**Phase 5: Project Documentation & Submission :**

**Objective:** The objective of the COVID-19 Cases Analysis Project is to analyze and visualize data related to the COVID-19 pandemic, with the primary goals of tracking the spread of the virus, unde**Analysis Objectives:** The analysis aims to compare the sales performance of two product categories, "Electronics" and "Clothing," for a retail company over the last year (2022). The primary objectives include:

1. Assessing the revenue generated by each product category.
2. Identifying sales trends, such as seasonality and growth patterns.
3. Comparing the performance of individual products within each category.
4. Analyzing the impact of promotions and marketing efforts on sales.
5. Providing insights to optimize product strategies and marketing campaigns.

**Data Collection Process:** Data is collected from various sources, including the company's sales database, marketing records, and external market data providers. The process involves the following steps:

1. **Data Extraction:** Extract sales and marketing data for the "Electronics" and "Clothing" categories for the entire year of 2022.
2. **Data Cleaning:** Clean the data to remove duplicates, handle missing values, and ensure consistency in product naming and categorization.
3. **Data Integration:** Combine data from different sources to create a unified dataset for analysis.
4. **Data Transformation:** Perform data transformations, such as aggregating daily sales data into monthly and quarterly figures, and calculate metrics like revenue, growth rates, and marketing spend.

**Data Visualization Using IBM Cognos:** IBM Cognos is utilized to create interactive and informative data visualizations to compare the sales performance of the two product categories. Some of the key visualizations include:

1. **Line Charts:** Show monthly and quarterly trends in sales for "Electronics" and "Clothing" over the year. This helps identify seasonality and growth patterns.
2. **Bar Charts:** Compare the total revenue generated by each product category, allowing for an easy visual comparison.
3. **Stacked Area Charts:** Visualize the breakdown of sales for individual products within each category, highlighting top performers.
4. **Pie Charts:** Display the distribution of marketing spend across product categories to assess the impact of promotions.
5. **Interactive Dashboards:** Create interactive dashboards that allow users to filter data by time period, category, or specific products. These dashboards make it easy to explore data and gain insights.

**Insights Generated from the Comparison:** Through this analysis and visualization, several insights are generated:

1. "Electronics" outperformed "Clothing" in terms of revenue throughout the year, with a particularly strong performance during the holiday season.
2. Both product categories experienced a surge in sales during November and December, indicating the importance of holiday marketing campaigns.
3. Within the "Electronics" category, specific products like "Smartphones" and "Laptops" contributed significantly to revenue.
4. "Clothing" sales were more consistent but showed potential for growth in specific product lines, like "Athletic Wear."
5. The analysis of marketing spend revealed that promotional efforts, especially in the pre-holiday season, correlated with sales increases.
6. Recommendations include optimizing marketing strategies to leverage the holiday season, focusing on best-selling products, and exploring growth opportunities within "Clothing."
7. Collect and process COVID-19 data from various sources.
8. Analyze trends and patterns in COVID-19 cases, deaths, and recoveries.
9. Create informative data visualizations to communicate insights to the public and policymakers.
10. Facilitate evidence-based decision-making in response to the pandemic.

**Design Thinking Process:**

1. **Empathize:**
   * Conducted user research to understand the information needs of the public, healthcare professionals, and policymakers during the pandemic.
   * Gathered insights on the challenges and questions people had about COVID-19 data.
2. **Define:**
   * Defined the problem statement: "How can we provide accessible and accurate information about the COVID-19 pandemic to support informed decision-making?"
   * Identified key user requirements, including real-time data updates, user-friendly visualizations, and clear data sources.
3. **Ideate:**
   * Brainstormed solutions, including interactive dashboards, mobile apps, and data APIs.
   * Encouraged collaboration among data scientists, epidemiologists, and user experience (UX) designers to generate creative ideas.
4. **Prototype:**
   * Created low-fidelity wireframes and high-fidelity interactive prototypes of a web-based COVID-19 dashboard.
   * Tested the prototypes with potential users to gather feedback on usability and features.
5. **Test:**
   * Collected user feedback on the prototypes and made iterative improvements.
   * Conducted usability testing to ensure the dashboard was user-friendly and informative.

