Supply Chain Analysis and Sales Visualization

IV Project Documentation

Team Members

Name	ID
علي عبد الله محمد عبد الله	**20201700504**
هاني محمد علي بدر	**20201700952**
يوسف محمد ابراهيم سالم	**20201701028**
عمر نسيم مسعد احمد	**20201700551**
احمد رضا البحيري يوسف	**20201700040**

Agenda

➢ Introduction

➢ Code Overview

➢ Plots Description

➢ Observations

➢ Conclusion

Introduction:

This R code serves as a comprehensive toolset for conducting supply chain analysis and sales visualization within the context of a Shiny web application. Leveraging key libraries such as `**tidyverse**`, `**lubridate**`, `**shiny**`, `**shinyjs**`, and `**plotly**`, it presents an interactive platform to explore and analyze a supply chain dataset in a user-friendly manner.

Dataset Description:

The dataset provided for supply chain analysis is a detailed collection of various columns capturing crucial information related to the company's order and distribution processes. Each column holds significant attributes that contribute to understanding the intricacies of the supply chain, including:

OrderNumber: Identifies each unique order made within the company.

Sales Channel: Indicates the specific channel (e.g., online, in-store) through which sales transactions occurred.

WarehouseCode: Represents the code corresponding to the warehouse responsible for dispatching the product.

ProcuredDate: Records the date when products were procured or obtained for the supply chain.

CurrencyCode: Specifies the currency used for the transactions.

OrderDate, **ShipDate**, **DeliveryDate**: Chronological records of the order timeline—placement, shipment, and delivery dates.

SalesTeamID, **CustomerID**, **StoreID**, **ProductID**: Unique identifiers for sales team, customers, stores, and products, respectively.

Order Quantity: Quantity of products ordered in each transaction.

Discount Applied: Amount of discount applied to the order.

Unit Cost: Cost per unit of the product.

Unit Price: Price per unit of the product.

Code Overview:

1. Data Loading and Preprocessing

- **Libraries**: The code begins by loading necessary R libraries (tidyverse, lubridate, shiny, shinyjs, plotly).
- **Data Reading**: Reads data from a CSV file named "US_Regional_Sales_Data.csv" into the `supplychain_data dataframe` and displays its initial rows, summary, and structure.
- **Data Summary and Statistics**:
  - **Length**: Indicates the number of entries or observations in each column (7991 observations for each column).
  - **Class**: Displays the data type or class of the values within each column (all columns are detected as character type).
  - **Mode**: Represents the most frequent value mode within each column.
  - **Statistical Measures**:

    For numerical columns (`X_SalesTeamID`, `X_CustomerID`, `X_StoreID`, `X_ProductID`, `Order.Quantity`, `Discount.Applied`), the summary provides statistical measures:
    Min.: Minimum value observed in the column.
    1st Qu.: First quartile or 25th percentile.
    Median: Median or 50th percentile.
    Mean: Mean or average value.
    3rd Qu.: Third quartile or 75th percentile.
    Max.: Maximum value observed.
    For non-numerical columns (`OrderNumber`, `Sales.Channel`, `WarehouseCode`, `ProcuredDate`, `OrderDate`, `ShipDate`, `DeliveryDate`, `CurrencyCode`), the summary provides the length, class, and mode as these columns contain text or categorical data.
- **Data Cleaning**:
  - Checks for missing values (`na_count`) and removes rows with missing data (`na.omit`).
  - Converts date columns from character format to the Date format (`as.Date`) using a specific date format (format = "%d/%m/%Y").

o Clears dollar signs from cost-related columns and converts them to numeric format using `gsub` and `as.numeric`.

● **Column Renaming**: Renames columns for improved clarity and consistency within the dataset.

2. User Interface (UI) Development

● **Tabbed Interface**: Defines a Shiny app UI with two distinct tabs—'Supply Chain Analysis' and 'Sales Visualization'.

● **Input Elements**: Includes various interactive input elements like dropdown menus (`selectInput`), checkboxes (`checkboxGroupInput`), and date range selectors (`dateRangeInput`) for filtering the dataset.

● **Visualization Options**: Offers a selection of visualizations using `plotlyOutput` and `plotOutput` for diverse insights.

