

ProductSalesAnalysisUsingPython

TEAMMEMBER

621521104098:Praveenkumar.A

Phase-3submissiondocument

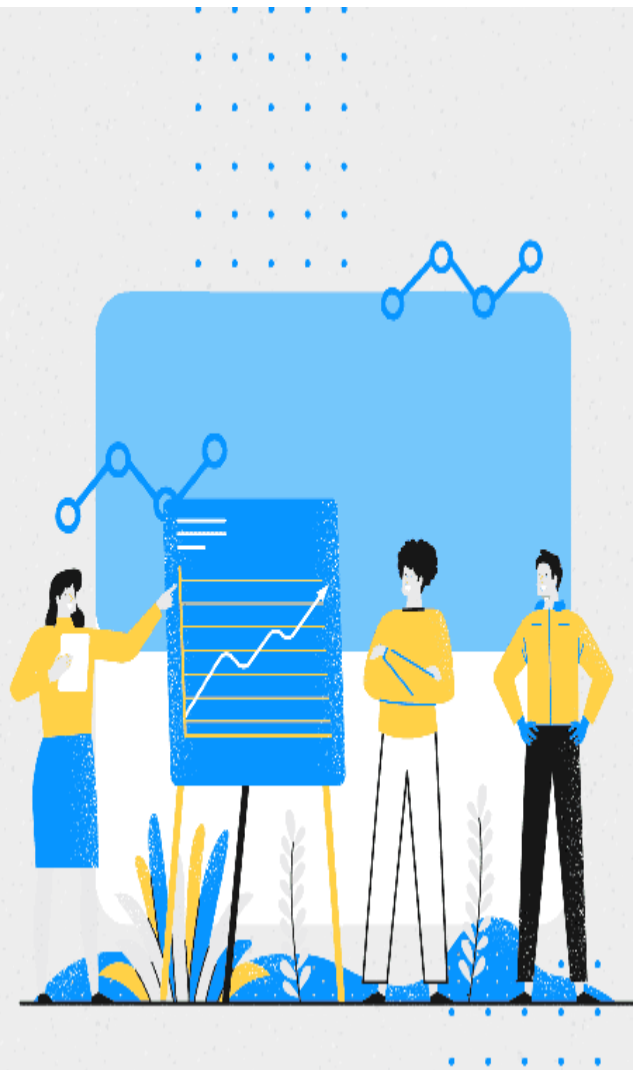
ProjectTitle:ProductSalesAnalysisUsingPython

Phase3: DevelopmentPart1

Topic: In this part you will begin building your project by loading and preprocessing the dataset.

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.



SampleDataBase:

sample.xls-file-for-testing [Compatibility Mode] - Microsoft Excel (Product Activation Failed)																							
Table Tools																							
File Home Insert Page Layout Formulas Data Review View Design																							
Clipboard Font Alignment Number Styles Cells Editing																							
N6 6																							
Segment	Country	Product	Discount Band	Units Sold	Manufactur	Sale Price	Gross Sale	Discount	Sales	COGS	Profit	Date	Month Number	Month Name	Year	Q	R	S	T	U	V		
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	\$ 35.00	\$ 52,955.00	\$ -	\$ 52,955.00	#####	\$ 136,170.00	2014-12-01	12	December	2014						
8	Midmarket	Germany	Montana	None	321	\$	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	\$ 17,554.00	\$ 22,662.00	2014-06-01	6	June	2014						
10	Government	France	Montana	None	1839	\$	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	\$ 4,835.00	\$ 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	26655	\$	5.00	\$ 125.00	\$ 3,331,875.00	\$ -	\$ 3,331,875.00	#####	\$ 13,327.50	2014-07-01	7	July	2014						
14	Small Business	Mexico	Montana	None	958	\$	5.00	\$ 300.00	\$ 2,874,000.00	\$ -	\$ 2,874,000.00	#####	\$ 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	\$ 3,920.00	\$ 2,920.00	2014-02-01	2	February	2014						
19	Midmarket	Mexico	Paseo	None	374	\$	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	\$ 17,554.00	\$ 22,662.00	2014-06-01	6	June	2014						
21	Government	Germany	Paseo	None	1006	\$	10.00	\$ 35.00	\$ 3,521,000.00	\$ -	\$ 3,521,000.00	#####	\$ 90,540.00	2014-06-01	6	June	2014						
22	Channel Partner:	Germany	Paseo	None	357	\$	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	\$ 2,364,000.00	\$ -	\$ 2,364,000.00	#####	\$ 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	\$ 35.00	\$ 6,037,500.00	\$ -	\$ 6,037,500.00	#####	\$ 155,250.00	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	\$ 7,336.00	\$ 3,608.00	2013-11-01	11	November	2013						
30	Midmarket	Canada	Paseo	None	2162	\$	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	\$ 35.00	\$ 52,955.00	\$ -	\$ 52,955.00	#####	\$ 136,170.00	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	\$ 2,255,000.00	\$ -	\$ 2,255,000.00	#####	\$ 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	\$ 8,493.00	\$ 19,439.00	2014-03-01	3	March	2014						
36	Government	Germany	Velo	None	1006	\$	120.00	\$ 35.00	\$ 3,521,000.00	\$ -	\$ 3,521,000.00	#####	\$ 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	\$ 4,835.00	\$ 13,905.00	2014-06-01	6	June	2014						
38	Enterprise	United States of America	Velo	None	2821	\$	120.00	\$ 125.00	\$ 3,526,250.00	\$ -	\$ 3,526,250.00	#####	\$ 14,105.00	2014-08-01	8	August	2014						
39	Enterprise	Canada	Velo	None	345	\$	120.00	\$ 43.125.00	\$ 14,925.00	\$ -	\$ 14,925.00	#####	\$ 1,725.00	2013-10-01	10	October	2013						
40	Small Business	Canada	VTT	None	2001	\$	250.00	\$ 300.00	\$ 6,003,000.00	\$ -	\$ 6,003,000.00	#####	\$ 100,050.00	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	\$ 35.00	\$ 534,450.00	\$ -	\$ 534,450.00	#####	\$ 177,450.00	2013-09-01	9	September	2013						

