E-Commerce Analysis Report

**About Dataset**

Sample Sales Data, Order Info, Sales, Customer, Shipping, etc., Used for Segmentation, Customer Analytics, Clustering and More. Inspired for retail analytics.

The data has 25 columns and 2823 rows

Objective: Analyse customer trends in an online retail setting to identify factors influencing

purchase behaviour and recommend strategies for improved sales performance.

• SQL (data exploration and manipulation)

• Python (data cleaning, analysis, and visualization)

• Power BI (data visualization and storytelling)

1-**Data Collection:**

I collect the data from Kaggle ([Sample Sales Data (kaggle.com)](https://www.kaggle.com/datasets/kyanyoga/sample-sales-data))

The dataset included information on sales transactions, such as order details, product information, customer demographics, and financial metrics.

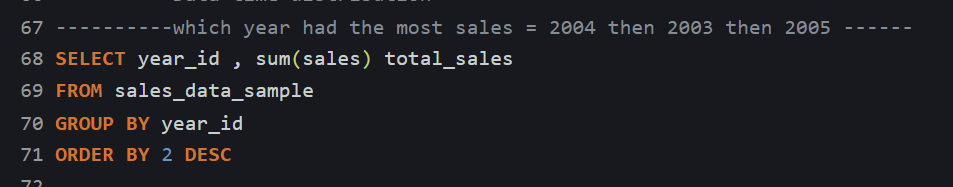
**SQL Part**

2-**Exploration Data:**

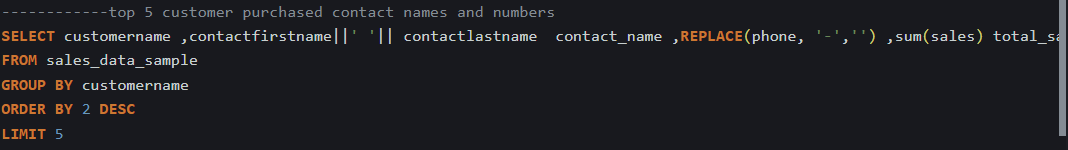
Using Sql, I explore the dataset and finding The Total Number of invoices, Unique Customers Top Five Customer Purchased, Exploring purchased pattern by country (demographic info), Purchased over years and other explorations explained by sql attached script.

Yearly Purchased:

|  |  |
| --- | --- |
| 2004 | 4724163 |
| 2003 | 3516980 |
| 2005 | 1791487 |



Top Five Customer Purchased:



|  |  |  |  |
| --- | --- | --- | --- |
| Customer Name | Contact Name | Phone | Sales |
| Canadian Gift Exchange Network | Yoshi Tannamuri | (604) 5553392 | 75238.92 |
| Mini Creations Ltd. | Wing C Tam | 5085559555 | 108951.13 |
| Tekni Collectables Inc. | William Brown | 2015559350 | 83228.19 |
| Handji Gifts& Co | Wendy Victorino | +65 224 1555 | 115498.73 |
| FunGiftIdeas.com | Violeta Benitez | 5085552555 | 98923.73 |

**Python Part**

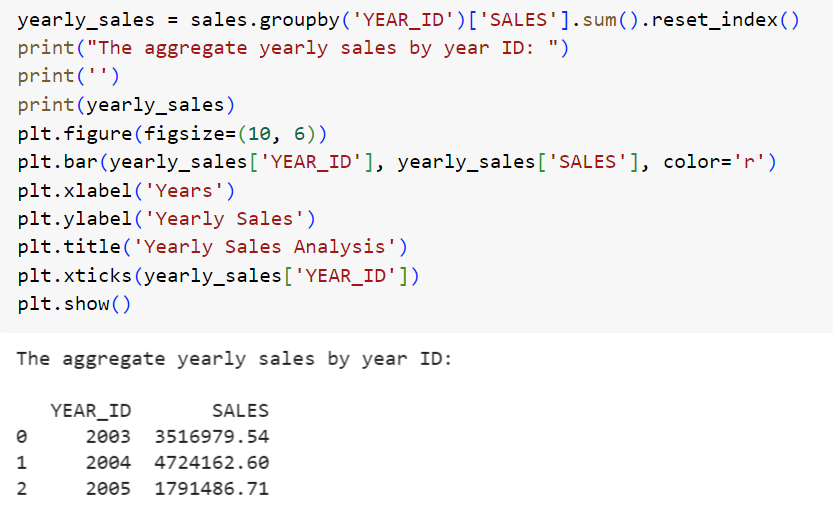
Data importing: using pandas library import data as csv file