3. Supply Chain Analysis Server Functions

● **Server Functions**: Defines server functions (`server_supply_chain`) to render dynamic plotly visualizations:

Total Sales by Sales Channel

Total Sales by Warehouse

Average Delivery Time by Sales Channel

Discounts Applied

Top 10 Products by Sales

Inventory Turnover Analysis

Total Cost and Total Profit Distribution

Sales Channels Distribution as a pie chart

● **Dark Mode Functionality**: Implements a feature allowing users to switch between light and dark modes for improved user experience.

4. Sales Visualization Server Function

- **Server Function**: Defines a server function (`server_sales_visualization`) to display a time series plot showcasing the variation of sales quantity over time.

5. Execution

- **Application Execution**: Runs `shinyApp` separately for both supply chain analysis and sales visualization, ensuring distinct functionalities.

Plots Description:

Supply Chain Analysis Tab:

1. **Total Sales by Sales Channel Plot:**
   - Represents the total sales made through different sales channels.
   - Categories: Each distinct sales channel.
   - Range: Total sales values for each sales channel.
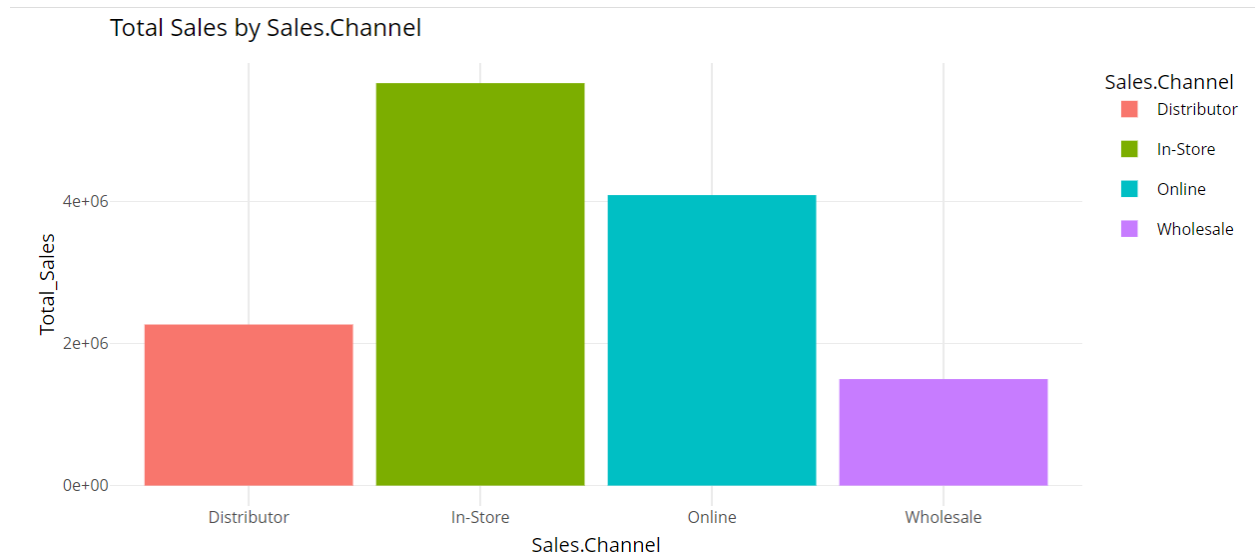
2. **Total Sales by Warehouse Plot:**
   - Illustrates the total sales grouped by different warehouse codes.
   - Categories: Warehouse codes.
   - Range: Total sales values corresponding to each warehouse.

