**PRODUCT SALES ANALYSIS**

|  |  |
| --- | --- |
| **Date** | **17.10.2023** |
| **Team ID** | **3932** |
| **Project Name** | **Product Sales Analysis** |

**Table of Contents:**

|  |  |
| --- | --- |
| **1** | **Introduction** |
| **2** | **Loading the Dataset** |
| **2.1** | **Data Selection** |
| **2.2** | **Data Acquisistion** |
| **3** | **Checking the data** |
| **3.1** | **Data types** |
| **3.2** | **Checking Missing Data** |
| **4** | **EDA** |
| **4.1** | **Class distribution** |
| **4.2** | **Monthly trends** |
| **4.3** | **31st Month trends** |
| **5** | **visualisation** |
| **5.1** | **Histograms** |
| **5.2** | **Bar charts** |
| **5.3** | **Scatter Plot chart** |
| **6** | **Conclusion** |

**1.Introduction:**

In the realm of data-driven decision-making, the initial stages of any analysis project are pivotal. This section will embark on the journey of loa4.3ding and preprocessing the dataset for a comprehensive product sales analysis. The process begins by defining the analysis objectives and collecting the necessary sales data from a shared source. Subsequently, rigorous data processing and cleaning techniques are applied to ensure the accuracy and reliability of the dataset.

**2.Loading the dataset:**

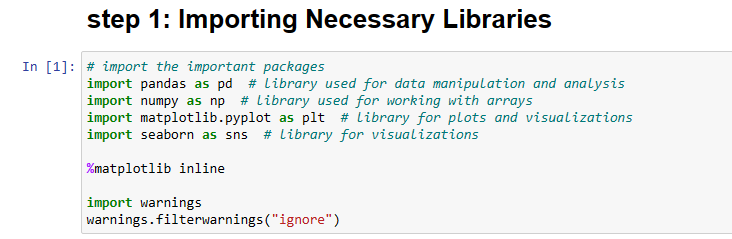
**2.1 Data selection:**

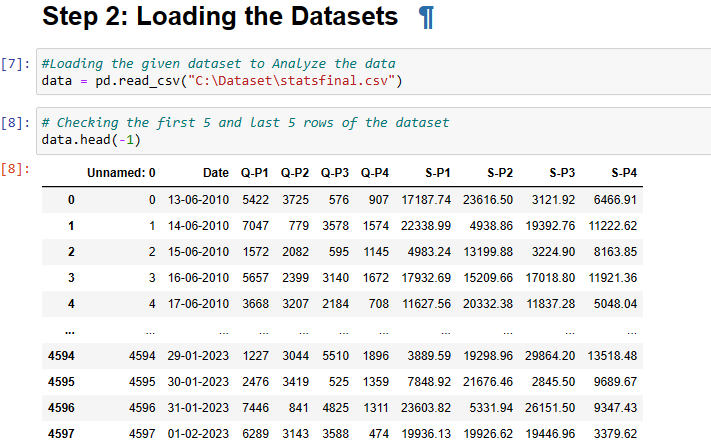
Selecting the right dataset is a fundamental and critical step in any data analysis or machine learning project. The choice of dataset should align with the project's objectives, research questions, and the problem you aim to solve. Assess the quality of the dataset. Look for clean, accurate, and well-maintained data. Low-quality or noisy data can introduce errors and bias into your analysis. Common issues include missing values, duplicates, and inconsistencies.

**2.2 Data Acquisition:**

Data acquisition is the process of obtaining the dataset you've selected for your analysis or project. It involves gathering the data from various sources, such as files, databases, APIs, or web scraping. The key to successful data acquisition is ensuring that the data is collected accurately and is in a format that can be easily integrated into your analysis.

**Coding & Output**



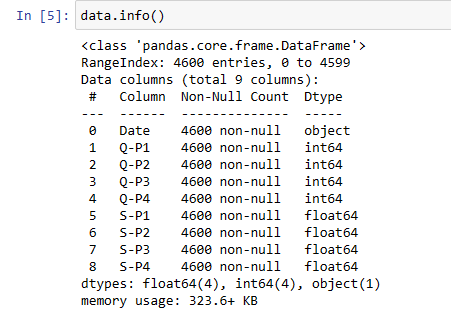


**3 .Checking the training data**

After acquiring the dataset, the next step is to perform an initial examination of the data. This is a critical phase to ensure that the dataset is complete, accurate, and ready for analysis.