**Data cleaning:**

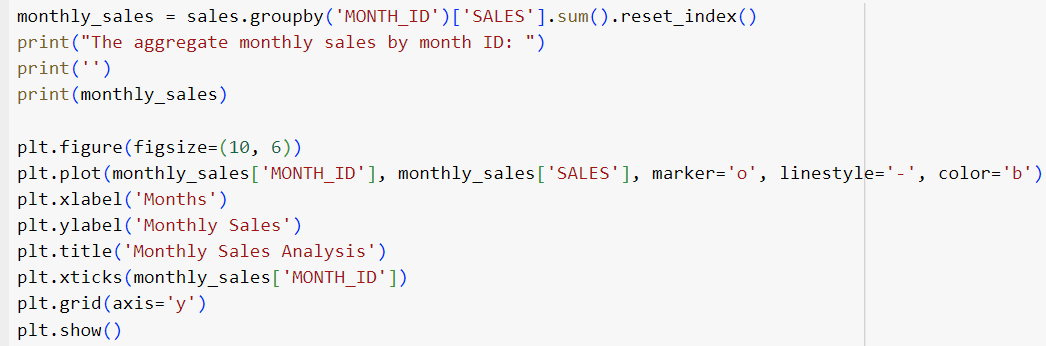
* Identifying and handling missing Data, After checking my first observation that 3columns have null values were correct (ADDRESLINE2 ,TERRITORY,STATE), Out of the three columns with Null values, I decided to drop Address Line 2, STATE and fill up Territory also drop unusable columns ['ADDRESSLINE2','POSTALCODE','CONTACTLASTNAME','CONTACTFIRSTNAME','PHONE','PRODUCTCODE']
* Standardizing Data Format, Convert order-date column type to datetime format , Ensure numerical columns are of appropriate type, Standardize text fields

**Data Analysis**

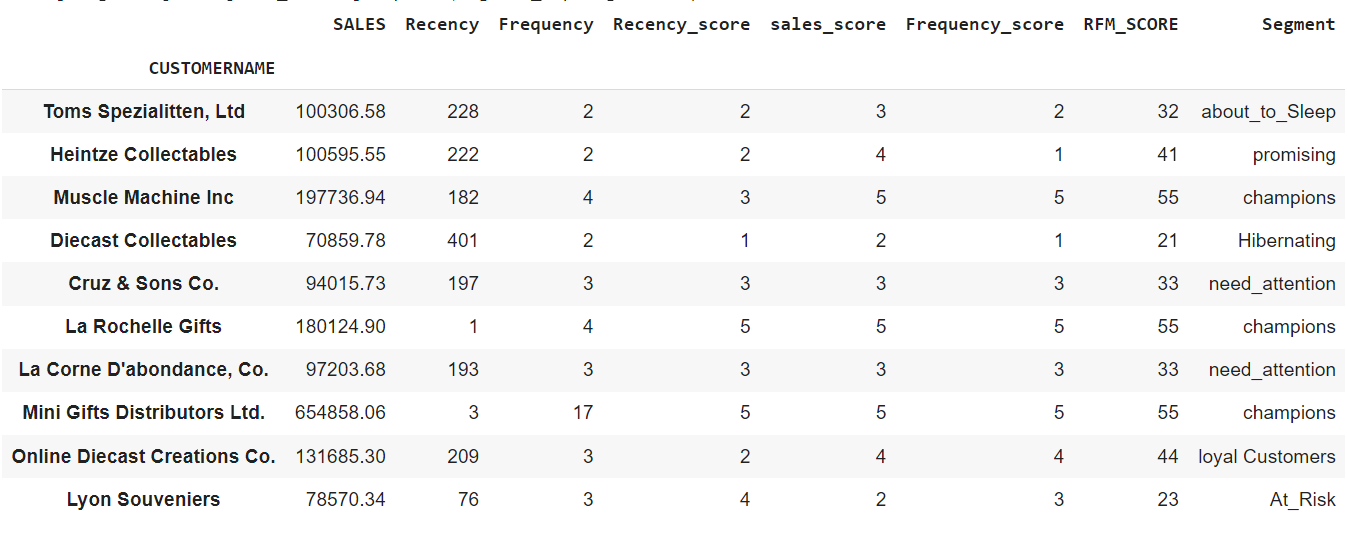
* Analys trends over time, observed that this data of 3 years (2003,2004.2005) but 2005 has only 5 month



