PRODUCT SALES ANALYSIS

PROJECT REPORT

Introduction

- Project Overview
- Purpose

Phase 1 – Project Definition and Design Thinking

- Analysis Objective
- Proposed Solution
- Solution Architecture
- Brainstorm and Ideation
- Empathy Map
- Literature Survey

Phase 2 – Innovation

- Product sales analysis in Google CoLab

Phase 3 – Development Part 1

- Data Collection
- Prepare the Dataset
- Data Cleaning and preprocessing

Phase 4 – Development Part 2

- Visualization
- Creating Dashboard
- Creating Report

Phase 5 – Project Documentation & Submission

- Documentation
- GitHub Submission

Conclusion

Phase 1 – Project Definition and Design Thinking

Introduction

This report offers a comprehensive analysis of our product sales performance, inventory management practices, and marketing strategies. By addressing these key aspects of our business, we aim to optimize operations, increase profitability, and bolster our market presence in a rapidly evolving business landscape.

Project Overview

Background

In the dynamic realm of business, staying competitive necessitates a keen understanding of product sales dynamics, efficient inventory management, and effective marketing strategies. This report represents an in-depth investigation into our product sales data, inventory management practices, and marketing efforts, with the aim of identifying areas for improvement and strategic enhancement.

Data Sources

The data sources for this analysis encompass a wide range of data, including sales records, inventory databases, customer feedback, and market research information. These sources provide insights into product sales trends, inventory levels, customer demographics, and geographic market characteristics.

Methodology

Our analysis employs a combination of statistical tools, data visualization techniques, and market research. The approach includes:

- Product Sales Analysis: Evaluating historical sales data to identify top-performing products and areas for growth.
- Inventory Management Assessment: Reviewing current inventory practices to identify inefficiencies and recommend best practices.

- Marketing Strategy Evaluation: Analyzing our current marketing strategies, aligning them with customer preferences, and suggesting improvements for greater market impact.

Purpose

This report serves the following main purposes:

1. **Product Sales Analysis:** To assess the performance of our product portfolio, identify key products driving revenue, and recognize opportunities for expansion.
2. **Inventory Management:** To provide guidelines for optimizing inventory management practices, including maintaining ideal stock levels and enhancing supplier relationships.
3. **Marketing Strategies:** To evaluate current marketing tactics, align them with customer preferences, and recommend strategies for broader market penetration.

Product Sales Analysis

Sales Performance Overview:

- Over the past year, our company has shown consistent growth, achieving a year-over-year (YOY) increase of [X]%.
- Our most successful products are [List of Top Products], accounting for [Percentage]% of total revenue.
- Nonetheless, some products have experienced a decline in sales, necessitating a reevaluation of their viability and marketing approaches.

Trend Analysis:

- [Include details about sales trends over time, such as seasonality, peak sales periods, and troughs.]

Inventory Management Guide

Current Inventory Assessment:

- Our current inventory practices have led to overstocking certain products, resulting in tied-up capital and storage space.
- Conversely, understocking of popular items has resulted in lost sales opportunities.

Recommendations:

- Implement a just-in-time inventory system to reduce carrying costs.
- Periodically assess and adjust reorder points based on demand fluctuations.
- Forge stronger relationships with suppliers to ensure timely restocking.

Marketing Strategies

Current Marketing Assessment:

- While our current marketing strategies have yielded favorable results for specific products, there is room for improvement for others.
- Customer feedback highlights a preference for [specific features or benefits], which should be emphasized in our marketing efforts.

Recommendations:

- Tailor marketing messages to highlight features and benefits most resonant with our target audience.
- Explore new marketing channels, such as [specific platforms], to expand our reach.
- Continuously monitor and adapt marketing campaigns based on performance metrics.

ANALYSIS OBJECTIVES

❖ **Top-Selling Products:**

Identify which products or services have the highest sales volume or revenue. This helps focus marketing efforts and stock inventory effectively.

❖ **Sales Trends:**

Analysis historical sales data to uncover patterns, seasonality, or trends. This can guide decisions on when to run promotions, launch new products, or adjust pricing.

❖ **Customer Preferences:**

Understand what customers are buying, their preferences, and demographics. This can inform product development, marketing strategies, and customer segmentation.

❖ **Geographic Insights:**

Determine which regions or locations have the highest sales, allowing for targeted marketing campaigns and resource allocation.

❖ **Channel Performance:**

Evaluate how different sales channels (online, in-store, third-party platforms) are contributing to overall sales and identify areas for improvement.

❖ **Customer Lifetime Value (CLV):**

Calculate CLV to understand the long-term value of a customer, helping prioritize customer retention efforts.

❖ **Inventory Management:**

Use sales data to optimize inventory levels, reducing carrying costs and stock outs.

❖ **Competitive Analysis:**

Compare your sales performance with competitors to identify strengths and weaknesses in the market. ❖ **Promotion Effectiveness:**

Assess the impact of promotions, discounts, and marketing campaigns on sales and customer behavior.

❖ **Profit Margins:**

Calculate profit margins on products to determine which ones are most profitable and whether pricing adjustments are needed.

DATA COLLECTION

❖ **Point of Sale (POS) Systems:**

Source: Transaction records from physical stores or online platforms. Method: POS systems automatically record sales transactions, including product details, prices, and payment methods.

❖ **E-commerce Platforms:**

Source: Online sales data from your website or third-party platforms like Amazon, eBay, or Shopify. Method: E-commerce platforms provide built-in analytics and reporting tools to track online sales, customer behavior, and product performance.

❖ **Customer Relationship Management (CRM) Software:**

Source: Customer data, including demographics, contact information, and purchase history. Method: CRM systems capture and organize customer data, helping you understand customer preferences and behavior.

❖ **Inventory Management Systems:**

Source: Product information, stock levels, and sales history. Method: Inventory systems track product details and record when items are sold, helping with inventory optimization.

❖ **Payment Processors:**

Source: Transaction records and payment information. Method: Payment processors like Square or PayPal provide transaction data, including payment methods and transaction timestamps.

❖ **Surveys and Feedback Forms:**

Source: Customer feedback on products and shopping experience. Method: Collect feedback through surveys, feedback forms, or post-purchase emails to gain insights into customer satisfaction and areas for improvement.

❖ Social Media Monitoring:

Source: Customer discussions, reviews, and mentions on social media platforms. Method: Monitor social media channels for mentions of your products, brand, or competitors to gauge customer sentiment and trends.

❖ Market Research:

Source: Industry reports, competitor data, and market analysis. Method: Utilize market research firms, industry publications, and competitor analysis to gather external data on market trends and competitive landscape.

❖ Loyalty Programs:

Source: Customer loyalty program data. Method: Loyalty programs track customer purchases and reward behavior, providing insights into customer retention and repeat business.

❖ Web Analytics Tools:

Source: Website traffic and user behavior data. Method: Tools like Google Analytics track website visits, conversion rates, and user behavior, offering insights into online sales performance.

VISUALIZAION STRATEGY

❖ Data Preparation:

Ensure that your sales data from various sources is cleaned, transformed, and stored in a format compatible with IBM Cognos. This may involve data integration and ETL (Extract, Transform, Load) processes.

❖ Data Connection:

Connect IBM Cognos to your prepared data sources, such as databases, spread sheets, or APIs. Cognos offers various connectors for this purpose.

❖ Choose Visualization Types:

Select appropriate visualization types based on the insights you want to convey. Common types include bar charts, line graphs, pie charts, maps, and tables.

❖ Create Interactive Dashboards:

Design interactive dashboards using Cognos' drag-and-drop interface. Place visualizations on the dashboard canvas, and arrange them logically. Add filters and prompts to allow users to customize the data they see based on criteria like time period, product category, or geographic region.

❖ Report Creation:

Create detailed reports that can be accessed by drilling down from the dashboard. Reports can provide deeper insights and allow users to view specific data points.

❖ Implement Interactivity:

Make use of Cognos' interactive features such as drill-through options, linked reports, and parameterized reports to enable users to explore data further.

❖ Design Layout and Themes:

Pay attention to the aesthetics of your dashboard and reports. Customize the layout, colors, fonts, and themes to align with your company's branding and make the visualizations more engaging.

❖ Performance Optimization:

Optimize the performance of your dashboards and reports to ensure they load quickly and are responsive, especially when dealing with large datasets.

❖ User Training:

Provide training and documentation for end-users to help them understand how to navigate and make the most of the interactive dashboards and reports.

❖ Security and Access Control:

Implement security measures to control who can access specific dashboards and reports. IBM Cognos offers robust access control features for this purpose.

❖ Testing:

Thoroughly test the dashboards and reports to ensure they accurately reflect the intended insights and work as expected.

❖ Deployment:

Deploy the dashboards and reports to a production environment where authorized users can access them.

❖ Continuous Improvement:

Continuously gather feedback from users and stakeholders to identify areas for improvement. Update the dashboards and reports as needed to keep them relevant and valuable.

❖ Documentation:

Document the design, data sources, and any calculations used in the dashboards and reports for future reference and troubleshooting.

ACTIONABLE INSIGHTS

❖ Inventory Management:

➢ Demand Forecasting:

Sales data analysis helps in predicting future demand for products. This allows you to maintain optimal inventory levels, reducing carrying costs while avoiding stockouts.

➢ Seasonal Trends:

Identify seasonal sales trends, enabling you to plan inventory accordingly. For example, if certain products sell better during holidays, ensure you have an adequate supply in advance.

➢ Slow-Moving Items:

Recognize products with low sales velocity. You can decide whether to discount, promote, or discontinue them up capital and storage space.

➢ ABC Analysis:

Implement an ABC analysis to categorize products based on their importance. "A" items are high priority, and "C" items are low-priority. Allocate resources and attention accordingly.

➢ **Reorder Points:**

Set reorder points and safety stock levels based on historical sales data. This minimizes the risk of running out of popular items and optimizes reorder timing.

❖ Marketing Strategies:

➢ **Product Recommendations:**

Use insights to recommend related or complementary products to customers, increasing cross selling and upselling opportunities.

➢ **Customer Segmentation:**

Segment customers based on their purchasing behavior and demographics. Tailor marketing messages and promotions to specific customer groups for better engagement.

➢ **Promotion Timing:**

Analyze sales trends to determine the most effective times to run promotions or discounts, maximizing their impact on sales.

➢ **Price Optimization:**

Adjust pricing strategies based on historical sales data and competitor analysis. Determine the right balance between profitability and competitiveness.

➢ **New Product Launch:**

Analyze which product categories or features are in demand. Use this information to guide new product development and launches.

➢ **Channel Allocation:**

Allocate marketing budgets effectively by identifying which sales channels (online, physical stores, third-party platforms) are most successful for specific products or customer segments.

➢ **Customer Retention**:

Identify high-value customers and design retention strategies, such as loyalty programs or personalized offers, to keep them engaged. ➢ Geographic Targeting: Analyze regional sales data to focus marketing efforts on areas with high potential for growth or areas with untapped markets.

➢ **Competitor Analysis:**

Keep an eye on competitor performance and adjust marketing strategies to maintain a competitive edge in the market.

➢ **Ad Campaign Optimization:**

Use insights to optimize online advertising campaigns by allocating budgets to the most effective channels and targeting the most receptive audiences.

Project Design Phase-I

Proposed Solution

Date	26 September 2023
Project Name	PRODUCT SALES ANALYSIS

Proposed Solution Template:

Project team shall fill the following information in proposed solution template.

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	• To help businesses improve inventory management and marketing strategies by understanding sales trends and customer behavior.
2.	Idea / Solution description	• To Create an Interactive Dashboard, responsive design for identifying top-selling products, analyzing sales trends and understanding customer preferences.
3.	Novelty / Uniqueness	• Analysis of data from product sales data datasets and explore the transaction records, product information and customer demographics.