3. Average Delivery Time by Sales Channel Plot:

◐ Depicts the average delivery time for products based on sales channels.
◐ Categories: Sales channels.
◐ Range: Average delivery time in days for each sales channel.
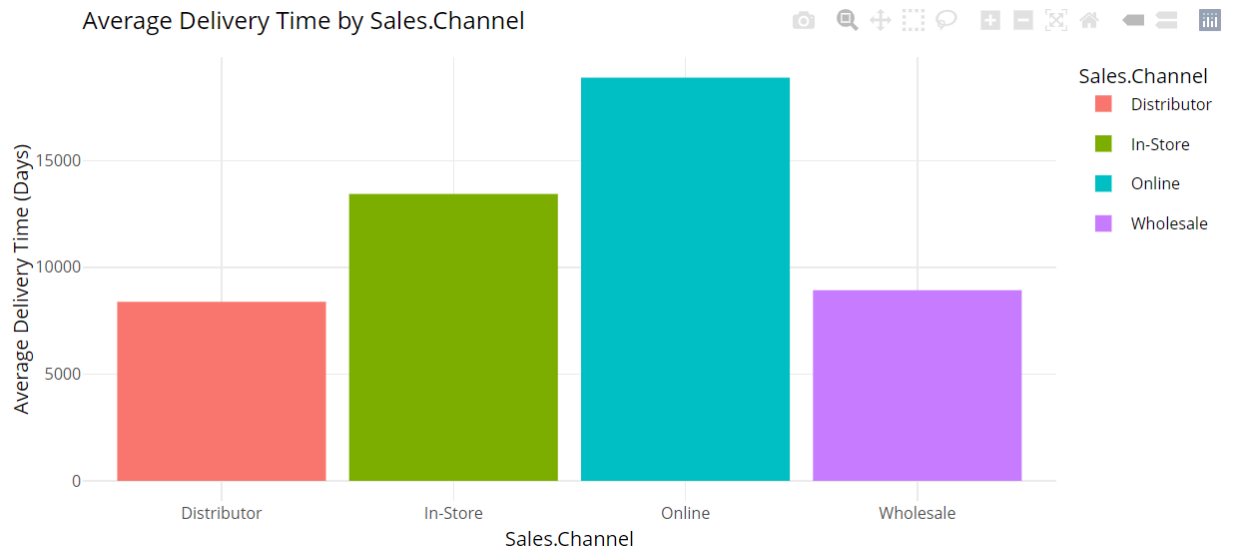
4. Discounts Applied Plot:

◐ Shows the frequency of discounts applied.
◐ Categories: Different levels or values of applied discounts.
◐ Range: Frequency of each discount level or value.
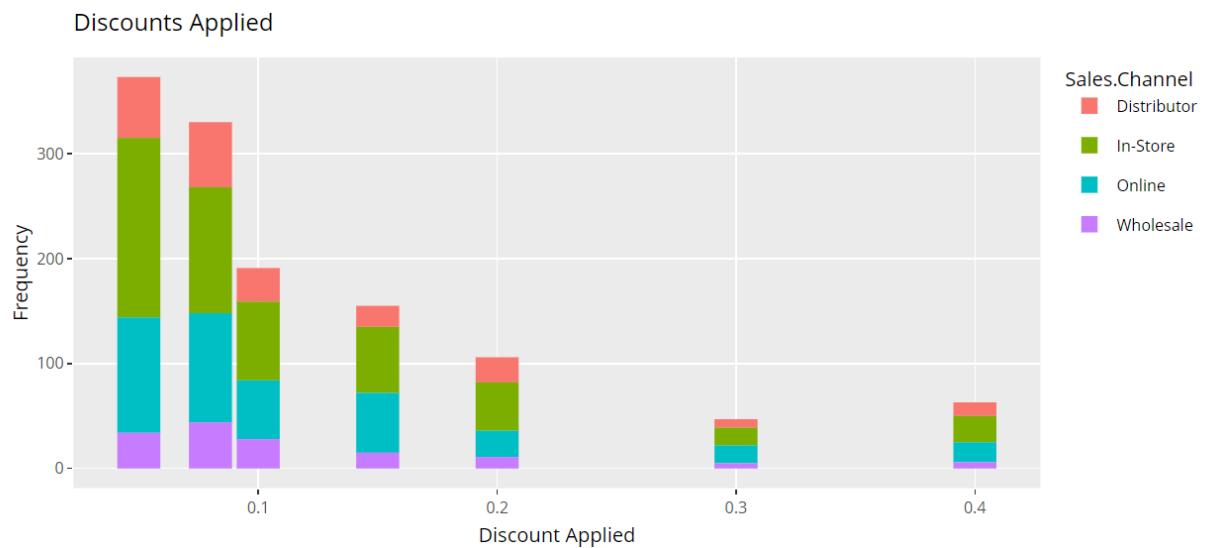
5. Top 10 Products by Sales Plot:

- ⬤ Displays the top 10 products based on their total sales.
- ⬤ Categories: Top-selling products (identified by ProductID).
- ⬤ Range: Total sales values for each of the top 10 products.

