

## PHASE 5: FINAL DOCUMENTATION

### Project Title: Product Sales Analysis

#### Team Members:

1. Rowena Annfiniya A (710621104038)
2. Saparna B (710621104044)
3. Carmal Princy A (710621104007)
4. Nandhini T (710621104028)
5. Ashiritha S (710621104003)

#### Introduction:

**1. Problem Statement:** Project Definition: The project involves using IBM Cognos to analyze sales data and extract insights about top selling products, peak sales periods, and customer preferences. The objective is to help businesses improve inventory management and marketing strategies by understanding sales trends and customer behavior. This project includes defining analysis objectives, collecting sales data, designing relevant visualizations in IBM Cognos, and deriving actionable insights.

**2. Understanding Stakeholder Needs:** We have identified several key stakeholders who have a vested interest in this project:

- Sales Team: They are concerned about declining sales and need strategies to boost their performance.
- Product Managers: They want to identify which product categories are underperforming and where potential growth opportunities lie.
- Executives: They are focused on overall company profitability and need data-driven insights to make informed decisions.

#### Dataset:

The dataset we will be using for this innovation phase is available at the following link:

<https://www.kaggle.com/datasets/ksabishek/product-sales-data>

#### Innovation:

##### Steps to Achieve Innovation:

##### 1. Data Exploration and Understanding

- **Data Loading:** Load the product sales dataset into a data analysis environment such as Python or R.
- **Data Overview:** Understand the structure of the dataset, including the number of rows and columns, data types, and missing values.
- **Exploratory Data Analysis (EDA):** Conduct initial EDA to get insights into the data. This may involve summary statistics, data visualization, and identifying patterns.

## **2. Problem Definition and Design Refinement:**

- **Review Initial Design:** Revisit the design concept from the previous phase and ensure it aligns with the insights gained during EDA.
- **Problem Definition:** Clearly define the problem or objective you aim to address using the dataset. This may involve refining your design or specifying the exact analysis required.
- **Hypothesis Formulation:** If applicable, formulate hypotheses that can guide your analysis and solution development.

## **3. Data Preprocessing:**

- **Data Cleaning:** Handle missing values, duplicate records, and outliers as needed.
- **Feature Engineering:** Create new features if they are necessary for your analysis or modeling.
- **Data Transformation:** Convert categorical variables into numerical formats if required.
- **Data Scaling or Normalization:** Scale the data to ensure consistency in units if you plan to use algorithms sensitive to feature scaling.

## **4. Analysis and Modeling:**

- **Select Appropriate Algorithms:** Choose the right machine learning or statistical algorithms based on the nature of the problem (e.g., regression, classification, time series analysis).
- **Model Building:** Develop models based on the selected algorithms.
- **Training and Testing:** Split the data into training and testing sets for model validation.
- **Parameter Tuning:** Fine-tune model hyperparameters to optimize performance.

## **5. Evaluation:**

- **Model Evaluation:** Assess the performance of the models using relevant metrics (e.g., accuracy, RMSE, MAE, ROC AUC) depending on the problem type.
- **Interpretability:** If applicable, interpret the model results to gain insights into the problem or solution.

## **6. Visualization and Reporting:**

- **Data Visualization:** Create visualizations to communicate key findings and insights effectively.

- **Report Generation:** Prepare a comprehensive report that includes problem definition, methodology, results, and recommendations.

## **7. Deployment:**

- **Implementation:** If the solution involves a software application or tool, develop it based on the results.
- **Testing:** Ensure that the implemented solution works correctly and meets the specified requirements.
- **Monitoring:** Set up monitoring to track the performance of the deployed solution over time.

## **Development:**

### **1. Data Preprocessing:**

- **Data Cleaning:** Clean the dataset to handle missing values, duplicates, and outliers. This step ensures that the data is of high quality and reliable for analysis.
- **Feature Engineering:** Create new features or derive relevant metrics that can enhance the analysis. For example, you might calculate monthly or quarterly aggregates for sales and expenses.

### **2. Exploratory Data Analysis (EDA):**

- **Data Visualization:** Visualize the data using libraries like Matplotlib, Seaborn, or Plotly to identify patterns and trends. Create plots such as time series plots, histograms, scatter plots, and box plots to gain insights.
- **KPI Calculation:** Calculate key performance indicators (KPIs) such as average monthly sales, growth rates, and seasonality indices to better understand the data.

### **3. Time Series Forecasting:**

- **Model Selection:** Choose an appropriate time series forecasting model, such as ARIMA, SARIMA, or Prophet, based on the characteristics of your sales data.
- **Model Training:** Split the data into training and testing sets. Train the chosen forecasting model on historical data to make future predictions.
- **Model Evaluation:** Assess the accuracy of the forecasting model using metrics like Mean Absolute Error (MAE) or Root Mean Square Error (RMSE).
- **Future Sales Prediction:** Generate forecasts for future sales, which will be used for inventory planning and decision-making.

