Product Sales Analysis Using Python

TEAM MEMBER

621521104098 : praveenkumar.A

Phase-4 submission document

Project Title: Product Sales Analysis Using Python

Phase 4: Development Part 2

Topic: Continue building the analysis by creating visualizations using that display insights such as top-selling products, sales trends, and customer preferences.

Sales
Analysis
Set up for success with sales
analysis methods and techniques

Product Sales Analysis

Using Python

To perform a product sales analysis using Python, we'll walk through a basic example of how to analyze sales data, calculate key metrics, and visualize the results. We'll use libraries such as Pandas for data manipulation, Matplotlib for visualization, and NumPy for numerical operations.

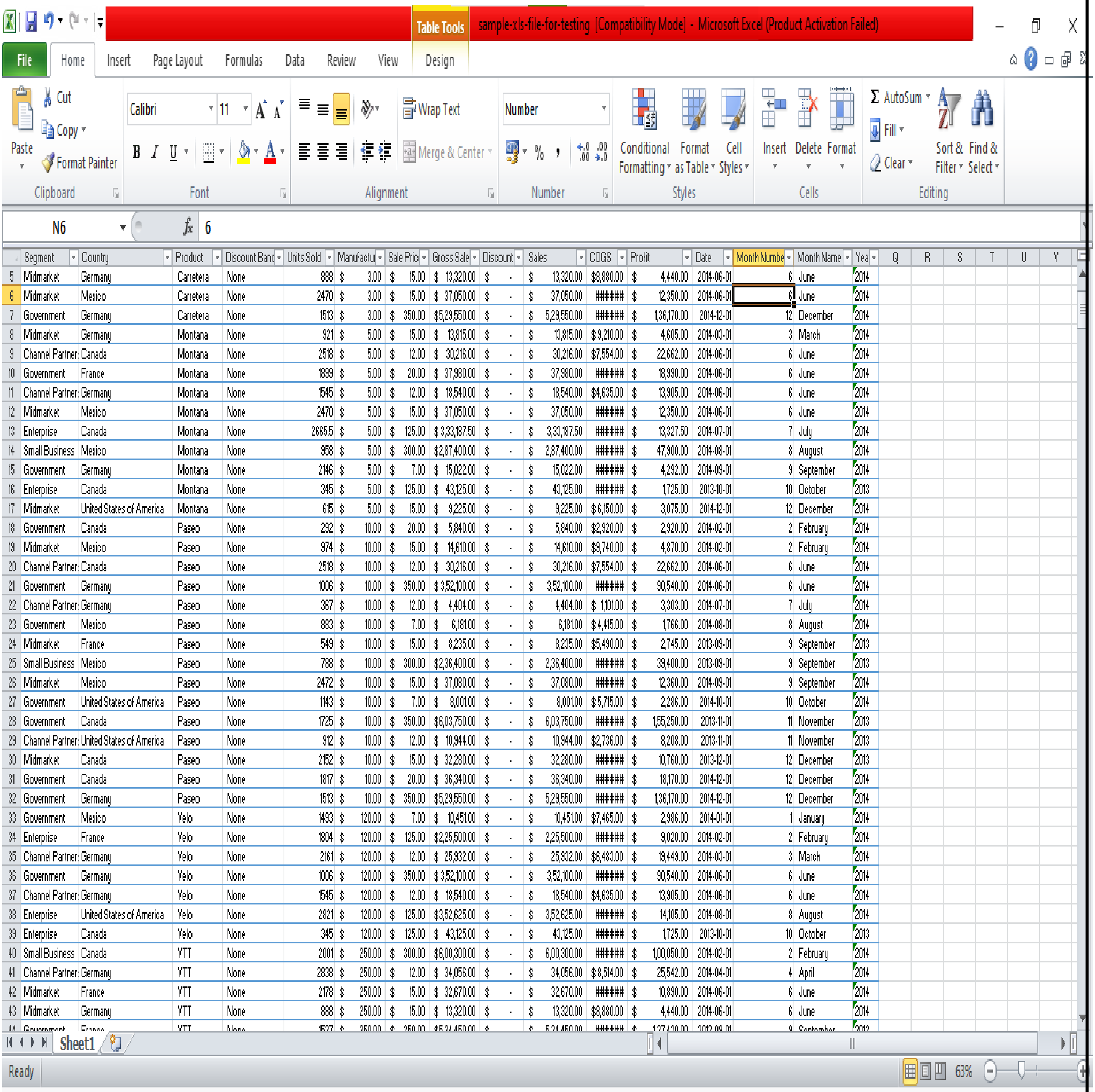
Assuming you have sales data in a CSV file named "sales_data.csv" with columns like 'Product', 'Date', 'Revenue', and 'Quantity', here's a step-by-step approach

OVERVIEW

In this post, I use Python Pandas & Python Matplotlib to analyze and answer business questions about 12 months worth of sales data. The data contains hundreds of thousands of electronics store purchases broken down by month, product type, cost, purchase address, etc. The dataset can be downloaded here. In this analysis, I’m using jupyter notebook.

