**IBM NAAN MUDHAVALVAN**

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

**TEAM MEMBERS:**

**01.Mohammed Dammem**

**02. Mohammed Ahsen**

**03.Mohammed Wasiq Mohideen**

**04.Thoufeeq Ahmed**

**05.Salman**

**Description:**

The COVID-19 Cases Analysis project involves utilizing IBM Cognos to perform a comprehensive analysis of COVID-19 cases and deaths data. The primary objective is to compare and contrast the mean values and standard deviations of cases and associated deaths per day and by country within the European Union/European Economic Area (EU/EEA). This project aims to gain insights into the COVID-19 situation, identify trends, and provide data-driven information to aid decision-makers in understanding the pandemic's impact.

**Problem Understanding:**

The global COVID-19 pandemic has had a significant impact on countries within the EU/EEA. To address this, we need to understand the specific challenges and questions this analysis aims to tackle:

* **Comparative Analysis:** We want to compare and contrast the mean values and standard deviations of daily COVID-19 cases and deaths across different EU/EEA countries. This comparison can reveal variations in the pandemic's impact.
* **Trend Identification:** Are there specific trends or patterns in the data over time? Identifying trends can help predict future developments and guide response efforts.
* **Regional Differences:** How do countries within the EU/EEA differ in terms of COVID-19 cases and deaths? Understanding regional variations is crucial for targeted interventions.

**Solution for Solving this Problem:**

To address the problem of analyzing COVID-19 cases and deaths data within the EU/EEA, we propose the following solution:

* **Data Collection:** Obtain the COVID-19 dataset from the provided link (<https://www.kaggle.com/datasets/chakradharmattapalli/covid-19-cases>). Ensure that the dataset is up-to-date and contains relevant fields, including date, country, daily cases, and daily deaths.
* **Analysis Objectives:** Define specific analysis objectives, including comparing mean values and standard deviations of daily cases and deaths, identifying trends, and analyzing regional variations.
* **Visualization Strategy:** Plan the visualization of mean values and standard deviations using IBM Cognos. Visualizations will include informative charts and graphs, such as line charts, bar charts, and heatmaps, to represent the data effectively.
* **Insights Generation:** Use the analysis results to derive insights:
  + Identify countries with high variations in daily cases and deaths.
  + Analyze significant trends or spikes in the data.
  + Explore correlations between case counts and death counts.
  + Assess the impact of time on the COVID-19 situation within the EU/EEA.

**Proposed System Designs:**

The proposed system for COVID-19 Cases Analysis using IBM Cognos will consist of the following key components:

* **Data Preparation:**

This phase involves data collection, cleaning, and preprocessing. The COVID-19 dataset will be obtained and transformed into a format suitable for analysis.

**Description:** The Data Preparation component involves collecting, cleaning, and preprocessing the COVID-19 dataset to make it suitable for analysis.

**Tasks:**

1. **Data Collection:** Obtain the COVID-19 dataset from the provided source (Kaggle) and ensure it is up-to-date.
2. **Data Cleaning:** Address data quality issues, such as missing values, duplicates, and inconsistencies. Handle missing data, detect and resolve duplicates, and standardize data formats.
3. **Data Validation:** Validate the dataset to ensure it meets analysis requirements, including the presence of necessary columns (date, country, daily cases, daily deaths).
4. **Data Preprocessing:** Transform the data as needed, including aggregating daily cases and deaths, calculating metrics (mean and standard deviation), and formatting for time series analysis.
5. **Data Integration:** Integrate additional data sources (if required) for contextual analysis, such as population data or healthcare system information.
6. **Data Export:** Export the cleaned and preprocessed dataset in a format compatible with IBM Cognos.

* **IBM Cognos Setup:**

We will configure IBM Cognos for data analysis. This includes connecting to the dataset, setting up data models, and preparing the environment for visualization.

**Description:** The IBM Cognos Setup component involves configuring the IBM Cognos environment for data analysis.

**Tasks:**

1. **Environment Configuration:** Set up the IBM Cognos environment, including installing and configuring the software as per system requirements.
2. **Data Connection:** Establish a connection between IBM Cognos and the prepared COVID-19 dataset. Ensure data is imported accurately.
3. **Data Models:** Create data models within IBM Cognos to facilitate analysis. Define relationships, hierarchies, and calculations as needed.

