Sales Data Analysis (2019)

Exploratory Data Analysis (EDA)

1. Data Overview
   - How many unique products are there in the dataset?
   - What is the range of order dates in the dataset?
   - Are there any missing or null values? If yes, how will you handle them?
   - Are there any duplicate rows in the dataset? If yes, should they be removed?

2. Basic Statistics
   - What is the average quantity ordered for each product?
   - What is the total revenue generated from all sales?
   - What is the minimum, maximum, and average price of the products sold?

3. Data Cleaning
   - Identify and clean invalid data in the Order Date column (e.g., non-datetime values).
   - Extract useful information from Order Date, such as month, day, and hour.
   - Split Purchase Address into components like city, state, and ZIP code for analysis.

Sales Performance Analysis

4. Best Month for Sales
   - Identify the month with the highest sales revenue.
   - Visualize sales revenue by month using a bar plot or line chart.

5. Best Time for Display Advertisement
   - Analyze the hours with the highest number of orders.
   - Suggest the best time to display advertisements to maximize sales.

6. Most Product-Selling City
   - Which city has the highest sales revenue?
   - Visualize city-wise sales using a bar plot or map-based visualization.

Product-Specific Analysis

7. Most Sold Products
   - Identify the top 5 most sold products by quantity.
   - What could be the possible reasons for these products' popularity? (e.g., price, seasonality, bundling)

8. Products Most Often Sold Together
    - Find product pairs that are frequently sold together in the same order.
    - Suggest potential bundling strategies based on the insights.

9. Price vs. Quantity Relationship
    - Analyze the relationship between product price and quantity ordered.
    - Are lower-priced items sold in higher quantities?

Customer Insights

10. Most Active Customers
    - Identify customers (based on Purchase Address) who have made the highest number of purchases or generated the highest revenue.
    - What percentage of revenue do the top 10% of customers contribute?

11. City-Wise Product Preferences
    - Which products are most popular in different cities?
    - Are there city-specific trends that might indicate localized preferences?

12. Time-Based Trends
    - Identify weekly and daily patterns in sales.
    - What are the busiest days of the week for sales?

Expected Deliverables

1. Jupyter Notebook/Script
    - Code for EDA, analysis, and visualization.

2. Report
    - Key insights and actionable recommendations.