E-commerce project

Eng: Ahmed Noaman
EpsilionAi

Project Manager

MOHAMED OSAMA RASHED

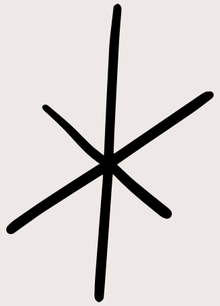
junior data analyst

Project Overview

provides a high-level summary of the project, including its purpose, scope, key objectives, methodology, and expected outcomes. Below is a template for a project overview based on your dataset and the context provided:

1. Project Purpose
2. Project Scope
3. Key Objectives
4. Methodology
5. Expected Outcomes
6. Tools and Technologies

Dataset Overview



RangeIndex: 1969 entries, 0 to 1968
Data columns (total 27 columns):

#	Column	Non-Null Count	Dtype
0	Row ID	1960 non-null	float64
1	Order Priority	1955 non-null	object
2	Discount	1960 non-null	float64
3	Unit Price	1960 non-null	float64
4	Shipping Cost	1960 non-null	float64
5	Customer ID	1960 non-null	float64
6	Customer Name	1958 non-null	object
7	Ship Mode	1954 non-null	object
8	Customer Segment	1960 non-null	object
9	Product Category	1960 non-null	object
10	Product Sub-Category	1960 non-null	object
11	Product Container	1960 non-null	object
12	Product Name	1960 non-null	object
13	Product Base Margin	1944 non-null	float64
14	Country	1960 non-null	object
15	Region	1960 non-null	object
16	State or Province	1955 non-null	object
17	City	1954 non-null	object
18	Postal Code	1952 non-null	float64
19	Order Date	1960 non-null	object
20	Ship Date	1960 non-null	object
21	Profit	1960 non-null	float64
22	Quantity ordered new	1960 non-null	float64
23	Sales	1960 non-null	float64
24	Order ID	1960 non-null	float64
25	Status	15 non-null	object
26	Manager	1960 non-null	object

dtypes: float64(11), object(16)

purpose

The purpose of this project is to analyze sales, customer, and product data to uncover insights that can help optimize business operations, improve customer satisfaction, and increase profitability. The analysis will focus on understanding key trends, identifying areas for improvement, and providing actionable recommendations.

Project Scope

The project will analyze a dataset containing 1960 rows and 26 columns, including:

- Sales Data: Sales, Profit, Discount, Unit Price, Shipping Cost, Quantity Ordered.
- Customer Data: Customer ID, Customer Name, Customer Segment, Region, Country, State, City, Postal Code.
- Product Data: Product Category, Product Sub-Category, Product Name, Product Container, Product Base Margin.
- Order Data: Order ID, Order Priority, Order Date, Ship Date, Ship Mode.
- Other Data: Manager, Status (mostly missing).

The analysis will cover:

- Sales performance by region, product category, and customer segment.
- Customer behavior and segmentation.
- Shipping and order fulfillment efficiency.
- Profitability analysis.
- Identification of key drivers of sales and profit.

Project Objectives

Sales Analysis

- Analyze total sales, profit, and discounts by region, product category, and customer segment.
- Identify top-performing products and regions.

Shipping and Order Fulfillment

- Evaluate shipping costs and delivery times by ship mode and region.
- Identify inefficiencies in the order fulfillment process.

Customer Analysis

- Segment customers based on purchasing behavior and demographics.
- Identify high-value customers and their preferences.

Profitability Analysis

- Analyze profit margins by product category and sub-category.
- Identify factors affecting profitability (e.g., discounts, shipping costs).

Data Cleaning and Preparation

- Handle missing values in columns like Order Priority, Customer Name, Ship Mode, and Product Base Margin.
- Convert data types (e.g., Order Date and Ship Date to datetime).

Methodology

1. Data Cleaning:
   - Handle missing values using imputation or placeholders.
   - Convert data types (e.g., dates to datetime, IDs to strings).
   - Drop irrelevant columns (e.g., Status).
2. Exploratory Data Analysis (EDA):
   - Perform univariate and multivariate analysis to understand distributions and relationships.
   - Visualize key metrics (e.g., sales, profit, discounts) using charts and graphs.
3. Customer Segmentation:
   - Use clustering or RFM (Recency, Frequency, Monetary) analysis to segment customers.
4. Sales and Profit Analysis:
   - Analyze sales and profit trends over time.
   - Identify correlations between discounts, shipping costs, and profitability.
5. Shipping and Order Analysis:
   - Evaluate shipping costs and delivery times by region and ship mode.
   - Identify opportunities to optimize shipping processes.
6. Reporting and Visualization:
   - Create dashboards and reports to present key insights.
   - Use tools like Tableau, Power BI, or Python libraries (e.g., plotly).

Expected Outcomes

1. Insights:
    - Identification of top-performing products, regions, and customer segments.
    - Understanding of key drivers of sales and profitability.
2. Recommendations:
    - Strategies to improve sales and profitability (e.g., targeted discounts, optimized shipping).
    - Actions to enhance customer satisfaction and retention.
3. Deliverables:
    - Cleaned and processed dataset.
    - Visualizations and dashboards.
    - Final report with actionable insights and recommendations.

Conclusion

Revenue: 1.9 M , Profit: 227.832 K , Cost: 25.434 K , N Of Customers: 1130
N Of Managers: 4 , N Of Orders: 1365 ,N Of Categories: 3 , N Of Sub-Category: 17
N Of Countries: 1 , N Of Regions: 4 , N Of Cities: 868 ,N Of States: 49

- Recommenditions to increase business performance, you can leverage insights from your dataset (e.g., sales, customer, and product data) to make data-driven decisions.

1. Focus on High-Value Customers
2. Optimize Product Offerings
3. Improve Customer Experience
4. Regional Performance Analysis
5. Discount and Pricing Strategy
6. Enhance Marketing Campaigns
7. Focus on Q3 and Q4 in year because there is no sales specially in 7,8,9,10,11 Months

Tools and Technologies

1. Programming Language: Python.
2. Libraries: Pandas, NumPy, plotly, streamlit.
3. Visualization Tools: plotly, streamlit.
4. Data Storage: CSV files.

Thank You

by Mohamed Osama Rashed