**Project : Covid-19 Cases Analysis**

**Phase 4 : project development part 2**

**Synopsis:**

**INTRODUCTION**

**CREATING VISUALIZATIONS**

**Create mean values visualization**

**Create standard deviation visualization**

**DERIVING INSIGHTS**

**Identifying Trends**

**Analyzing Variations**

**Correlation Analysis**

**CONCLUSION**

**INTRODUCTION**

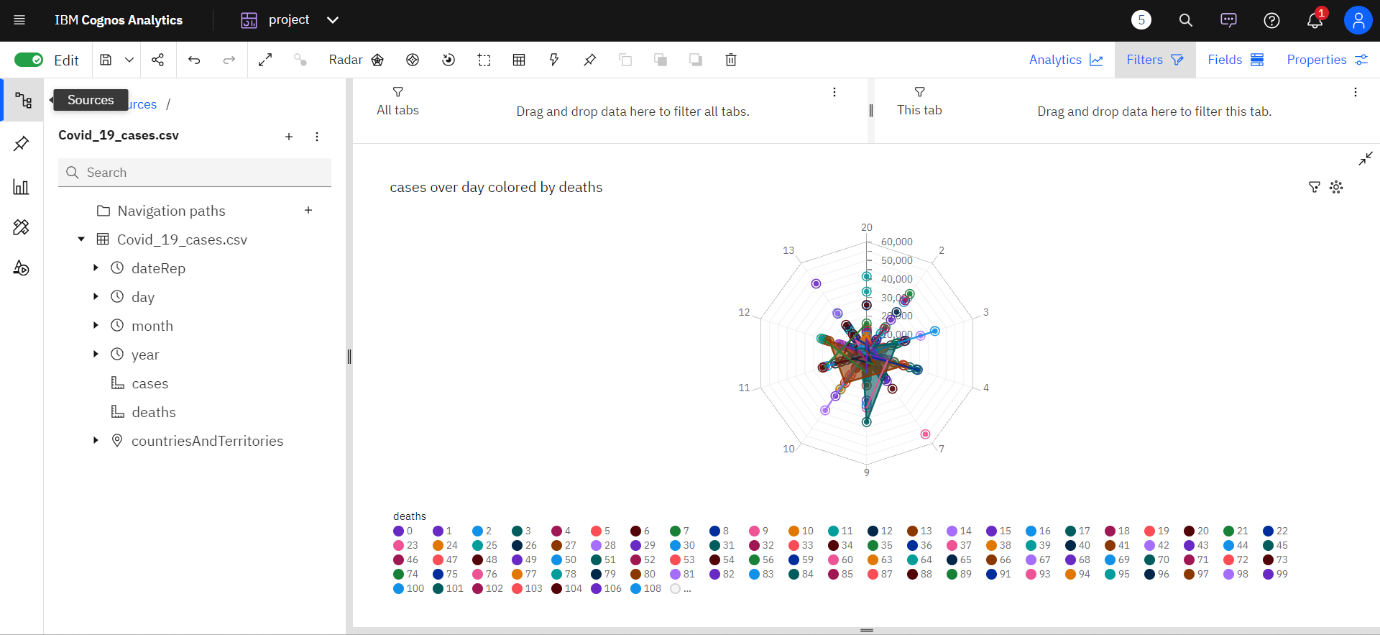
Continue building the analysis by creating visualizations using IBM Cognos and deriving insights from the data.

Create charts and graphs in IBM Cognos to visualize and compare the mean values and standard deviations of COVID-19 cases and associated deaths.