### **4. Interactive Dashboards:**

- **Dashboard Design:** Design interactive dashboards using tools like Tableau, Power BI, or Python libraries such as Plotly and Dash. These dashboards should present the analysis results in a visually engaging and user-friendly format.

- Data Visualization: Incorporate various visualizations, including line charts, bar charts, heatmaps, and slicers, to allow users to interact with the data and gain insights.

## 5. Recommendations:

- Actionable Insights: Based on the analysis and forecasting results, provide actionable recommendations to address the business challenges outlined in the problem statement.
- Presentation: Summarize the recommendations in a clear and concise manner, making it easy for stakeholders to understand and implement changes in sales strategies.

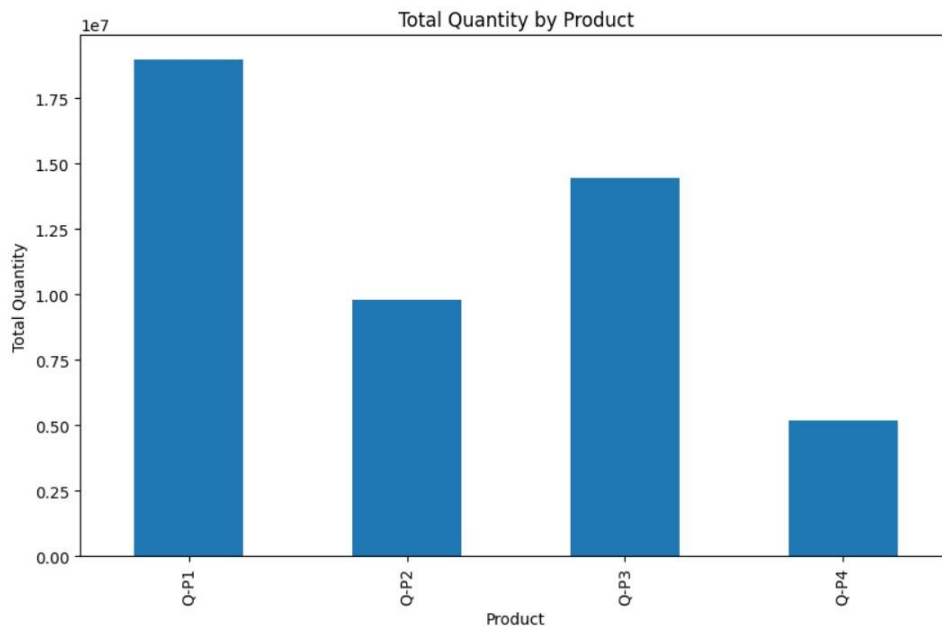
## Code:

```
import pandas as pd
import matplotlib.pyplot as plt

# Load the cleaned dataset
df = pd.read_csv(r'C:\Users\deepa\Downloads\archive (2)\statsfinal.csv')

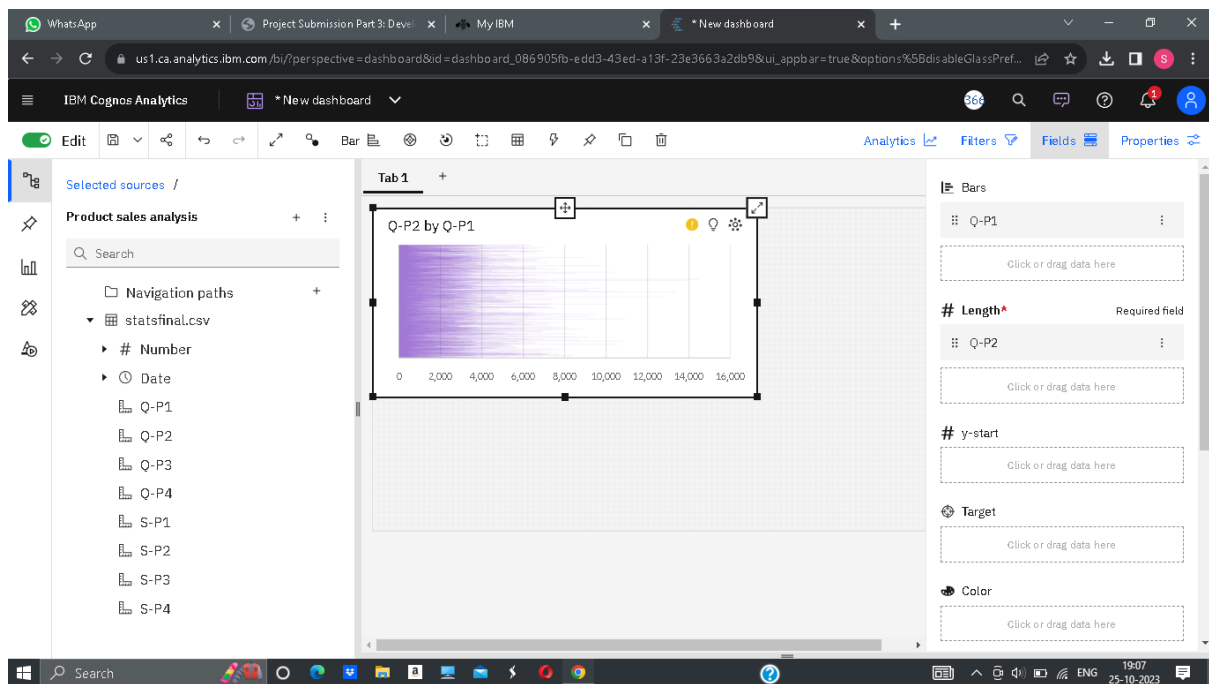
# Example 1: Line Plot
plt.figure(figsize=(10, 6))
plt.plot(df['Date'], df['Q-P1'], label='Quantity P1')
plt.plot(df['Date'], df['Q-P2'], label='Quantity P2')
plt.plot(df['Date'], df['Q-P3'], label='Quantity P3')
plt.plot(df['Date'], df['Q-P4'], label='Quantity P4')
plt.xlabel('Date')
plt.ylabel('Quantity')
plt.title('Quantity vs. Date')
plt.legend()
plt.show()

# Example 2: Bar Plot
total_quantity = df[['Q-P1', 'Q-P2', 'Q-P3', 'Q-P4']].sum()
total_quantity.plot(kind='bar', figsize=(10, 6))
plt.xlabel('Product')
plt.ylabel('Total Quantity')
plt.title('Total Quantity by Product')
plt.show()
```