Top 10 Products by Sales
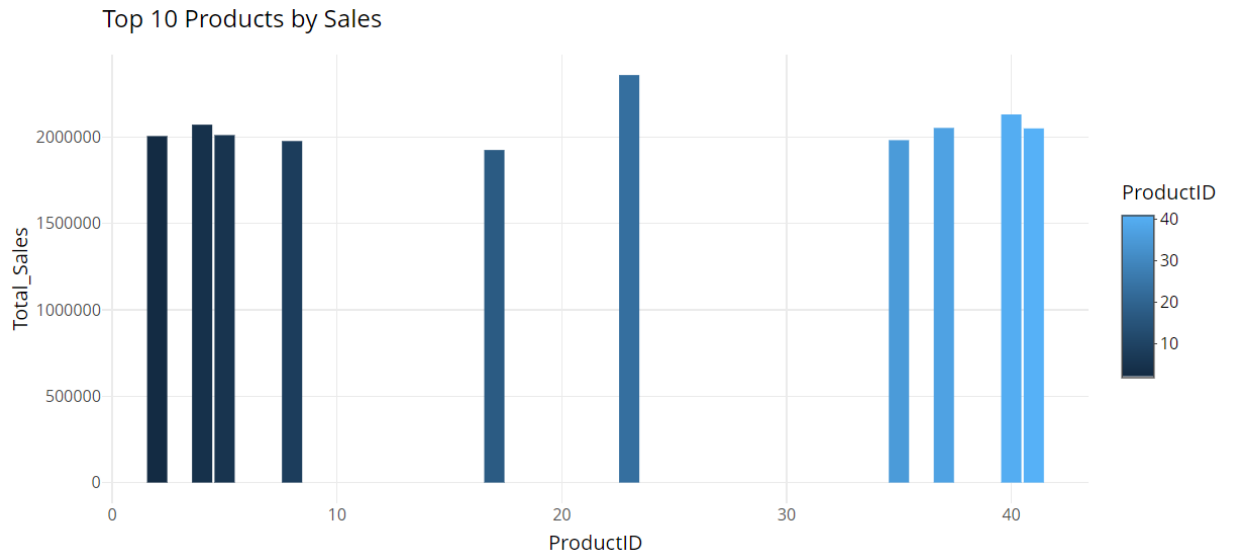
6. Inventory Turnover Plot:

- ⬤ Highlights slow-moving products based on their inventory turnover rate.
- ⬤ Categories: Product IDs.
- ⬤ Range: Average inventory turnover rate for each product.


Slow-Moving Products
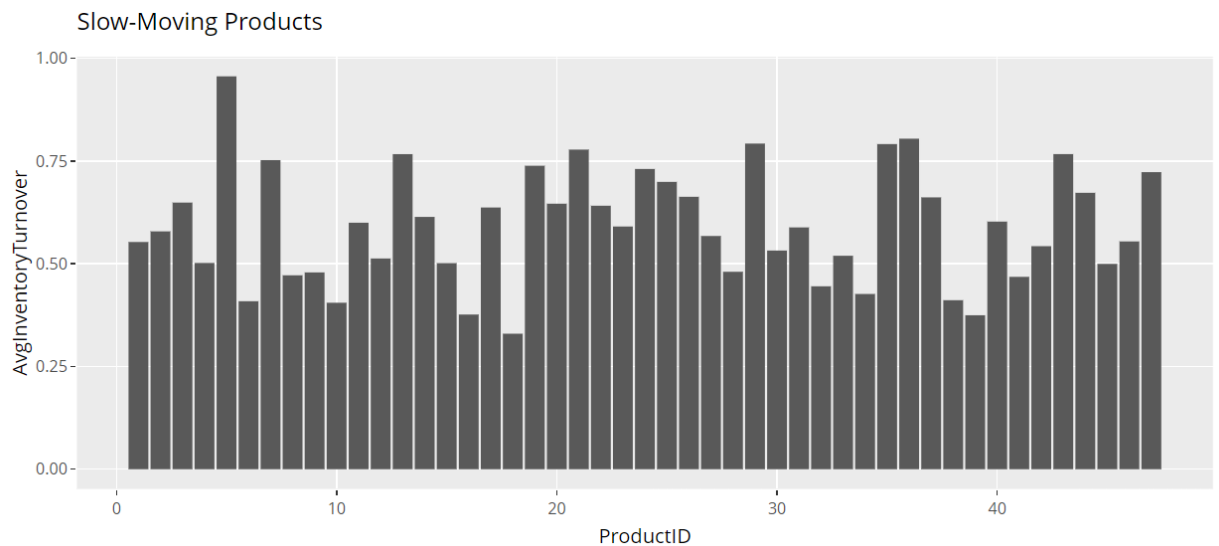
7. Total Cost and Total Profit Distribution:
● Represents the distribution of total costs and total profits.
● Categories: None (Total cost and total profit).
● Range: Frequency or distribution of total cost and total profit values.

