# Data Science Assignment: Ecommerce Transactions Dataset

## 1. Introduction

**Dataset Overview**:
The E-Commerce Transactions dataset consists of three key files: Customers.csv, Products.csv, and Transactions.csv. The goal of the analysis is to extract meaningful insights that can aid business decision-making and help optimize marketing, sales strategies, and inventory management.

- **Customers.csv**: Contains customer demographic information such as unique ID, name, region, and signup date.
- **Products.csv**: Contains details about the products sold, such as product ID, name, category, and price.
- **Transactions.csv**: Contains transactional data, linking customers to products purchased, with details like transaction ID, customer ID, product ID, quantity, price, and total value.

**Purpose of Analysis**:
This report focuses on **Exploratory Data Analysis (EDA)** to gain actionable business insights. These insights will assist the business in understanding customer behaviours, product sales trends, and the overall Ecommerce environment.

---

## 2. Methodology

To analyze the data, we followed these main steps:

1. **Data Cleaning**:
   - We checked for and removed any duplicate or missing data.
   - We made sure the dates were correctly formatted (like customer signup dates and transaction dates).
2. **Exploratory Data Analysis (EDA)**:
   - We calculated basic statistics like total sales and average prices.
   - We created visual representations of the data, such as charts and graphs, to better understand patterns and trends.
   - We looked at customer distribution by region, spending behavior, product preferences, and how sales varied by month.
3. **Insights Extraction**:
   - We reviewed the data and came up with five key insights that could help the business make informed decisions.

---

## 3. Key Insights

Here are the five key insights from our analysis:

**Insight 1: Region-based Customer Distribution**

- **What we found**: Most customers are concentrated in **Region X** and **Region Y**, while other regions have fewer customers.

- **What it means for the business**: These two regions are key markets. We could focus marketing efforts here to get more sales. At the same time, we should also try to attract customers from other regions by offering targeted promotions.

**Insight 2: Best-Selling Products**

- **What we found**: **Product Z**, **Product A**, and **Product B** are the top-selling products.
- **What it means for the business**: These products should be prioritized in terms of restocking and promoting. We need to keep these products available and market them regularly to drive sales.

**Insight 3: Seasonal Revenue Trends**

- **What we found**: Sales tend to increase in **November** and **December**, likely due to the holiday season and special shopping events.
- **What it means for the business**: We should prepare for these sales spikes by increasing inventory and planning marketing campaigns for these months to capture more sales.

**Insight 4: Customer Spending Behaviour**

- **What we found**: Most customers spend only a small amount, but a small group of high-value customers contribute significantly to the total revenue.
- **What it means for the business**: We need to identify and focus on high-value customers. Offering them special deals or loyalty programs could increase their spending and ensure they keep buying from us.

**Insight 5: Product Categories with High Demand**

- **What we found**: The **Electronics** category has the highest demand, followed by **Home Appliances**.
- **What it means for the business**: These categories should be prioritized when purchasing inventory. We should focus on offering more products in these categories, and tailor promotions around them.

---

## 4. Conclusion

The **Exploratory Data Analysis (EDA)** provided several important insights for the business:

1. **Focus on key regions**: Regions X and Y are vital markets, and we should try to grow the customer base in other regions.
2. **Prioritize top-selling products**: Product Z, A, and B should be the main focus for inventory and promotions.
3. **Plan for seasonal trends**: Sales spike in November and December, so we need to prepare in advance for these periods.
4. **Segment high-value customers**: A small group of customers generate most of the revenue, so they should be offered special loyalty programs.
5. **Focus on popular categories**: Electronics and Home Appliances are in high demand, so we should focus our efforts there.

These insights can guide the business in making more strategic decisions, improving sales, and focusing on key areas for growth.