**PHASE:3**

**Product Sales Analysis**

**TEAM MEMBERS:**

NAME: Bavithraraj V

REG NO:721221104012

NAME: Yogesh M

REG NO:721221104063

NAME: Gokul R

REG NO:721221104017

NAME: Kishore Kannan R

REG NO:721221104029

**Introduction:**

Product sales analysis is the systematic examination of sales data to gain insights into product performance, trends, and customer behaviour. It informs strategic decisions such as pricing, inventory management, and marketing strategies, helping businesses enhance profitability and adapt to market dynamics.

**Data Collection:**

* Collect product sales data from the provided source.

**Dataset Link :** [**https://www.kaggle.com/datasets/ksabishek/product-sales-data**](https://www.kaggle.com/datasets/ksabishek/product-sales-data)

**Data Preprocessing and Cleaning:**

* Clean the collected data to ensure its quality and accuracy.

**#importing data set**

import pandas as pd

import numpy as np

import sklearn

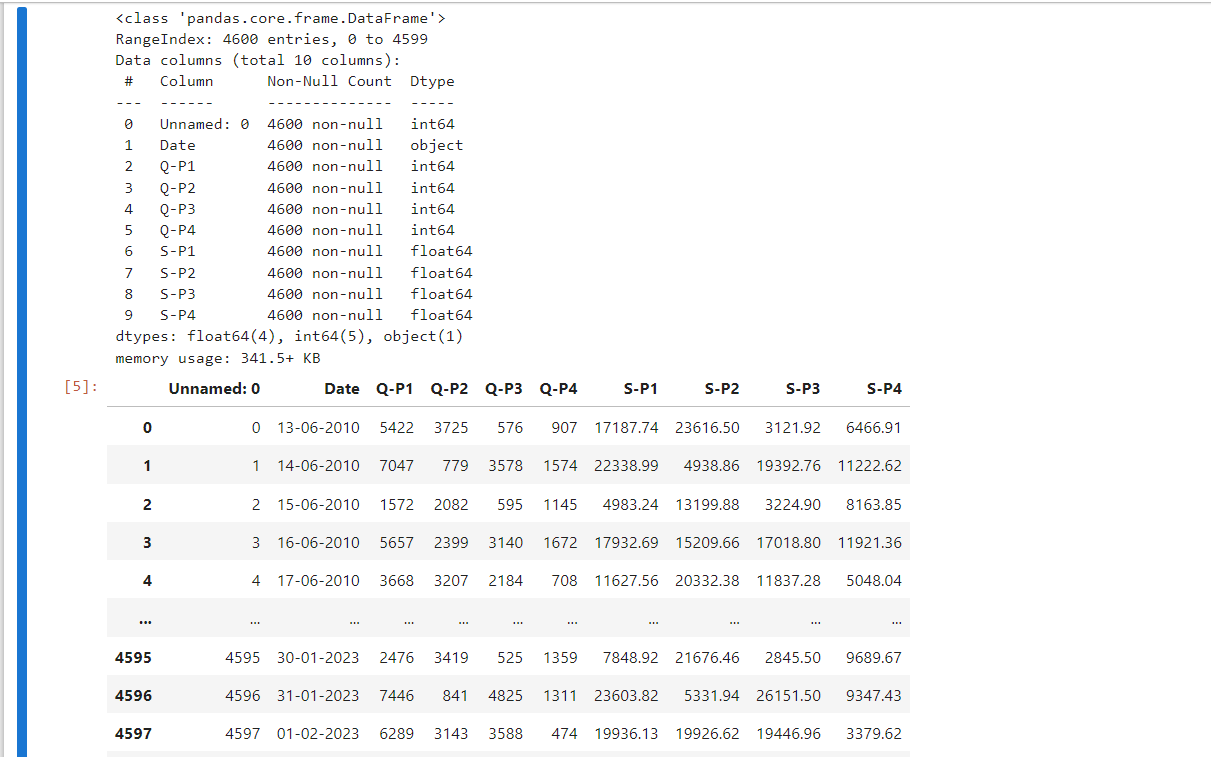
from sklearn.preprocessing import StandardScaler

data = pd.read\_csv("statsfinal.csv")

datadata.info()

data

**output:**



**#cleansing the data set**

data['Date'].fillna('Unknown', inplace=True)

data.drop\_duplicates(subset=['Q-P1', 'Q-P2', 'Q-P3','Q-P4','S-P1', 'S-P2', 'S-P3','S-P4',], keep='first', inplace=True)

data['Q-P1'] = data['Q-P1'].astype(int)

data['Q-P2'] = data['Q-P2'].astype(int)

data['Q-P3'] = data['Q-P3'].astype(int)

data['Q-P4'] = data['Q-P4'].astype(int)

data['S-P1'] = data['S-P1'].astype(float)

data['S-P2'] = data['S-P2'].astype(float)

data['S-P3'] = data['S-P3'].astype(float)

data['S-P4'] = data['S-P4'].astype(float)