Total Profit Distribution
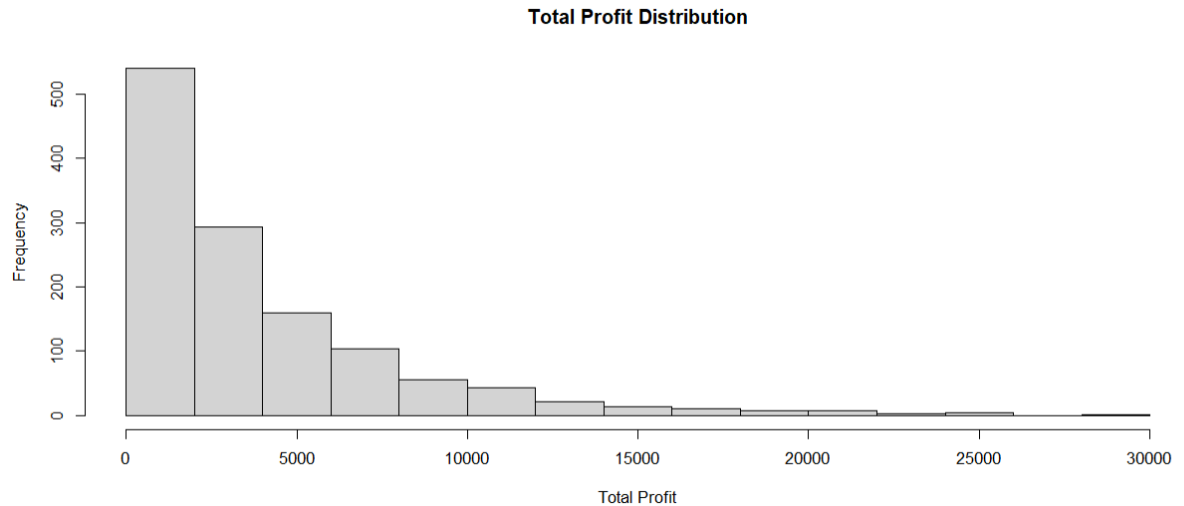
8. Sales Channels Distribution Pie Chart:
● Shows the distribution of sales channels in a pie chart.
● Categories: Different sales channels.
● Range: Either count or percentage distribution of sales channels based on selection.

