Phase 2: Analysis Questions Phase

Objective:

The goal of this phase is to identify relevant analysis questions that can be answered using the available dataset. These questions will provide insights to the organization's decision-makers regarding sales performance, customer behavior, and profitability.

Dataset Overview:

The dataset contains various attributes related to orders, customers, products, sales, discounts, and profits. The key columns in the dataset are:

- **OrderID**: Unique identifier for each order.

- **OrderDate**: Date when the order was placed.

- **ShipDate**: Date when the order was shipped.

- **ShipMode**: Shipping mode (e.g., Standard Class, Second Class).

- **CustomerID**: Unique identifier for each customer.

- **Segment**: Customer segment (e.g., Consumer, Corporate).

- **Region**: Geographical region of the customer (e.g., West, South).

- **Category**: Product category (e.g., Furniture, Office Supplies).

- **Sales**: Total sales amount for each order.

- **Discount**: Discount applied to each order.

- **Profit**: Profit generated from each sale.

Areas of Analysis:

To guide the organization's strategy and decision-making, the following areas of analysis have been identified:

---

1. Sales Performance:

This analysis focuses on understanding which products, categories, and regions are driving the most sales.

Potential Questions:

- **What product categories and sub-categories generate the highest and lowest sales?**

    - Investigating sales trends by **Category** and **Sub-Category**.

- **Which regions are the top contributors to total sales?**

    - Analyzing the **Region** and **Sales** data to see geographic performance.

- **What are the top-performing products based on sales in each region?**

    - Analyzing the sales distribution across products in different regions.

Insights Expected:

- Identifying the best-selling products and regions will help allocate resources and focus marketing efforts where they are most effective.

2. Profitability Analysis:

Profit is one of the most crucial indicators of success. This analysis looks into what factors contribute to high or low profits.

Potential Questions:

- **Which products, categories, or sub-categories are the most/least profitable?**

  - Analyzing **Profit** with respect to different product groups.

- **What impact do discounts have on profitability?**

  - Investigating the relationship between **Discount** and **Profit** to see if excessive discounts reduce profitability.

- **Are there any regions or states where profitability is consistently low?**

  - Analyzing profit trends by **Region** and **State** to identify areas of concern.

Insights Expected:

- Understanding profitability by category and region will help in making strategic decisions around pricing, discounting, and resource allocation.

3. Customer Behavior:

Analyzing customer behavior helps the organization understand which customer segments are the most valuable.

Potential Questions:

- **Which customer segments contribute the most to overall sales?**
  - Investigating the contribution of different **Segments** (e.g., Consumer, Corporate, Home Office) to total sales.

- **Which customers are the most valuable in terms of repeat orders and revenue?**
  - Using **CustomerID** and **Sales** to identify high-value customers.

- **Do different customer segments have different purchasing habits?**
  - Analyzing which segments buy certain product categories and how much they spend.

Insights Expected:

- Insights from customer behavior will assist in targeted marketing efforts, retention strategies, and customer segmentation improvements.

---

4. Shipping and Delivery Performance:

The speed and cost of shipping are key factors that influence customer satisfaction and profitability.

Potential Questions:

- **What impact does the shipping mode have on profitability and sales?**
    - Analyzing the relationship between **Ship Mode**, **Sales**, and **Profit**.
- **Is there a significant relationship between shipping delays and lower profits?**
    - Investigating the difference between **OrderDate** and **ShipDate** and its correlation with **Profit**.

Insights Expected:

- Understanding how shipping affects sales and profitability will help optimize delivery processes and reduce potential losses.

---

5. Time-Series Analysis:

This analysis focuses on understanding how sales and profit trends evolve over time, helping to predict future performance.

Potential Questions:

- **What are the sales and profit trends over time (monthly, quarterly, yearly)?**
    - Using **OrderDate** to analyze how sales and profits fluctuate over time.
- **Are there specific periods (e.g., holidays, quarters) when sales or profits spike?**
    - Identifying trends and peaks in sales or profit during certain time periods.

Insights Expected:

- Time-series insights will help the organization plan inventory, promotions, and pricing strategies according to seasonal trends.

Deliverables:

1. **Set of Analysis Questions**:

   - A comprehensive set of analysis questions derived from the dataset that will guide further analysis and reporting. These questions should cover multiple areas such as sales performance, profitability, customer behavior, and shipping.

2. **Data to Support Analysis**:

   - A breakdown of the data attributes that will be used to answer each analysis question. For example, sales by product category, profit by region, etc.

3. **Visualization Plans (Optional)**:

   - Potential visualizations to represent the analysis questions (e.g., bar charts, scatter plots, time-series graphs). These visualizations can be created in later phases to communicate findings effectively.

Conclusion:

In this phase, we have identified key areas of analysis that can provide actionable insights to decision-makers. These questions will guide the next steps of data analysis using SQL

and Python, helping to uncover trends and patterns that will influence the company's strategies related to sales, profitability, and customer engagement.