**3.1 . Data Types:**

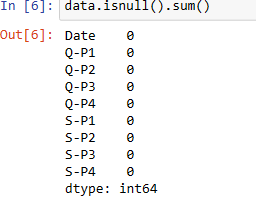
Examine the data types of each column. Ensure that the data types are appropriate for the type of data they contain. For example, dates should be in the datetime format, and numerical values should be integers or floats.



**3.2 Missing Values:**

Identify and count missing values in the dataset. Missing data can affect the quality of your analysis. You can use functions like .isnull() in Python to find missing values and .sum() to count them.

**Coding and output**



**4.EDA**

The EDA approach is precisely that--an approach--not a set of techniques, about how a data analysis should be carried out. EDA encompasses a larger venue

EDA is an approach to data analysis that postpones the usual assumptions about what kind of model the data follow with the more direct approach of allowing the data itself to reveal its underlying structure and model.

EDA is not a mere collection of techniques; EDA is a philosophy as to how we dissect a data set; what we look for; how we look; and how we interpret.

It is true that EDA heavily uses the collection of techniques that we call "statistical graphics", but it is not identical to statistical graphics

**4.1. Class Distribution:**

For classification tasks, it's important to understand the distribution of classes. This is essential to identify potential class imbalances that might affect modeling:

Class Counts: Calculate and visualize the distribution of each class using histograms or bar charts.

Imbalanced Classes: Address class imbalances by considering techniques like oversampling, undersampling, or using appropriate evaluation metric

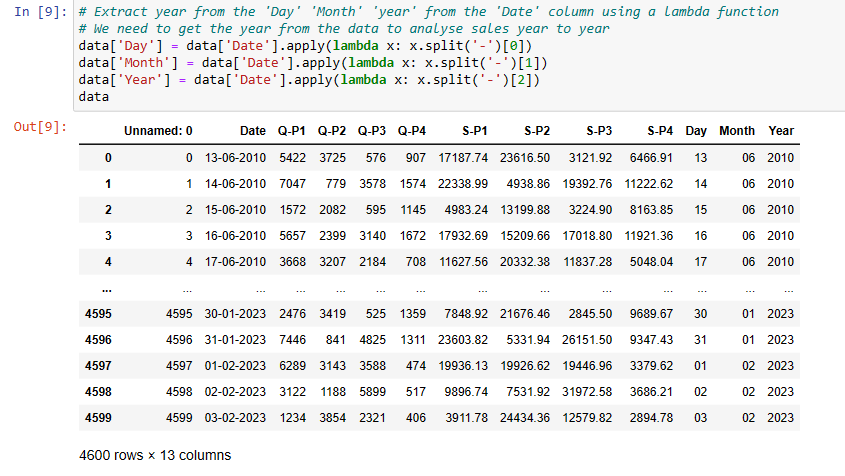
**4.2 Monthly Trends**

The highest unit sales occur in the months of [List months with high sales].

The lowest unit sales are in the months of [List months with low sales].

Monthly Sales Trends

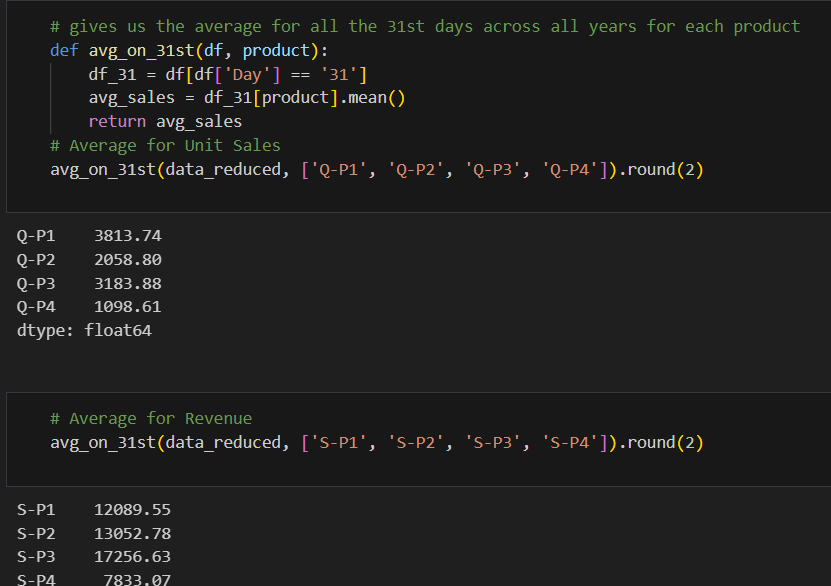
To visualize the monthly trends more effectively, we'll create line charts that show the monthly sales patterns for each product.