* **Data Visualization:**

Utilizing IBM Cognos, we will design and create visualizations to represent mean values and standard deviations effectively. These visualizations will be dynamic and interactive, allowing users to explore the data.

**Description:** Data Visualization in IBM Cognos focuses on designing and creating effective visualizations to represent mean values and standard deviations.

**Tasks:**

1. **Visualization Design:** Plan the design of visualizations that effectively communicate mean values and standard deviations. Consider using line charts, bar charts, and other relevant chart types.
2. **Dynamic Interactivity:** Create dynamic and interactive visualizations that allow users to explore data by adjusting parameters, filtering, and zooming in on specific time periods or countries.

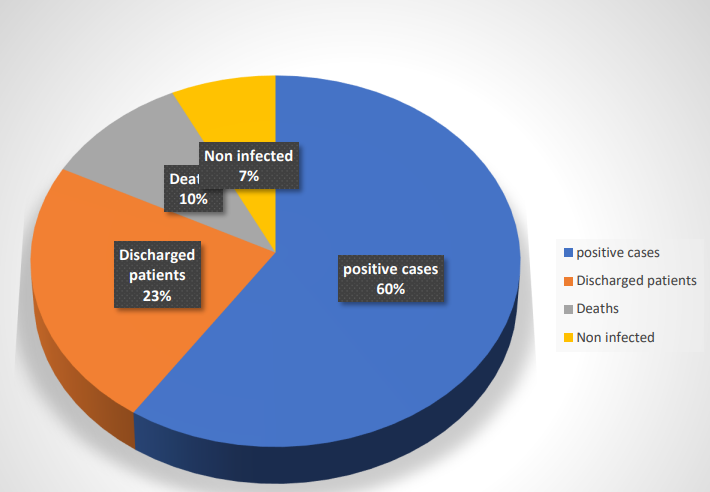
* **Insights Dashboard:**

We will develop an insights dashboard within IBM Cognos, presenting key findings, trends, and correlations. Users can interact with the dashboard to gain deeper insights into the COVID-19 data.

**Description:** The Insights Dashboard component involves developing a dashboard within IBM Cognos to present key findings, trends, and correlations.

**Tasks:**

1. **Dashboard Creation:** Design and create an interactive dashboard that provides an overview of the COVID-19 data analysis.
2. **Key Findings:** Present key findings, including mean values, standard deviations, trends over time, regional variations, and any identified correlations between cases and deaths.
3. **User Interaction:** Ensure that the dashboard allows users to interact with the data, explore insights, and customize their views.
4. **Dashboard Creation:** Design and create an interactive dashboard that provides an overview of the COVID-19 data analysis.
5. **Key Findings:** Present key findings, including mean values, standard deviations, trends over time, regional variations, and any identified correlations between cases and deaths.
6. **User Interaction:** Ensure that the dashboard allows users to interact with the data, explore insights, and customize their views.



* **Documentation:**

A comprehensive documentation process will accompany the project. This includes documenting data sources, data transformation steps, visualization choices, and insights derived.

**Description:** Comprehensive documentation is crucial to capture the project's details and processes.

**Tasks:**

1. **Data Source Documentation:** Provide detailed documentation of the COVID-19 dataset, including its source, structure, and any relevant metadata.
2. **Data Transformation Documentation:** Document all data cleaning, preprocessing, and integration steps, including any data corrections or transformations.
3. **Visualization Documentation:** Describe the choices made for visualization types, parameters, and configurations.
4. **Insights Documentation:** Record and explain the insights derived from the analysis, including any trends, correlations, or notable findings.

By following this proposed system design, we aim to provide a robust and insightful analysis of COVID-19 cases and deaths data in the EU/EEA, ultimately contributing to better decision-making during these challenging times.

This documentation serves as a roadmap for the COVID-19 Cases Analysis project and outlines the key steps and objectives for successfully analyzing and visualizing the data using IBM Cognos.

Dataset Link:

[Covid-19 Case Dataset on Kaggle](https://www.kaggle.com/datasets/chakradharmattapalli/covid-19-cases)