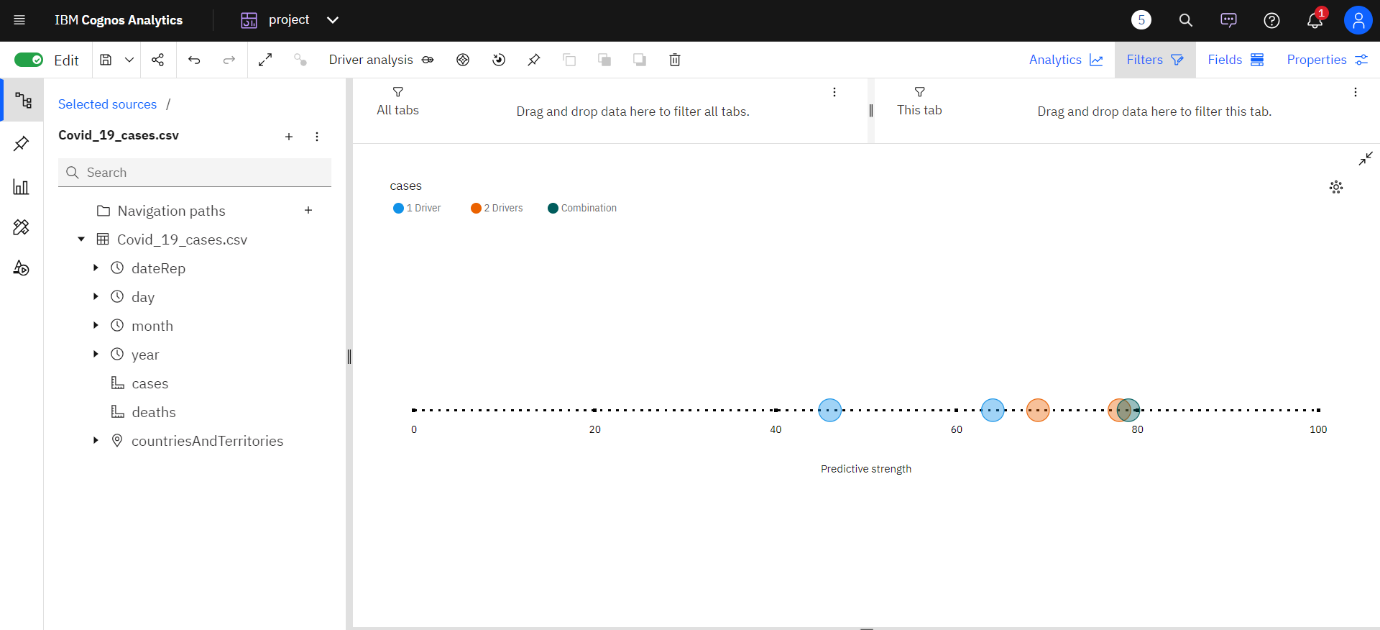
Analyze the visualizations to identify trends, variations, and potential correlations between cases and deaths.

**CREATING VISUALIZATIONS**

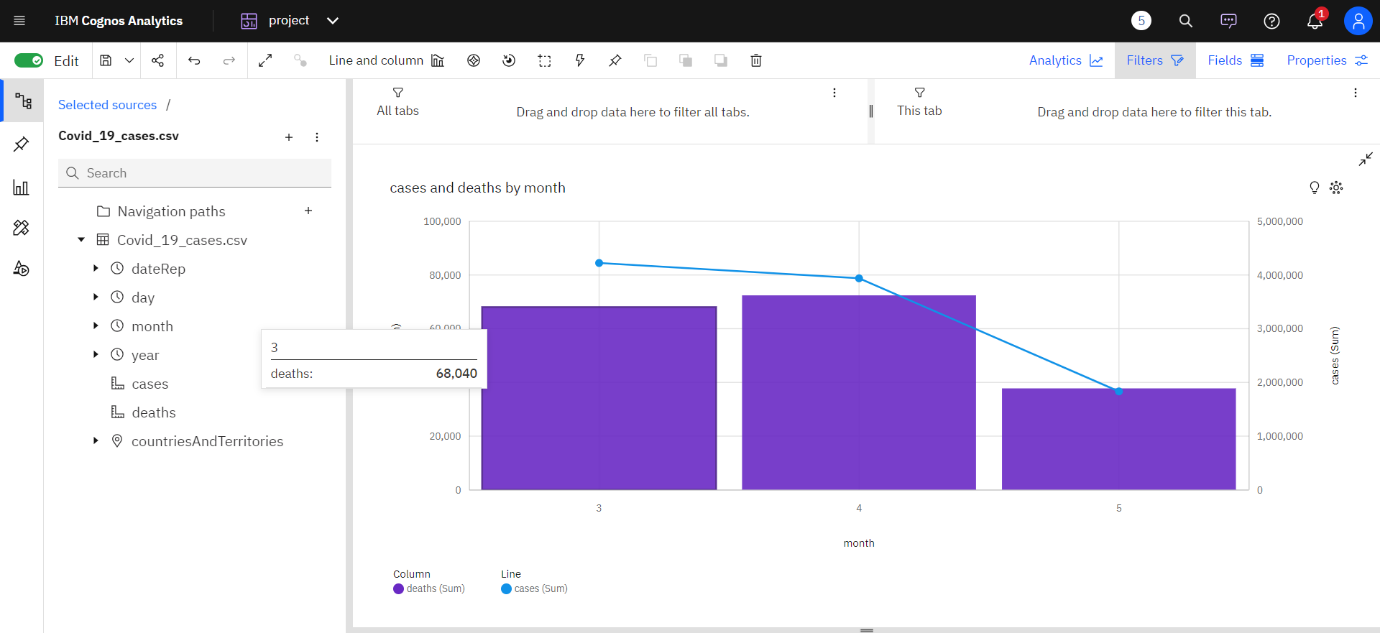
**Visualization of cases over day colored by deaths**

