**COVID 19 Cases Analysis**

Design Document

**Phase 1: Problem Definition and Design Thinking**

In this part you will need to understand the problem statement and create a document on what have you understood and how will you proceed ahead with solving the problem. Please think on a design and present in form of a document.

Project Definition:

The project involves analysing COVID-19 cases and deaths data using IBM Cognos. The objective is to compare and contrast the mean values and standard deviations of cases and associated deaths per day and by country in the EU/EEA. This project encompasses defining analysis objectives, collecting COVID-19 data, designing relevant visualizations in IBM Cognos, and deriving insights from the data.

**Design Thinking:**

1.Analysis Objectives:

* Tracking the daily, weekly, or monthly trends in cases and deaths to understand the progression of the pandemic. Comparing data across regions, countries, or states to identify areas with higher infection rates or mortality rate.
* Monitoring the strain on healthcare systems by examining hospitalization rates and ICU admissions. Evaluating the impact of vaccination campaigns by comparing data before and after vaccine rollout.
* Validating and refining epidemiological models by comparing their predictions to actual data. Identifying and protecting vulnerable populations, such as the elderly and those with pre-existing conditions.
* Calculate the mean (average) number of cases or deaths over a specific period.

These comparisons help in understanding the central tendency and the dispersion of COVID-19 data, which can be crucial for decision-making, policy formulation, and public health responses.

2.Data Collection:

To achieve objectives of covid 19 cases the data is provided from the provided data set which includes the various parameters, such as mean and standard deviations. This dataset is crucial for our analysis and modelling efforts. It is essential to ensure that data is complete, reliable and well-documented.

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

3. Visualization Strategy:

1.Data preparation:

* Ensure your data source is connected to IBM Cognos, and the data is structured with columns for values of interest and any grouping variables. Create a report with open IBM Cognos.

2.Calculate Mean and Standard Deviation:

* Create new calculated items or use Cognos functions to calculate the mean and standard deviation for your data. For mean, you can use the AVG () function, and for standard deviation, you can use the STDEV () function.

3.Create Charts and Graphs:

* + Select the type of chart or graph that best suits your data. Common choices include:
  + Bar Chart: Suitable for comparing means of different categories.
  + Line Chart: Useful for tracking changes in means over time.
  + Error Bar Chart: Shows means with error bars representing standard deviations.
  + Box Plot: Visualizes the distribution of data including mean and standard deviation.
  + Scatter Plot: Helpful for visualizing the relationship between two variables along with their means and standard deviations.

4.Group and Filter Data (if needed):

* Use grouping and filtering options in Cognos to segment your data based on different criteria. This can be particularly useful when comparing means and standard deviations across subgroups.

4.Insights Generation:

Comparing the values and standard deviations of cases and deaths in a dataset, such as COVID-19 statistics, can provide several potential insights:

* Variability in Outcomes: If the standard deviation of deaths is much lower than that of cases, it suggests that outcomes (deaths) are more consistent or less variable compared to the number of reported cases. This could be due to better healthcare or early detection.
* Effectiveness of Interventions: A decrease in the standard deviation of cases and deaths over time may suggest that public health interventions, such as vaccination campaigns or social distancing measures, are becoming more effective at reducing both cases and deaths.
* 5. Early Warning Signals: A sudden increase in the standard deviation of cases while the standard deviation of deaths remains stable could be an early warning sign of a potential surge in fatalities, prompting the need for proactive measures.
* 6. Long-term Trends: Tracking changes in the standard deviations over time can help identify long-term trends in the pandemic's impact, shedding light on whether it's becoming more or less predictable.
* 7. Geographical Variations: Comparing standard deviations across different regions can highlight geographical variations in the spread and impact of the disease. High variation in some areas may require targeted interventions.

With a well-defined project scope and design thinking, we are poised to embark on a structured and data-driven analysis of COVID 19 cases, ultimately providing valuable insights into COVID 19 cases analysis. The subsequent phases of the project will build upon this foundation, including data preprocessing, exploratory data analysis, model development, and documentation.

Submit by:

ACHYUTH KUMAR.V

AU720921243059

JCT College Of Engineering And

Technology