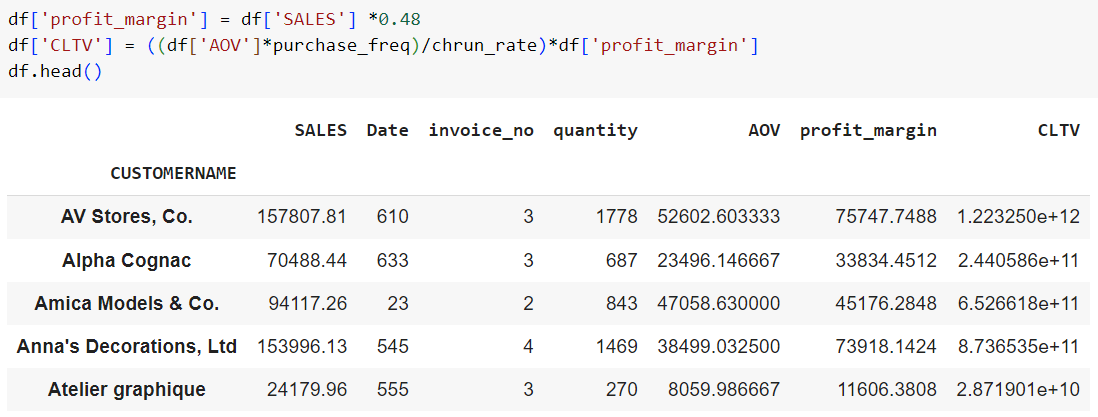
* Sales monthly, plotted a graph using matplot library, observed that high sales month is (November, October) the 4th Quarter of year



* Calculating Customer Segmentation by RFM (RECENCY , FREQUANCY, MONEYTARY)



* Calculating Profit margin ratio (Assumption Profit = PRICEEACH – MSRP)
* Calculating Customer Life Time Value CLTV



