**PHASE 5**

**SUBMISSION**

1. **Share the GitHub repository link containing the project's code and files:**

**Link:**

**B. Provide instructions on how to replicate the analysis and generate visualizations using IBM Cognos:**

Replicating the analysis and generating visualizations using IBM Cognos involves several steps. Below are the instructions on how to perform this analysis:

**1. Data Source Connection:**

- Ensuring that we have access to our dataset in IBM Cognos. Connecting to our data source, where you have columns for four unit columns, four revenue columns and a date column.

**2. Data Preparation:**

- Checking our data is clean and structured. Verify that our date column is in the proper date format. If necessary, transform or clean the data to remove duplicates or inconsistencies.

**3. Create a New Report:**

- Opening IBM Cognos and creating a new report to begin our analysis.

**4. Select Data:**

- Choosing the appropriate data source and selecting the relevant data columns:

- Date: Place this on the X-axis of your visualizations.

- Unit Columns (Q-P1,Q-P2,Q-P3,Q-P4): These will be used to group and segment our data.

- Revenue Columns (S-P1,S-P2,S-P3,S-P4): These will be used for our measures.

**5. Build Visualizations:**

- Creating the visualizations that we need to analyze our data. For each analysis task:

*Examples:*

**a. Top-Selling Products:**

- Using a bar chart or table.

- Placing the product (unit) on the X-axis and total revenue on the Y-axis.

- Sorting the products by revenue in descending order.

**b. Peak Sales Periods:**

- Creating a line chart or time series chart.

- Placing the date on the X-axis and total revenue on the Y-axis.

- Adjust the chart settings to emphasize peaks, if necessary.

**c. Customer Preferences:**

- Utilize pie charts, bar charts, or crosstabs.

- Displaying customer preferences by unit.

- Allowing users to filter by customer segments or specific products.

**6. Filters and Parameters:**

- Implementing filters and parameters to allow users to customize their analysis. Creating parameters for date ranges, units, and other relevant dimensions, as needed.

**7. Aggregations and Calculations:**

- Calculating total revenue for each unit or product and any other metrics (e.g., average revenue per unit) using IBM Cognos functions.

**8. Sorting and Grouping:**

- Grouping the data by date, product, or unit to provide insights based on the user's preferences. Adjust sorting and grouping based on our analysis objectives.

**9. Interactive Elements:**

- Enabling interactivity in our visualizations, allowing users to drill down, filter, and explore the data further.

**10. Security and Permissions:**

- Setting up user permissions and access control to ensure data security and restrict access to sensitive information.

**11. Testing and Validation:**

- Thoroughly testing our reports and visualizations to ensure accuracy and user-friendliness.

**12. User Training:**

- Providing training to end-users on how to access, filter, and interpret the reports effectively.

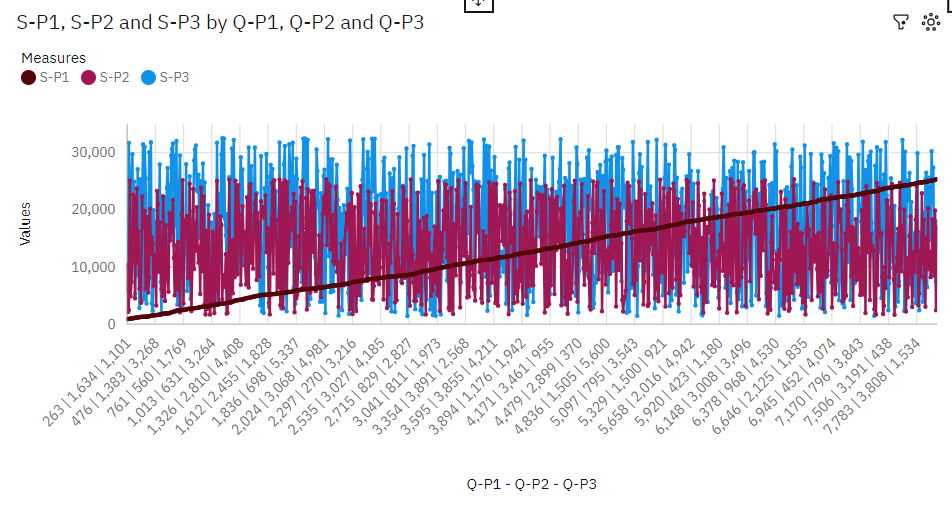
**13. Documentation:**

- Creating a documentation detailing data sources, report definitions, and instructions for users to reference.

**14. Maintenance and Updates:**

- Regularly updating the reports to reflect changing data and evolving business needs.

