PROJECT TITLE: COVID-19 USING COGNOS

PROBLEM STATEMENT:

The project involves analyzing 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.

ANALYSIS OBJECTIVES:

Defining the specific objectives of analyzing COVID-19 cases and deaths data, such as comparing mean values and standard deviations.

1)Mean Values Comparison:

Calculate and compare the mean daily COVID-19 cases and deaths for each EU/EEA country over a specified time period. This will help identify countries with higher or lower average case and death rates.

2)Standard Deviations Comparison:

Calculate and compare the standard deviations of daily COVID-19 cases and deaths for each EU/EEA country. Analyze whether some countries exhibit greater variability in their case and death counts over time.

3)Temporal Analysis and Impact of Interventions:

Examine the trends in COVID-19 cases and deaths over time for each EU/EEA country. Identify any spikes or significant fluctuations and correlate them with events or interventions.Investigate the impact of different public health interventions or policy measures on the trajectory of COVID-19 cases and deaths in various EU/EEA countries.

4)Geographical Comparison and Population Adjusted Analysis:

Explore variations in COVID-19 cases and deaths by visualizing the data on a geographical map. Identify any regional patterns or clusters of high or low incidence. Normalize the data by considering population size for each country. Calculate and compare the per capita COVID-19 case and death rates to account for differences in population density and demographics.

5)Identification of Outliers and Data Quality Assessment:

Detect and analyze outliers in the data. Determine if any EU/EEA countries deviate significantly from the overall trends in cases and deaths.Evaluate the quality and consistency of the COVID-19 data for each EU/EEA country, including checking for missing or inconsistent values, and propose data cleansing and enhancement if necessary.

6)Forecasting and Communication of Results:

Utilize historical data to create predictive models for future COVID-19 cases and deaths within the EU/EEA countries, considering factors such as vaccination rates, seasonality, and emerging variants.: Present the findings and insights in a clear and accessible manner, using visualizations, dashboards, and reports generated with IBM Cognos. Ensure that the analysis results are easily understandable for policymakers and the general public.

7)Recommendations and Continuous Monitoring:

Based on the analysis, provide recommendations for public health actions, policy adjustments, or further research that can help mitigate the impact of COVID-19 in the EU/EEA countries.Set up a system for ongoing monitoring of COVID-19 data in the EU/EEA, allowing for timely updates and adjustments to strategies as the situation evolves.

By defining these specific objectives, the project will provide valuable insights into the dynamics of COVID-19 within the EU/EEA and support evidence-based decision-making to manage the pandemic effectively.

DATA COLLECTION

Obtain the provided data file containing COVID-19 cases and deaths information per day and by country in the EU/EEA

Date:

In the context of the EU/EEA COVID-19 dataset, the "Date" column records the specific date when COVID-19 data was collected. It follows the standard format "YYYY-MM-DD", allowing for precise tracking of the pandemic's progression over time.

Day:

This column may represent the information about the day of the week corresponding to each recorded date. This is particularly valuable for understanding weekly variations and patterns in COVID-19 cases and deaths within the EU/EEA.

Month:

Specifies the numerical month of the year, ranging from 1 to 12, allowing for the analysis of monthly variations.

Year:

This column represents the specific year when the COVID-19 data was recorded.

Cases:

Contains the count of confirmed COVID-19 cases reported on each date within the EU/EEA region, reflecting the virus's spread over time.It represents the count of individuals who tested positive for the virus. Tracking this data over time is essential for understanding the pandemic's progression.

Deaths:

Records the number of deaths attributed to COVID-19 on each date within the EU/EEA region, offering insights into the pandemic's impact on public health in this specific geographical context.

Country:

Specifies the name of the EU or EEA member country or region to which the COVID-19 data pertains, enabling geographical analysis and comparisons within the EU/EEA.

This dataset serves as a valuable resource for tracking and analyzing COVID-19 data trends within the EU and EEA regions. It allows researchers, analysts, and policymakers to gain insights into the pandemic's progression within this specific geographical area, identify patterns, and assess its impact on different EU and EEA member states.

VISUALIZATION STRATEGY

Planning how to visualize the mean values and standard deviations using IBM Cognos to create informative charts and graphs.

Designing an effective visualization strategy is crucial to convey our analysis results clearly and efficiently. Here are some considerations for visualizing mean values and standard deviations using IBM Cognos:

1)Chart Selection:

- Line charts: to show trends over time or continuous variables.
- Bar charts: for comparisons between different categories or groups .Error bars on bar charts: to represent standard deviations.
- Box plots: to display the distribution of data, including medians , quartiles, and outliers.

2)Grouping and Aggregation:

We determine how we want to group and aggregate our data. For example, we may want to show mean values and standard deviations per day, per week, or per month, depending on your analysis objectives

3)Color Coding labels and Legends:

We use color strategically to highlight key information. Provide clear labels for axes, data points, and legends. Ensure that our audience can easily interpret the visualizations without confusion.

4)Tooltips and Interactivity:

We utilize tooltips and interactive features. Tooltips can display specific values when users hover over data points.

5)Annotations:

We add annotations or text labels to the charts to highlight significant events or observations. Annotations can provide context to the data.

6)Normalization:

Consider normalizing data if you're comparing countries with significantly different populations or sizes. Per capita values can make comparisons fairer.

7)Trend Lines:

If we are using line charts to show trends, we could consider adding trend lines or moving averages to help them identify patterns and trends more easily.

8)Axis Scaling and Documentation:

We choose appropriate scaling for our axes.we ensure that the scaling does not exaggerate or minimize differences in the data.We include explanations and

descriptions alongside your visualizations to help users understand what they are seeing and the significance of mean values and standard deviations.

9)Consistency and Accessibility

We maintain a consistent visual style and color scheme throughout our visualizations and reports. Consistency makes it easier for users to understand and interpret the information.We ensure that our visualizations are accessible to a wide audience, including individuals with disabilities. We use accessible colors, provide alternative text for images, and consider readability for all users.

10)Testing and Feedback:

We test our visualizations with potential users or colleagues to gather feedback. Make improvements based on their input to enhance the effectiveness of our visualizations.

By carefully planning and implementing our visualization strategy in IBM Cognos, we can create informative and insightful charts and graphs that effectively communicate the mean values and standard deviations in your COVID-19 data analysis.

INSIGHT GENERATION:

Identify potential insights from the comparison of mean values and standard derivations of cases and deaths.

Mean Comparison Insights:

Present insights derived from the comparison of mean values for COVID-19 cases and deaths. Discuss trends and variations observed. Examples include,The mean number of COVID-19 cases increased steadily from [start date] to [end date], indicating a growing trend.While the mean number of cases increased, the mean number of deaths remained relatively stable during the same period.

Standard Deviation Insights:

Share insights based on the comparison of standard deviations for COVID-19 cases and deaths. Interpret standard deviation values in the context of data variability.The standard deviation of cases was highest during [specific time period], suggesting increased variability in case counts.The standard deviation of deaths decreased steadily over the course of the pandemic, indicating more consistent reporting and response.

In conclusion, this project employs IBM Cognos to analyze COVID-19 data, facilitating a deeper understanding of the pandemic's dynamics within the EU/EEA. The insights derived from mean values, standard deviations, and visualizations aid evidence-based decision-making to effectively manage the impact of COVID-19.