**Data Exporting:**

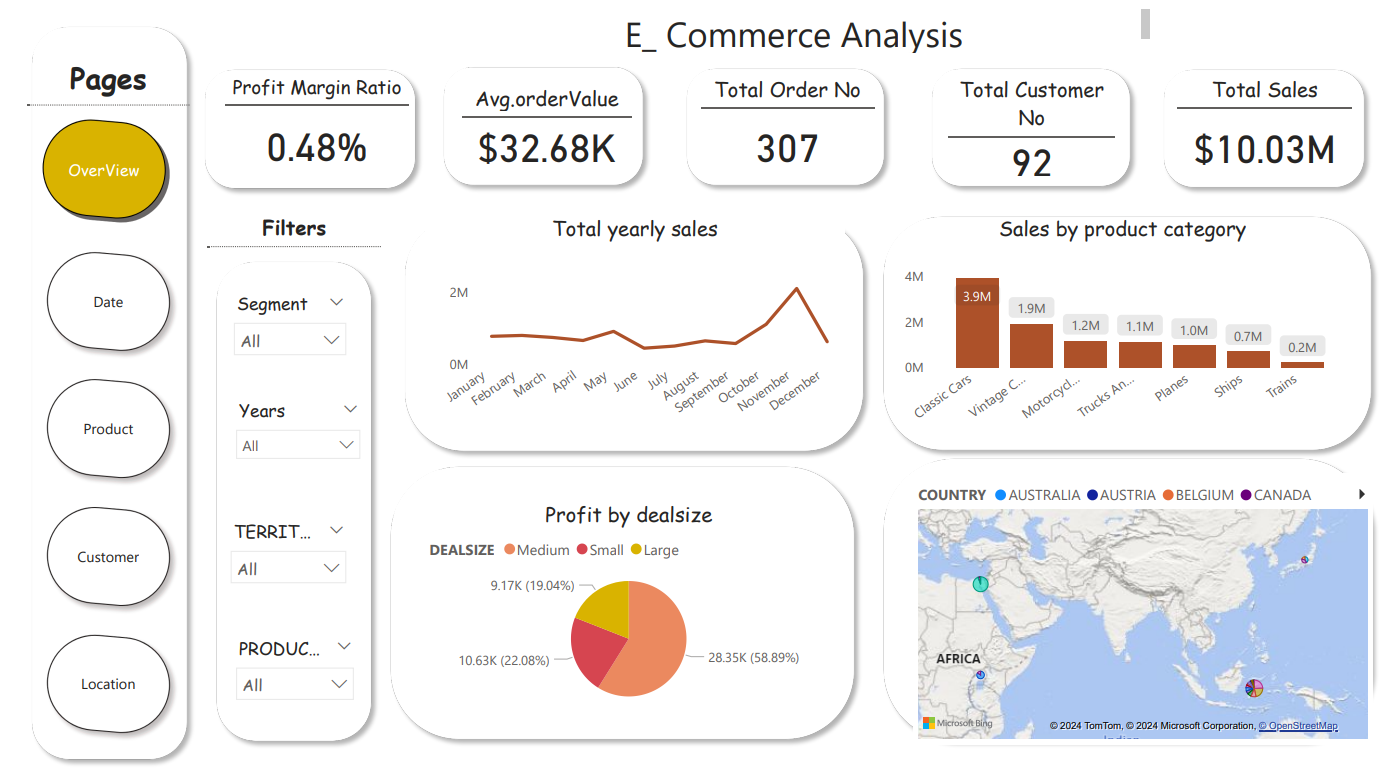
BY Using Pandas export data as excel file to be visualized by Power bi

Data Visualization:

