Customer Shopping Behavior Analysis

1. Project Overview

This project analyzes customer shopping behavior using transactional data from 3,900 purchases across various product categories. The goal is to uncover insights into spending patterns, customer segments, product preferences, and subscription behavior to guide strategic business decisions.

2. Dataset Summary

- Rows: 1000
- Columns: 14
- Key Features:
  - Customer Details (Transaction ID ,Customer_name , Age,)
  - Product details (Product_name , Category , Quantity , Rating)
  - Order Details (Unit_price , Sales_Channel ,Payment_type,Order_id,Order_date ,Unit_cost , Sales)

3. Exploratory Data Analysis using Python

We began with data preparation and cleaning in Python:

Data Loading: Imported the dataset using pandas.

Initial Exploration: Used df.info() to check structure and .describe() for summary statistics.

```
RangeIndex: 1000 entries, 0 to 999
Data columns (total 14 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   Transaction_ID   1000 non-null   object
 1   Customer_Name    1000 non-null   object
 2   Age              1000 non-null   int64
 3   product_name     1000 non-null   object
 4   Category         1000 non-null   object
 5   Rating           1000 non-null   float64
 6   Quantity         1000 non-null   int64
 7   Unit_Price       1000 non-null   float64
 8   Sales_Channel    1000 non-null   object
 9   Payment_Type     1000 non-null   object
 10  Order_Date       1000 non-null   object
 11  Order_ID         1000 non-null   int64
 12  Sales            1000 non-null   float64
 13  Unit_Cost        1000 non-null   float64
dtypes: float64(4), int64(3), object(7)
memory usage: 109.5+ KB
```

	Age	Rating	Quantity	Unit_Price	Order_ID	Sales	Unit_Cost
count	1000.000000	1000.000000	1000.000000	1000.00000	1.000000e+03	1000.000000	1000.000000
mean	40.978000	2.982400	4.899000	262.10684	5.496813e+08	1306.464680	184.965110
std	11.398923	1.155346	2.548719	216.82106	2.571334e+08	1419.232477	175.289311
min	22.000000	1.000000	1.000000	9.33000	1.029280e+08	9.330000	6.920000
25%	31.000000	2.000000	3.000000	81.73000	3.280740e+08	308.120000	56.670000
50%	41.000000	3.000000	5.000000	154.06000	5.566097e+08	770.300000	97.440000
75%	51.000000	4.000000	7.000000	421.89000	7.696945e+08	1748.800000	263.330000
max	60.000000	5.000000	9.000000	668.27000	9.955298e+08	6014.430000	524.960000

---GETTING INSIGHTS USING PYTHON

-Peak sales period in an year :

```
Month with Highest Sales: August
Peak Sales Period Month Wise:
Month
August       145564.06
February     132834.20
July         126921.27
January      110765.39
October      109346.71
December     105600.18
June         103713.79
September    100151.50
March         97910.39
November      95466.76
April         91800.77
May           86389.66
```

— Top 10 Customers who contributed in the sales

```
Top 10 Customers by Revenues:
Customer_Name
Faiyaz Ahuja       6116.46
Sumer Talwar       6014.43
Anahita Tella      6014.43
Jiya Vora          6014.43
Parinaaz Golla     6014.43
Mahika Sharma      6014.43
Tarini Bhatt       6014.43
Bhavin Bains       6014.43
Pari Lalla         6014.43
Samaira Mahal      5860.89
```

— Revenue by product category and Product Contribution % to Total Sales :

```
Revenue by Product Category:
Category
Electronics    823525.89
Fashion        482938.79
Name: Sales, dtype: float64
Product Contribution % to Total Sales:
product_name
Laptop         13.214534
Smartphone     12.090273
Keyboard       11.280409
Shoes          10.413928
Backpack        9.535616
```

4. Data Analysis using SQL (Business Transactions)

I performed structured analysis in PostgreSQL to answer key business questions:

---Top 5 products based on average rating :

	product_name character varying (64)	avg_rating numeric
1	Keyboard	3.23
2	Mouse	3.12
3	Laptop	3.08
4	Camera	3.03
5	Backpack	2.97

--- Most prefered payment type by the customer :

	payment_type character varying (64)	count bigint
1	Cash	225
2	Net Banking	204
3	UPI	199
4	Credit Card	188
5	Debit Card	184

--- Which product category contributed more to the company :

	category character varying (64)	sum numeric
1	Electronics	823525.89
2	Fashion	482938.79

5. Dashboard in Power BI

Finally, we built an interactive dashboard in **Power BI** to present insights visually

--Key Takeaway

This project demonstrates my ability to work across the full data analytics lifecycle, combining Python, SQL, and Power BI to deliver actionable business insights.