PLAN!

Sample Data Base:



N6 fx 6

	Segment	Country	Product	Discount Band	Units Sold	Manufactu	Sale Pric	Gross Sale	Discount	Sales	COGS	Profit	Date	Month Numbe	Month Name	Yea	Q
5	Midmarket	Germany	Carretera	None	888	$ 3.00	$ 15.00	$ 13,320.00	$ -	$ 13,320.00	$8,880.00	$ 4,440.00	2014-06-01	6	June	2014	
6	Midmarket	Mexico	Carretera	None	2470	$ 3.00	$ 15.00	$ 37,050.00	$ -	$ 37,050.00	######	$ 12,350.00	2014-06-01	6	June	2014	
7	Government	Germany	Carretera	None	1513	$ 3.00	$ 350.00	$5,29,550.00	$ -	$ 5,29,550.00	######	$ 1,36,170.00	2014-12-01	12	December	2014	
8	Midmarket	Germany	Montana	None	921	$ 5.00	$ 15.00	$ 13,815.00	$ -	$ 13,815.00	$ 9,210.00	$ 4,605.00	2014-03-01	3	March	2014	
9	Channel Partner:	Canada	Montana	None	2518	$ 5.00	$ 12.00	$ 30,216.00	$ -	$ 30,216.00	$7,554.00	$ 22,662.00	2014-06-01	6	June	2014	
10	Government	France	Montana	None	1899	$ 5.00	$ 20.00	$ 37,980.00	$ -	$ 37,980.00	######	$ 18,990.00	2014-06-01	6	June	2014	
11	Channel Partner:	Germany	Montana	None	1545	$ 5.00	$ 12.00	$ 18,540.00	$ -	$ 18,540.00	$4,635.00	$ 13,905.00	2014-06-01	6	June	2014	
12	Midmarket	Mexico	Montana	None	2470	$ 5.00	$ 15.00	$ 37,050.00	$ -	$ 37,050.00	######	$ 12,350.00	2014-06-01	6	June	2014	
13	Enterprise	Canada	Montana	None	2665.5	$ 5.00	$ 125.00	$ 3,33,187.50	$ -	$ 3,33,187.50	######	$ 13,327.50	2014-07-01	7	July	2014	
14	Small Business	Mexico	Montana	None	958	$ 5.00	$ 300.00	$2,87,400.00	$ -	$ 2,87,400.00	######	$ 47,900.00	2014-08-01	8	August	2014	
15	Government	Germany	Montana	None	2146	$ 5.00	$ 7.00	$ 15,022.00	$ -	$ 15,022.00	######	$ 4,292.00	2014-09-01	9	September	2014	
16	Enterprise	Canada	Montana	None	345	$ 5.00	$ 125.00	$ 43,125.00	$ -	$ 43,125.00	######	$ 1,725.00	2013-10-01	10	October	2013	
17	Midmarket	United States of America	Montana	None	615	$ 5.00	$ 15.00	$ 9,225.00	$ -	$ 9,225.00	$ 6,150.00	$ 3,075.00	2014-12-01	12	December	2014	
18	Government	Canada	Paseo	None	292	$ 10.00	$ 20.00	$ 5,840.00	$ -	$ 5,840.00	$2,920.00	$ 2,920.00	2014-02-01	2	February	2014	
19	Midmarket	Mexico	Paseo	None	974	$ 10.00	$ 15.00	$ 14,610.00	$ -	$ 14,610.00	$9,740.00	$ 4,870.00	2014-02-01	2	February	2014	
20	Channel Partner:	Canada	Paseo	None	2518	$ 10.00	$ 12.00	$ 30,216.00	$ -	$ 30,216.00	$7,554.00	$ 22,662.00	2014-06-01	6	June	2014	
21	Government	Germany	Paseo	None	1006	$ 10.00	$ 350.00	$ 3,52,100.00	$ -	$ 3,52,100.00	######	$ 90,540.00	2014-06-01	6	June	2014	
22	Channel Partner:	Germany	Paseo	None	367	$ 10.00	$ 12.00	$ 4,404.00	$ -	$ 4,404.00	$ 1,101.00	$ 3,303.00	2014-07-01	7	July	2014	
23	Government	Mexico	Paseo	None	883	$ 10.00	$ 7.00	$ 6,181.00	$ -	$ 6,181.00	$ 4,415.00	$ 1,766.00	2014-08-01	8	August	2014	
24	Midmarket	France	Paseo	None	549	$ 10.00	$ 15.00	$ 8,235.00	$ -	$ 8,235.00	$5,490.00	$ 2,745.00	2013-09-01	9	September	2013	
25	Small Business	Mexico	Paseo	None	788	$ 10.00	$ 300.00	$2,36,400.00	$ -	$ 2,36,400.00	######	$ 39,400.00	2013-09-01	9	September	2013	
