deaths and socio-economic and health systems. We organized qualitative information from the literature and conducted a thematic analysis to recognize patterns of cases and deaths and explain the findings from the quantitative data.

**COVID-19 case analysis using IBM cognos visualization :**

* Identifying the most suitable machine learning technique for prediction, toperform on clinical reports of patients.
* Preparing a machine learning model that could make accurate predictions of COVID-19 in patients.
* Identifying the features that affects the prediction of COVID-19 in patients.

**Visualization for deaths :**

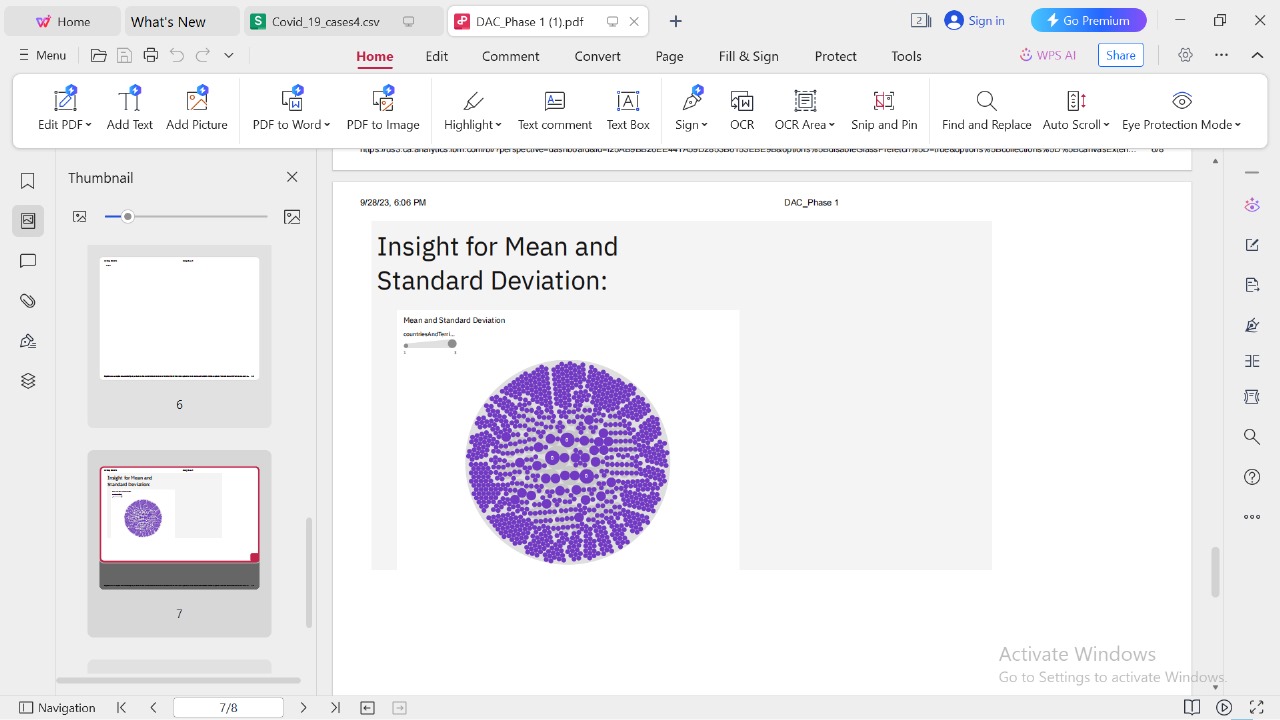
**Visualization for cases :**

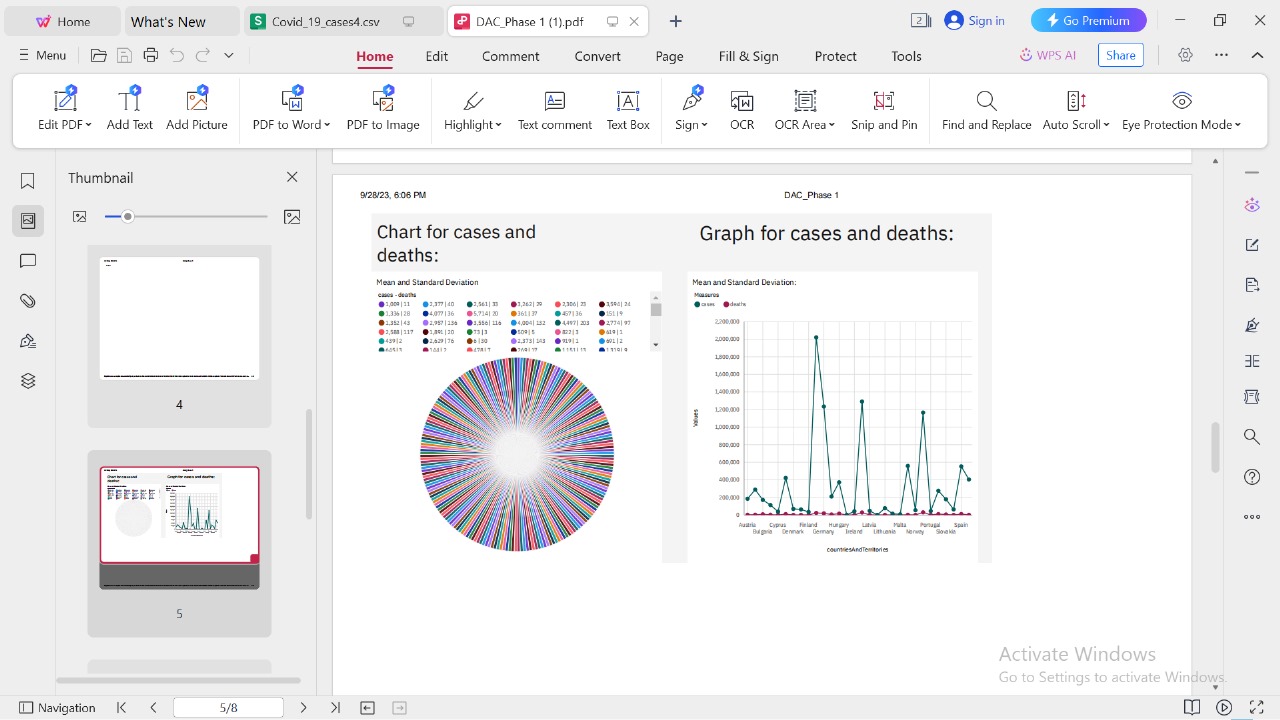
**Deriving insights from data :**

* Covid-19 pandemic has been a global threat since its outbreak in Wuhan, China, in 2019. Surveys, interviews have been carried out to get data from certain Cities in different Countries and Continents in other to draw insights.
* To a data scientist or anyone who is enthusiastic about data science, data are characteristics or information either quantitative(numerical) or qualitative(categorical) that are collected from persons, organisations etcetera in other to draw insights, improve sales, predict future occurencies and so on.
* Before drawing insights from a data set, always remember that there is no one size fits it all. Data exploration which includes sorting(rearranging data in ascending or descending order, data filtration(creating a subset of available data), data processing(aggregation and statistical operations), data cleaning and preparation takes up to 75% of the total time for the project.