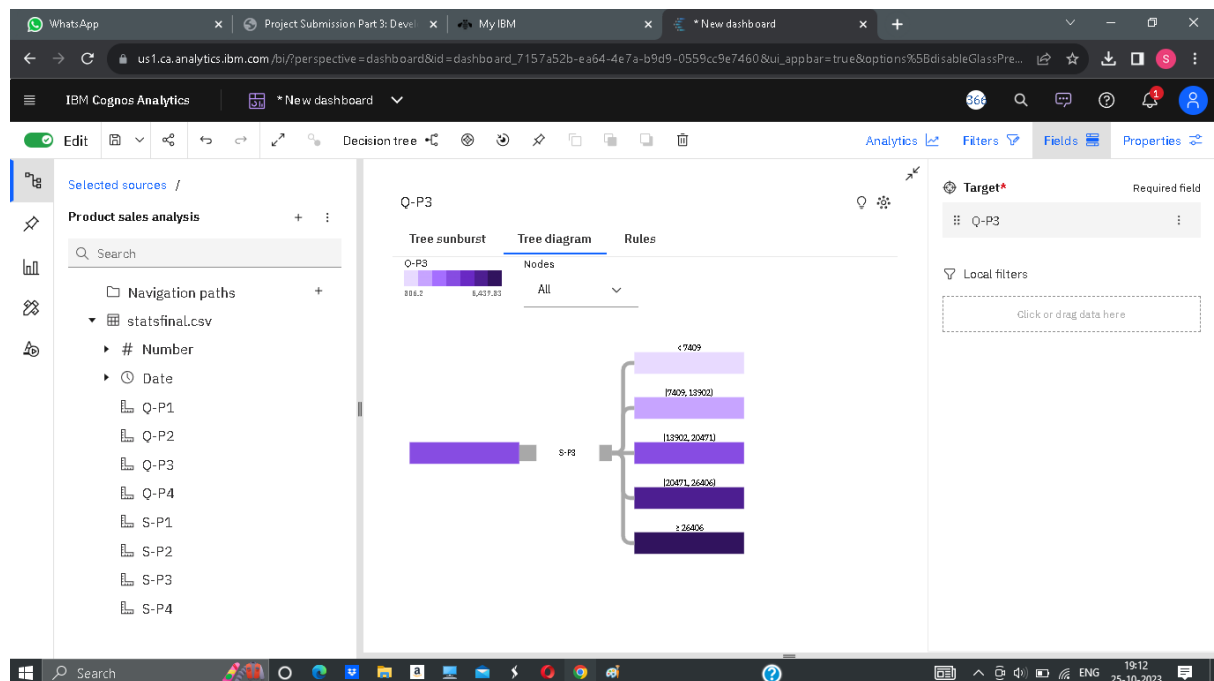
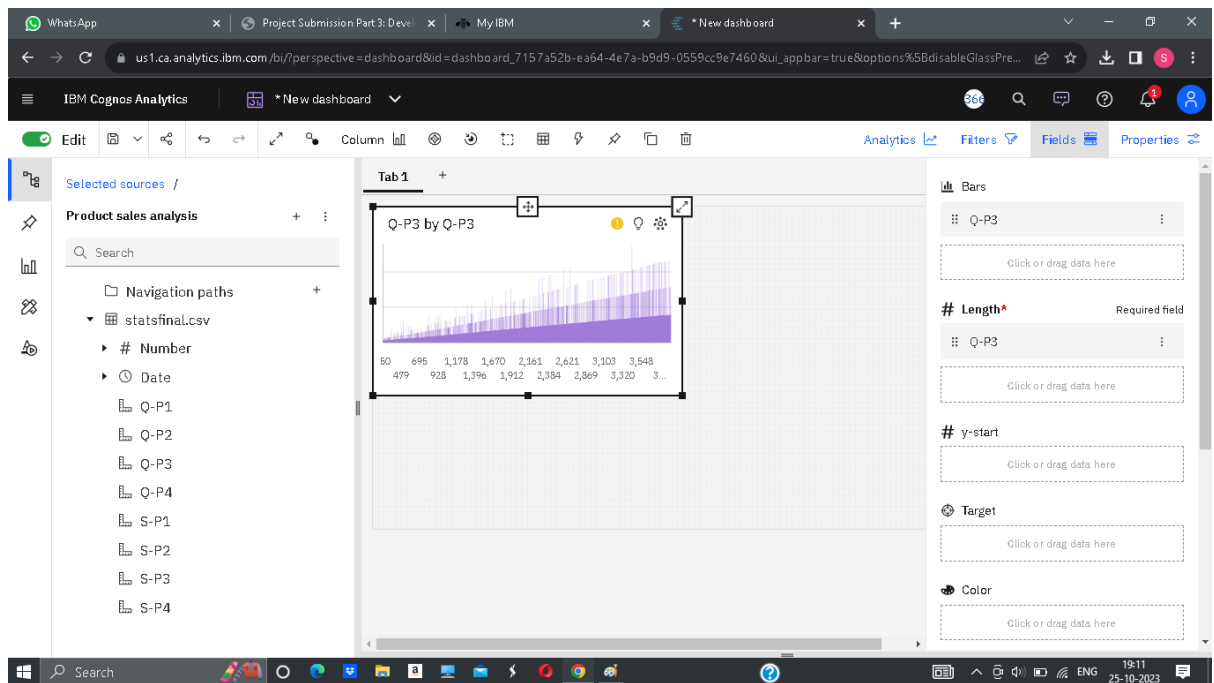


