**SALES - EDA TASK SHEET**

Loading the Data:

1. How do you load the dataset into a pandas DataFrame?

Basic Data Exploration:

1. What is the structure of the dataset (e.g., number of rows and columns)?
2. What are the data types of each column?
3. Are there any missing values in the dataset?

Descriptive Statistics:

1. What are the summary statistics (mean, median, mode, standard deviation) for the Price Each and Sales columns?
2. What are the total sales for each month?

Data Cleaning:

1. How do you handle missing values in the dataset?
2. Are there any duplicate rows in the dataset? If so, how can you remove them?

Date and Time Analysis:

1. How can you extract the year and day from the Order Date column?
2. What is the distribution of orders by hour of the day?

Sales Analysis:

1. What are the top 5 products by sales?
2. How does the quantity ordered vary by product?

City and Location Analysis:

1. What are the total sales by city?
2. Which city has the highest number of orders?

Monthly Trends:

1. How do sales vary by month?
2. Can you identify any seasonal patterns in the sales data?

Product Analysis:

1. What is the average order quantity for each product?
2. What are the most and least expensive products?

Correlation Analysis:

1. Is there a correlation between Quantity Ordered and Price Each?
2. How do Sales and Quantity Ordered correlate?

Visualization:

1. How can you visualize the total sales per month using a bar plot?
2. How can you visualize the distribution of order quantities using a histogram?

Sales Trends:

1. How do sales vary by hour of the day?
2. How do sales vary by day of the week?

Address Analysis:

1. How many unique purchase addresses are there in the dataset?
2. What is the most common purchase address?

Time Series Analysis:

1. How can you visualize sales trends over time using a line plot?

Group By Analysis:

1. How can you group data by City and calculate the total sales for each city?
2. How can you group data by Product and calculate the average sales price for each product?

Pivot Tables:

1. How can you create a pivot table to summarize total sales by month and city?
2. How can you create a pivot table to summarize average quantity ordered by product and month?

Anomaly Detection:

1. Are there any anomalies or outliers in the Sales data?
2. How can you identify orders with unusually high quantities?

Customer Insights:

1. What is the distribution of orders by customer (purchase address)?
2. How many unique customers are there in the dataset?

Categorical Analysis:

1. What is the proportion of orders from each city?