****

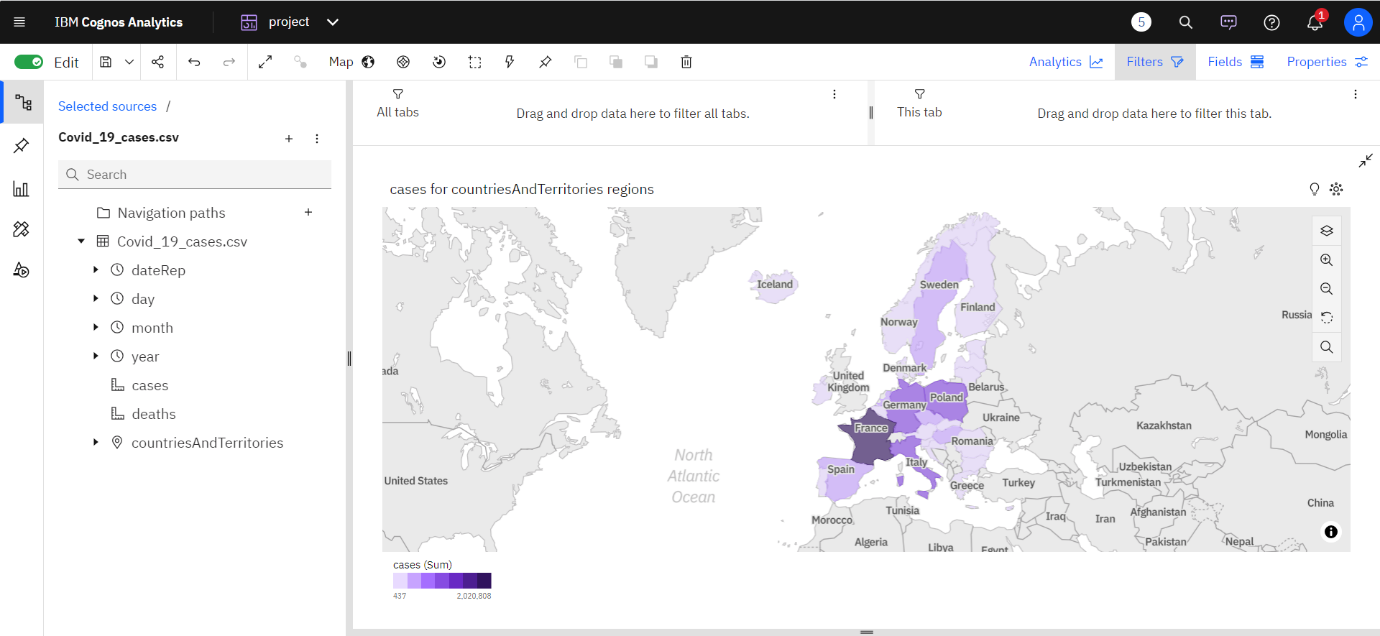
**Visualization by cases**

****

**Visualization cases and deaths by month**

****

**visualization of cases for countries and territories regions**

****

**Create mean values visualization:**

* Create a line chart or bar chart with dates on the x-axis and the mean values of covid-19 cases and deaths on the y-axis.
* Aggregate the data to calculate the mean values for cases and deaths for each date.

**Steps to aggregate the data and calculate the mean values for covid-19 cases and deaths for each date in ibm cognos :**

1. **Create a query:**

* Start by creating a new query in ibm cognos based on your imported dataset.
* Select date, cases, and deaths
* In the query, select the date column, the column representing covid-19 cases, and the column representing deaths.

1. **Aggregate the data:**

To calculate the mean values, you need to aggregate the data. In ibm cognos, you can do this using aggregation functions. Here’s how you can do it:

**For cases:**

* Click on the "cases" column.
* In the properties pane, click on the ellipsis (...) Next to "aggregate".
* Choose "mean" or "average" from the list of aggregation functions.

**For deaths:**

* Click on the "deaths" column.
* In the properties pane, click on the ellipsis (...) Next to "aggregate".
* Choose "mean" or "average" from the list of aggregation functions.

**Group by date:**

* To calculate mean values for each date, you need to group the data by the date column.
* Drag the "date" column to the "group by" area of the query.