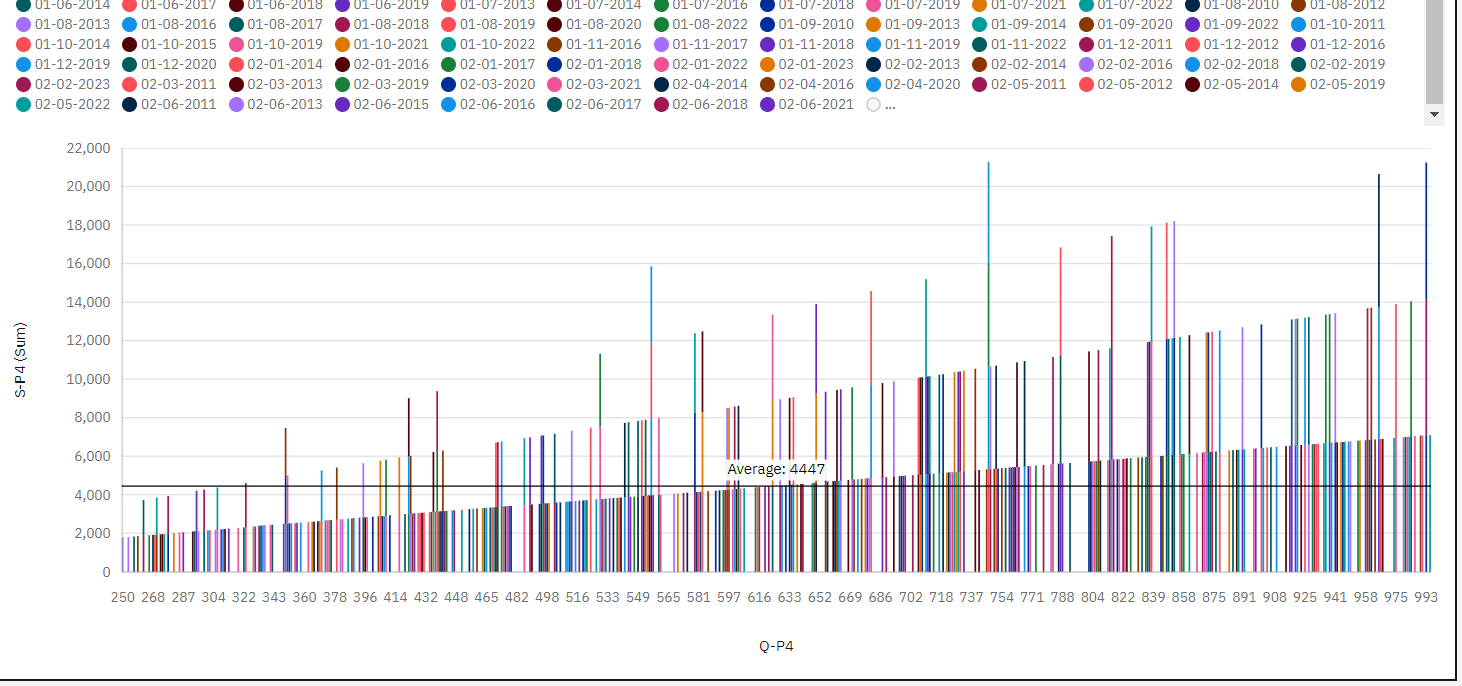
1. **Include example outputs of the visualizations and derived insights :**
2. **Line :**

****

**Insights:**

* Q-P1 7910 has the highest total S-P1 due to Q-P2 1290.
* Q-P3 5703 has the highest values of both S-P1 and Q-P4.
* 7910 S-P1 at over 50 thousand is 97% higher than the Q-P4 of over 1500.
* 3407 S-P1 at over 67 thousand is 96% higher than the Q-P4 of over 2500
* 5703 S-P1 at over 54 thousand is 93% higher than the Q-P4 of nearly 4 thousand.
* S-P1 and Q-P4 diverged the most when Q-P1 is 7826, and when S-P1 was nearly 49 thousand higher than the Q-P4.
* S-P1 and Q-P4 diverged the most when Q-P2 is 3407, and when S-P1 was nearly 65 thousand higher than the Q-P4.
* S-P1 and Q-P4 diverged the most when Q-P3 is 5703, and when S-P1 was nearly 51 thousand higher than the Q-P4.
* Q-P1 1846 has the highest Total Q-P4 but is ranked #471 in Total S-P1.
* Q-P2 3407 has the highest Total S-P1 but is ranked #37 in Total Q-P4.
* Q-P2 3855 has the highest Total Q-P4 but is ranked #22 in Total S-P1.
* Q-P2 3407 has the highest S-P1 at over 67 thousand, out of which Q-P1 7501 contributed the most at almost 24 thousand.
* 7996 has a S-P1 of over 25 thousand for Q-P2 2672.
* 4430|1942|3211 (0.2 %), 4351|276|2337 (0.2 %), 5704|808|891 (0.2 %), 3941|3098|282 (0.2 %), and 3855|1015|1746 (0.2 %) are the most frequently occurring categories of Q-P1 - Q-P2 - Q-P3 with a combined count of 10 items with S-P1 values (1.1 % of the total) .
* 4430|1942|3211 (0.2 %), 4351|276|2337 (0.2 %), 5704|808|891 (0.2 %), 3941|3098|282 (0.2 %), and 3855|1015|1746 (0.2 %) are the most frequently occurring categories of Q-P1 - Q-P2 - Q-P3 with a combined count of 10 items with S-P2 values (1.1 % of the total) .