## **The steps I took to analyse the data set are thus :**

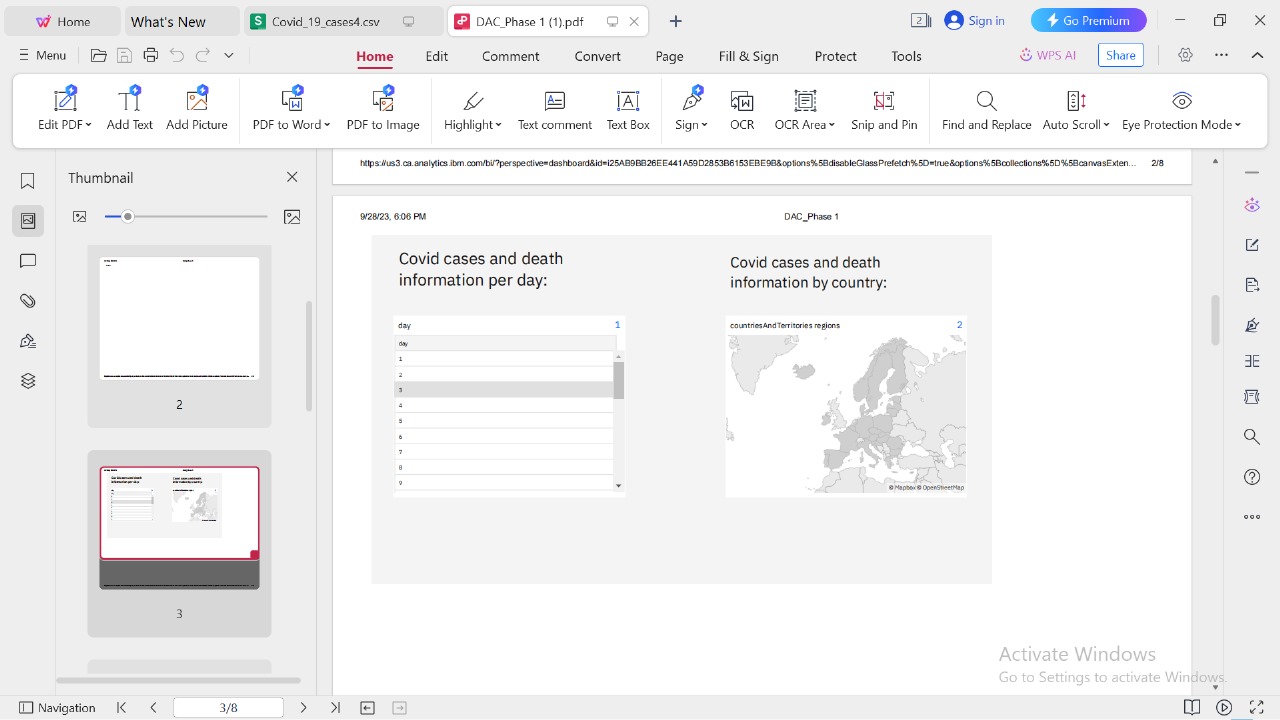
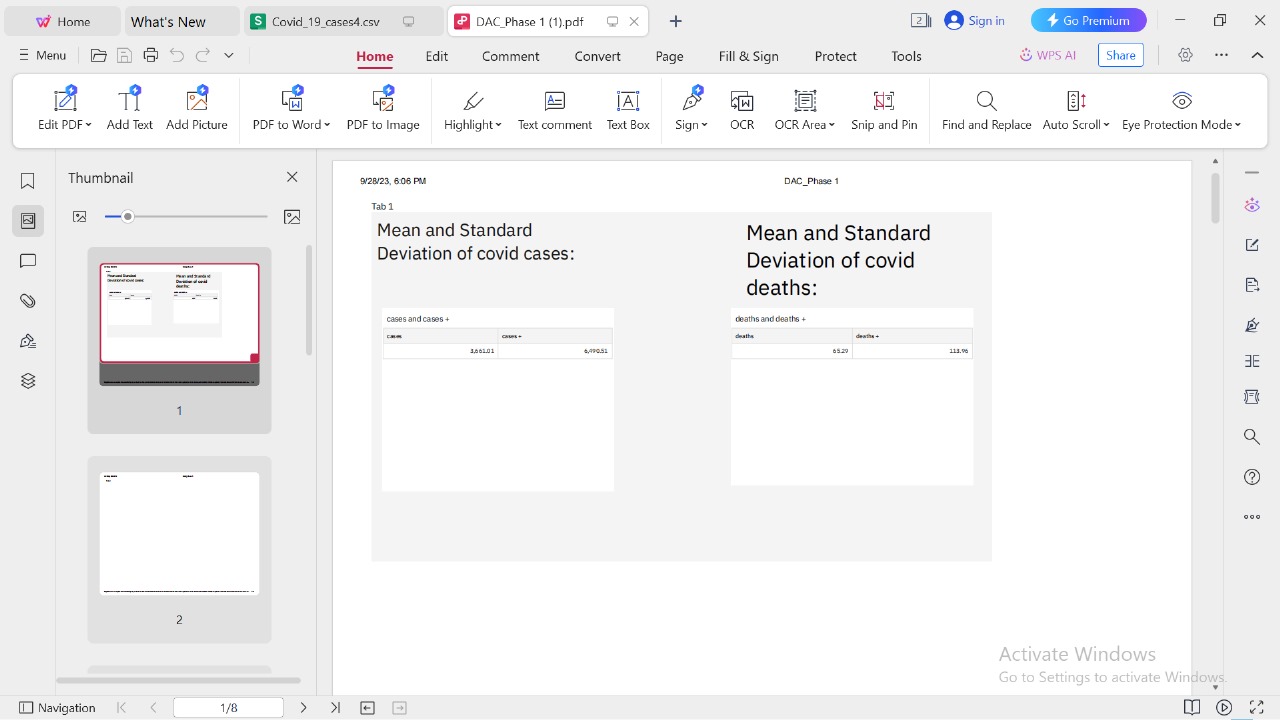
* Definitions
* Data Cleaning
* Visualization
* Drawing Insights