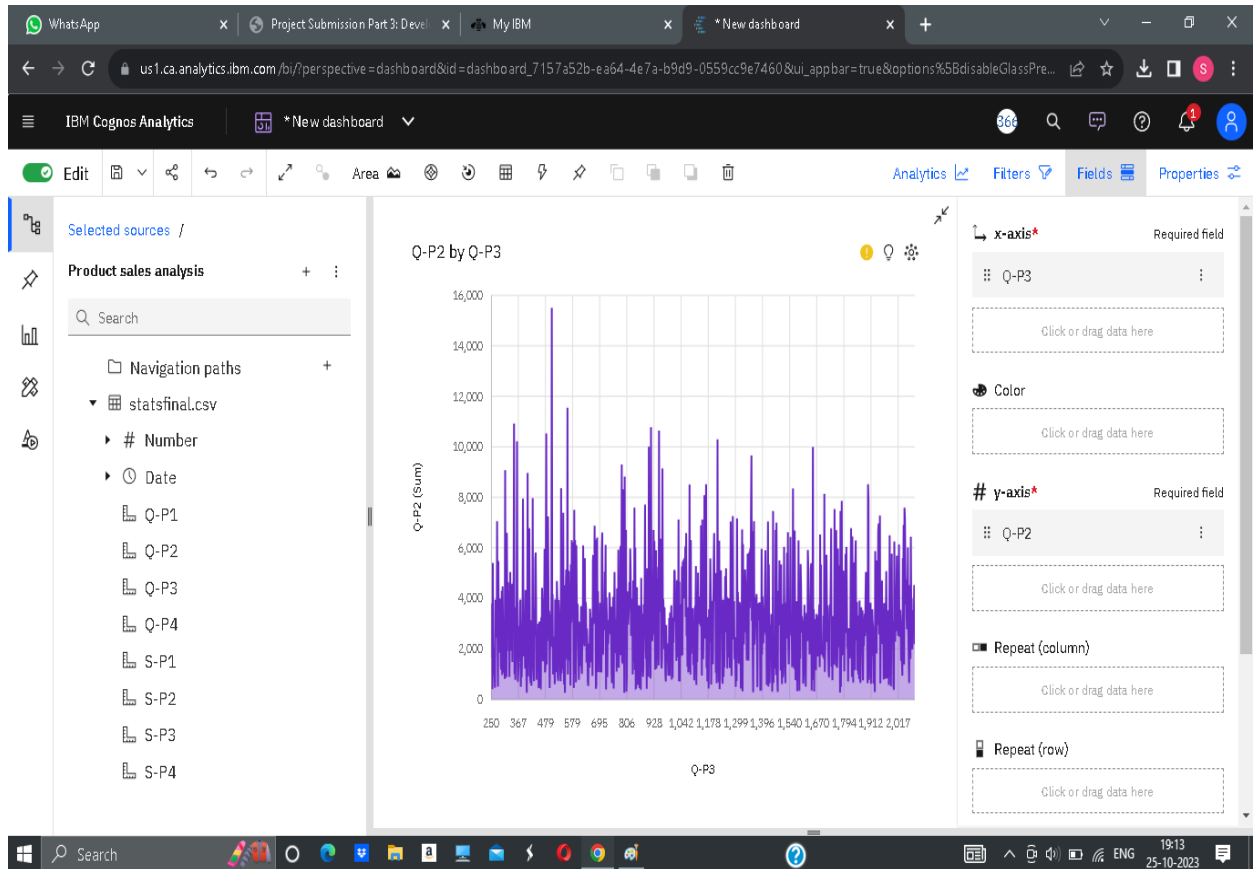
In [ ]:

## Visualization using IBM Cognos(Part-1):



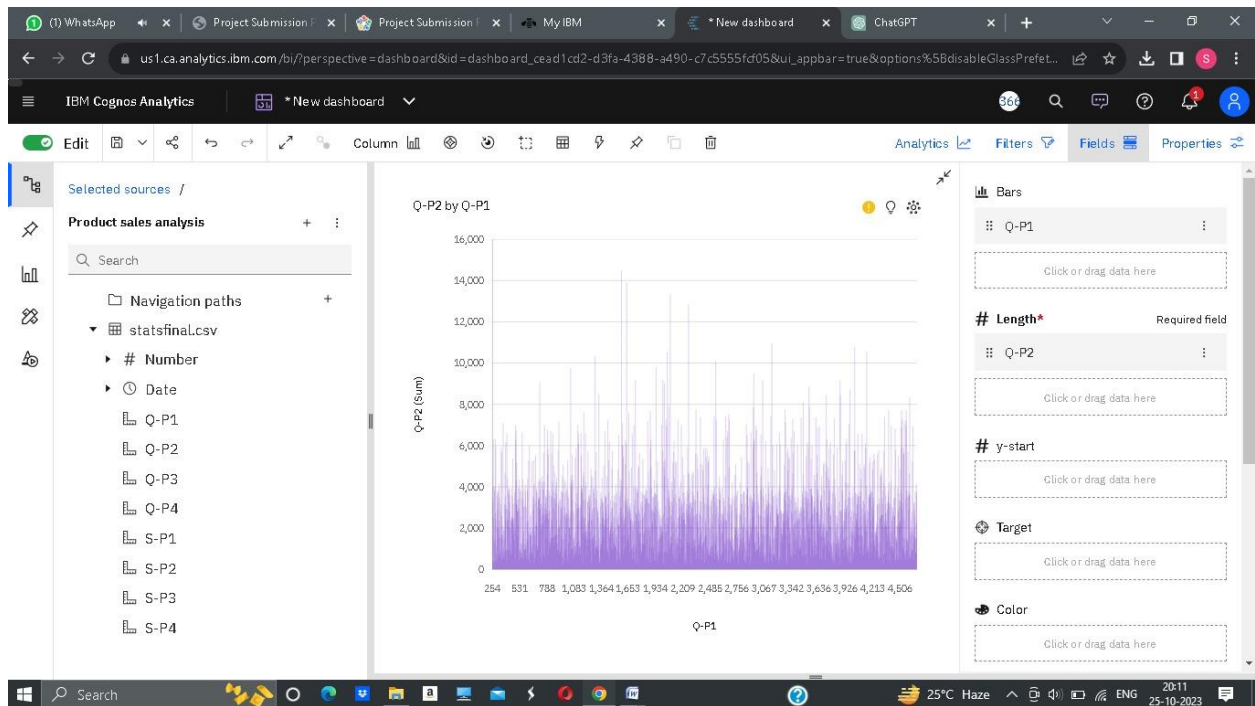