#stored the cleaned data into another file

data.to\_csv('cleanddataset.csv', index=False)

**Exploratory Data Analysis (EDA):**

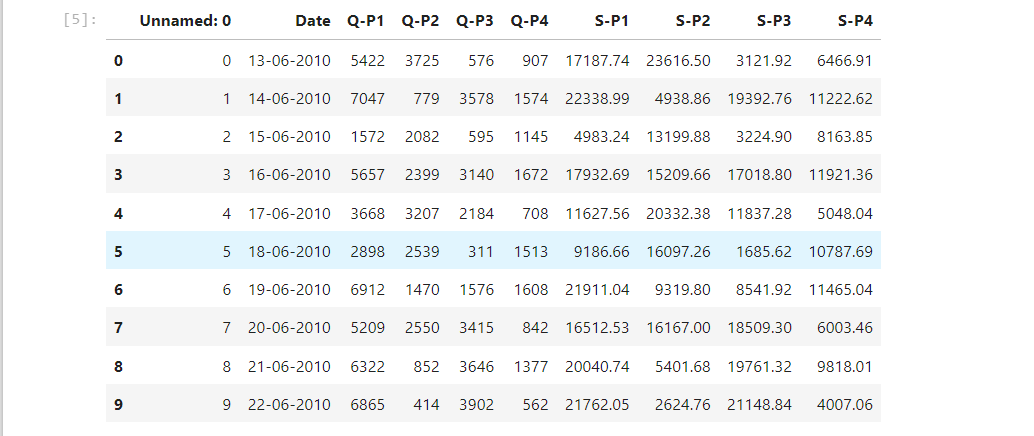
data = pd.read\_csv('cleanddataset.csv')

data.shape

data.head(10)

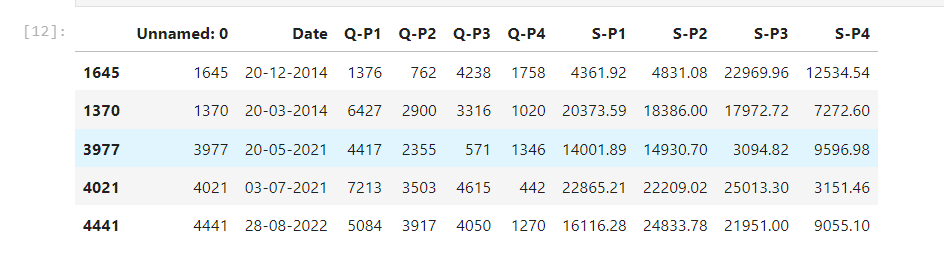
data.sample(5)

**output:**



data.sample

**output:**



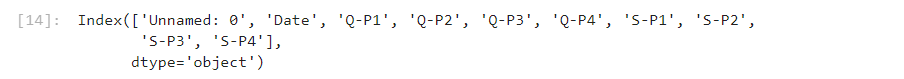
data.shape

**output:**

****

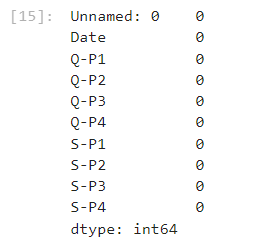
data.columns

**output:**



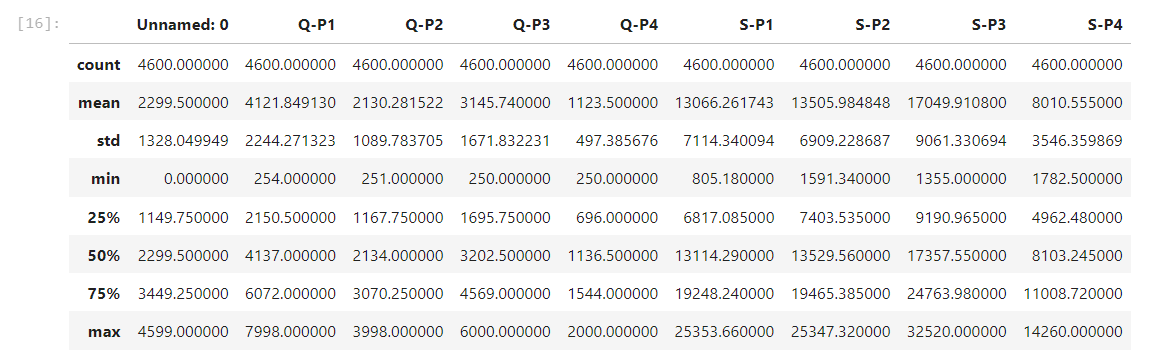
pd.isnull(data).sum()