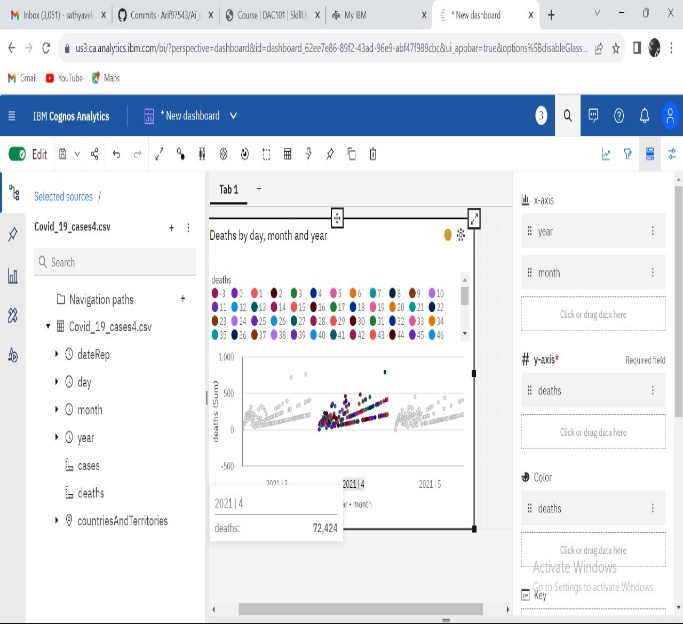
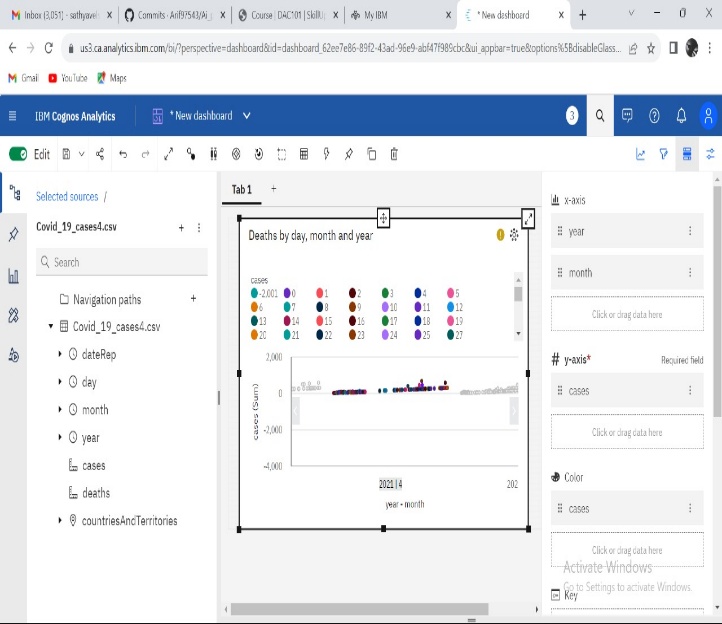
**Insights from the data :**

**Charts and graphs in IBM cognos to visualize and compare the mean values and standard deviation :**

**Analyze the visualizations to identify trends, variations, and potential correlations between cases and deaths :**

* Visualization is an essential tool for summarizing and making sense of data. During the COVID-19 pandemic, time-series plots of cases and deaths have become fixtures of news reports, social media posts, and dashboards.
* In time-series plots, the x-axis is time and the y-axis is the variable of interest, for example daily new cases per million population. Information about a second variable of interest can be presented as a second line plotted against the same x-axis, with or without a secondary y-axis, but this is not the only way to present two variables. For example, if the desire is to draw the viewer’s attention to the ratio of the two variables, the ratio can be plotted over time; however, the absolute magnitudes are lost if only the ratio is plotted.
* An alternative visualization approach is to plot the two variables against each other as a scatterplot, one on the x-axis, the other on the y-. In such a plot, time has no axis of its own, but can be represented visually as a gradient of weight, width, or color along the line, and/or can be conveyed using arrows or arrowheads along the line (similar to how vector fields are often displayed). Such plots are standard in the study of dynamical systems, leading for example to the phase diagrams of differential equations.

****

****

**Dataset link :**

[**https://www.kaggle.com/datasets/chakradharmattapalli/covid-19-cases**](https://www.kaggle.com/datasets/chakradharmattapalli/covid-19-cases)

**Conclusion :**

The coronavirus disease continues to spread across the world following a trajectory that is difficult to predict. The health, humanitarian and socio-economic policies adopted by countries will determine the speed and strength of the recovery.