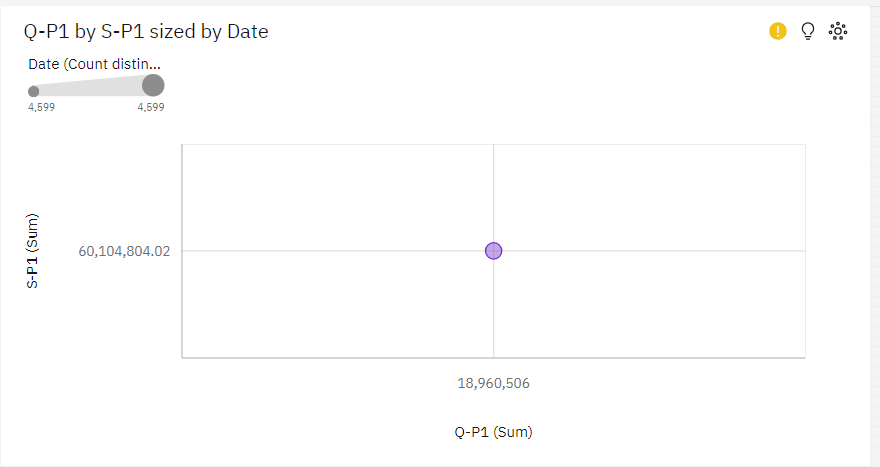
1. **Stacked column:**

****

**Insights:**

* Q-P4 934 has the highest total S-P4 due to Date 12-08-2014.
* 25-12-2018 Q-P1 at nearly eight thousand is 61% higher than the S-P4 of over three thousand.
* 257 Q-P1 at over 17 thousand is 68% higher than the S-P4 of nearly 5500.

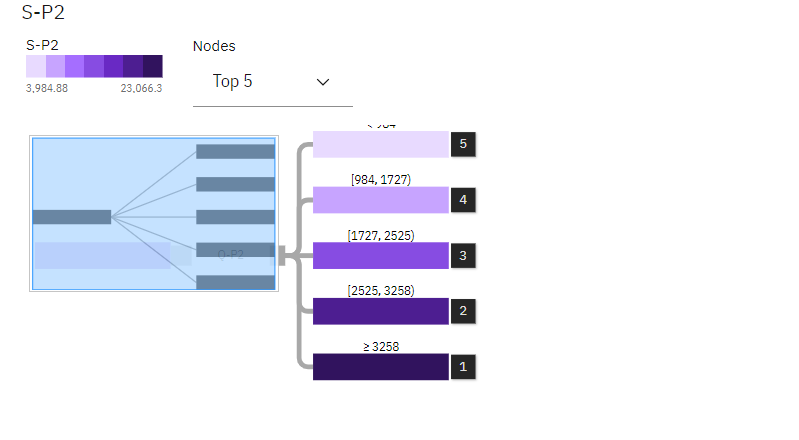
**3.Bubble:**



**Insights:**

* The total number of results for S-P1, across all Q-P1, is over 4500.

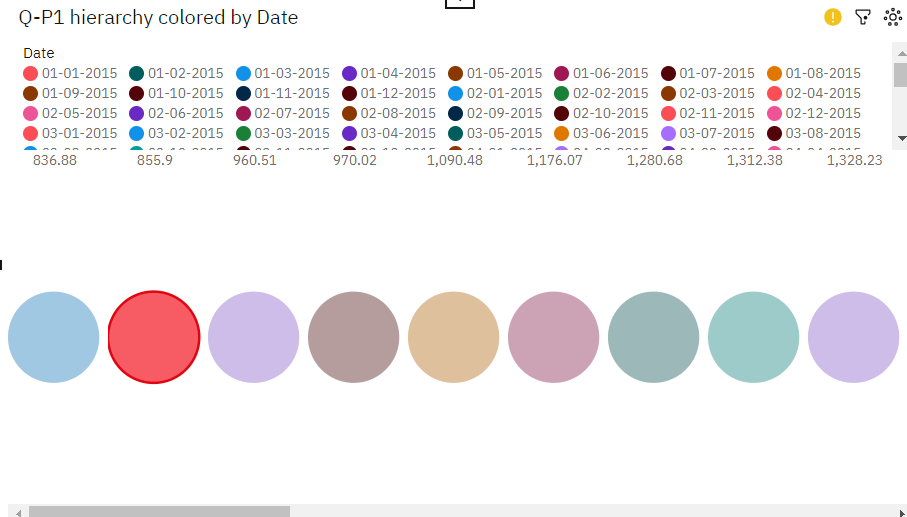
**4.Decision Tree:**



**Insights:**

* Q-P2 Strongly derives S-P2(96.1%)

**5.Hierachy Bubble:**

****

**Insights:**

* Date 31-9-2015 has the highest Unaggregated Q-P1 but is ranked #125 in Total Q-P4.
* Date 27-03-2015 has the highest Total Q-P4 but is ranked #49 in Unaggregated Q-P1.