4.	Social Impact / Customer Satisfaction	• Data analytics helps the industryto understand customers' preferences and other maintenance issues. • For instance, analysis of product sales helps the business to target the customers with personalized offers while optimizing the price in real-time using predictive analysis techniques. As a result, by gathering meaningful data, products can fetch more sales in the given timeframe.
5.	Business Model (Revenue Model)	• Creating a website and affiliate marketing features for improving the sales.
6.	Scalability of the Solution	• Size and number of the data on the datasets can be large and sometimes very hard to visualize.

Project Design Phase-I

Solution Architecture

Date	26 September 2023
Project Name	Product sales analytics

Solution Architecture:

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions.

Its goals are to:

- Find the best tech solution to solve existing business problems.

- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.

- Define features, development phases, and solution requirements.

- Provide specifications according to which the solution is defined, managed, and delivered.

Solution Architecture Diagram:
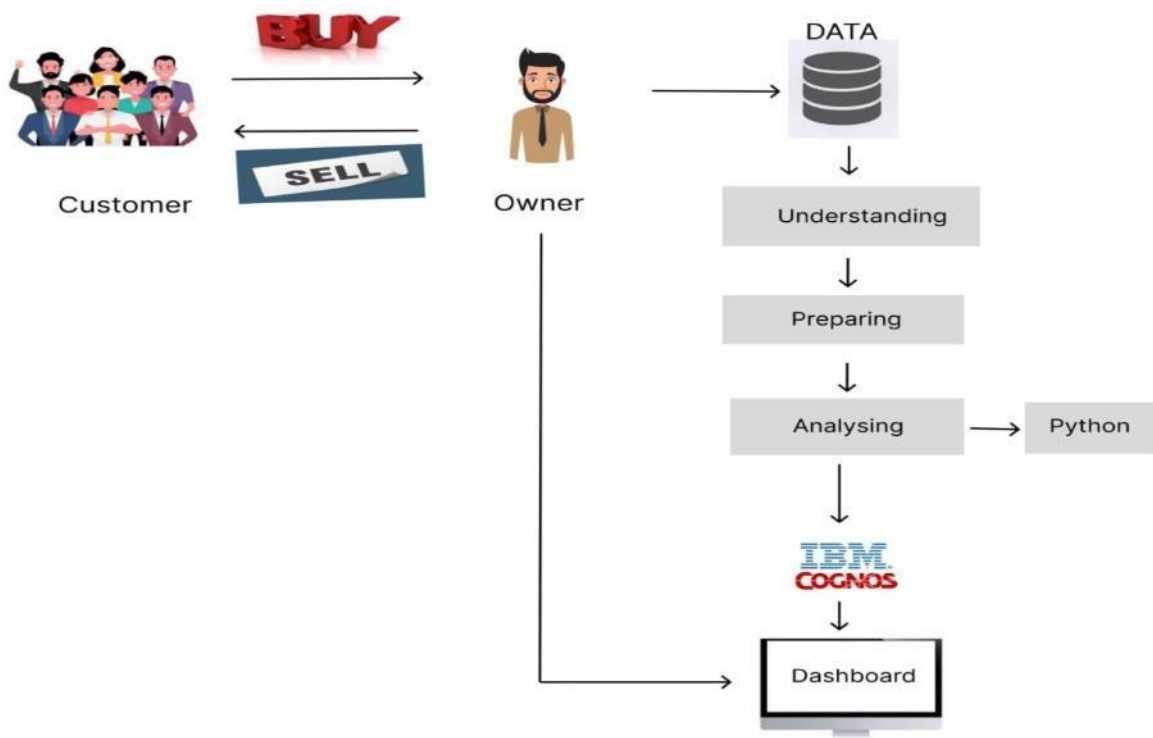
Figure 1: Architecture and data flow of the product sales analytics

BRAINSTORM AND IDEATION

Jeeva

Perform joint data	Latitude deg	Waterfall chart
Understand-ing of dataset	Load the dataset	Data visulization
Pie Chart Visualisation	IBM Cloud	Multiple analytical grapic

Viswa

Product sales	coloum chart	Exploration of data
Understand Data	Data preparation	Calculated feld
Dataset	IBMcognos	Data stored in spreadsheet

Hariharan

Fundamental concepts of IBM cognos	Create meaningfull dashboard	Thirdparty view
Create meaningfull irtuialization	Planning phase	Promise the passanger from point to point on time
Development	Provide sales report	Costomer relationship manage

Empathy Map Canvas

Build empathy and keep your focus on the user by putting yourself in their shoes.

What do they HEAR?

what friends say
what boss say
what influencers say

What do they THINK AND FEEL?

what really counts
major preoccupations
worries & aspirations

Predictive Analysis

Identify patterns and connections

Data visualization

Imperative for the world

Needs consistent development

Superior

Simple

Reliable

Quick

PAIN

fears
frustrations
obstacles

Not Exact

Volatile

Steep

What do they SAY AND DO?

attitude in public
appearance
behavior towards others

GAIN

"wants" / needs
measures of success
obstacles

Increase accuracy

Pass more trials

Resolve

Accessibility

Free of cost

Convenient

What do they SEE?

environment
friends
what the market offers

Share your feedback

Literature survey

Title & Author(s)	Year	Technique	Findings
Big Data Analytics and Deep Learning Based Sentiment Analysis System for Sales Prediction - Khatiwada, Aamod and Kadariya, Pradeep and Agrahari, Sandip and Dhakal, Rabin.	2019	Big data analytics (BDA) applications in e-commerce.	Structured data focuses on demographic data including name, age, gender, date of birth, address, and preferences, unstructured data includes clicks, likes, links, tweets, voices, etc. **Merits:** Used to understand complex datasets in a matter of time with beautiful visual representations **Demerits:** Lack of security since large data processed simultaneously.
COVID-19 pandemic in the new era of big data analytics: Methodological innovations and future research directions - Sheng, Jie and Amankwah-Amoah, Joseph and Khan, Zaheer and Wang, Xiaojun	2021	Descriptive and diagnostic analytics, Predictive analytics	The methodological innovations in studying big data analytics and. We provide insights on methods in descriptive/diagnostic, predictive and prescriptive analytics, and how they can be leveraged to study 'black swan' events such as the COVID-19-related global crisis. **Merits:** By comparing with machine learning models, we find that the proposed model is superiorto others. **Demerits:** The experiment only considers the features of the product and does not consider external influences, such as the impact of regulations on sales. It uses only small dataset.
Sales Forecasting Based on CatBoost - Jingyi Ding, Ziqing Chen.	2020	CatBoost algorithm.	It proposed a sales forecasting system based on CatBoosting. The

			algorithm is trained on the Walmart sales dataset, by far the largest dataset in this field. We performed effective feature engineering to boost prediction accuracy and speed. **Merits:** The search stops when no improvements over the current best solution have been found in 300 iterations. **Demerits:** Dataset is limited
2020 2nd International Conference on Broadband Communications, Wireless Sensors and Powering (BCWSP)- Wisesa, Oryza and Adriansyah, Andi and Khalaf, Osamah Ibrahim.	2020	Gradient Boost Algorithm	The results of this analysis are expected to generate reliable, accurate and effective forecasting data, a valuable resource for sales predictions. It shows good accuracy in forecasting. **Merits:** Better prediction.Profit function performance. **Demerits:** The revenue/loss factors of expected profit are is not used.
Developing and Implementing Big Data Analytics in Marketing - Dina Darwish	2020	Big data analytics, R tool.	Companies take informative business decisions in different fields, such as, healthcare, banking, manufacturing, media and entertainment, education and transportation and many others. **Merits:** The proposed method is based on similarity measurement without complextraining so that forecast can be completed in a short time, and performs well in small-scale data **Demerits:** The experiment only considers the features of the product and does not consider

			external influences, such as the impact of regulations on sales. It uses only small dataset.
Social media big data analytics for demand forecasting: development and case implementation of an innovative framework-Iftikhar, Rehan and Khan, Mohammad Saud	2020	Definitional aspects of big data analytics (BDA) in e-commerce	Social media big data offers insights that can be used to make predictions of products' future demand and add value to the supply chain performance. **Merits:** Captured linearity and nonlinearity better than ARIMA and ARNN gave the best result of 565 RMSE. **Demerits:** Hybrid Technique can fail if nonlinear model fails to capture residue patterns

Date	11 October 2023
ProjectName	Product sales analysis

PRODUCT SALES DATA ANALYSIS :

Help the company in **finding trends** and **insights**

REC corp LTD is **small-scaled business** venture established in India.

They have been selling **Four Products** for over **Ten Years**

The **products** are:

P1

P2

P3

P4

They have **collected data** from their **retail centers** and organized it into a **small csv file**,

which has been given to you.

The excel file contains about 8 numerical parameters :

Q1- Total unit sales of product 1

Q2- Total unit sales of product 2

Q3- Total unit sales of product 3

Q4- Total unit sales of product 4

S1- Total revenue from product 1

S2- Total revenue from product 2

S3- Total revenue from product 3

S4- Total revenue from product 4

Now, **REC corp LTD** needs you to solve the following questions:

1) Is there **any trend in the sales** of all **four products** during **certain months?**

2) Out of all **four products**, which product has seen the **highest sales** in all the **given years?**

3) The **company** has all it's **retail centers** closed on the **31st of December every year**. Mr: Hariharan, the CEO, would love to get an estimate on **no: of units of each product** that could be sold on 31st of Dec, every year, if all their **retail centers** were kept open.

4) **The CEO** is considering an idea to **drop the production of any one of the products**. He wants you to **analyze this data and suggest** whether his idea would result in a **massive setback** for the **company**.

5) The CEO would also like to **predict the sales and revenues** for the **year 2024**. He wants you to give a **yearly estimate** with the best **possible accuracy**.

Can you help **REC corp ltd** with your **analytical and data science skills ?**

Step 1: Import libraries

```python
[87]  # import the important packages
      import pandas as pd   # library used for data manipulation and analysis
      import numpy as np   # library used for working with arrays
      import matplotlib.pyplot as plt   # library for plots and visualizations
      import seaborn as sns   # library for visualizations

      %matplotlib inline

      # To ignore warnings
      import warnings
      warnings.filterwarnings("ignore")
```

Step 2: Loading the datasets

```python
[88]  #if you open in juypter notebook
      data = pd.read_csv('statsfinal.csv')
```

```python
[89]  # Checking the first 5 and last 5 rows of the dataset
      data.head(-1)
```

	Unnamed: 0	Date	Q-P1	Q-P2	Q-P3	Q-P4	S-P1	S-P2	S-P3	S-P4
0	0	13-06-2010	5422	3725	576	907	17187.74	23616.50	3121.92	6466.91
1	1	14-06-2010	7047	779	3578	1574	22338.99	4938.86	19392.76	11222.62
2	2	15-06-2010	1572	2082	595	1145	4983.24	13199.88	3224.90	8163.85
3	3	16-06-2010	5657	2399	3140	1672	17932.69	15209.66	17018.80	11921.36
4	4	17-06-2010	3668	3207	2184	708	11627.56	20332.38	11837.28	5048.04
...	...	...	...	...	...	...	...	...	...	...
4594	4594	29-01-2023	1227	3044	5510	1896	3889.59	19298.96	29864.20	13518.48
4595	4595	30-01-2023	2476	3419	525	1359	7848.92	21676.46	2845.50	9689.67
4596	4596	31-01-2023	7446	841	4825	1311	23603.82	5331.94	26151.50	9347.43
4597	4597	01-02-2023	6289	3143	3588	474	19936.13	19926.62	19446.96	3379.62
4598	4598	02-02-2023	3122	1188	5899	517	9896.74	7531.92	31972.58	3686.21

4599 rows × 10 columns

OBSERVATIONS:

- We can observe the first entry in the data, starts at 13-06-2010. This means the data for year 2010 is not complete.
- We can observe the last entry in the data, ends at 02-02-2023. This means the data for year 2023 is also not complete. it will be best to drop year 2010 and year 2023.

```python
[90]  # drop the first column
      data = data.drop(columns=['Unnamed: 0'])
```

Step 3: Checking the info of the training data

```
[91] data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4600 entries, 0 to 4599
Data columns (total 9 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Date    4600 non-null   object
 1   Q-P1    4600 non-null   int64
 2   Q-P2    4600 non-null   int64
 3   Q-P3    4600 non-null   int64
 4   Q-P4    4600 non-null   int64
 5   S-P1    4600 non-null   float64
 6   S-P2    4600 non-null   float64
 7   S-P3    4600 non-null   float64
 8   S-P4    4600 non-null   float64
dtypes: float64(4), int64(4), object(1)
memory usage: 323.6+ KB
```

OBSERVATIONS:

- The train dataset has 4600 entries(rows) and 9 columns. (we dropped one column)
- Date is an object data type. the rest of numerical in nature.

