(EDA) questions

* Stage 1: Basic Descriptive Analysis

Focus on understanding the dataset’s structure, identifying basic trends, and preparing the data for more in-depth analysis.

1. What are the basic statistics for sales, profit, discount, and quantity?
2. What is the distribution of categorical variables (e.g., Ship Mode, Segment, Category)?
3. What are the overall sales trends over time (monthly, quarterly, yearly)?
4. Which products generate the most and least sales?
5. Which regions, states, and cities have the highest and lowest sales?
6. What are the most common shipping modes used?

* Stage 2: Intermediate Analysis

Start looking at trends, correlations, and customer behaviors

1. What is the relationship between sales and profit (Sales vs. Profit analysis)?
2. How do sales differ across customer segments (Consumer, Corporate, Home Office)?
3. How do sales trends vary by region, category, and sub-category?
4. How does the frequency of orders change over time for different customer segments?
5. What is the average order processing time (Order Date vs. Ship Date), and how does it vary by shipping mode?
6. How do discounts impact profit margins (Discount vs. Profit analysis)?
7. Are there any outliers in sales, profit, or discount data that skew the overall analysis?

* Stage 3: Advanced Analysis

Focus on deeper insights, such as customer segmentation, product affinity

1. What are the top customers based on Recency, Frequency, and Monetary (RFM) analysis?
2. Which customer segments (using RFM) are the most valuable, and how can they be targeted for retention?
3. Which products are frequently bought together (Market Basket Analysis), and what cross-selling opportunities exist?
4. What are the most profitable product categories and sub-categories?
5. How do shipping delays impact customer satisfaction or sales performance?