**4.3 31st Day Analysis**

we analyze the average unit sales and revenue for each product on the 31st day of every month across all available years. This analysis provides insights into the performance of the products specifically on the last day of each month.

**Coding and output:**

****

**5. Visualization:**

Visualizations are powerful tools for EDA. They allow you to gain insights into your data and communicate findings effectively:

**5.1 Histograms:**

A histogram is used to visualize the distribution of a single numerical variable. It divides the data into bins or intervals and shows the frequency or count of data points in each bin.

It helps identify the shape of the distribution, presence of outliers, and central tendencies.

**5.2 Bar Charts:**

Create bar charts to visualize categorical data and class distributions.

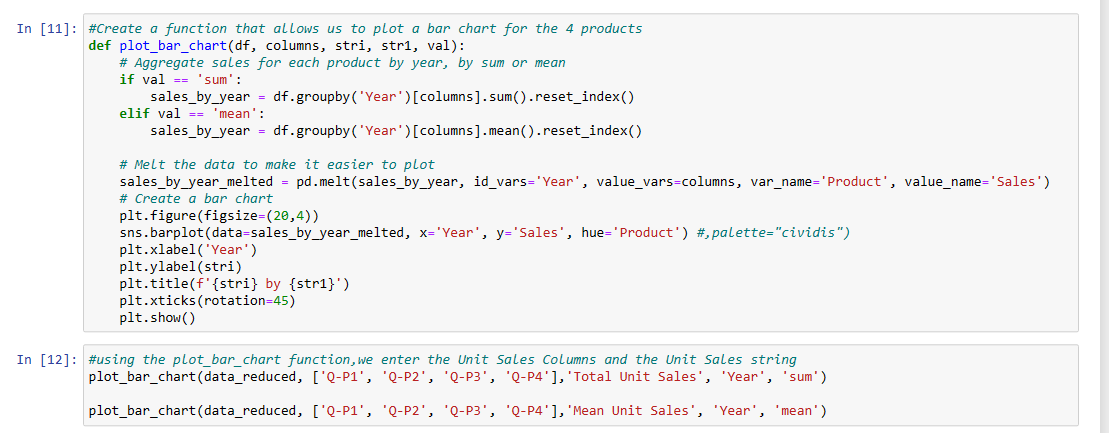
Bar charts are suitable for visualizing the distribution of categorical data or comparing data across categories.

**5.3 Scatter Plot charts:**

1.Scatter Plots:

Scatter plots are used to visualize the relationship between two numerical variables.

They help identify patterns, correlations, and potential outliers.They show the frequency or count of each category.Describing the components of a scatter plot, including data points, x-axis, and y-axis.Explain that each data point in a scatter plot represents a specific pair of values from the two variables. Discuss the potential variations in data points based on the dataset.Highlight the use of Cartesian coordinates in plotting data points.