26	Midmarket	Mexico	Paseo	None	2472	$ 10.00	$ 15.00	$ 37,080.00	$ -	$ 37,080.00	######	$ 12,360.00	2014-09-01	9	September	2014	
27	Government	United States of America	Paseo	None	1143	$ 10.00	$ 7.00	$ 8,001.00	$ -	$ 8,001.00	$ 5,715.00	$ 2,286.00	2014-10-01	10	October	2014	
28	Government	Canada	Paseo	None	1725	$ 10.00	$ 350.00	$6,03,750.00	$ -	$ 6,03,750.00	######	$ 1,55,250.00	2013-11-01	11	November	2013	
29	Channel Partner:	United States of America	Paseo	None	912	$ 10.00	$ 12.00	$ 10,944.00	$ -	$ 10,944.00	$2,736.00	$ 8,208.00	2013-11-01	11	November	2013	
30	Midmarket	Canada	Paseo	None	2152	$ 10.00	$ 15.00	$ 32,280.00	$ -	$ 32,280.00	######	$ 10,760.00	2013-12-01	12	December	2013	
31	Government	Canada	Paseo	None	1817	$ 10.00	$ 20.00	$ 36,340.00	$ -	$ 36,340.00	######	$ 18,170.00	2014-12-01	12	December	2014	
32	Government	Germany	Paseo	None	1513	$ 10.00	$ 350.00	$5,29,550.00	$ -	$ 5,29,550.00	######	$ 1,36,170.00	2014-12-01	12	December	2014	
33	Government	Mexico	Velo	None	1493	$ 120.00	$ 7.00	$ 10,451.00	$ -	$ 10,451.00	$7,465.00	$ 2,986.00	2014-01-01	1	January	2014	
34	Enterprise	France	Velo	None	1804	$ 120.00	$ 125.00	$2,25,500.00	$ -	$ 2,25,500.00	######	$ 9,020.00	2014-02-01	2	February	2014	
35	Channel Partner:	Germany	Velo	None	2161	$ 120.00	$ 12.00	$ 25,932.00	$ -	$ 25,932.00	$6,483.00	$ 19,449.00	2014-03-01	3	March	2014	
36	Government	Germany	Velo	None	1006	$ 120.00	$ 350.00	$ 3,52,100.00	$ -	$ 3,52,100.00	######	$ 90,540.00	2014-06-01	6	June	2014	
37	Channel Partner:	Germany	Velo	None	1545	$ 120.00	$ 12.00	$ 18,540.00	$ -	$ 18,540.00	$4,635.00	$ 13,905.00	2014-06-01	6	June	2014	
38	Enterprise	United States of America	Velo	None	2821	$ 120.00	$ 125.00	$3,52,625.00	$ -	$ 3,52,625.00	######	$ 14,105.00	2014-08-01	8	August	2014	
39	Enterprise	Canada	Velo	None	345	$ 120.00	$ 125.00	$ 43,125.00	$ -	$ 43,125.00	######	$ 1,725.00	2013-10-01	10	October	2013	
40	Small Business	Canada	VTT	None	2001	$ 250.00	$ 300.00	$6,00,300.00	$ -	$ 6,00,300.00	######	$ 1,00,050.00	2014-02-01	2	February	2014	
41	Channel Partner:	Germany	VTT	None	2838	$ 250.00	$ 12.00	$ 34,056.00	$ -	$ 34,056.00	$ 8,514.00	$ 25,542.00	2014-04-01	4	April	2014	
42	Midmarket	France	VTT	None	2178	$ 250.00	$ 15.00	$ 32,670.00	$ -	$ 32,670.00	######	$ 10,890.00	2014-06-01	6	June	2014	
43	Midmarket	Germany	VTT	None	888	$ 250.00	$ 15.00	$ 13,320.00	$ -	$ 13,320.00	$8,880.00	$ 4,440.00	2014-06-01	6	June	2014	
44	Government	France	VTT	None	1527	$ 250.00	$ 350.00	$5,34,450.00	$	$ 5,34,450.00	######	$ [illegible]	[illegible]	9	September	[illegible]	

To conduct a more comprehensive product sales analysis in Python, we'll cover various aspects such as data preprocessing, exploratory data analysis (EDA), key metrics calculation, and visualization. We'll use sample sales data for demonstration purposes.