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
DataFramesales_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())
```

```
#Droprowswithmissingvalues
sales_data.dropna(inplace=True)
```

4.KeyMetricsCalculation:

Calculatekeymetricssuchastotalrevenue,totalquantitiesold,andaverage sellingprice.

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

#Totalquantitiesold
total_quantity_sold=sales_data['Quantity'].sum()

#Averagesellingprice
average_selling_price=total_revenue/total_quantity_sold

print('TotalRevenue:',total_revenue)
print('TotalQuantitySold:',total_quantity_sold)
print('AverageSellingPrice:',average_selling_price)
```

5.ExploratoryDataAnalysis(EDA):

Exploredatato understandthedistributionandrelationshipsbetweenvariables.

```
#Summarystatistics
print(sales_data.describe())

#Visualizethedistributionofrevenueandquantitiesold
```

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

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

6.ProductPerformanceAnalysis:

Analyzetheperformanceofproductsbasedonrevenueandquantitiesold.

```
#Groupdatabyproductandcalculatetotalrevenueandtotalquantitiesold
foreachproduct
product_performance=sales_data.groupby('Product').agg({'Revenue':'sum',
'Quantity':'sum'}).reset_index()
```

```
#Sortproductsbyrevenueindescendingorder
product_performance=product_performance.sort_values(by='Revenue',
ascending=False)
```

```
#Displaythetop-performingproducts
print('TopPerformingProducts:')
print(product_performance.head())
```

```
#Visualizetopperformingproducts
```



```
plt.figure(figsize=(12,6))
```

```
sns.barplot(x='Product',y='Revenue',data=product_performance.head(10))
```

```
plt.xticks(rotation=45)
```

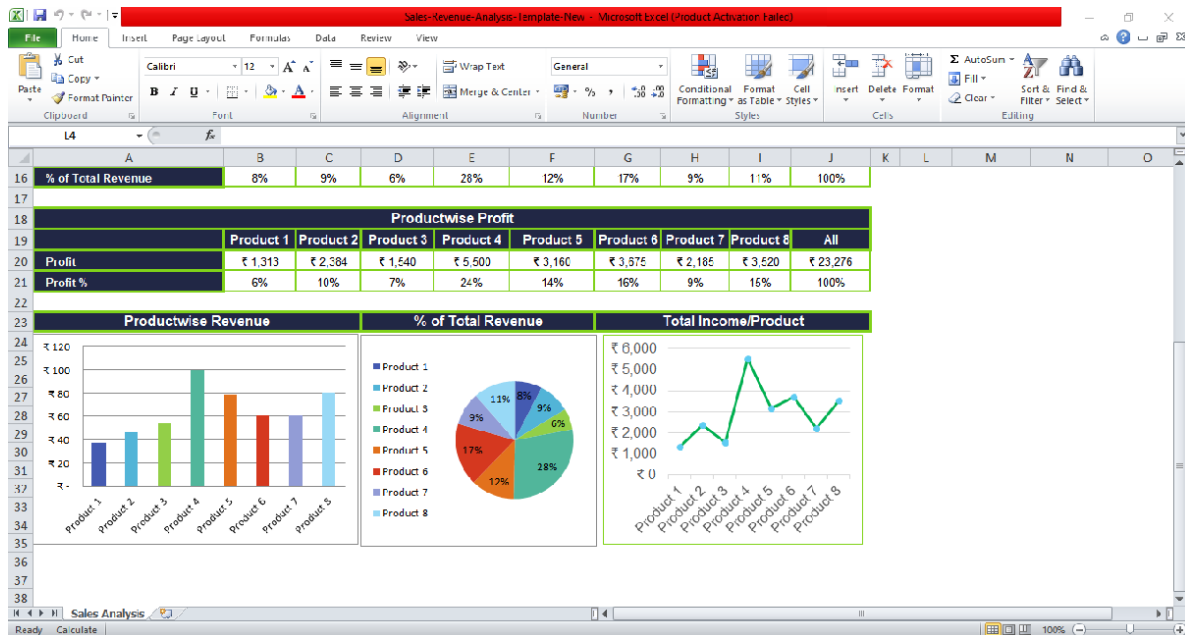
```
plt.title('TopPerformingProductsbyRevenue')
```

```
plt.xlabel('Product')
```

```
plt.ylabel('TotalRevenue')
```

```
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 specific requirements of your dataset and business needs.



Advantages of Sales Revenue Analysis

- It can be an effective tool for marketing and sales teams for achieving and defining targets.
- This analytics can be useful for new startups, online retail sales, or any other small business to track their sales and profits.
- From such analysis, you get insight to improve in areas where products and services aren't performing well. This helps to make informed decisions.
- Sales Revenue analysis helps us to determine profitability.
- You can design detailed and feasible plans for the future based on these data.
- Moreover, it helps the business to know where to invest and how to invest.
- Design marketing campaigns and allocate an appropriate budget for these activities.