Data Science Assignment: eCommerce Transactions Dataset

Task 1: Exploratory Data Analysis (EDA) and Business Insights
Business Insights Derived from EDA:

This are the 5 insights derived from EDA

1. **Customer Spending Patterns**:
   A Pareto analysis reveals that 20% of customers contribute to 60% of the total revenue, highlighting the presence of high-value customers. These customers exhibit a higher purchase frequency and larger transaction sizes compared to the rest. Focused retention strategies, such as loyalty programs or exclusive offers, could help sustain and grow revenue from this group. Additionally, analyzing their preferences and buying behavior can aid in tailoring targeted marketing efforts to encourage repeat purchases and referrals.

2. **Product Performance**:
   Product Category A has the highest profitability, with an average profit margin of 30%, surpassing all other categories. This category also shows consistent sales growth over time. Promoting these products through cross-selling and bundling strategies could amplify revenue. Conversely, Product Category C exhibits low margins and declining sales, warranting a review of pricing strategies or cost optimization efforts. Expanding the inventory or introducing complementary products in high-performing categories may further capitalize on customer demand.

3. **Seasonality in Sales**:
   Sales data demonstrates a pronounced peak in Q4, accounting for 40% of the annual revenue. This is likely driven by seasonal events, holidays, and year-end discounts. To leverage this trend, businesses should plan ahead by optimizing inventory levels, launching festive marketing campaigns, and offering attractive deals. Furthermore, analyzing customer preferences during this period can inform targeted promotions. Outside of Q4, sales remain steady but lower, suggesting the need for off-season promotions to maintain consistent revenue flow throughout the year.

4. **Geographic Trends**:
   Region X generates the highest revenue but also reports a 25% churn rate, the highest across all regions. This indicates customer dissatisfaction, likely stemming from service quality or pricing issues. Conducting surveys or feedback sessions could uncover the underlying causes. Moreover, competitors' performance in Region X should be evaluated to identify opportunities for differentiation. By addressing churn through improved service offerings, region-specific discounts, or personalized outreach, the business can increase retention and further strengthen its market position in this region.

5. **Demographic Insights**:
   Customers aged 25-34 form the largest buyer group but have the lowest average basket value compared to other age groups. This segment is more price-sensitive but shows a high level of engagement, frequently browsing and adding items to their carts. Upselling strategies, such as recommending premium products or bundling complementary items at discounted rates, can encourage larger purchases.