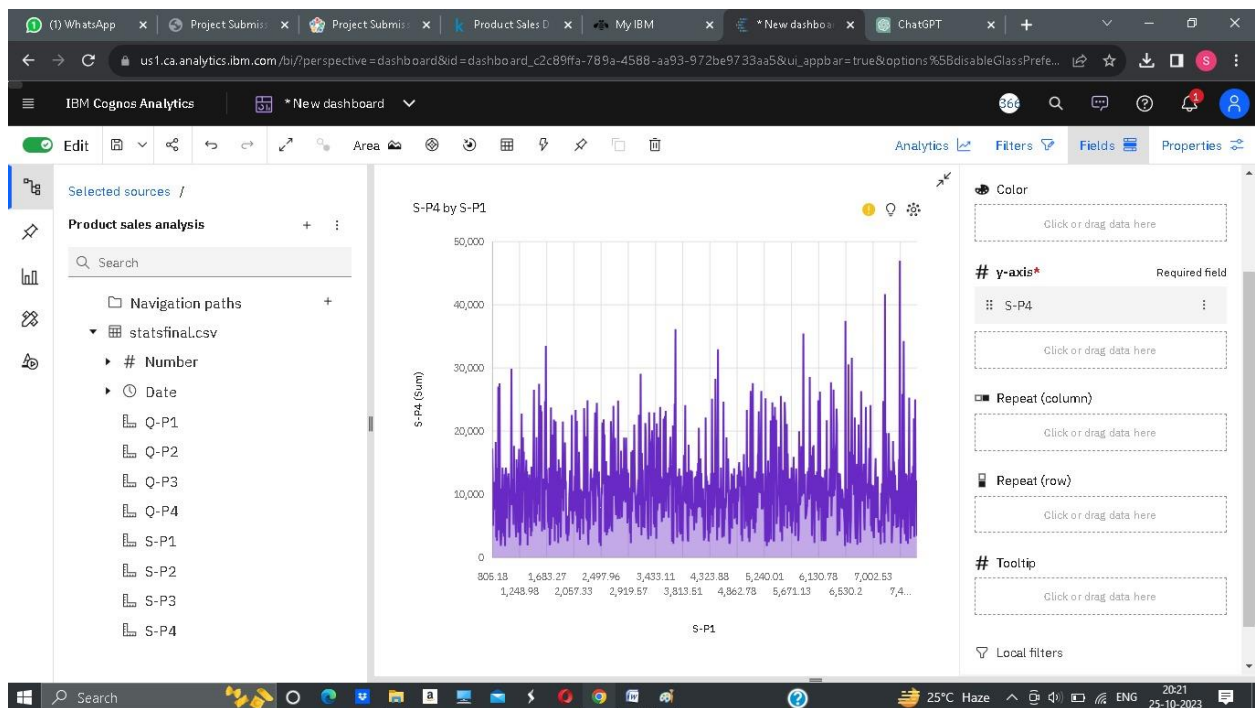
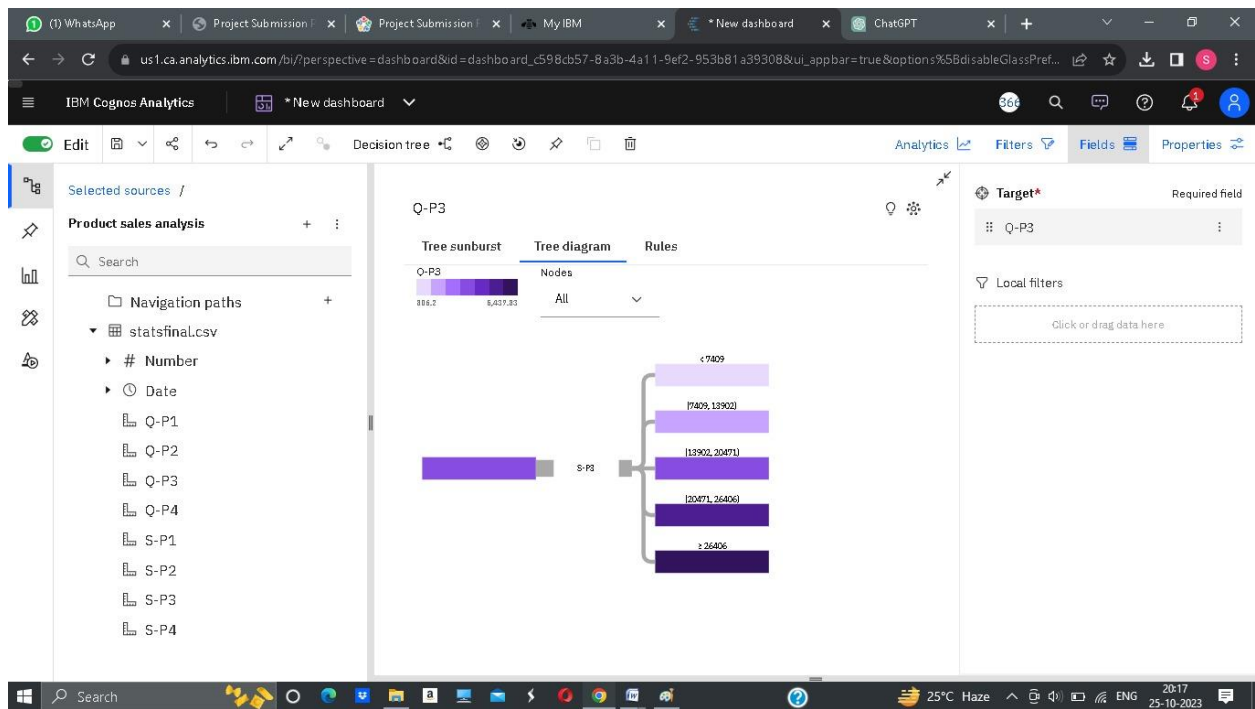