Step 4: Check for missing values

```
[92] data.isnull().sum()
```

```
Date    0
Q-P1    0
Q-P2    0
Q-P3    0
Q-P4    0
S-P1    0
S-P2    0
S-P3    0
S-P4    0
dtype: int64
```

OBSERVATIONS:

- we have no missing data

Step 5: EDA

EDA: Exploratory data analysis

Lets extract the year, month and Day from the date

```
[93]  # Extract year from the 'Day' 'Month' 'year' from the 'Date' column using a lambda function
      # We need to get the year from the data to analyse sales year to year
      data['Day'] = data['Date'].apply(lambda x: x.split('-')[0])
      data['Month'] = data['Date'].apply(lambda x: x.split('-')[1])
      data['Year'] = data['Date'].apply(lambda x: x.split('-')[2])
      data
```

	Date	Q-P1	Q-P2	Q-P3	Q-P4	S-P1	S-P2	S-P3	S-P4	Day	Month	Year
0	13-06-2010	5422	3725	576	907	17187.74	23616.50	3121.92	6466.91	13	06	2010
1	14-06-2010	7047	779	3578	1574	22338.99	4938.86	19392.76	11222.62	14	06	2010
2	15-06-2010	1572	2082	595	1145	4983.24	13199.88	3224.90	8163.85	15	06	2010
3	16-06-2010	5657	2399	3140	1672	17932.69	15209.66	17018.80	11921.36	16	06	2010
4	17-06-2010	3668	3207	2184	708	11627.56	20332.38	11837.28	5048.04	17	06	2010
...	...	...	...	...	...	...	...	...	...	...	...	...
4595	30-01-2023	2476	3419	525	1359	7848.92	21676.46	2845.50	9689.67	30	01	2023
4596	31-01-2023	7446	841	4825	1311	23603.82	5331.94	26151.50	9347.43	31	01	2023
4597	01-02-2023	6289	3143	3588	474	19936.13	19926.62	19446.96	3379.62	01	02	2023
4598	02-02-2023	3122	1188	5899	517	9896.74	7531.92	31972.58	3686.21	02	02	2023
4599	03-02-2023	1234	3854	2321	406	3911.78	24434.36	12579.82	2894.78	03	02	2023

4600 rows × 12 columns

- Lets drop rows for years 2010 and year 2023

```
[94]  data_reduced = data.query("Year != '2010' and Year != '2023'")
```

Graph our TOTAL & MEAN unit sold for each product using a histogram

```python
[95] Create a function that allows us to plot a bar chart for the 4 products
     ef plot_bar_chart(df, columns, stri, str1, val):
         # Aggregate sales for each product by year, by sum or mean
         if val == 'sum':
             sales_by_year = df.groupby('Year')[columns].sum().reset_index()
         elif val == 'mean':
             sales_by_year = df.groupby('Year')[columns].mean().reset_index()

         # Melt the data to make it easier to plot
         sales_by_year_melted = pd.melt(sales_by_year, id_vars='Year', value_vars=columns, var_name='Produc

         # Create a bar chart
         plt.figure(figsize=(20,4))
         sns.barplot(data=sales_by_year_melted, x='Year', y='Sales', hue='Product') #,palette="cividis")
         plt.xlabel('Year')
         plt.ylabel(stri)
         plt.title(f'{stri} by {str1}')
         plt.xticks(rotation=45)
         plt.show()
```

```python
[96] #use the plot_bar_chart function, enter the Unit Sales Columns and the Unit Sales string
     plot_bar_chart(data_reduced, ['Q-P1', 'Q-P2', 'Q-P3', 'Q-P4'],'Total Unit Sales', 'Year', 'sum')

     plot_bar_chart(data_reduced, ['Q-P1', 'Q-P2', 'Q-P3', 'Q-P4'],'Mean Unit Sales', 'Year', 'mean')
```
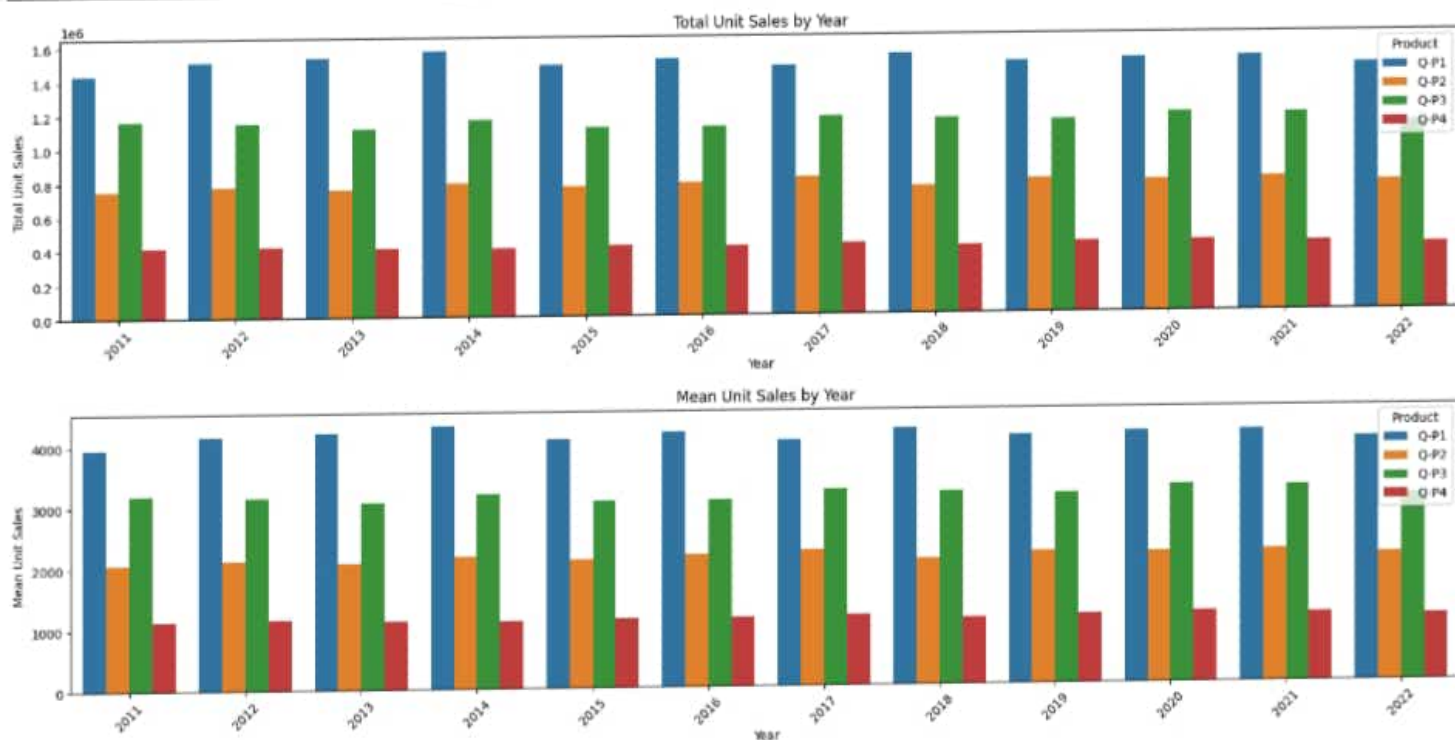
```
[97] #use the plot_bar_chart function, enter the Revenue Columns and the Revenue string
     plot_bar_chart(data_reduced, ['S-P1', 'S-P2', 'S-P3', 'S-P4'], 'Total Revenue', 'Year', 'sum')

     plot_bar_chart(data_reduced, ['S-P1', 'S-P2', 'S-P3', 'S-P4'], 'Mean Revenue', 'Year', 'mean')
```
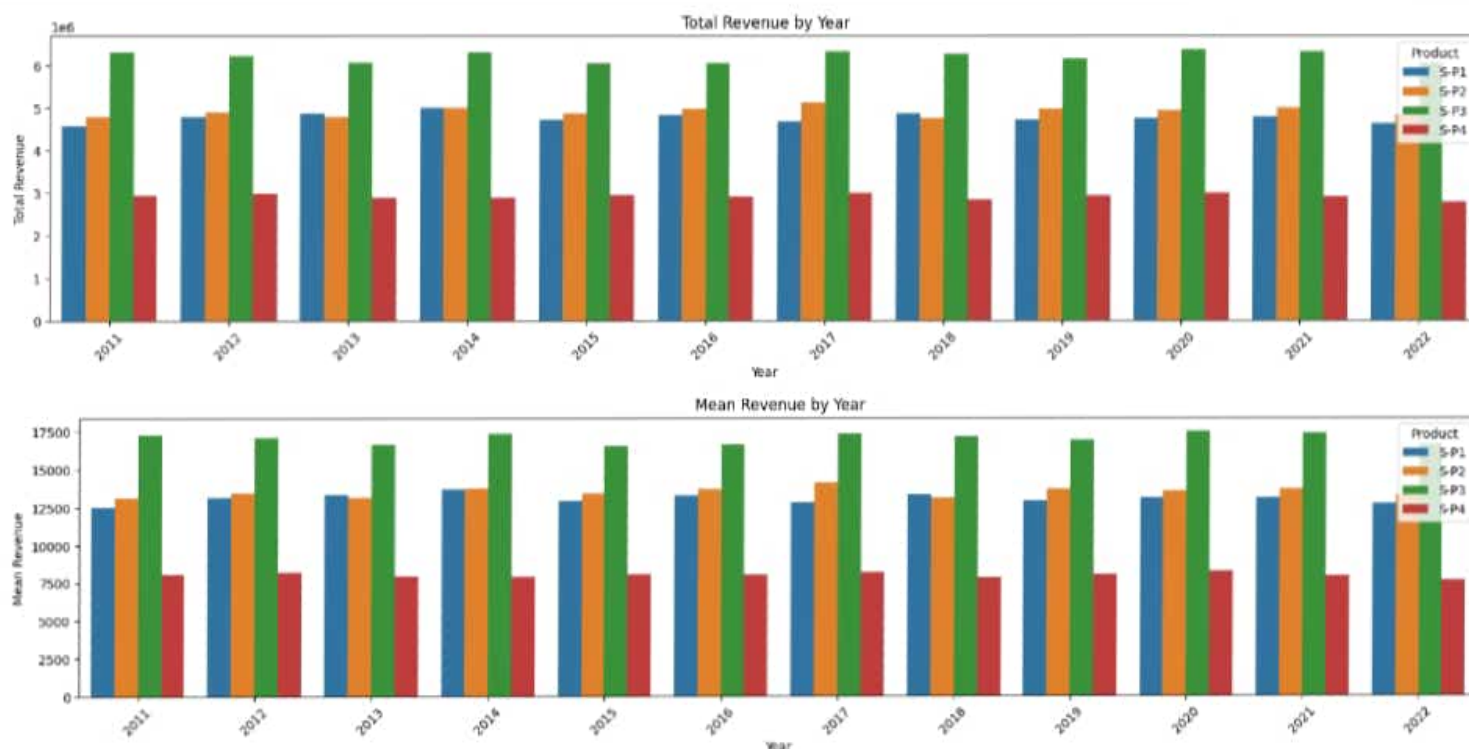

OBSERVATIONS:

- We can observe that P1 has the highest unit sales for each year. And it's highest is in year 2014.
- We can observe taht P4 has the lowest unit sales of all the products. Note

REC corp LTD needs you to solve the following questions:

1) Is there **any trend** in the **sales of all four products** during **certain months?**

"Trend in sales of all four products during certain months"

OBSERVATIONS:

- We can observe that all products drop in Feb. There also appears a very drastic drop after 12th month.
- The value show 9, which must be part of month 09. We need to rename this column to match with the 09. Before doing further analysis.