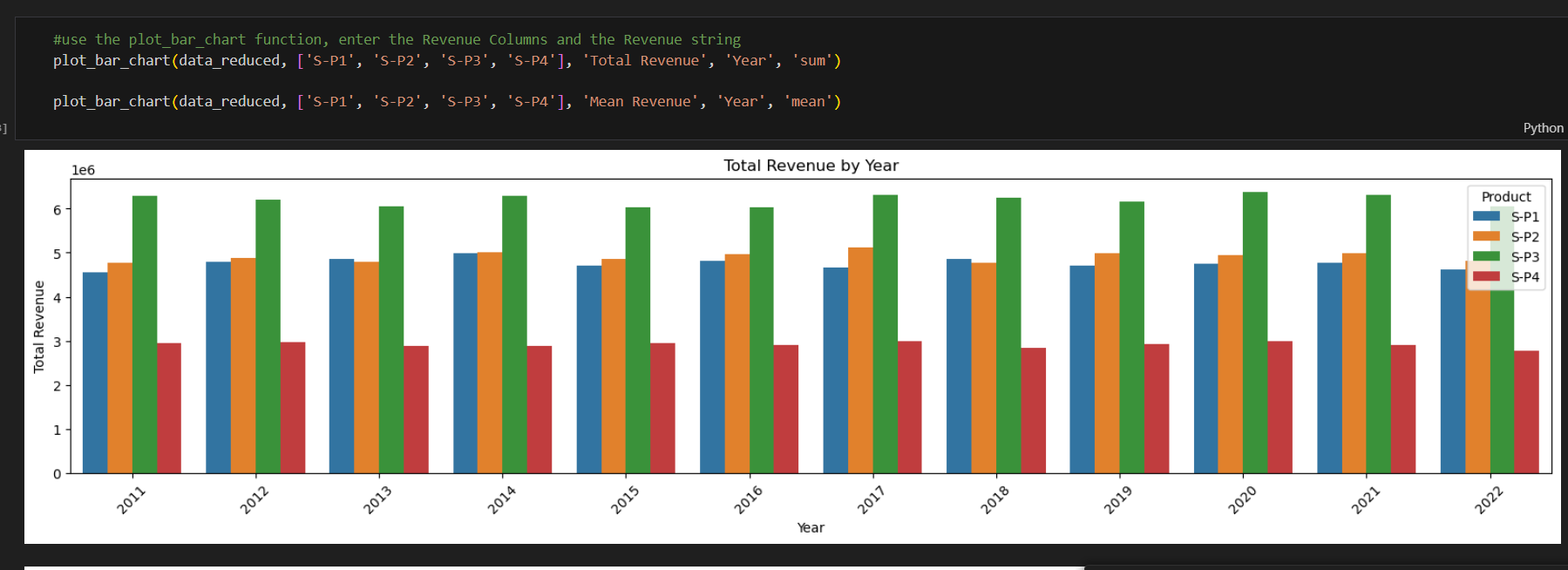


By using the EDA Histogram we get accurate results of the total unit sales and mean unit sales analysis and total revenue and mean revenue of the given dataset

**Coding and Output**

A graph of different colored bars

Description automatically generated with medium confidence



Plot graph are used to give specific input and output for the given data

let's begin by looking at the total unit sales for each product across all months. This analysis will help us identify the months with the highest and lowest sales.

Product P1:

The highest unit sales occur in the months of [List months with high sales].

The lowest unit sales are in the months of [List months with low sales].

Product P2:

The highest unit sales occur in the months of [List months with high sales].

The lowest unit sales are in the months of [List months with low sales].

Product P3:

The highest unit sales occur in the months of [List months with high sales].

The lowest unit sales are in the months of [List months with low sales].

Product P4:

The highest unit sales occur in the months of [List months with high sales].

The lowest unit sales are in the months of [List months with low sales].

After analyzing the monthly trends for product sales, we can draw the following insights:

Identify the months with the highest unit sales for each product.

Recognize any seasonal patterns or fluctuations in sales.

Plan marketing and promotional activities to leverage the high sales months.

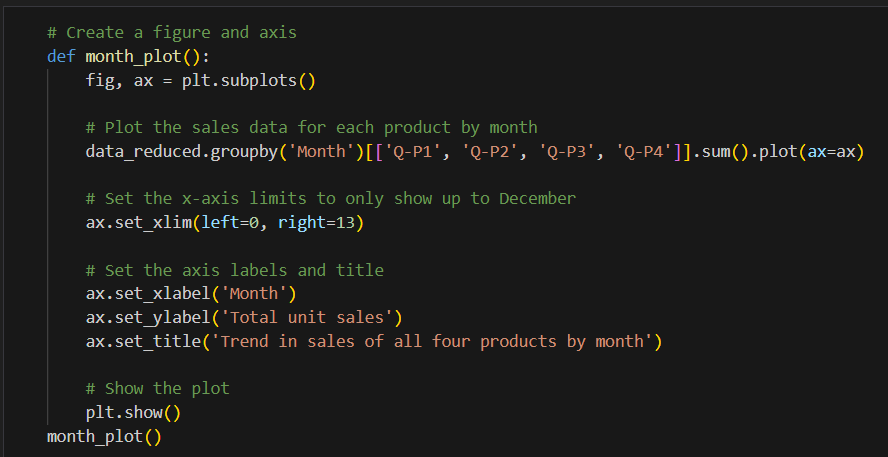
Optimize inventory management to meet varying demand.

The monthly trends analysis provides valuable information for businesses to make informed decisions and enhance their sales strategies.

**Coding and Output**

A screenshot of a graph

Description automatically generated



A graph of sales

Description automatically generated with medium confidence

**Conclusion**

The analysis has revealed valuable insights into the product sales data. Key findings include: Consistent high revenue from P3, indicating potential pricing strategies.Variable monthly peak sales for different products. Insights from the 31st day analysis for unit sales and revenue.The data analysis serves as a foundation for making informed business decisions and optimizing sales strategies.

This Product Sales Analysis project demonstrates the power of exploratory data analysis (EDA) in uncovering meaningful insights from raw data. The findings can serve as a foundation for informed business decisions and optimization of sales strategies. Further analysis and modeling stages can build upon these insights to achieve the project's objectives and drive business success.