**Data Exploration & Understanding (SQL):**

**1. Total Number of Invoices and Unique Customers:** I found that the total number of invoices is 307, and the number of unique customers is 92. There are 7 product categories.

**2. Total Sales Amount and Average Order Value:** I calculated the total sales amount and found it to be 10 million. The average order value is 33K and the most frequently purchased item is classic cars.

**3. Top Customers by Order Number:** I wanted to see the customers with the highest order number and found (Euro shopping channel, Mini Gifts, Reims collectables).

**4. Countries with the Highest Total Sales Amount:** I ran a query to see the countries with the highest total sales amount and found that the USA, followed by Spain, France, Australia, and the UK, are the top countries in order of total sales amount.

**5. Top Products by Total Sales Amount:**  I wanted to see the top products in terms of total sales amount and found that classic cars generate 3,919,615.66 in total sales, followed by vintage cars. The product with the lowest sales is trains.

**6. Trends Over Time:** I ran a query to analyze trends over time and found that the highest sales were recorded in 2004, and the month with the highest sales is November.

**7. Order Status:**  I wanted to see the order status and found that 286 orders were shipped and only 4 were cancelled.

**8. Deal size Analysis:** I wanted to explore the deal size column and found that most deals are medium.

**9. Top Cities by Total Sales Amount:** I found that the city with the highest total sales amount is Madrid, with 1,082,551, followed by San Rafael and NYC.

**10. Countries and cities with the Largest Number of Customers:**  The countries with the largest number of customers are the USA, followed by France, and then the UK and the cities with the largest number of customers are NYC then Paris then Madrid.

**Data Preparation & Analysis (Python):**

**1. Libraries used**: I utilized the pandas, matplotlib, and seaborn libraries**.**

**2. Data Import**: import the E-Commerce Sales dataset.

**3**. I changed columns name into lowercase for easy typing**.**

**4.Handling Null Values:** When I was checking for null values, I found a few columns with null values (addressline2, postalcode, state, territory). Then, I deleted the columns that were not important to me and that I believed would not be useful for the analysis and dashboard (addressline2, addressline1, phone, postalcode).

**5. Dropping the States Column:** I found that many countries do not have states, so I decided to drop the states column.

**6. Replacing Null Values in the Territory Column:** When I found that the USA and Canada are the countries with null values in the territory column, I replaced the null values with 'North America'.

**7.** Change orderdate from object type to Datetime type

**8. Combine similar product descriptions for consistency:** To combine similar product descriptions, I split the products in the product line column into 2 product categories. I grouped classic cars, motorcycles, and vintage cars into 'vehicles', and planes, trucks and buses, trains, and ships into 'heavy vehicles'.

**9. Customer Segmentation Based on Purchase Frequency and Spending:** To segment customers based on purchase frequency and total spending, I created a dataframe that included customer name, total spending, and frequency. Then, I assigned a score for both spending and frequency from 1 to 5. Next, I combined the spending score and frequency score into a score code, and then I Assume segment levels, then I added segment column to dataframe df.

**10. Calculating Customer Lifetime Value (CLTV):** to calculate related values such as customer\_lifetime\_value(CLTV)(Average\_Purchase\_value\*Average\_Frequency rate\*Average customer lifespan) I first calculated Average\_purchase\_value and average\_frequency Then, I calculated Average\_customer\_lifespan(Average of years a customer stays active/total of customer )and the CLTV, then I added the CLTV column to the dataframe.

11. Analyzing trends over time (monthly, yearly, quarterly).

**Data Visualization (Power BI)**

The Power BI Dashboard was structured into four pages to provide comprehensive insights

1. **Overview**: provide an overall snapshot of sales performance, trends and key metrics.
2. **Customer**: analyze customer behavior, segments and purchasing patterns.
3. **Product:** assess product performance.
4. **Location:** visualize regional sales distribution.

**Business Question:**

1. What are the total sales over different time?
2. Which productsubcategory generate the most revenue?
3. Who are top customers based on total sales?
4. What is the distribution of customer segments?
5. How do purchasing behaviors differ among customer segments?
6. Which Product category have the highest and lowest sales volume?
7. What are the seasonal trends in product sales?
8. Which territories and countries generate the most sales?

**Insights:**

1. Total Sales Trend: the total sales have shown a consistent increase from 2003 to 2005, with the highest sales recorded in 2004, and the month with the highest sales is November. And the quarter with the highest sales is 4.
2. Regional Sales Performance: The EMMA region has consistently outperformed the north America region in terms of total sales, USA has the highest total sales amount and Madrid has the highest total sales amount.
3. Top Customer: the top 5 customers based on total sales are Euro Shopping Channel, Mini Gifts Distributors, Australian Collectors, Muscle Machine Inc and La Rochelle Gifts.
4. The medium deal size category has the highest contribution to total sales, followed by small and large deal sizes.
5. Vehicles category has the highest sales amount, Classic Cars and vintage cars were the top-performing product sub category
6. High value customers who made frequent, large purchases represent 41 of total number of customers and have spent 6Million out of the total sales and low value customers who had infrequent, smaller purchases represent 26 of total number of customers and have spent 1Million from total sales

**Recommendation:**

1. **Focus on High value Customers:** Implement targeted marketing campaigns and loyalty programs or excusive promotions to strengthen the relationship and increase their sales contribution and enhance customer service and support.
2. **Optimize inventory management:** Ensure high-demand products are well stocked during peal seasons.
3. **Strengthen Presence in Top-Performing Regions:** Allocate more resources and marketing efforts towards the EMMA region and identify the factors contributing to its success and replicate them in other regions.
4. **Expand Regional Presence (Market Expansion):** Consider exploring opportunities in the APAC region, conduct market research to understand the potential demand to cater to that market.