**. Dataset Overview**

The supply chain dataset contains information about orders, products, suppliers, and delivery timelines. The key columns include:

* **Order ID**: Unique identifier for each order.
* **Product Category**: Categories of the products sold.
* **Quantity**: Number of items ordered.
* **Unit Price**: Price per unit of the product.
* **Total Price**: Total cost of the order.
* **Order Date and Delivery Date**: Dates of order placement and delivery.
* **Location**: The location of the delivery.
* **Supplier**: Supplier name.
* **Status**: Order status (e.g., completed, canceled).

**2. Descriptive Statistics**

The dataset includes 200 records. Below are the key statistics for numerical columns:

|  |  |  |  |
| --- | --- | --- | --- |
| **Metric** | **Quantity** | **Unit Price** | **Total Price** |
| **Count** | 200 | 200 | 200 |
| **Mean** | 255.29 | $530.37 | $137,078.88 |
| **Standard Dev.** | 150.47 | $285.60 | $116,149.96 |
| **Minimum** | 2 | $14.62 | $302.56 |
| **25th Percentile** | 123.75 | $299.07 | $42,327.75 |
| **Median** | 258.50 | $534.30 | $104,811.88 |
| **75th Percentile** | 387.75 | $779.65 | $215,655.54 |
| **Maximum** | 500 | $996.14 | $460,417.32 |

**3. Key Findings**

**a) Product Category Distribution**

* The dataset contains six product categories:
  + Pharmaceuticals
  + Automotive
  + Clothing
  + Food
  + Electronics
  + Furniture
* The most ordered categories are **Pharmaceuticals** and **Automotive**, each with around 40 orders.

**b) Order Trends**

* A time series analysis of the "Total Price" shows significant fluctuations over the timeline. Some months experience peak orders, possibly indicating seasonal demand.

**c) Delivery Performance**

* The average delivery time is analyzed using "Days to Deliver".
* A distribution of delivery times reveals a wide range, with a notable peak around the average delivery days.

**d) Correlation Analysis**

* **Quantity and Total Price**: A strong positive correlation (0.71) indicates larger orders lead to higher total prices.
* **Unit Price and Total Price**: Moderate positive correlation (0.62).
* **Quantity and Unit Price**: Minimal correlation (0.039).

**4. Visual Insights**

The following visualizations provide a clear understanding of the dataset:

1. **Product Category Distribution**: Highlights the most popular categories by the number of orders.
2. **Total Price Over Time**: Visualizes revenue trends over the order timeline.
3. **Delivery Time Distribution**: Shows the spread of delivery times, with peaks and delays identified.
4. **Correlation Heatmap**: Demonstrates relationships between numerical variables.

**5. Recommendations**

* **Optimize Inventory Management**: Focus on high-demand categories like Pharmaceuticals and Automotive.
* **Improve Delivery Performance**: Address delays by analyzing suppliers and locations with high delivery times.
* **Seasonal Planning**: Utilize time-series trends to prepare for peak demand periods.
* **Pricing Strategies**: Analyze low-correlation products for potential pricing optimizations.

**6. Next Steps**

* Conduct deeper analysis on supplier performance and specific locations.
* Identify patterns in canceled orders for mitigation strategies.
* Use advanced predictive models for demand forecasting.

This analysis aims to enhance supply chain efficiency, minimize delays, and boost profitability.