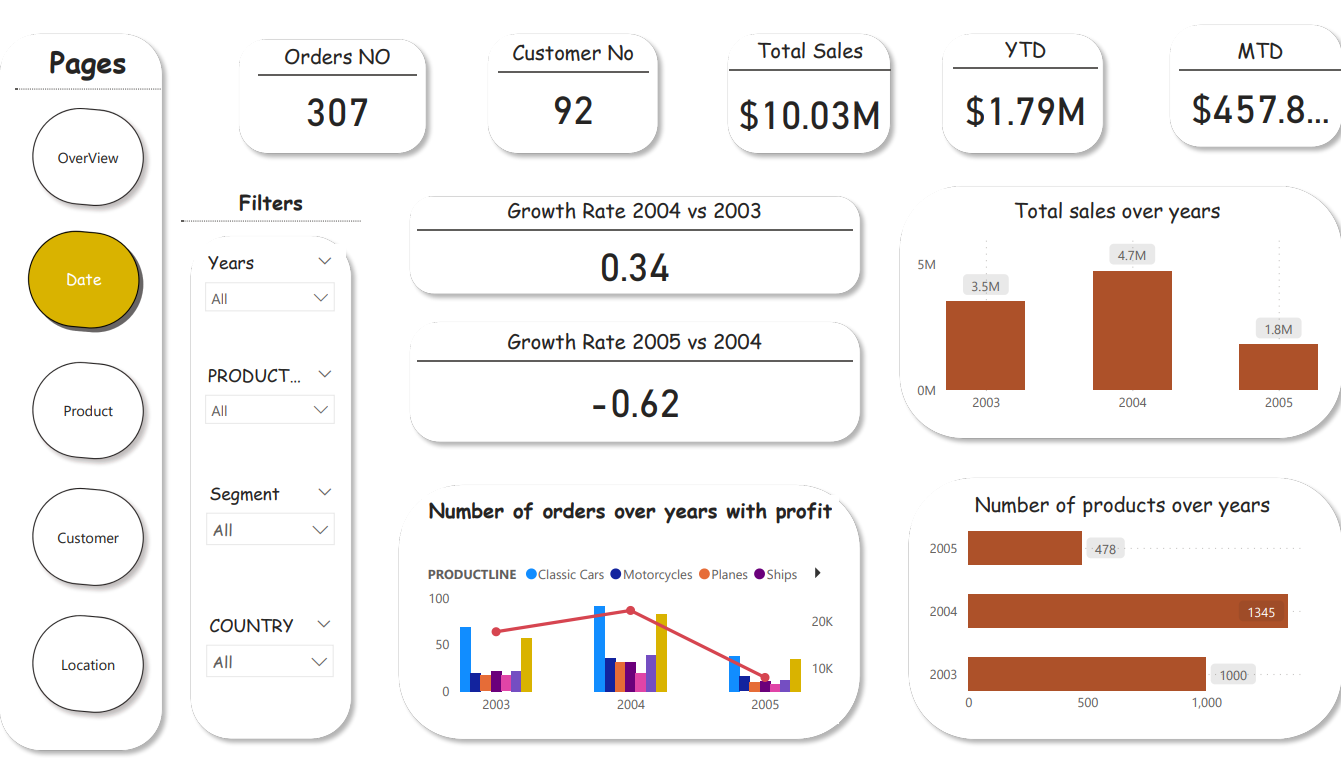
**Power BI Part**

Creating An Interactive Dashboard & Storytelling about customer purchased

