AMAZON SALES REPORT

Problem Statement: Analyzing order trends, shipping service levels, and sales distribution to gain insights into business performance. The goal is to identify key patterns in order volume, shipping preferences, and relationships between quantity and amount.

1. Introduction

This document presents a detailed analysis and visualization of the given dataset, focusing on orders, shipping service levels, and various categorical and numerical distributions.

2. Data Preparation

- Converted the 'Date' column to datetime format for time series analysis.
- Converted numeric columns ('Qty', 'Amount', 'ship-postal-code') to numeric types while handling errors.

3. Univariate Analysis

3.1 Numerical Columns

- **Quantity Distribution**: A histogram with KDE was plotted to analyze the distribution of 'Qty'.
- **Amount Distribution**: A histogram with KDE was plotted to analyze the distribution of 'Amount'.

3.2 Categorical Columns

- **Shipping Service Level Distribution**: A count plot was generated to show the frequency of different shipping service levels.
- **Sales Channel Distribution**: A count plot was created to visualize the sales channels.
- **Category Distribution**: A count plot was created for product categories.
- **Ship-Country Distribution**: A count plot was created to analyze the distribution of shipments by country.

4. Time Series Analysis

- A line plot was created to show the trend of the number of orders over time.
- The top 10 order dates were displayed.
- The top 10 shipping service levels were identified and displayed.

5. Bivariate Analysis

- **Scatter Plot**: A scatter plot was created to examine the relationship between 'Qty' and 'Amount'.
- **Box Plot**: A box plot was created to analyze the distribution of 'Amount' across different sales channels.
- **Heatmap**: A heatmap was generated to show the correlation between 'Qty' and 'Amount'.

6. Multivariate Analysis

- **Pair Plot**: A pair plot was created to visualize relationships between 'Qty' and 'Amount'.
- **Correlation Matrix**: A heatmap was generated for the correlation matrix of 'Qty', 'Amount', and 'ship-postal-code'.

7. Conclusion

- The analysis provides insights into order trends, shipping service levels, and sales distribution.
- The correlation analysis highlights relationships between numerical features.
- The visualizations help in understanding patterns and making data-driven decisions.

8. Recommendations

- Further analysis on seasonal trends in orders.
- Exploring additional features to improve correlation insights.
- Investigating outliers in the data for better anomaly detection.