**Graduation Project Documentation**  
*Dataset Used: Super Store Dataset*

**Business Objectives:**

Doing a complete analysis of dataset to identify trends, pattern and maps. And determine the impact of key performance indicators (KPIs) on sales performance and provide the target audience with recommendations.

**The target audience:**

Sales, marketing and production teams and stakeholders like investors or partners.

**PHASE ONE: Data Preparation**

1. **Tables Created**:
   * **Customers Table**:  
     Columns include:
     + Order ID (Foreign Key)
     + Customer ID (Primary Key)
     + Customer Name
     + Segment
     + Country, City, State, and Region  
       *Actions*: Removed duplicates and corrected data types.
   * **Orders Table**:  
     Columns include:
     + Order ID (Foreign Key)
     + Customer ID (Foreign Key)
     + Product ID (Primary Key)
     + Category, Sub-category, Product Name, Sales, Quantity, Profit, Returns, and Payment Mode
     + Conditional Column: Created to identify whether an order results in a profit or loss.  
       *Actions*: Removed duplicates and corrected data types.
   * **Shipping Table**:  
     Columns include:
     + Order ID (Foreign Key)
     + Shipping Mode
     + Order Date
     + Ship Date
     + Calculated Column: Added a column that calculates the shipping duration by subtracting the Order Date from the Ship Date.  
       *Actions*: Removed duplicates and corrected data types.
   * **Dates Table**:
     + This table contains a single column for dates to be used in analysis.

**PHASE TWO: Data Modeling**

* **Star Schema**:  
  The *Orders* table serves as the fact table and connects to:
  + **Shipping Table** via Order ID.
  + **Customers Table** via Order ID.
* **Relationships**:
  + Established relationships with the appropriate cardinality between the tables.

**PHASE THREE: Key Business Questions**

We aimed to answer the following business questions:

1. How many orders were placed?
2. How many customers are there?
3. What is the total profit?
4. What are the total sales?
5. Which payment types are used?
6. What is the percentage of profit and loss?
7. Which states place the most orders?
8. What are the top 10 products sold?
9. Which states have the highest order volume?
10. Who are the top customers?
11. What are the most popular product categories?
12. Sales by region.
13. What is the percentage of returns?

**PHASE FOUR: Visualizations**

We created the following visualizations to answer the above questions:

1. **Total Orders and Customers**: Displays the total number of orders and customers.
2. **Profit and Sales**: A breakdown of total profit and sales.
3. **Payment Type Analysis**: Shows the distribution of payment types.
4. **Profit/Loss Analysis**: Visualizing the percentage of profit and loss.
5. **Top States**: Bar charts and maps highlighting states with the most orders.
6. **Top Products and Customers**: A ranking of top-selling products and top-buying customers.
7. **Category Analysis**: Identifies the most frequently sold product categories.
8. **Sales by Region**: A map visualization showing sales distribution by region.
9. **Return Percentage**: A chart illustrating the percentage of returns.