3) The company has all it's retail centers closed on the 31st of December every year. Mr: Hariharan , the CEO , would love to get an estimate on no: of units of each product that could be sold on 31st of Dec , every year , if all their retail centers were kept open.

[Estimate for each product the unit of sales that could be sold on 31st of Dec, if all their retail centers were kept open.]

```
[98] #get the 31st day for each month in each year. Note: not every month has 31 days
     def month_31_data(df, months):
         m31_data = df[df['Month'].isin(months) & (df['Day'] == '31')]
         return m31_data

     _31_months = month_31_data(data_reduced, ['01', '02', '03', '04', '05', '06', '07', '08', '09', '10',
     _31_months
```

	Date	Q-P1	Q-P2	Q-P3	Q-P4	S-P1	S-P2	S-P3	S-P4	Day	Month	Year
231	31-01-2011	939	3325	1863	1612	2976.63	21080.50	10097.46	11493.56	31	01	2011
290	31-03-2011	464	2220	421	1663	1470.88	14074.80	2281.82	11857.19	31	03	2011
351	31-05-2011	1507	2980	3816	1202	4777.19	18893.20	20682.72	8570.26	31	05	2011
412	31-07-2011	4336	744	4717	667	13745.12	4716.96	25566.14	4755.71	31	07	2011
442	31-08-2011	4548	1484	1596	1974	14417.16	9408.56	8650.32	14074.62	31	08	2011
...	...	...	...	...	...	...	...	...	...	...	...	...
4291	31-03-2022	3092	1645	4823	1864	9801.64	10429.30	26140.66	13290.32	31	03	2022
4352	31-05-2022	3669	2710	3067	1593	11630.73	17181.40	16623.14	11358.09	31	05	2022
4413	31-07-2022	1437	833	1867	1270	4555.29	5281.22	10119.14	9055.10	31	07	2022
4443	31-08-2022	1035	1639	3658	841	3280.95	10391.26	19826.36	5996.33	31	08	2022
4535	31-11-2022	4600	2006	3796	1426	14582.00	12718.04	20574.32	10167.38	31	11	2022

72 rows × 12 columns

```
[99] plot_bar_chart(_31_months, ['Q-P1', 'Q-P2', 'Q-P3', 'Q-P4'], 'Average Units', 'each Month, for 31st',
```


Average Units by each Month, for 31st
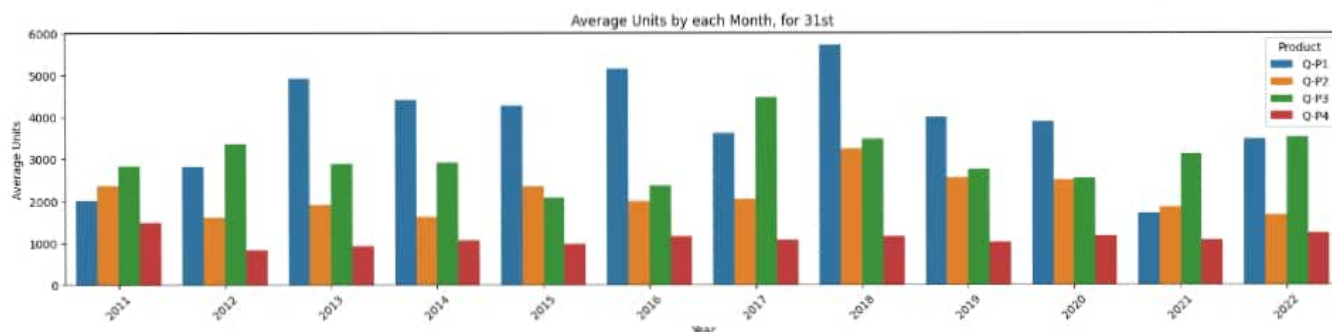
```
[100] plot_bar_chart(_31_months, ['S-P1', 'S-P2', 'S-P3', 'S-P4'], 'Average Revenue', 'each Month, for 31st
```


Average Revenue by each Month, for 31st
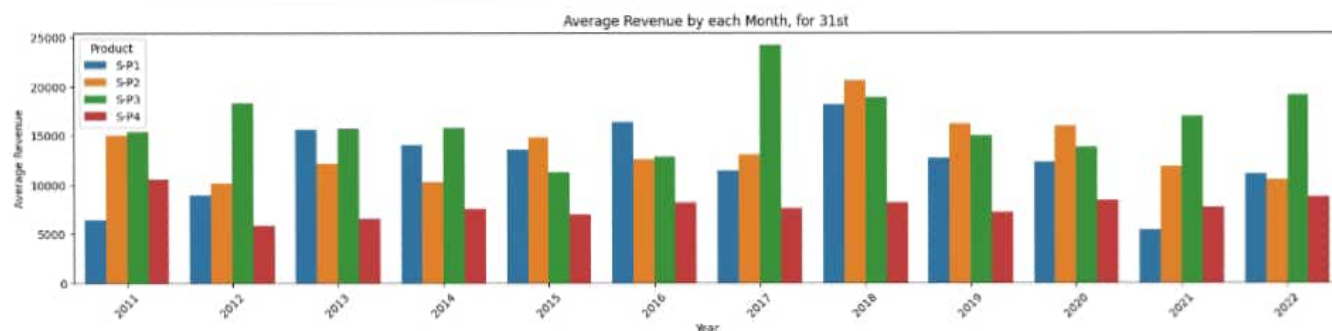
OBSERVATIONS:

- Overall we can see that P1 has the highest unit sales on the 31st for each year, except for 2021 and 2022. (These could be as a result to Covid and other economy issues.)
- P3 has the second highest unit sales for all the 31st in each year.

5)The CEO would also like to predict the sales and revenues for the year 2024. He wants you to give a yearly estimate with the best possible accuracy.

```
[101]  # gives us the average for all the 31st days across all years for each product
       def avg_on_31st(df, product):
           df_31 = df[df['Day'] == '31']
           avg_sales = df_31[product].mean()
           return avg_sales

       # Average for Unit Sales
       avg_on_31st(data_reduced, ['Q-P1', 'Q-P2', 'Q-P3', 'Q-P4']).round(2)
```

```
Q-P1     3813.74
Q-P2     2058.80
Q-P3     3183.88
Q-P4     1098.61
dtype: float64
```

```
[102]  # Average for Revenue
       avg_on_31st(data_reduced, ['S-P1', 'S-P2', 'S-P3', 'S-P4']).round(2)
```

```
S-P1     12089.55
S-P2     13052.78
S-P3     17256.63
S-P4      7833.07
dtype: float64
```

CONCLUSION:

Unit Sales 2011 - 2022

- P1 has the highest unit sales for each year. And it's highest is in year 2014.
- We can observe that P4 has the lowest unit sales of all the products. Revenues 2011 - 2022
- We can observe that P3 brought in the most revenue. This could be as a result of multiple things:
- P3 was sold for higher than the rest, as it had the second highest unit sales for each year.
- We can observe than P1 and P2 brought in similar revenues for each year. With P2 bringing in slightly more.
- P1 despite having the most unit sold, brought in the second lowest revenue each year.

Average Month Sales 2011 - 2022

- We can observe that all Products unit sales drop in Feb.
- We can observe that Feb and Dec have the lowest sales for each product For P1
- We can observe Mar - Jul having the highest unit sales For P2
- We can observe Jan, Mar - Aug having the highest unit sales For P3
- We can observe May & Sep having the highest unit sales For P4
- We can observe uniform sales from Jan - Dec

Project Development Part – 1

Product Sales Analysis
Phase 3

Date	24 October 2023
Project Name	Product Sales Analysis

Phase 3:

- Data Collection
- Prepare the dataset
- Data Cleaning and preprocessing

Data Collection:

Data collection is the pivotal first step in constructing a dataset for analysis, machine learning, or any data-driven project. It involves systematically gathering information from diverse sources, which can encompass structured data from databases and spreadsheets, as well as unstructured data from text, images, and more. The process begins with clearly defining project objectives to guide the selection of appropriate data sources. Subsequently, data is extracted, transformed, and cleansed to ensure high quality and accuracy. Attention to data privacy and security is essential to safeguard sensitive information during collection. Detailed documentation of the process is maintained, and if data collection is ongoing, continuous monitoring is established to keep the dataset current and reliable. Effective data collection forms the bedrock upon which robust datasets are built, enabling organizations and individuals to derive meaningful insights from their data.

Download the Dataset:

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

Load the Dataset:

Tool used: IBM Cognos Analytics

IBM Cognos Analytics:

IBM Cognos analysis is a powerful tool for extracting actionable insights from data. This platform offers a wide range of analytical features, including data visualization, reporting, and predictive analytics, to help organizations make informed decisions. With Cognos, users can explore data, create interactive dashboards, and generate comprehensive reports that provide a clear understanding of business performance. Its ability to seamlessly integrate with various data sources, along with features for data exploration, makes it a valuable asset for data-driven organizations seeking to uncover trends, patterns, and opportunities in their data.