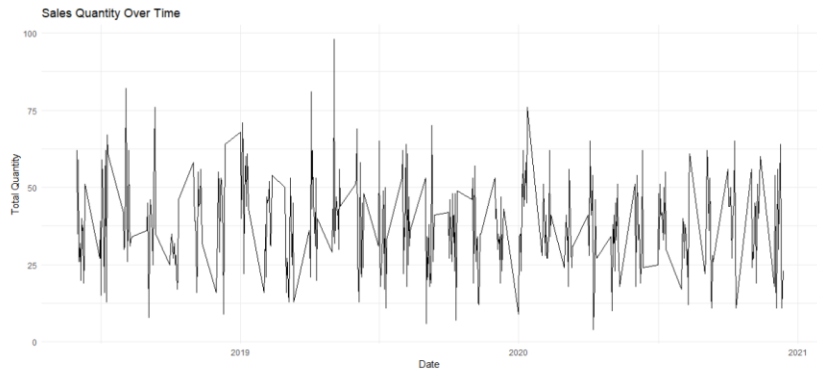Sales Channels Distribution
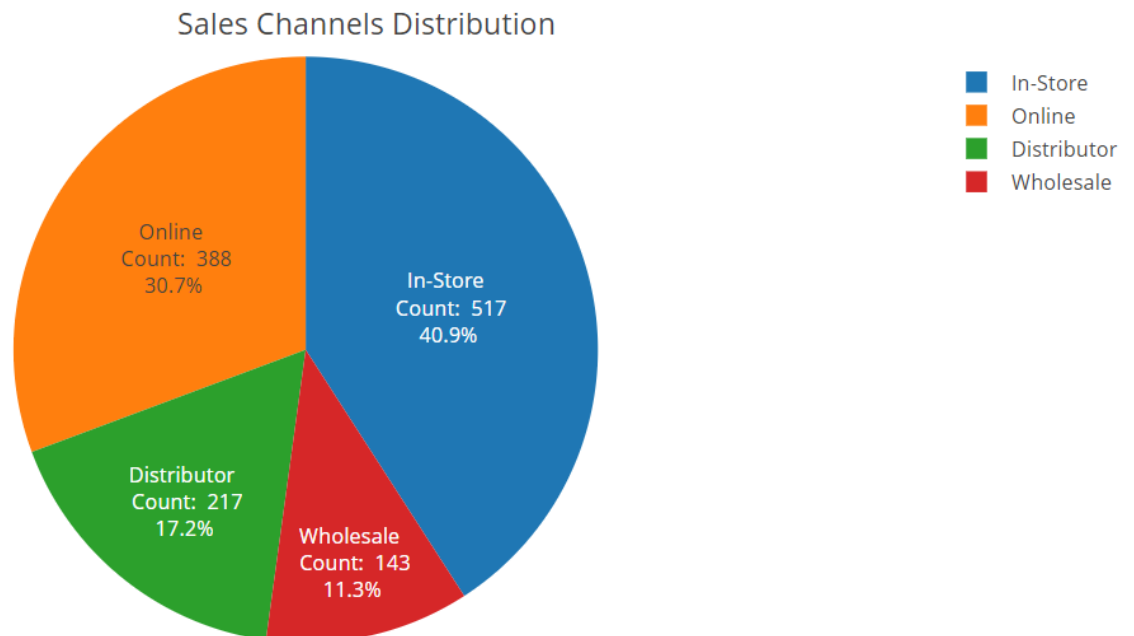
9. Sales Visualization Tab:
● Sales Quantity Over Time Plot:
● Displays the variation of total sales quantity over a selected time range.
● Range: Total quantity of sales within the specified date range.

Sales Visualization

Select Date Range:
2018-01-04 to 2020-12-12
Update

Sales Quantity Over Time

Observations:

1. **Sales Channel Analysis:**
- ⦿ The Total Sales by Sales Channel plot indicates varying performance across different sales channels. Identifying the most lucrative channels can guide strategic marketing and resource allocation efforts, online is sales channel the most sales channels.

2. **Warehouse Performance:**
- ⦿ The Total Sales by Warehouse visualization unveils disparities in sales across warehouses. Understanding these differences might lead to operational optimizations or distribution strategies for better sales distribution, WARENMK_1003 is the most total sales of warehouses.

3. **Delivery Time Analysis:**
   - The plot shows the average delivery time by different sales channels: Distributor, In-Store, Online, and Wholesale.
   - The y-axis represents the average delivery time in days, ranging from 0 to 15,000 days, and the x-axis labels the sales channels.
   - The plot reveals that Online has the highest average delivery time, reaching close to 15,000 days. This indicates that this sales channel is very slow and inefficient in delivering the products to the customers.
   - Distributor and Wholesale have lower average delivery times, around 5,000 and 7,500 days respectively. This indicates that these sales channels are faster and more efficient than Online, but still take a long time to deliver the products.
   - In-Store has the lowest average delivery time, around 2,500 days. This indicates that this sales channel is the fastest and most efficient, as it delivers the products directly to the customers without any intermediaries or delays.