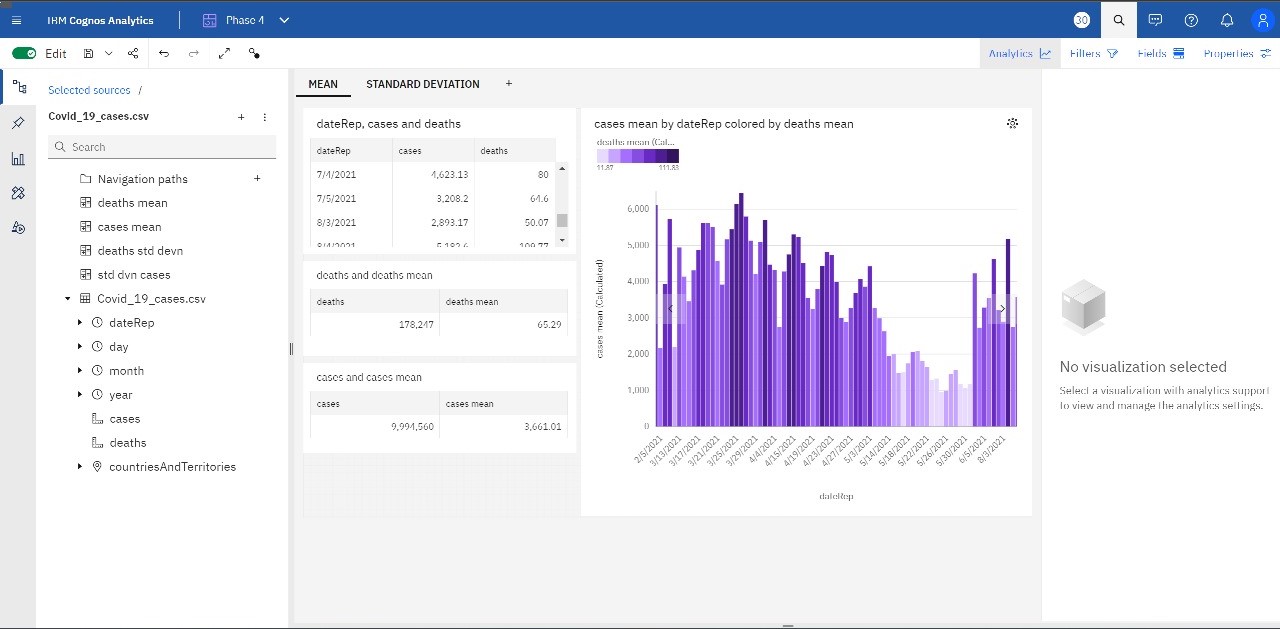
**Run the query:**

Run the query to see the aggregated data. You should now have mean values for covid-19 cases and deaths for each date in the result set.

1. **Create visualizations:**

With the aggregated data, you can create visualizations like line charts or Bar Charts With Dates On The X-Axis And Mean Values For Cases And Deaths On The Y-Axis.

**visualize and compare the mean values**

****

**Create standard deviation visualization:**

* Create a bar chart with dates on the x-axis and standard deviation values of covid-19 cases and deaths on the y-axis.
* Aggregate the data to calculate standard deviation values for cases and deaths for each date.

**To aggregate the data and calculate the standard deviations values for covid-19 cases and deaths for each date in ibm cognos :**

**Create a Query:**

* Start by creating a new query in IBM Cognos based on your imported dataset.
* Select Date, Cases, and Deaths:
* In the query, select the date column, the column representing COVID-19 cases, and the column representing deaths.

**Aggregate the Data:**

To calculate the standard deviation values, you need to aggregate the data. In IBM Cognos, you can do this using aggregation functions. Here’s how you can do it:

**For Cases:**

* Click on the "Cases" column.
* In the Properties pane, click on the ellipsis (...) next to "Aggregate".
* Choose "Standard Deviation" from the list of aggregation functions.

**For Deaths:**

* Click on the "Deaths" column.
* In the Properties pane, click on the ellipsis (...) next to "Aggregate".
* Choose "Standard Deviation" from the list of aggregation functions.

**Group by Date:**

* To calculate standard deviation values for each date, you need to group the data by the date column.
* Drag the "Date" column to the "Group By" area of the query.