Load the dataset:

Loading datasets in IBM Cognos analysis is a straightforward yet crucial step for data-driven insights. This process involves importing and connecting data from various sources, including databases, spreadsheets, and cloud services, into the Cognos environment. Users can then prepare, transform, and clean the data to create a unified, structured dataset suitable for analysis. By providing a seamless integration with multiple data sources, Cognos streamlines the process, ensuring that data is readily available for reporting, visualization, and advanced analytics. This efficient dataset loading feature empowers organizations to harness the full potential of their data for informed decision-making and actionable insights.
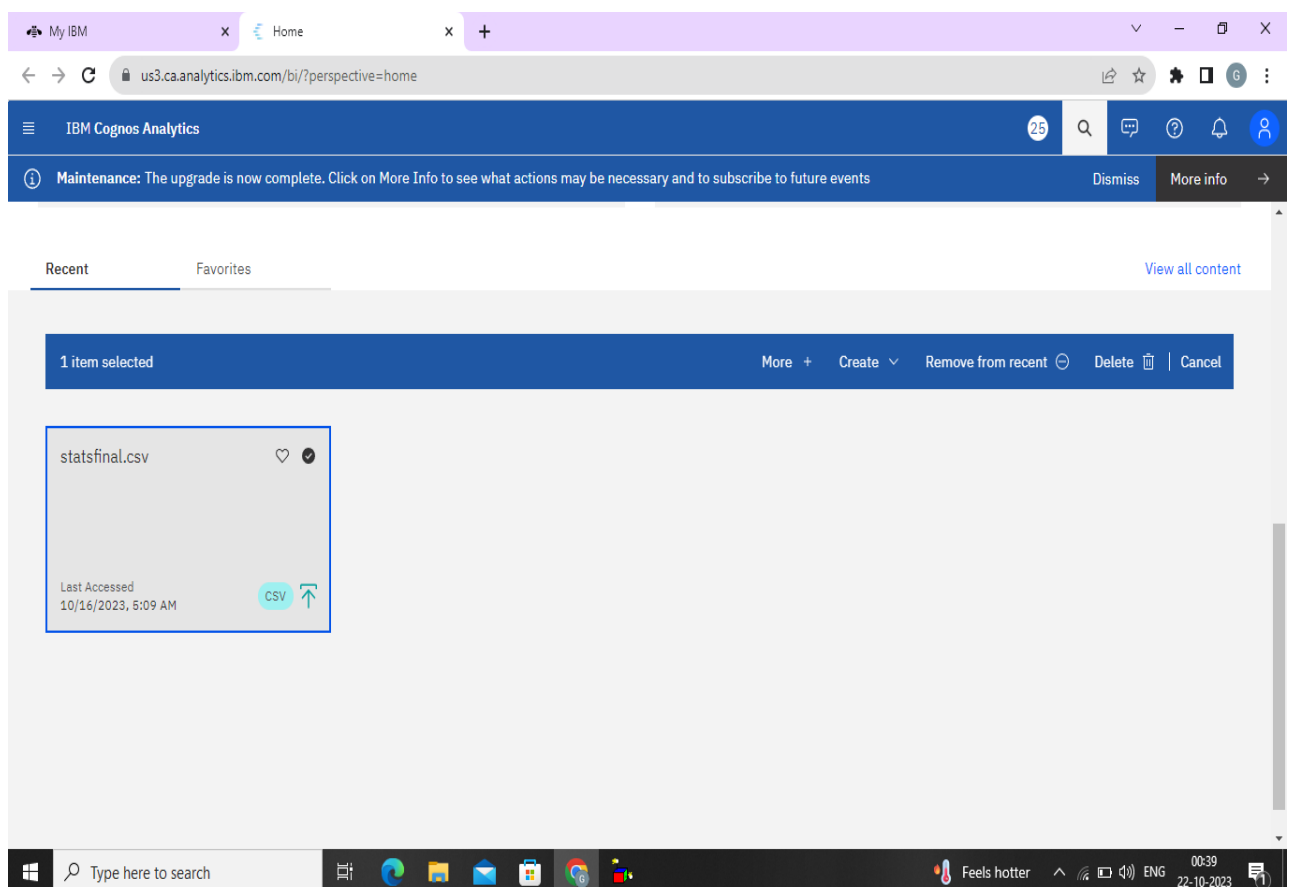
Data Preparation:

Data preparation in IBM Cognos analysis is a crucial phase where raw data is refined into a usable and valuable asset. This process involves tasks such as cleansing, transformation, and structuring of data to ensure accuracy and consistency. With Cognos, users can easily reshape data, handle missing values, and create calculated fields, making it ready for analysis. Data preparation in Cognos enables organizations to work with cleaner, more meaningful datasets, thus improving the quality and reliability of insights and reports generated within the platform. It's a vital step that empowers users to uncover patterns, trends, and actionable information from their data with confidence.

Data Cleaning:

Data cleaning in IBM Cognos analysis is the process of refining and enhancing the quality of your data. It involves identifying and rectifying errors, inconsistencies, and inaccuracies within your datasets. With Cognos, users can detect and address issues such as duplicate records, missing values, outliers, and data discrepancies. By employing various data cleaning techniques, Cognos ensures that your data is accurate, consistent, and reliable. Clean data is the foundation for producing trustworthy reports and insights, ultimately enabling more informed decision-making and analysis within the Cognos platform.

Product Sales Analysis data:

About Dataset

REC corp LTD. is small-scaled business venture established in India.

- They have been selling FOUR PRODUCTS for OVER TEN YEARS.

- The products are P1, P2, P3 and P4.

- They have collected data from their retail centers and organized it

  into a small csv file.

The excel file contains about 8 numerical parameters:

- Q1- Total unit sales of product 1
- Q2- Total unit sales of product 2
- Q3- Total unit sales of product 3
- Q4- Total unit sales of product 4

- S1- Total revenue from product 1
- S2- Total revenue from product 2
- S3- Total revenue from product 3
- S4- Total revenue from product 4

Project Development Part – 2

Product Sales Analysis
Phase - 4

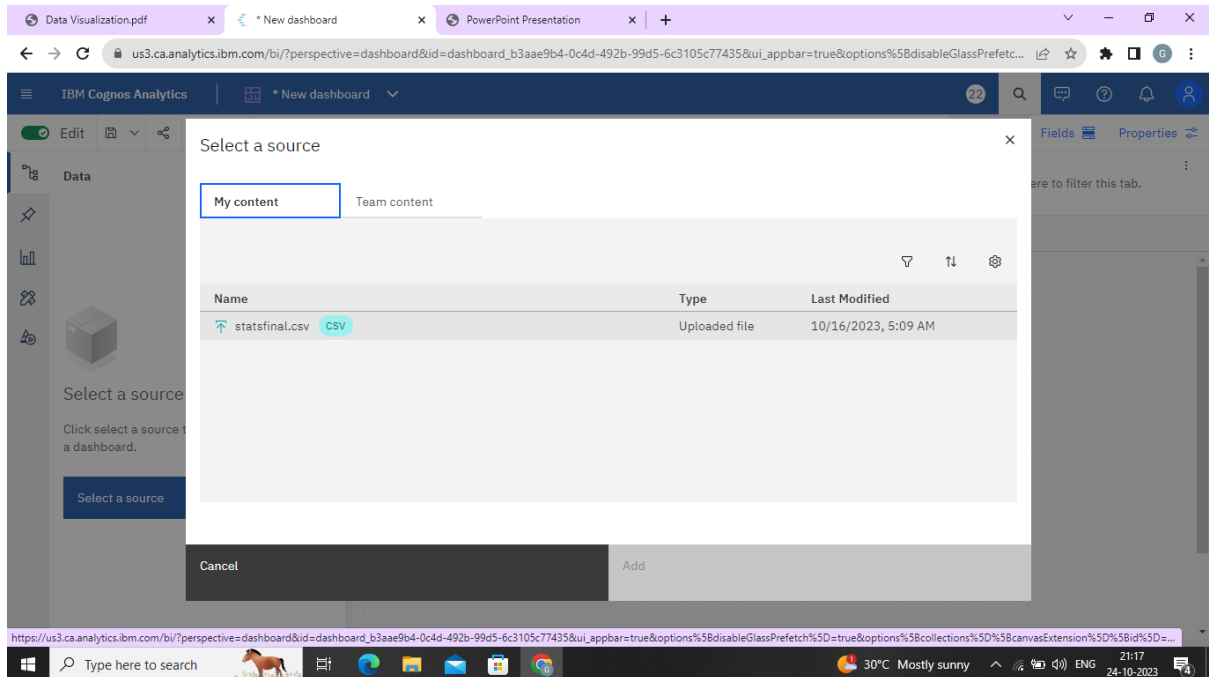| Date	26 October 2023
Project Name	Product Sales Analysis

Tasks:

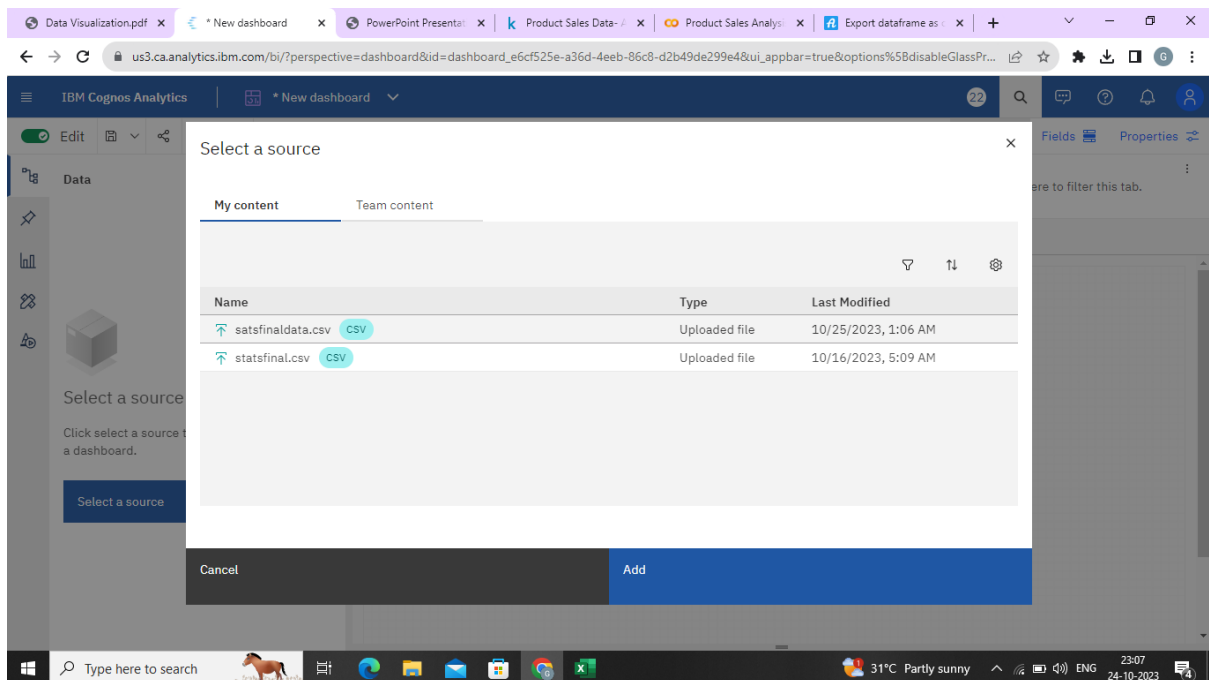- Visualizations

- Creating Dashboard

- Creating Report

Creating Dashboard using IBM Cognos:
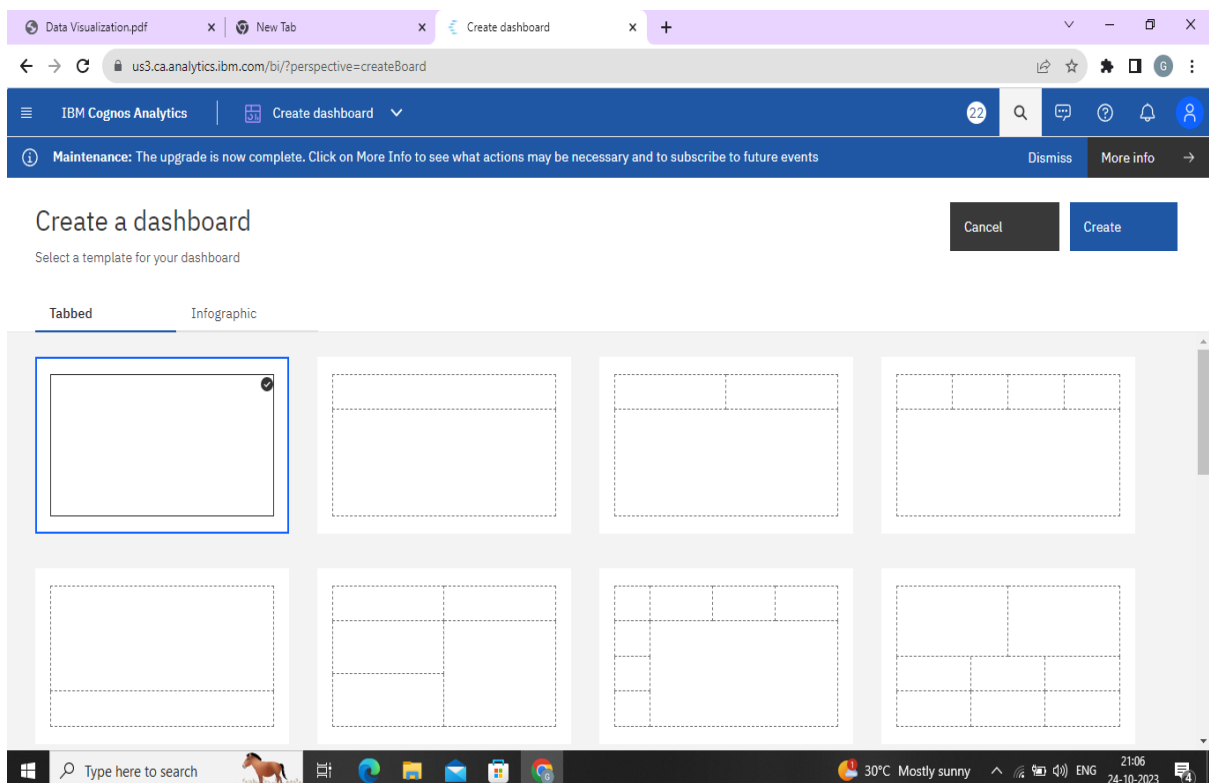
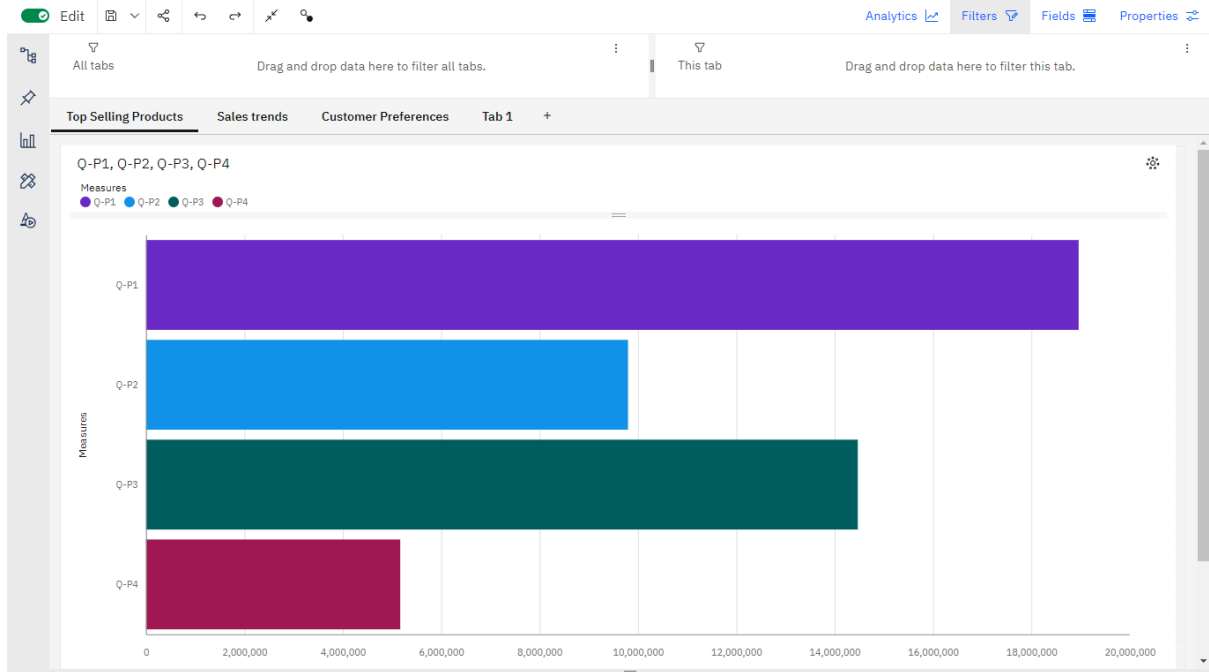**Loading the dataset to create a dashboard:**

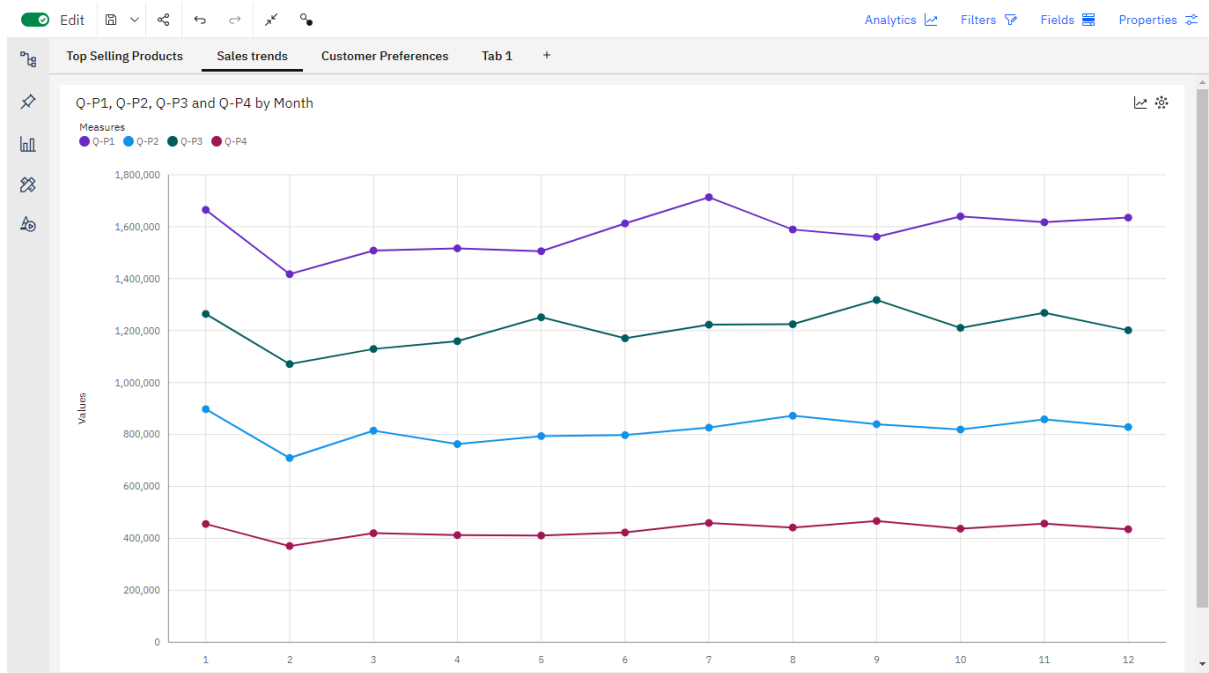

Selecting EDA processed data for visualizations:
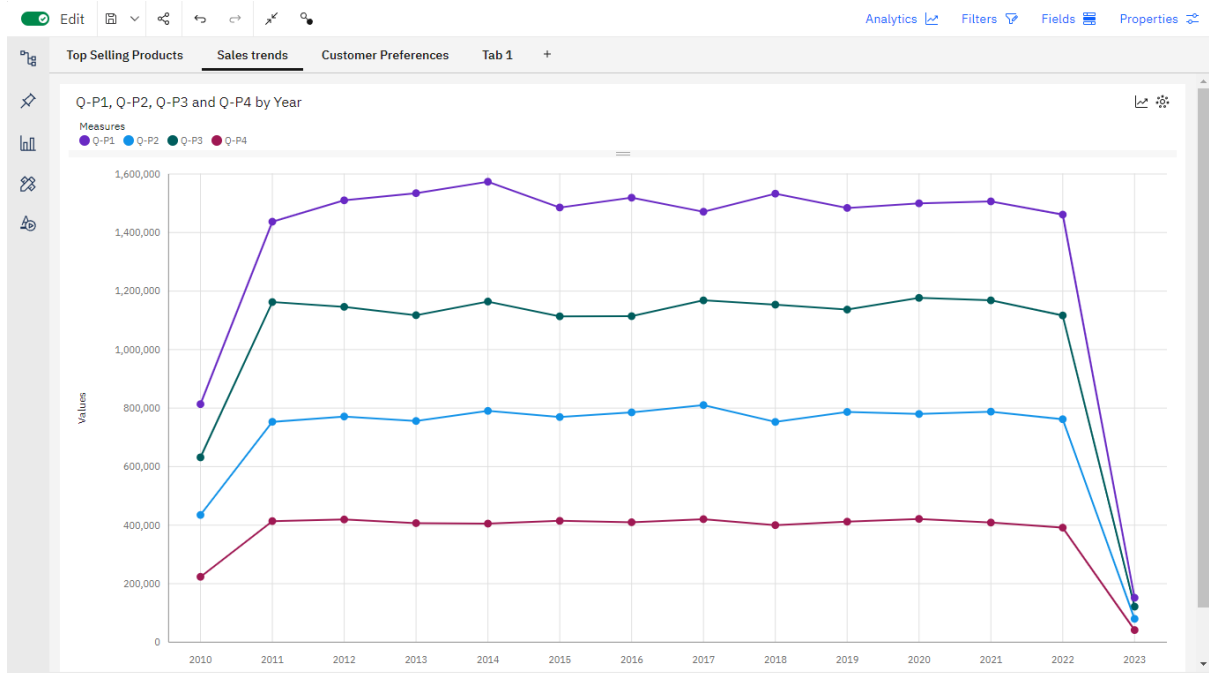
Visualization:

