**Regular project Questions**

**Project General Description**

The project involves performing data analysis on an e-commerce sales dataset. It covers data cleaning, exploratory data analysis (EDA), customer segmentation, and forecasting.

**Task 1: Data Cleaning**

**Description:** Clean and preprocess the dataset by handling missing values, formatting data types, and removing incorrect records.

**Requirements:**  Handle missing values in CustomerID  
- Convert InvoiceDate to DateTime type  
- Remove rows with negative Quantity or UnitPrice  
- Create a TotalPrice column.

**Task 2: Exploratory Data Analysis (EDA)**

**Description:**Perform basic descriptive statistics and identify insights from the dataset. Analyze top-selling products and calculate total revenue and transactions.values. You will also check the data types of each column to understand how the data is structured.  
**Requirements:** - Use .describe() for statistical summary.  
- Identify top 10 selling products by Quantity.  
- Calculate total revenue and number of transactions.

**Task 3: Time Series Analysis**

**Description:**Analyze sales trends over time and visualize monthly sales. Identify months with highest and lowest sales.

**Requirements:** - Resample the data to obtain monthly sales.  
- Plot monthly sales trends using a line plot.  
- Identify the month with the highest and lowest sales.

**Task 4: RFM Analysis (Customer Segmentation)**

**Description:**Segment customers based on recency, frequency, and monetary value. Visualize customer distribution in segments.

**Requirements:** - Calculate recency (days since last purchase).  
- Calculate frequency (number of purchases).  
- Calculate monetary value (total spend).  
- Segment customers into High, Medium, Low-value groups.

**Task 5: Product Category Analysis**

**Description:**Analyze sales and revenue by product category and visualize the top categories by revenue.

**Requirements:** - Extract product category from Description.  
- Calculate sales and revenue by category.  
- Create a bar plot of the top 5 categories by revenue.

**Task 6: Geographical Analysis**

**Description:**Analyze sales distribution by country and identify top revenue-generating countries.

**Requirements:** - Calculate total revenue by country.  
- Create a bar plot of the top 10 countries by revenue.  
- Calculate the percentage of sales from the top 3 countries.

**Task 7: Customer Behavior Analysis**

**Description:**Analyze customer behavior based on order quantity, sales, and purchasing patterns.

**Requirements:** - Plot the distribution of order quantities.  
- Create a scatter plot of Quantity vs. TotalPrice.  
- Calculate and plot average daily sales throughout the week.

**Task 8: Moving Average Forecast**

**Description:**Implement a basic moving average forecast for sales based on past data.

**Requirements** - Prepare daily sales data.  
- Calculate a 7-day moving average of sales.  
- Plot actual sales vs. moving average for the last 3 months.

**Task 9: Summary Dashboard Creation**

**Description:**Create a dashboard that summarizes key insights through visualizations.

**Requirements** :- Create a 2x2 subplot with:  
1. Monthly sales trend  
2. Top 5 products by revenue  
3. Customer segment distribution  
4. Top 5 countries by revenue.

**Task 10: Optimize Data Processing**

**Description:**Optimize a computationally intensive task using vectorized operations.

**Requirements:**- Implement a task using loops.  
- Implement the same task using vectorized operations.  
- Compare and report performance differences.

**Task 11: Report Generation**

**Description:**Generate a summary report with key insights and recommendations.

**Requirements:**- Summarize overall revenue, top-selling products, best customer segments, and countries.  
- Provide insights from time series analysis and recommendations.