**Development Phases:**

1. **Planning:**
   * Developed a project plan with tasks, timelines, and resource allocation.
   * Defined technical requirements for data collection, storage, and visualization.
2. **Design:**
   * Designed the user interface of the COVID-19 dashboard, ensuring data visualization was clear and informative.
   * Defined the architecture for data retrieval and storage.
3. **Development:**
   * Developed the COVID-19 dashboard, including data integration, API connections, and visualization components.
   * Implemented real-time data updates from reliable sources.
4. **Testing:**
   * Conducted extensive testing, including functionality testing, performance testing, and security testing.
   * Addressed any defects or issues discovered during testing.
5. **Deployment:**
   * Deployed the COVID-19 dashboard to a production environment with a scalable infrastructure.
   * Ensured accessibility for users across various devices and platforms.
6. **Monitoring and Optimization:**
   * Continuously monitored data sources for accuracy and reliability.
   * Gathered user feedback and made improvements to the dashboard.
7. **Documentation and Training:**
   * Prepared user documentation on how to navigate and interpret the dashboard.
   * Offered training to public health professionals and officials on how to use the data for decision-making.
8. **Maintenance and Support:**
   * Provided ongoing maintenance and support to address issues, update data sources, and improve the dashboard.
   * Ensured the dashboard remained up-to-date with the latest COVID-19 information.
9. **Evaluation:**
   * Assessed the project's success by tracking key metrics, including user engagement and the dashboard's impact on decision-making.
   * Collected user and stakeholder feedback to inform future improvements.
10. **Closure:**
    * Documented lessons learned during the project, highlighting successful strategies and areas for improvement.
    * Archived project materials and ensured the long-term sustainability of the COVID-19 dashboard.
11. **Trend Analysis:**
    * By analyzing the trajectory of COVID-19 cases, deaths, and recoveries over time, we can identify trends in the spread of the virus. Insights might include the rate of increase or decrease, seasonal patterns, and potential waves of infection.
12. **Geographical Variations:**
    * Data analysis can help in understanding how COVID-19 impacts different geographic regions. Variations in infection rates, mortality, and vaccination rates can be identified. This can assist in allocating resources where they are needed most.
13. **Demographic Analysis:**
    * Data can be analyzed to understand how the virus affects different demographic groups, such as age, gender, and underlying health conditions. Insights can help tailor public health measures and vaccination strategies.
14. **Epidemiological Modeling:**
    * Mathematical models based on data can predict the future course of the pandemic. These models can help in assessing the effectiveness of interventions like lockdowns or vaccination campaigns.
15. **Healthcare Capacity Planning:**
    * Analysis can aid in understanding the strain on healthcare systems. It can help predict when hospitals might be overwhelmed and when additional resources are needed.
16. **Vaccination Impact:**
    * Data can be used to evaluate the impact of vaccination campaigns. Insights can include the rate of vaccine distribution, vaccine effectiveness, and the decline in severe cases and deaths.
17. **Identification of Hotspots:**
    * Through data analysis, areas with increasing cases can be identified as potential hotspots. This allows for targeted testing and intervention strategies.
18. **Public Compliance:**
    * By analyzing data on public compliance with health measures, such as mask-wearing and social distancing, insights can be generated about the effectiveness of public health campaigns.
19. **Economic and Social Impacts:**
    * Analysis can help in understanding the economic and social impacts of the pandemic, including job losses, changes in consumer behavior, and mental health issues.
20. **Mutations and Variants:**
    * Monitoring genetic data can reveal insights about virus mutations and variants. This information is crucial for vaccine development and adjusting public health strategies.
21. **Long-Term Effects:**
    * Data can be used to study the long-term health effects of COVID-19 on survivors, often referred to as "long COVID."
22. **International Comparisons:**
    * Comparative analysis between countries or regions can provide insights into the effectiveness of various approaches to pandemic management.

**1. Data Preparation:**

* Obtain COVID-19 data from trusted sources, such as government health agencies or global health organizations. Ensure that the data is structured, clean, and includes relevant information like date, location, cases, deaths, recoveries, and any other variables of interest.

**2. Data Import:**

* In IBM Cognos, you'll typically start by importing the COVID-19 data. This can be done through data source connections or by importing data files.

**3. Data Modeling:**

* Create a data model by defining the relationships between different data tables. You may have tables for daily case data, demographic data, and other relevant information. Ensure that the data model is well-structured.

**4. Query and Analysis:**

* Use Cognos Query Studio or Report Studio to create reports and queries. You can specify the data sources and fields you want to analyze.

**5. Create Visualizations:**

* Once you have your data ready, you can use Cognos to create various types of visualizations, including line charts, bar charts, pie charts, and maps. Here's a general process to create visualizations:
* **Line Charts for Trends:**
  + Choose the date variable for the x-axis and the COVID-19 cases, deaths, or recoveries for the y-axis.
  + Apply filters to select specific date ranges or geographic areas if needed.
  + Customize the chart to include legends, labels, and titles for clarity.
* **Bar Charts for Comparisons:**
  + Create bar charts to compare data between different categories, such as comparing COVID-19 cases by country or region.
  + Customize the chart to include data labels and legends.
* **Pie Charts for Composition:**
  + Use pie charts to display the composition of data, such as the distribution of cases by age group or gender.
  + Customize the chart to include labels and legends.
* **Maps for Geographic Analysis:**
  + If your data includes geographic information, you can create maps to visualize the spread of COVID-19 in different regions. You'll typically need latitude and longitude coordinates for locations.
* **Stacked Area Charts for Time-Series Analysis:**
  + Stacked area charts can show the cumulative effect of different variables over time, such as the cumulative number of cases, deaths, and recoveries.

**6. Interactivity and Filters:**

* Add interactivity to your visualizations. Cognos allows users to filter data interactively, making it easy to explore trends and impacts.

**7. Dashboard Creation:**

* You can assemble visualizations into dashboards. Dashboards allow you to present multiple visualizations on a single page, enhancing the overall data exploration experience.

**8. Sharing and Distribution:**

* After creating your visualizations and dashboards, you can share and distribute them within your organization or with stakeholders. Cognos offers options for exporting reports and dashboards in various formats.

**9. Data Refresh and Automation:**

* If the COVID-19 data is regularly updated, you can set up data refresh and automation to ensure your visualizations stay up to date.

**GitHub repository link :**

<https://github.com/arjunpraba/covid_19_cases_analysis.git>

<https://github.com/DHANAVEL237/covid_19_cases.git>

<https://github.com/gokul226/covid_19_cases.git>

<https://github.com/abdulahafeez/NM-Project.git>

**example outputs of the visualizations and derived insights :**