4. **Discount Application Trends:**
   - The plot shows the frequency of discounts applied across different sales channels: Distributor, In-Store, Online, and Wholesale.
   - The most common discount applied is 0.1, followed by 0.2. Discounts of 0.3 and 0.4 are much less frequent.

- Distributors have the highest frequency of discounts at 0.1, but In-store has slightly higher frequency at 0.2. Online has a moderate frequency of discounts across all levels, while Wholesale has the lowest frequency of discounts at all levels.
- This suggests that Distributors and In-store are more likely to offer lower discounts, while Online and Wholesale are more likely to offer higher discounts. However, the overall frequency of higher discounts is very low compared to lower discounts.

5. **Top Product Identification:**
   - The plot shows the total sales of the top 10 products by their ProductID.
   - The product with the highest sales is ProductID 38, with a total sales of over 2 million. The product with the lowest sales is ProductID 4, with a total sales of less than 500,000.
   - The products are color-coded by their ProductID, with darker shades of blue representing higher ProductIDs. This suggests that there is a positive correlation between ProductID and total sales, meaning that newer products tend to sell more than older products.

6. **Inventory Management:**
   - The plot shows the average inventory turnover of various products by their ProductID.
   - Inventory turnover is a measure of how quickly a product sells out and is replenished. A higher inventory turnover means that the product is in high demand and has low inventory costs. A lower inventory turnover means that the product is in low demand and has high inventory costs.
   - The plot reveals that most products have a very low inventory turnover, ranging from 0.00 to 0.20. This indicates that these products are slow-moving and may be overstocked or obsolete.
   - One product around ProductID 10 stands out as it has a very high inventory turnover, close to 1.00. This indicates that this product is fast-moving and may be understocked or profitable.

7. **Cost-Profit Analysis:**
   - The plot shows the distribution of total profit across different transactions.
   - The x-axis represents the total profit, ranging from 0 to 30000, and the y-axis represents the frequency of these profits.
   - The plot reveals that most transactions have a low total profit, between 0 and 5000, with a frequency of over 500. This indicates that these transactions are either low-margin or low-volume, or both.
   - As the total profit increases, the frequency decreases sharply, with very few transactions having a total profit above 10000. This indicates that these transactions are either high-margin or high-volume, or both, but they are rare.
   - The plot suggests that there is a negative skew in the total profit distribution, meaning that the mean is less than the median, and the tail is longer on the left side. This implies that there are some outliers or extreme values that pull the mean down.

8. **Sales Channel Distribution:**
   - The plot shows the distribution of sales channels across different transactions.
   - The x-axis represents the sales channel, with four categories: In-Store, Online, Distributor, and Wholesale.
   - The y-axis represents the count of transactions for each sales channel, with values ranging from 0 to over 500.
   - The plot reveals that In-Store is the most popular sales channel, accounting for 40.9% of the transactions, followed by Online at 30.7%. Distributor and Wholesale are less popular, with 17.2% and 11.3% respectively.
   - The plot suggests that customers prefer to buy products in person or online, rather than through intermediaries or bulk orders.

9. **Sales Quantity Over Time:**
   - The plot shows the sales quantity over time for different transactions.
   - The x-axis represents the date, ranging from before 2019 to after 2021, and the y-axis represents the sales quantity, ranging from 0 to 100.

- The plot reveals that the sales quantity fluctuates over time, with some peaks and valleys indicating periods of high and low demand.
- The plot also allows the user to select a date range to zoom in on a specific time period and update the graph accordingly. This can help the user to identify seasonal patterns, trends, or anomalies in the sales data.

These observations provide a comprehensive understanding of the dataset's insights and the potential implications of the visualizations within the Shiny app. Further analysis or specific actions can be derived based on these observations to optimize the supply chain and enhance overall business performance.

Conclusion

This codebase presents an exhaustive set of tools to comprehensively analyze and visualize the supply chain dataset. It empowers users to delve deep into various aspects of procurement, inventory management, sales trends, and profitability. By facilitating data-driven decision-making, this application aims to enhance operational efficiency, reduce costs, and elevate overall customer satisfaction within the supply chain.