- **Top Selling Products:**

- **Sales Trends – Monthly:**

- **Sales Trends – Yearly:**

- **Sales Trends – Monthly & Yearly:**
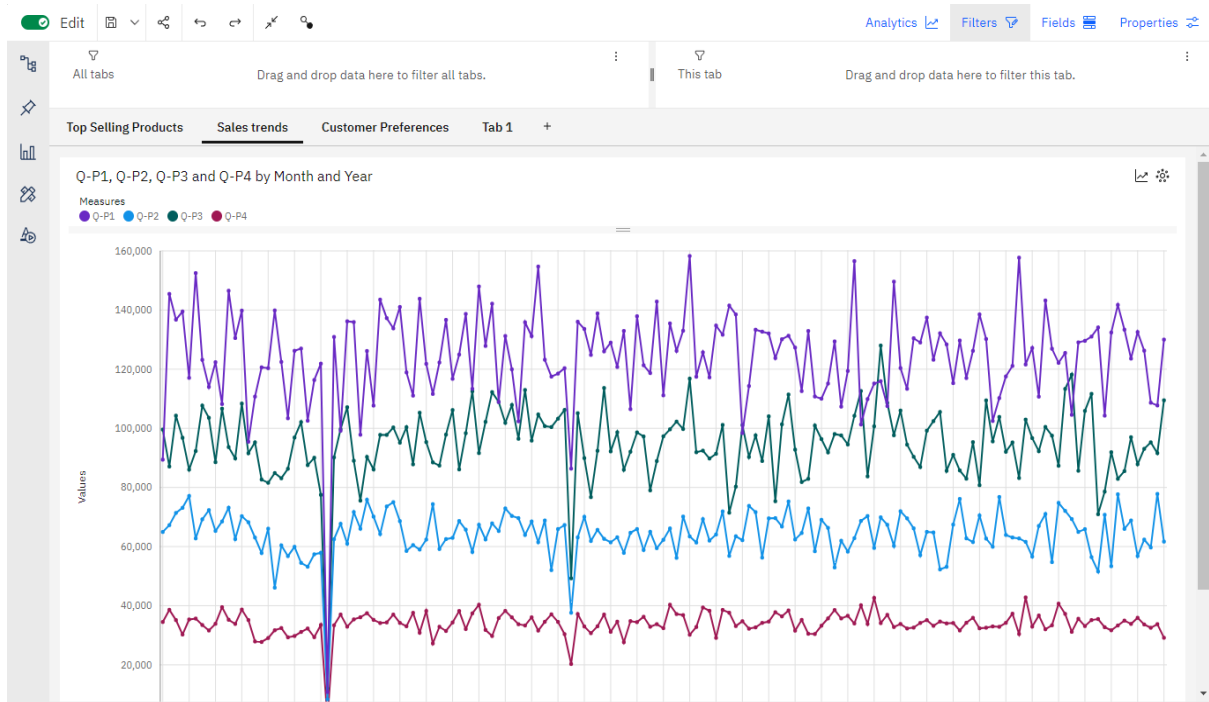
- **Customer Preferences:**
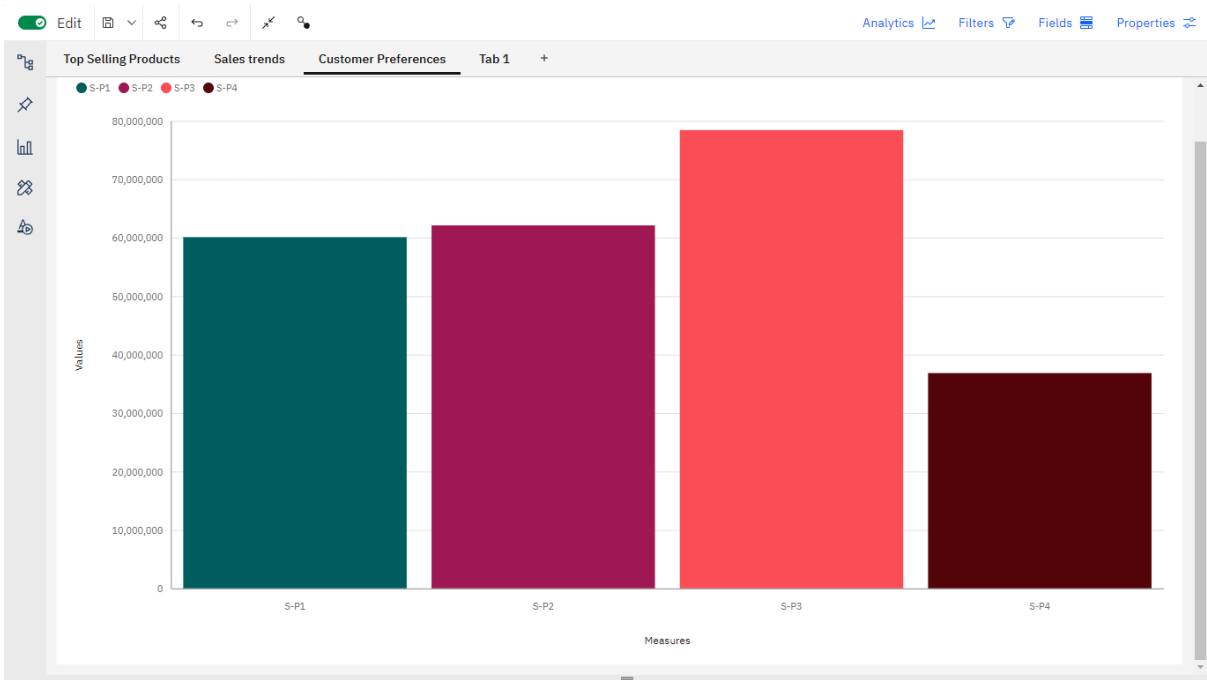
Creating Dashboard:

- **Top Selling Products**
- **Sales Trends**
- **Customer Preferences**

Creating Report:

- Click on the share button from the menu tab.
- Click on the export button from the share dialog box.
- Select the letter option from the page size dropdown.
- Select the required orientation (Landscape) from the orientation dropdown.
- Click on the export button from the dialog box

Conclusion:

Product Sales Analysis:

P1 (Product 1):

Unit Sales 2011 - 2022:
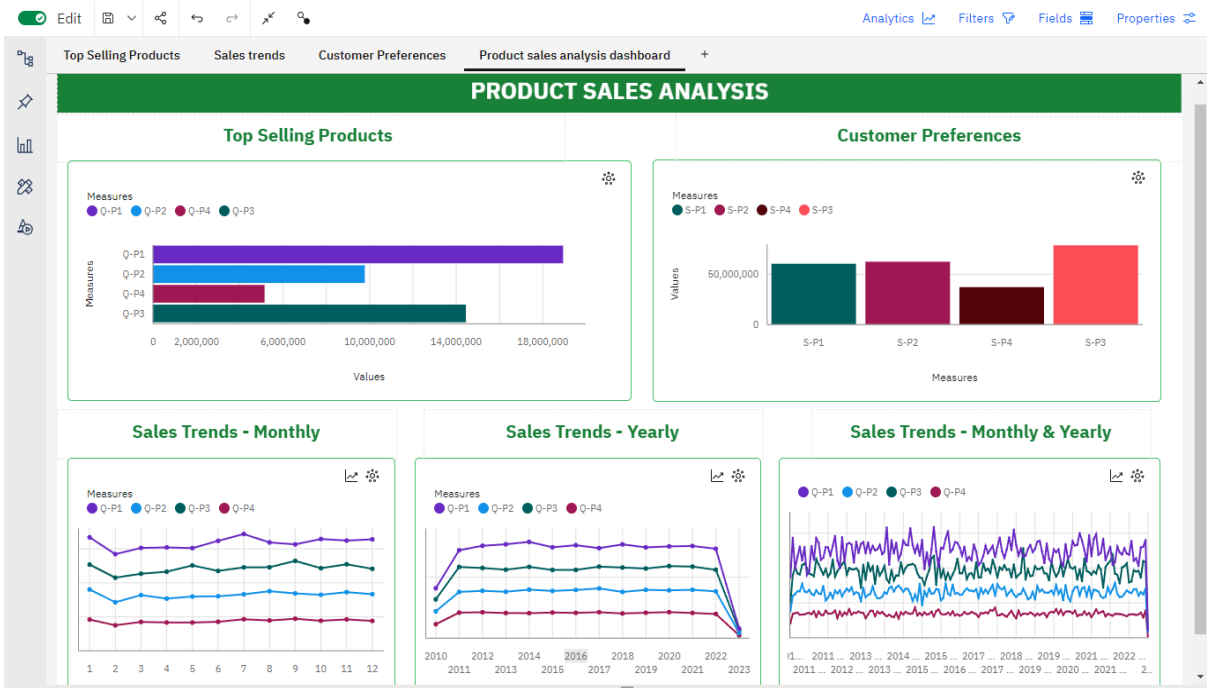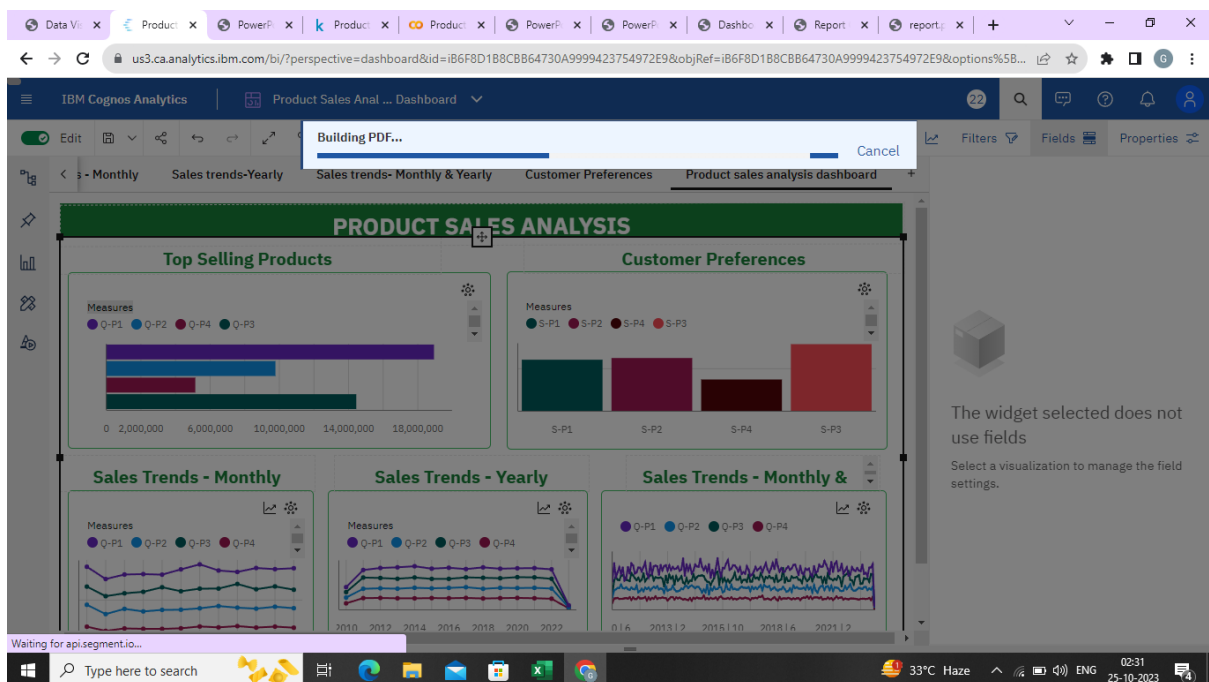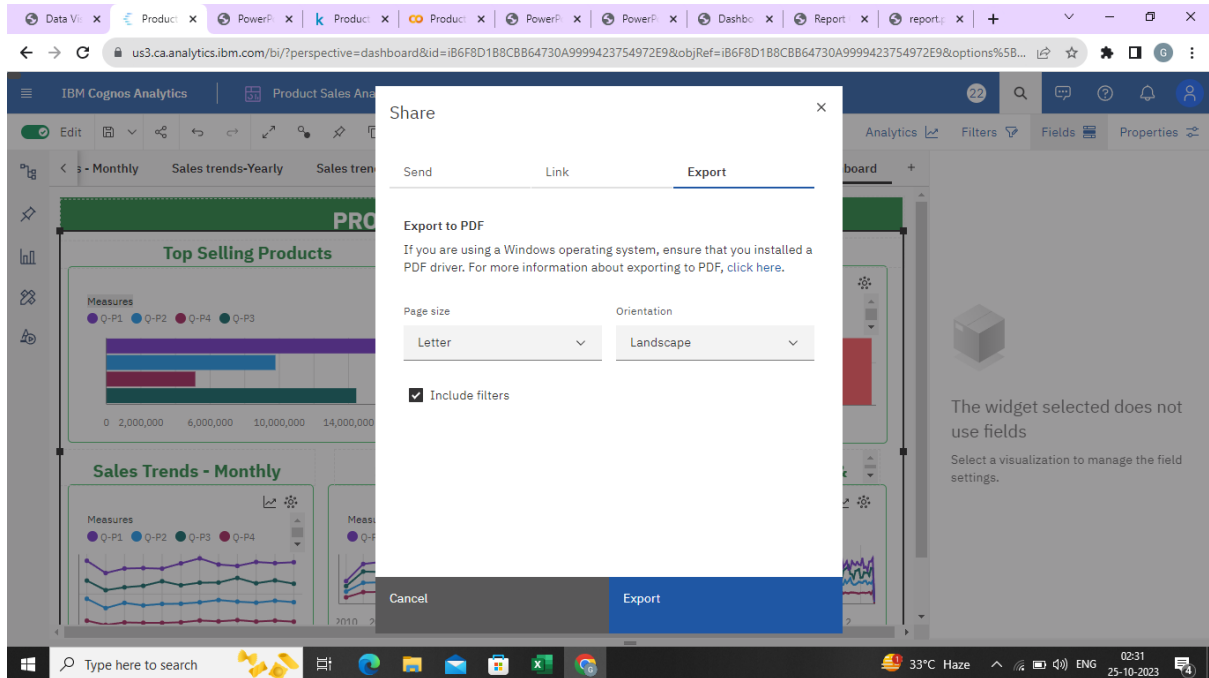- P1 consistently had the highest unit sales each year, with its peak in 2014. This indicates a strong demand for P1, especially during specific periods.
- **Recommendation**: Given the seasonality in P1's sales, it's vital to manage inventory effectively. Ensure higher stock levels in preparation for peak months, primarily from March to July. Also, capitalize on its high unit sales to drive revenue by offering complementary products or package deals. Consider special promotions during its peak season to further boost sales.