**output:**

****

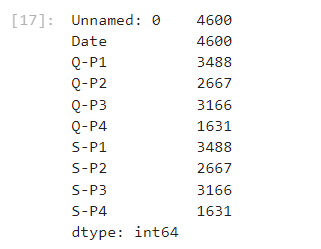
data.describe()

**output:**



data.nunique()

**output:**



**Visualization and Analysis:**

##can assign the each chart to one axes at a time

fig,axrr=plt.subplots(4,2,figsize=(20,20))

ax=axrr[0][0]

ax.set\_title("Q-P1")

data['Q-P1'].value\_counts().plot.area(ax=axrr[0][0])

ax=axrr[0][1]

ax.set\_title("Q-P2")

data['Q-P2'].value\_counts().plot.area(ax=axrr[0][1])

ax=axrr[1][0]

ax.set\_title("Q-P3")

data['Q-P3'].value\_counts().plot.area(ax=axrr[1][0])

ax=axrr[1][1]

ax.set\_title("Q-P4")

data['Q-P4'].value\_counts().plot.area(ax=axrr[1][1])

ax=axrr[2][0]

ax.set\_title("S-P1")

data['S-P1'].value\_counts().plot.area(ax=axrr[2][0])

ax=axrr[2][0]

ax.set\_title("S-P2")

data['S-P2'].value\_counts().plot.area(ax=axrr[2][1])

ax=axrr[2][0]

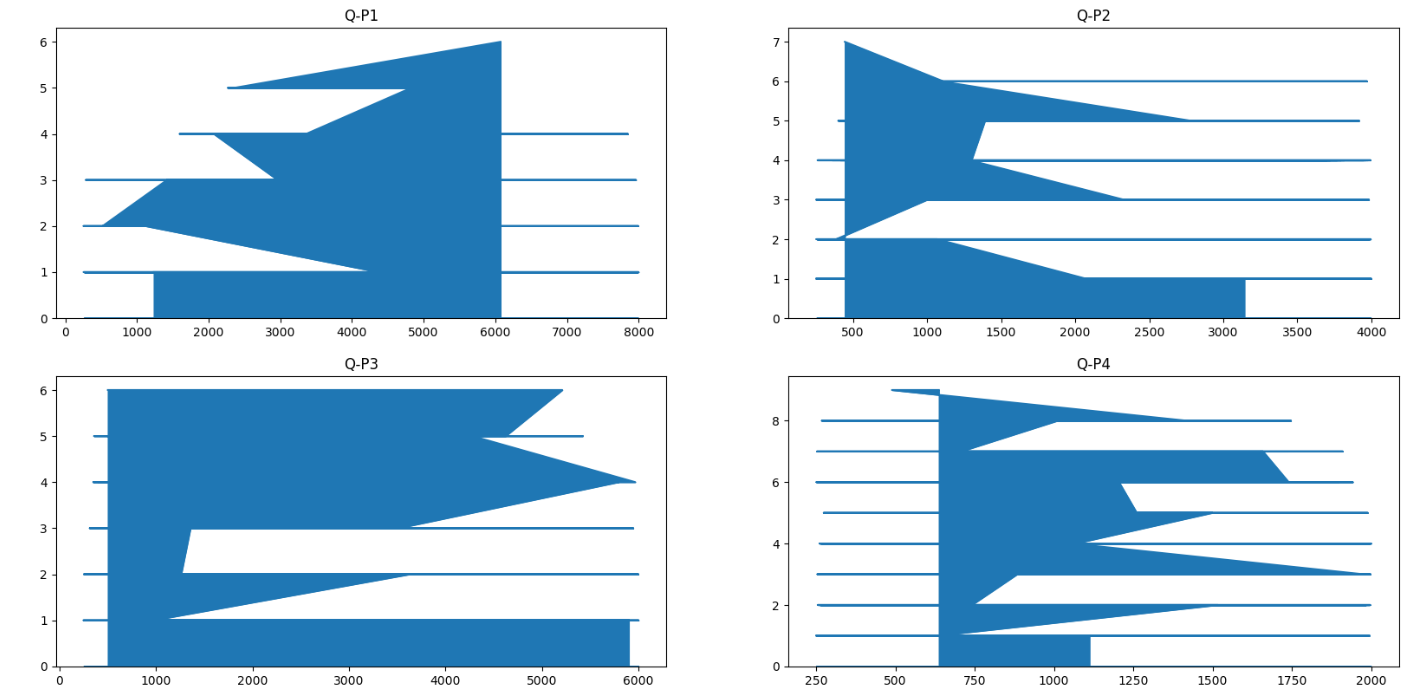
ax.set\_title("S-P3")