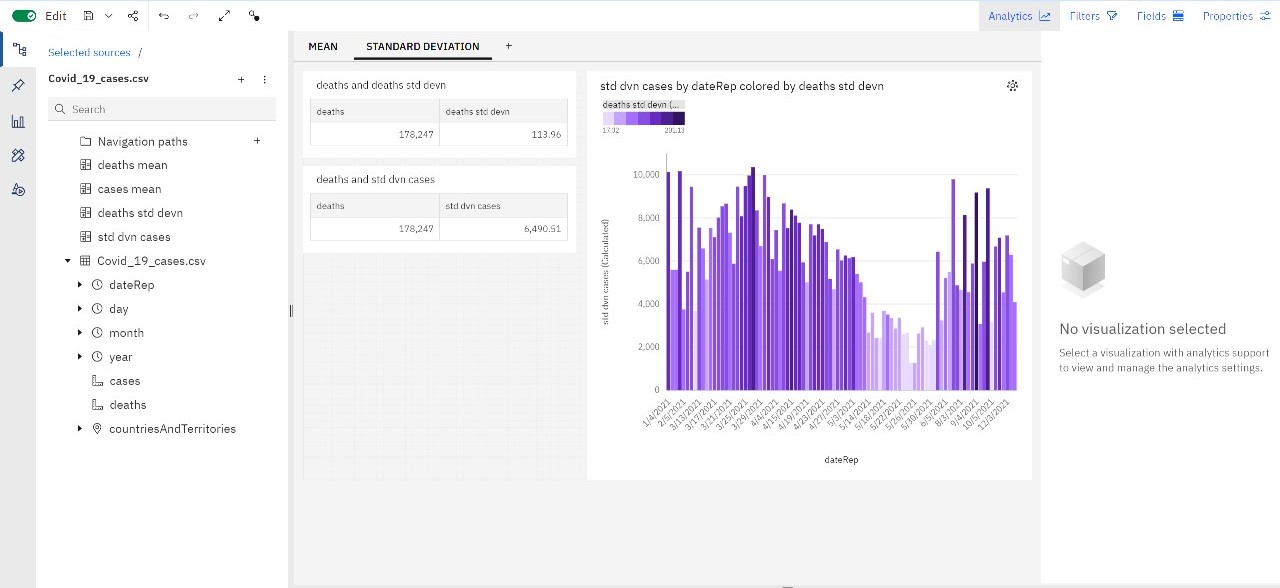
**Run the Query:**

Run the query to see the aggregated data. You should now have standard deviation values for COVID-19 cases and deaths for each date in the result set.

**Create Visualizations:**

With the aggregated data, you can create visualizations like bar charts or line charts with dates on the x-axis and standard deviation values for cases and deaths on the y-axis.

**visualize and compare the standard deviation values**

****

**DERIVING INSIGHTS**

Deriving insights for COVID-19 cases and death analysis involves analyzing trends,variations, and correlations related to the pandemic.

By extracting valuable information from the data, such as identifying hotspots, understanding mortality rates, and predicting future trends, insights aid in informed decision-making for healthcare strategies, resource allocation, and public health interventions.

**Identifying Trends:**

**Cases:**

Top Peak Dates for Cases:

1. **Timestamp('2021-12-05 00:00:00')**
2. **Timestamp('2021-12-05 00:00:00')**
3. **Timestamp('2021-12-04 00:00:00')**

**Deaths:**

Top Peak Dates for Deaths:

1. **Timestamp('2021-12-05 00:00:00')**
2. **Timestamp('2021-12-05 00:00:00')**
3. **Timestamp('2021-12-05 00:00:00')**

**Analyzing Variations:**

**Cases:**

Top Sudden Rise Dates for Cases:

**839 2021-11-05**

**846 2021-04-05**

**849 2021-01-05**

**850 2021-04-30**

**851 2021-04-29**

**...**

**2167 2021-03-18**

**2171 2021-03-14**

**2173 2021-12-03**

**2590 2021-04-19**

**2597 2021-12-04**

**Deaths:**

Top Sudden Rise Dates for Deaths:

**853 2021-04-27**

**859 2021-04-21**

**860 2021-04-20**

**867 2021-04-13**

**872 2021-08-04**

**...**

**2174 2021-11-03**

**2614 2021-03-26**

**2618 2021-03-22**

**2635 2021-05-03**

**2637 2021-03-03**

**Correlation Analysis:**

Correlation Coefficient between Cases and Deaths is **0.766308878657635**

0.766308878657635 (indicating a strong positive correlation).

**Conclusion:**

Summarize the main insights derived from the visualization analysis.

Emphasize the importance of ongoing monitoring and analysis in understanding the COVID-19 situation.

Remember, the format and depth of analysis can be adjusted based on the complexity of the dataset and the specific requirements of the project. It's crucial to provide clear and concise explanations, supported by evidence from the visualizations, to convey the findings effectively.