Revenues 2011 - 2022:
- Despite having the highest unit sales, P1 generated the second-lowest revenue each year, implying that the product may be underpriced or not effectively monetized.
- **Recommendation**: Evaluate the pricing strategy for P1. While it has high unit sales, consider raising its price slightly, especially during its peak season. Focus on upselling and cross-selling to increase the average transaction value. Identify opportunities to increase revenue without solely relying on increasing unit sales.

Average Month Sales 2011 - 2022:
- P1 experiences a drop in unit sales in February, but the highest sales occur from March to July.
- **Recommendation**: Mitigate the February sales drop by running special promotions or discounts to stimulate sales during this period. During peak months, leverage marketing strategies to highlight P1's features and advantages to maximize sales.

Estimated Unit Sales for 31st of Dec:
- Historically, P1 has high unit sales on December 31st, with exceptions in 2021 and 2022.
- **Recommendation**: For December 31st, maintain higher inventory levels and run special promotions or sales events to take advantage of this historically strong sales date. Also, closely monitor external factors and be ready to adapt to challenges like those experienced in 2021 and 2022.

P2 (Product 2):

- **Unit Sales 2011 - 2022**:
  - P2 consistently had strong unit sales each year, generating slightly more revenue than P1.
  - **Recommendation**: Given P2's consistent sales performance, focus on expanding its product line or offering complementary items. Highlight the unique features or benefits of P2 in marketing campaigns to further boost unit sales and revenue.

- **Revenues 2011 - 2022**:
  - P2 brought in slightly more revenue than P1, despite lower unit sales.
  - **Recommendation**: Capitalize on P2's revenue growth potential by designing marketing campaigns that emphasize its value proposition. Create bundles or promotions that encourage customers to choose P2. Analyze customer data to identify segments that prefer P2, and target them with tailored marketing strategies.

- **Average Month Sales 2011 - 2022**:
  - P2's highest unit sales were in January and from March to August.
  - **Recommendation**: During the high-sales months, focus marketing efforts and promotions to maximize revenue. Identify and promote unique selling points of P2 to attract customers during the peak season.

P3 (Product 3):

Unit Sales 2011 - 2022:
- P3 consistently ranked second in unit sales each year and was the highest revenue generator.
- **Recommendation**: P3's performance in unit sales and revenue is impressive. Maintain its premium pricing strategy if it's justified by its features and quality. Market P3 as a high-value product, and consider expanding the product line or offering complementary products that align with its customer base.

Revenues 2011 - 2022:
- P3 generated the highest revenue, likely due to its higher price and consistently strong unit sales.
- **Recommendation**: Continue to position P3 as a premium product. Develop marketing campaigns highlighting its quality and unique features. Offer special promotions or bundles for loyal customers to maintain and potentially increase revenue.

Average Month Sales 2011 - 2022:
- P3's highest unit sales occurred in May and September.
- **Recommendation**: Leverage these peak months to the fullest by running targeted promotions or campaigns. Focus on the key selling points of P3 to attract customers and maximize sales during these periods.

P4 (Product 4):

Unit Sales 2011 - 2022:
- P4 consistently had the lowest unit sales among all products.
- **Recommendation**: Given the consistently low demand for P4, maintain minimal inventory levels and monitor sales trends closely. Consider whether it's cost-effective to continue offering P4 and evaluate the potential impact of discontinuing it on your overall business.

Revenues 2011 - 2022:
- P4 brought in the lowest revenue among all products.
- **Recommendation**: Due to low sales and revenue, carefully assess the viability of continuing to offer P4. If it's essential to your product line, explore marketing strategies or potential product enhancements to improve its performance.

Average Month Sales 2011 - 2022:
- P4 had uniform sales throughout the year.
- **Recommendation**: Given consistent low demand for P4, focus marketing efforts on other products with higher sales potential. Consider alternative uses or ways to position P4 if discontinuation is not an option.

Inventory Management Strategies:

1. P1's Seasonal Variation:

- **Insight**: Product P1 consistently has the highest unit sales each year, with its peak in 2014. This suggests a strong seasonality in demand for P1, with the highest sales occurring during specific months.
- **Recommendation**: To effectively manage inventory for P1, it's crucial to anticipate and prepare for the peak demand. You should maintain higher inventory levels in the months leading up to the historically strong sales period (March to July). This strategy helps to prevent stockouts and maximize sales during the high-demand months. You can also adjust your procurement and production schedules accordingly.

2. P4's Consistent Low Sales:

- **Insight**: Product P4 consistently had the lowest unit sales among all products, indicating a consistently lower demand.
- **Recommendation**: Given the consistently low sales for P4, you may consider keeping lower inventory levels for this product. This will free up valuable warehouse space and reduce carrying costs associated with excess inventory. Regularly monitor the sales trend for P4 and adjust inventory levels accordingly to minimize overstock.

3. Estimating Inventory for December 31st:

- **Insight**: Estimating unit sales for December 31st is challenging, and you used the historical average as an estimate. P1 had the highest unit sales on December 31st in most years, except for 2021 and 2022.
- **Recommendation**: While December 31st sales are unpredictable, you can use historical averages as a starting point for inventory planning. For products like P1 and P3, which have historically performed well on December 31st, it's advisable to ensure higher inventory levels for this date. However, remain agile and be prepared to adjust inventory levels as external factors or market conditions change.

4. Mitigating February Sales Drop:

- **Insight**: All products experienced a drop in unit sales in February, suggesting a seasonal decline in demand during this month.
- **Recommendation**: In response to the February sales drop, consider implementing inventory management strategies that account for this seasonally low demand. You might reduce inventory levels during February to minimize carrying costs. Alternatively, plan special promotions, discounts, or marketing campaigns during this period to stimulate sales and mitigate the dip in revenue.

Marketing Strategies:

1. Pricing for P3:

- **Insight**: Product P3 generated the highest revenue, possibly due to its higher price and consistently strong unit sales.
- **Recommendation**: Given P3's ability to generate high revenue, you can consider maintaining a premium price for this product if there's a demand for it. Highlight the unique features or value proposition of P3 to justify the price. Market P3 as a premium product and emphasize its quality or features that make it stand out.

2. P1's High Sales, Low Revenue:

- **Insight**: Despite consistently having the highest unit sales, P1 doesn't generate the highest revenue each year.

- **Recommendation**: To optimize revenue for P1, implement marketing strategies that encourage customers to purchase more units or add complementary items to their P1 purchase. This cross-selling approach can boost the average transaction value, leading to higher revenue without the need to increase unit prices.

3. P2's Revenue Growth:

- **Insight**: P2 generates slightly more revenue than P1 despite having slightly lower unit sales.
- **Recommendation**: Recognize the potential for P2 to be a revenue driver. Focus marketing efforts on P2, highlighting its value proposition and potential areas of growth. Consider expanding the product line or offering complementary products to further increase revenue.

4. Seasonal Promotions:

- **Insight**: Sales patterns for each product exhibit seasonality. For example, P1's highest unit sales occur from March to July.
- **Recommendation**: Plan marketing campaigns around these seasonal trends. During the peak sales months for each product, implement targeted promotions, discounts, or special offers to capitalize on the higher demand. Effective timing of these promotions can significantly boost sales and revenue.

5. December 31st Sales:

- **Insight**: P1 and P3 historically perform well on December 31st, with high unit sales.
- **Recommendation**: Leverage historical data to plan specific promotions or sales events on December 31st for P1 and P3. These campaigns can take advantage of the expected high unit sales on that date. Consider offering limited-time discounts or bundling products to encourage higher sales on this day.

6. Adapt to External Factors:

- **Insight**: External factors like COVID-19 and economic issues can impact sales.
- **Recommendation**: Maintain flexibility in your marketing strategies. Be prepared to adjust your tactics in response to changing external factors. This might involve shifting resources, messaging, or campaigns to align with evolving consumer behavior and economic conditions.

7. Inventory Clearance Sales:

- **Insight**: December is historically a period of lower sales.
- **Recommendation**: Consider conducting year-end inventory clearance sales in December. This approach can help reduce carrying costs, clear out excess inventory, and generate revenue during a typically slow period.

8. Customer Segmentation:

- **Insight**: Understanding customer demographics and preferences is critical.
- **Recommendation**: Analyze customer data to segment your audience and tailor marketing campaigns for each product to specific target groups. Understanding what drives sales for different customer segments can help you create more effective and personalized marketing strategies.

Summary:

The analysis of unit sales, revenues, average month sales, and estimated sales for December 31st across four products (P1, P2, P3, and P4) provides valuable insights for inventory management and marketing strategies:

Inventory Management:

1. **P1**: Recognize the strong seasonality in P1's unit sales, with peaks from March to July. Maintain higher inventory levels in anticipation of peak periods to prevent stockouts.
2. **P4**: Given consistently low demand for P4, consider keeping minimal inventory levels to minimize carrying costs.
3. **December 31st Estimations**: Use historical averages as a baseline for inventory planning. Ensure higher inventory for products like P1 and P3, which historically perform well on December 31st.
4. **February Sales Drop**: Implement strategies to address the February sales drop, such as promotions or discounts, to stimulate sales during this seasonally low-demand period.

Marketing Strategies:

1. **P1**: Focus on upselling and cross-selling to boost revenue without solely relying on increasing unit sales. Implement promotions during peak seasons and December 31st.

2. **P2**: Capitalize on P2's consistent performance by expanding its product line and offering complementary items. Tailor marketing campaigns to highlight its value proposition.
3. **P3**: Position P3 as a premium product with a higher price, emphasizing quality and unique features. Maintain premium pricing and expand marketing efforts during peak sales months.
4. **P4**: Evaluate the viability of P4 due to consistently low demand. Consider discontinuation or explore strategies to improve its performance.
5. **Seasonal Promotions**: Leverage seasonal sales patterns for each product by planning targeted promotions during peak months. Offer discounts or bundles to maximize sales and revenue.
6. **Adapt to External Factors**: Stay agile and adapt marketing strategies in response to external factors like COVID-19 and economic issues.
7. **Customer Segmentation**: Analyze customer data to tailor marketing campaigns for each product to specific target groups, maximizing effectiveness.