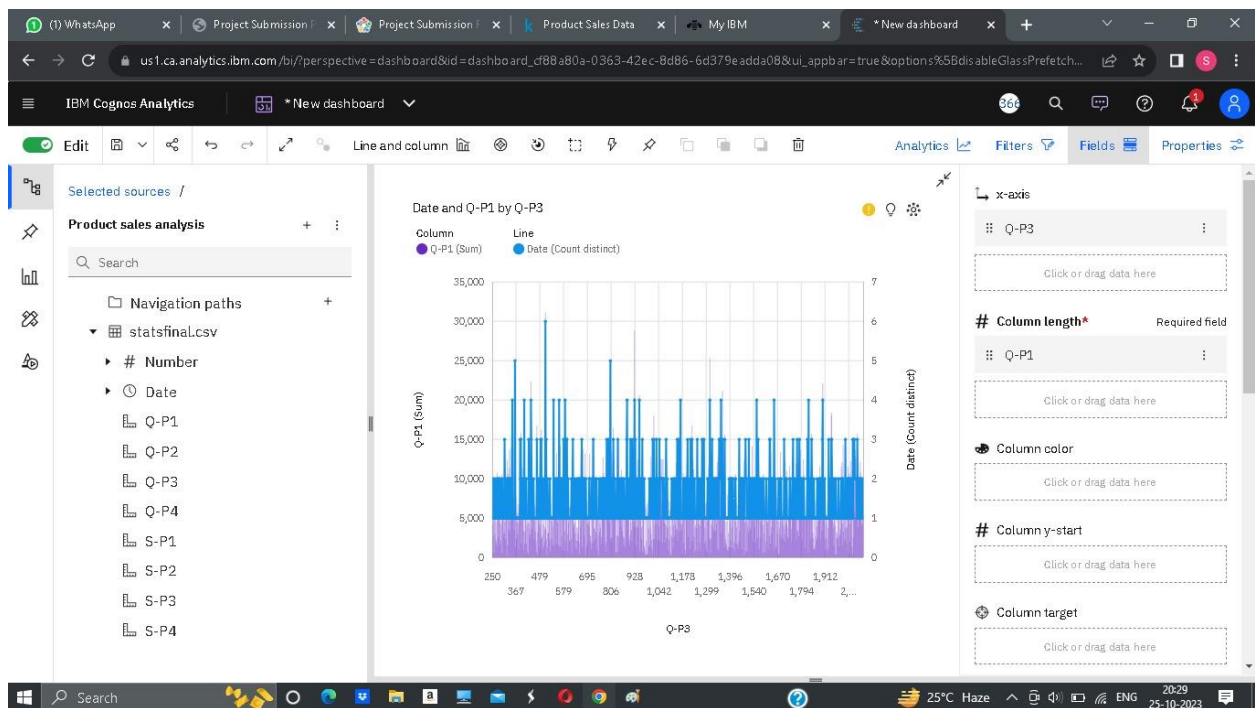
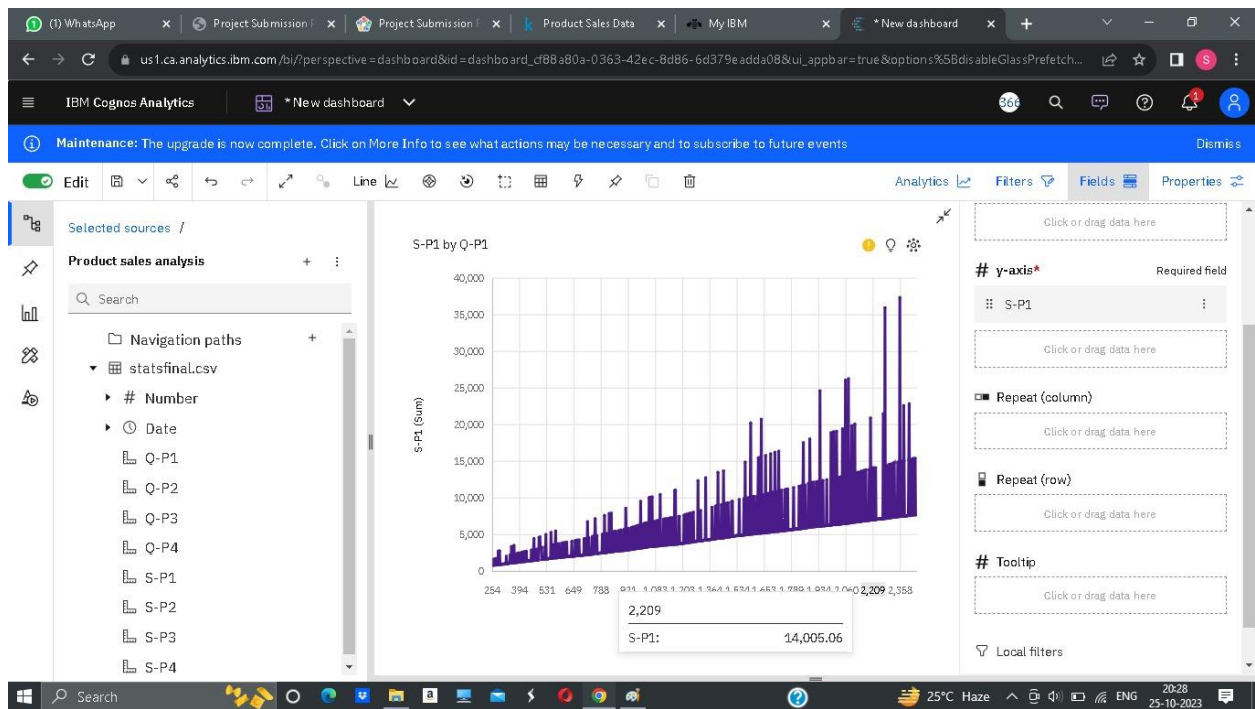


