

AFAME TECHNOLOGIES

Project 1: SALES DATA ANALYSIS

Goal:

Use sales data analysis to find patterns, best-selling items, and revenue indicators to help in business decision making.

Approach:

1. Data Collection and Cleaning:

- Gather the sales dataset, ensuring it includes relevant information such as item sold, quantity, price, and date of sale.
- Check for missing values, duplicates, and inconsistencies in the data. Clean the dataset to ensure accuracy in analysis.

2. Total Sales Calculation:

- Calculate the total sales revenue over the entire dataset period. This will provide an overview of the business's performance.

3. Sales Trends Analysis:

- Analyse sales trends over time by plotting a time series graph. This will help identify seasonal patterns, spikes, and dips in sales.
- Consider using moving averages or other smoothing techniques to visualise trends more clearly.

4. Best-Selling Products Identification:

- Determine the best-selling products by aggregating sales data for each item.
- Calculate the total quantity sold and revenue generated for each product.
- Rank products based on their sales performance.

5. Visualisation

- Create visualisations such as bar charts, pie charts, or line graphs to represent the data.
- Visualise total sales revenue, sales trends over time, and the distribution of sales across different products.
- Use colour coding or annotations to highlight key insights.

6. Insights and Recommendations:

- Interpret the findings from the analysis.
- Identify patterns, correlations, and outliers in the data.
- Provide actionable recommendations based on the insights gained.
- Suggest strategies to improve sales tactics, such as promoting top-selling products, optimising pricing, or adjusting inventory levels.

7. Tools and Technologies:

- Utilise programming languages like Python or R for data analysis and visualisation.
- Libraries such as pandas, matplotlib, and seaborn can be helpful for data manipulation and plotting.

SOLUTION:

We can follow these steps using Python and popular data analysis libraries like Pandas, Matplotlib, and Seaborn:

1. Load the sales dataset.
2. Explore the data to understand its structure and content.
3. Compute total sales.
4. Analyse sales trends over time.
5. Determine the best-selling products.
6. Build visualisations to convey the findings effectively.

Let's start by loading the dataset and exploring its structure:

```
1 import pandas as pd
2
3 # Load the dataset
4 ECOMM_DATA = pd.ExcelFile("ECOMM DATA.xlsx")
5
6 # Display the first few rows of the dataset
7 print(ECOMM_DATA.head())
8
9 # Get information about the dataset
10 print(ECOMM_DATA.info())
```

Once we understand the structure of the dataset, we can proceed with computing total sales:

```
12 # Compute total sales
13 total_sales = ECOMM_DATA['Revenue'].sum()
14 print("Total Sales: ${:,.2f}".format(total_sales))
```

Next, we can analyse sales trends over time. The dataset contains a Order date column, we can plot the sales over time:

```

16 import matplotlib.pyplot as plt
17
18 # Convert the date column to datetime format
19 ECOMM_DATA['Date'] = pd.to_datetime(ECOMM_DATA['Date'])
20
21 # Group sales by date and sum the revenue
22 daily_sales = ECOMM_DATA.groupby('Date')['Revenue'].sum()
23
24 # Plot sales trends over time
25 plt.figure(figsize=(10, 6))
26 plt.plot(daily_sales.index, daily_sales.values, marker='o', linestyle='-')
27 plt.title('Daily Sales Trends')
28 plt.xlabel('Date')
29 plt.ylabel('Revenue ($)')
30 plt.grid(True)
31 plt.show()

```

Then, we can determine the best-selling products by analyzing the sales quantity:

```

33 # Group sales by product and sum the quantity sold
34 best_selling_products = ECOMM_DATA.groupby('Product Name')['Quantity'].sum().sort_values(ascending=False)
35
36 print("Best-selling Products:")
37 print(best_selling_products.head(10))
38

```

Finally, let's build visualizations to convey our findings effectively. We can create a bar plot to visualize the best-selling products:

```

39 import seaborn as sns
40
41 # Plot the best-selling products
42 plt.figure(figsize=(10, 6))
43 sns.barplot(x=best_selling_products.head(10).values, y=best_selling_products.head(10).index, palette='viridis')
44 plt.title('Top 10 Best-selling Products')
45 plt.xlabel('Quantity Sold')
46 plt.ylabel('Product Name')
47 plt.show()
48

```

This process will provide insightful information about the sales data, helping in business decision-making.

