Customer Shopping Behaviour Analysis Report

1. Overview

This project analyzes customer shopping behaviour using a combination of **Python**, **PostgreSQL**, and **Power BI** to uncover patterns in purchasing, revenue generation, subscriptions, and customer demographics. The objective of the analysis is to transform raw transactional and customer data into actionable business insights that support strategic decision-making.

The project demonstrates an end-to-end analytics workflow, including data extraction using SQL, data processing and validation using Python, and interactive visualization using Power BI. The results are presented in a professional format suitable for portfolio review and technical interviews.

2. Business Questions

The analysis was structured around the following key business questions:

- How many customers are in the dataset, and what is their overall purchasing behaviour?
- What is the average purchase amount and average customer review rating?
- What proportion of customers are subscribed versus non-subscribed?
- Which product categories generate the highest revenue and sales volume?
- How does customer behaviour vary across different age groups?
- What role do subscription status, gender, and shipping type play in purchasing patterns?

These questions reflect real-world concerns related to customer segmentation, revenue optimization, and marketing strategy.

3. Methodology

Data Extraction (PostgreSQL)

- Customer and transaction data were queried using **PostgreSQL**.
- SQL queries utilized `JOIN`, `GROUP BY`, `ORDER BY`, and aggregate functions such as `SUM()`, `COUNT()`, and `AVG()` to calculate KPIs and segment-level metrics.
- Query outputs were validated and exported for further analysis.

Data Processing (Python)

- Python was used for data inspection, cleaning, and validation.
- Libraries such as **pandas** and **numpy** were applied to ensure data consistency and accuracy.
- Python scripts helped verify SQL results and prepare structured datasets for visualization.

Data Visualization (Power BI)

- A Power BI dashboard was created to present insights interactively.

- Key KPIs (number of customers, average purchase amount, average rating) were displayed using KPI cards.
- Bar charts, donut charts, and slicers were used to analyze revenue, sales, subscription status, age groups, and product categories.
- Filters enabled dynamic analysis by gender, category, subscription status, and shipping type.

4. Key Findings

- **Customer Base**: The dataset contains approximately 3.9K customers, indicating a sizable and diverse customer population.
- **Purchase Behaviour**: The average purchase amount is around $59.76, with an average review rating of 3.75, suggesting moderate customer satisfaction.
- **Subscription Status**: A majority of customers are non-subscribers, while a smaller but significant segment holds active subscriptions.
- **Category Performance**: Clothing generates the highest revenue and sales volume, followed by accessories. Footwear and outerwear contribute comparatively less.
- **Age Group Trends**: Young adults and middle-aged customers drive the majority of revenue and sales, indicating strong purchasing power in these segments.

5. Insights

- **Subscription Opportunity**: Since most customers are non-subscribers, targeted subscription-based incentives could significantly improve recurring revenue.
- **Category Focus**: High-performing categories such as clothing should be prioritized for promotions and inventory planning.
- **Customer Segmentation**: Young adults represent a high-value segment and should be the focus of personalized marketing campaigns.
- **Operational Strategy**: Shipping preferences and category-level performance can guide logistics optimization and pricing strategies.
- **Dashboard Value**: The Power BI dashboard enables stakeholders to quickly identify trends and make data-driven decisions without deep technical knowledge.

6. Conclusion

This customer shopping behaviour analysis project highlights the effective use of **PostgreSQL, Python, and Power BI** to deliver business-relevant insights. The project demonstrates strong technical skills in SQL querying, data manipulation, and dashboard development, along with the ability to communicate insights clearly.

For interviewers and recruiters, this project showcases: - End-to-end data analytics capability - Practical experience with relational databases and Python - Strong data visualization and storytelling skills - A business-oriented approach to analytics

Overall, this project serves as a comprehensive example of how data analytics can support customer understanding and strategic decision-making in a retail context.