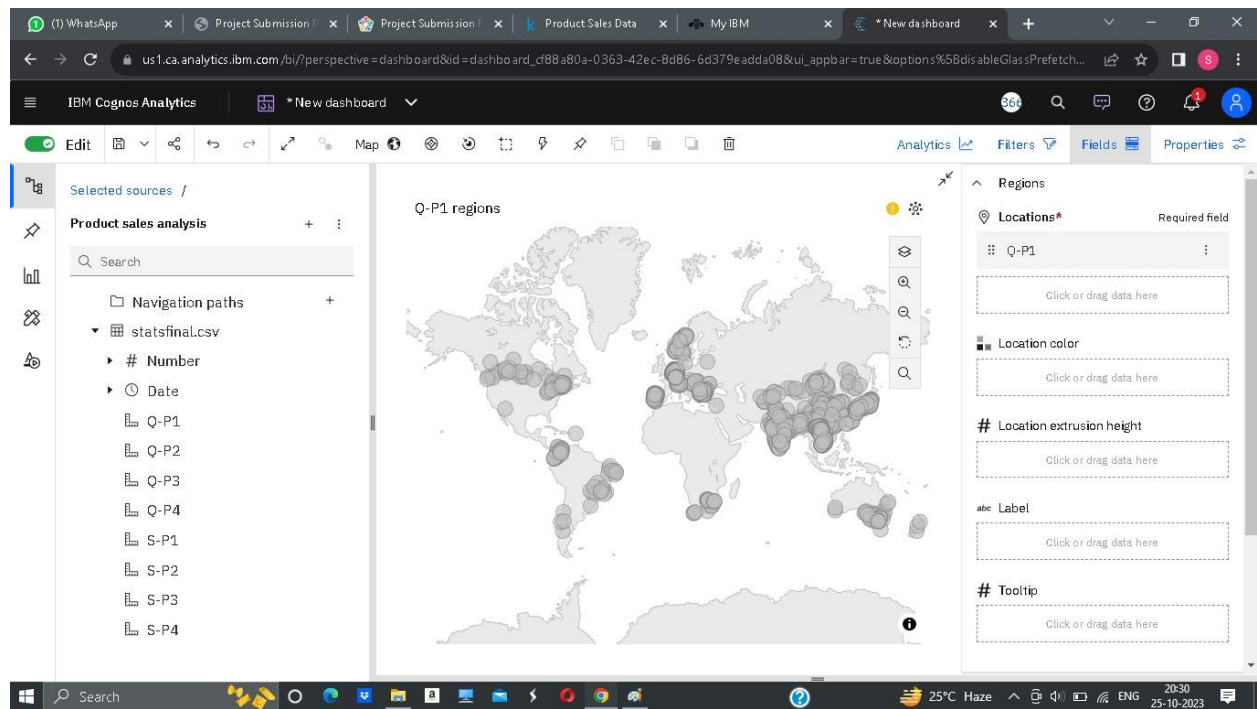


## Visualizations using IBM Cognos(Part-2):









## Required Tools & Libraries:

- IBM Cognos: For creating interactive dashboards and reports.
- Python (optional): For data preprocessing and analysis.
- Python (programming language)
- Jupyter Notebook (optional ,but recommended for code presentation)
- Pandas :For data manipulation and analysis.
- Matplotlib and Seaborn : For data visualization.
- Scikit-Learn: For machine learning if advanced analysis required.

## Conclusion:

The "Product Sales Analysis" project delivers valuable insights and actionable recommendations to optimize sales strategies, improve product performance, and increase profitability. With accurate sales trend predictions, a focus on top-performing products, and a deeper understanding of marketing ROI and seasonality, the project equips the organization to make data-driven decisions and achieve business growth.