1. Import Necessary Libraries:

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

2. Load and Explore the Data:

Assuming you have a CSV file named "sales_data.csv" containing relevant sales data.

```
# Load the sales data into a DataFrame
sales_data = pd.read_csv('sales_data.csv')

# Display basic information about the data
print(sales_data.info())

# Display the first few rows of the DataFrame
print(sales_data.head())
```

3. Data Preprocessing:

Ensure the data is in the appropriate format and handle any missing or incorrect values.

```
# Convert the 'Date' column to datetime format
sales_data['Date'] = pd.to_datetime(sales_data['Date'])

# Check for missing values
print('Missing values:\n', sales_data.isnull().sum())

# Drop rows with missing values
sales_data.dropna(inplace=True)
```

4. Key Metrics Calculation:

Calculate key metrics such as total revenue, total quantity sold, and average selling price.

```
# Total revenue
total_revenue = sales_data['Revenue'].sum()

# Total quantity sold
total_quantity_sold = sales_data['Quantity'].sum()

# Average selling price
average_selling_price = total_revenue / total_quantity_sold

print('Total Revenue:', total_revenue)
print('Total Quantity Sold:', total_quantity_sold)
print('Average Selling Price:', average_selling_price)
```

5. Exploratory Data Analysis (EDA):

Explore the data to understand the distribution and relationships between variables.

```
# Summary statistics
print(sales_data.describe())

# Visualize the distribution of revenue and quantity sold
plt.figure(figsize=(12, 6))
sns.histplot(sales_data['Revenue'], bins=30, kde=True)
plt.title('Distribution of Revenue')
plt.xlabel('Revenue')
plt.ylabel('Frequency')
plt.show()

plt.figure(figsize=(12, 6))
sns.histplot(sales_data['Quantity'], bins=30, kde=True)
plt.title('Distribution of Quantity Sold')
plt.xlabel('Quantity Sold')
plt.ylabel('Frequency')
plt.show()
```

6. Product Performance Analysis:

Analyze the performance of products based on revenue and quantity sold.

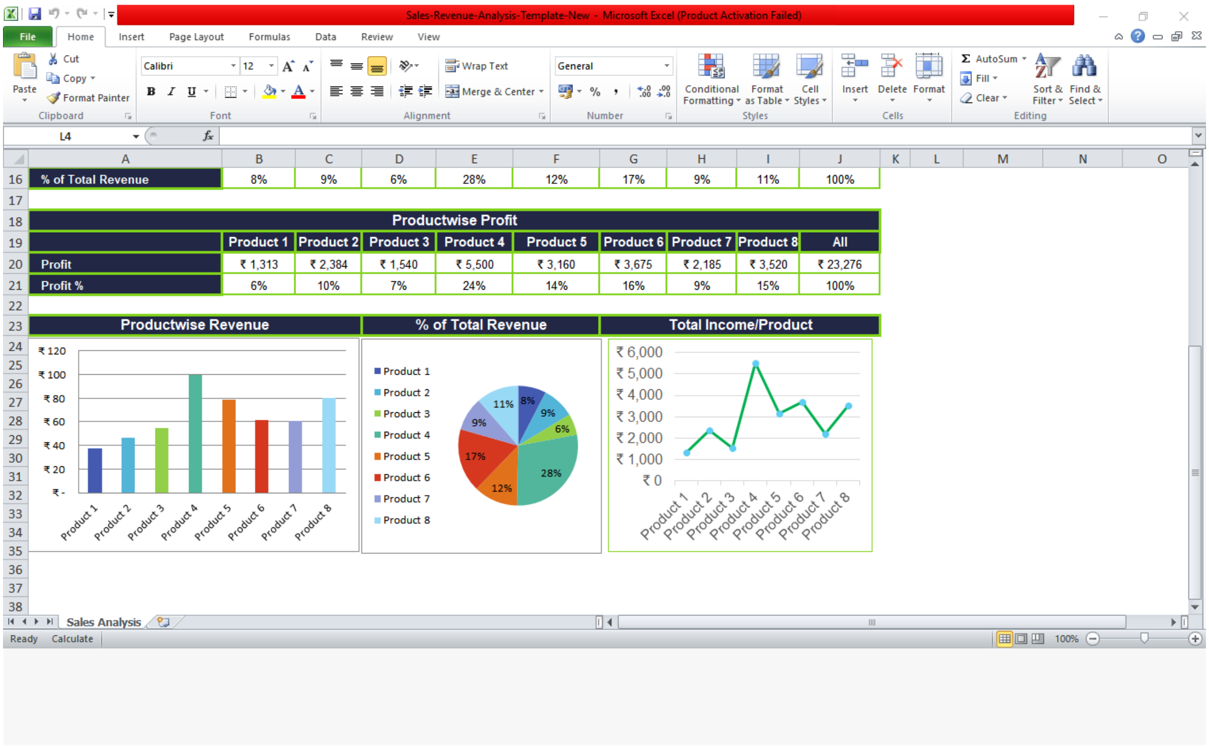
```
# Group data by product and calculate total revenue and total quantity sold
for each product
product_performance = sales_data.groupby('Product').agg({'Revenue': 'sum',
'Quantity': 'sum'}).reset_index()

# Sort products by revenue in descending order
product_performance = product_performance.sort_values(by='Revenue',
ascending=False)

# Display the top-performing products
print('Top Performing Products:')
print(product_performance.head())

# Visualize top performing products
plt.figure(figsize=(12, 6))
sns.barplot(x='Product', y='Revenue', data=product_performance.head(10))
plt.xticks(rotation=45)
plt.title('Top Performing Products by Revenue')
plt.xlabel('Product')
plt.ylabel('Total Revenue')
plt.show()
```