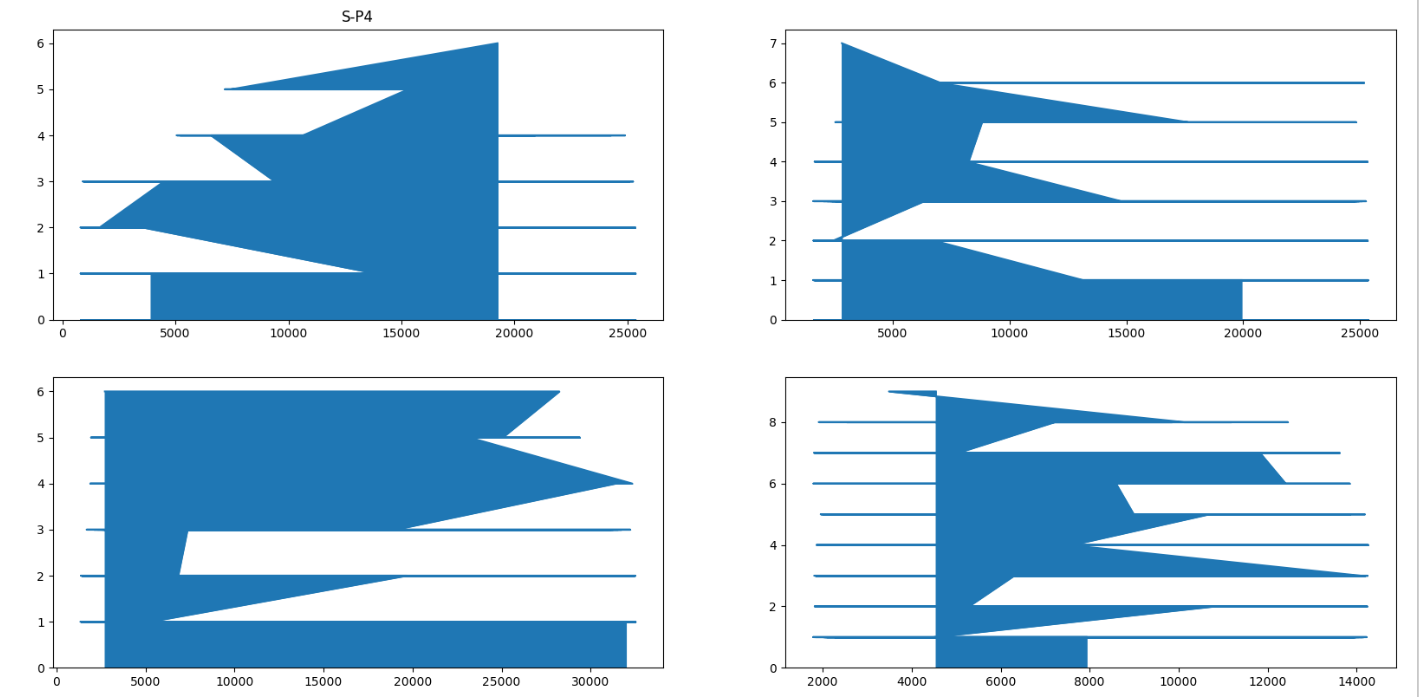
data['S-P3'].value\_counts().plot.area(ax=axrr[3][0])

ax=axrr[2][0]

ax.set\_title("S-P4")

data['S-P4'].value\_counts().plot.area(ax=axrr[3][1])**output:**





**Project Conclusion:**

In conclusion, product sales analysis is an indispensable tool for modern businesses, enabling them to harness the power of data to make informed decisions. By delving into sales trends, customer preferences, and performance metrics, organizations can optimize their strategies, enhance customer satisfaction, and ultimately achieve sustained growth and profitability.