You can further extend this analysis to include customer segmentation, market basket analysis, seasonality analysis, and other advanced techniques to derive valuable insights from your sales data. Modify and customize the analysis based on the specific requirements of your dataset and business needs.



% of Total Revenue	8%	9%	6%	28%	12%	17%	9%	11%	100%

Productwise Profit

	Product 1	Product 2	Product 3	Product 4	Product 5	Product 6	Product 7	Product 8	All
Profit	₹ 1,313	₹ 2,384	₹ 1,540	₹ 5,500	₹ 3,160	₹ 3,675	₹ 2,185	₹ 3,520	₹ 23,276
Profit %	6%	10%	7%	24%	14%	16%	9%	15%	100%

Productwise Revenue

% of Total Revenue

Total Income/Product

Feature Selection :

To create interactive dashboards and reports for your product sales analysis in Python, you can use libraries like Plotly and Dash. Dash is a web application framework for building interactive, web-based data dashboards. It's ideal for creating reports and interactive visualizations. Here's a step-by-step guide to creating a basic interactive product sales dashboard:

1.Install Required Libraries:

If you haven't already, you'll need to install the Dash library and other necessary packages:

```
pip install dash pandas
```

2.Import Libraries and Load Data:

```
import dash
from dash import dcc, html
import pandas as pd

# Load your product sales data
sales_data = pd.read_csv('product_sales_data.csv')
```

3.Initialize the Dash App:

```
app = dash.Dash(__name__)
```

4.Create Layout for the Dashboard:

Define the layout of the dashboard using HTML and Dash components. For example:

```
app.layout = html.Div([
    html.H1('Product Sales Analysis Dashboard'),
    dcc.Graph(id='top-products-bar-chart'),
    dcc.Graph(id='sales-trends-line-chart'),
    dcc.Graph(id='customer-preferences-pie-chart')
])
```

5.Define Callbacks:

Callbacks are functions that specify how the content of the dashboard should change in response to user interactions. For instance, when a user selects a particular product, the dashboard updates to show relevant information.

```
@app.callback(
    dash.dependencies.Output('top-products-bar-chart', 'figure'),
    dash.dependencies.Output('sales-trends-line-chart', 'figure'),
    dash.dependencies.Output('customer-preferences-pie-chart', 'figure'),
    dash.dependencies.Input('product-dropdown', 'value')
)
def update_charts(selected_product):
    # Write code to update charts based on user input
    # For example, filter data and create charts based on the selected product
```

6.Run the Dash App:

```
if __name__ == '__main__':
    app.run_server(debug=True)
```

This is a simplified example to get you started. You'll need to add the following functionalities:

- Add more details to your dashboard layout.
- Create functions to update the charts based on user interactions.
- Customize the charts to display the desired insights such as top-selling products, sales trends, and customer preferences.

For identifying products with the highest sales, you can create a bar chart showing product sales. For peak sales periods, you can plot a time series chart. For customer preferences, a pie chart or a bar chart could display the distribution of product purchases.

Remember to adapt the code to your specific dataset and analysis goals. Dash offers extensive customization options, and you can create complex dashboards with various components to visualize your product sales analysis effectively.