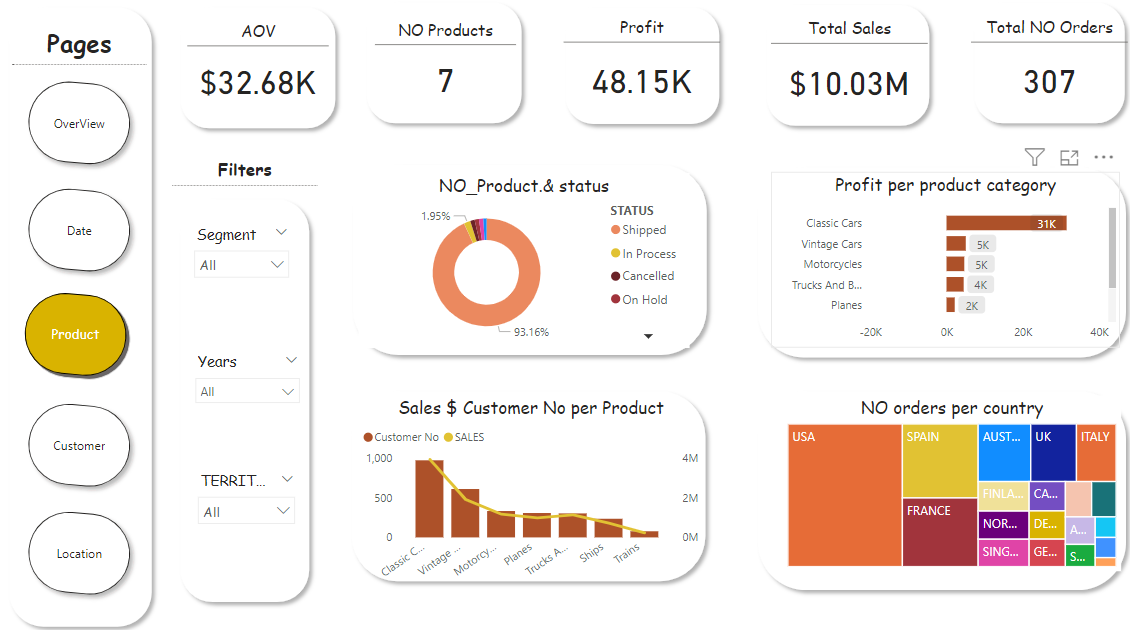
**Overview Page**



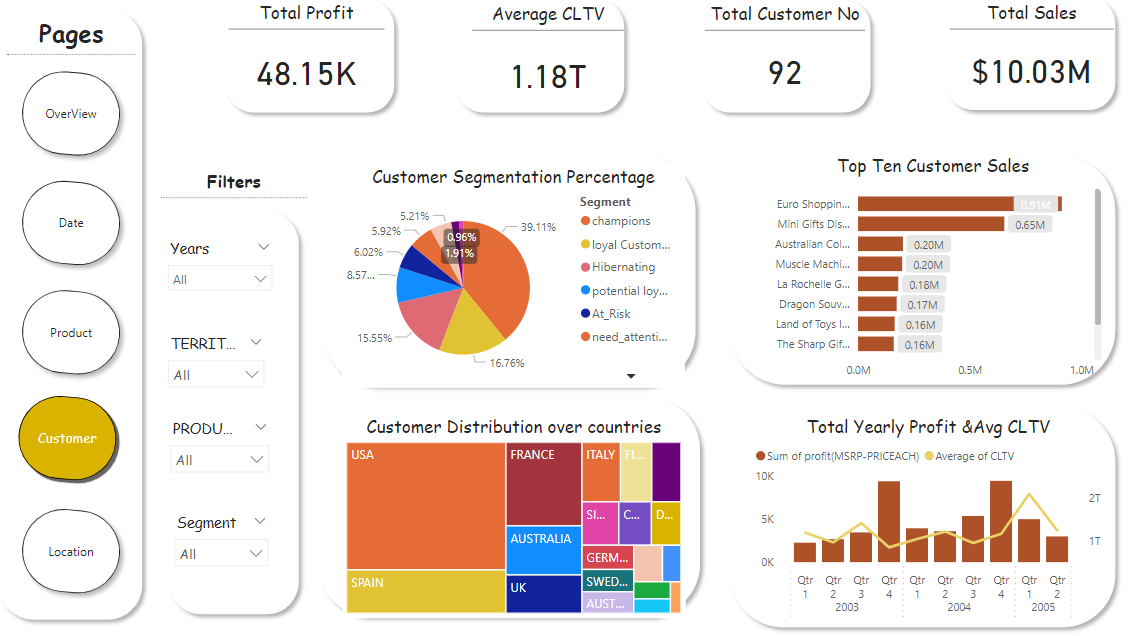
**Date Page**



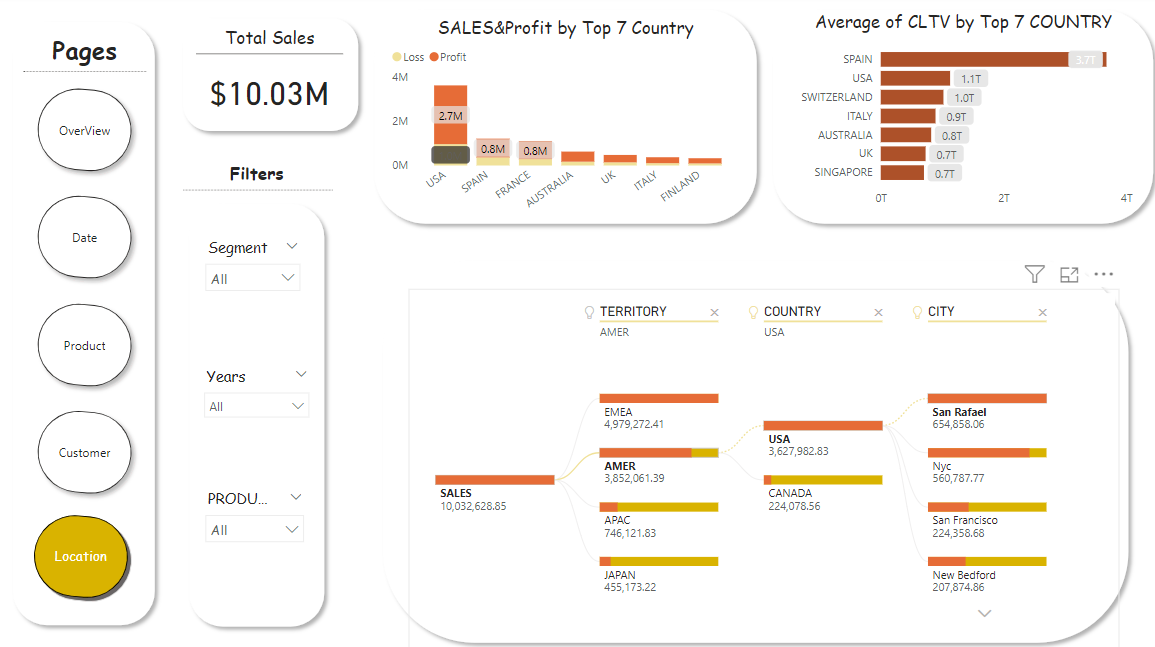
**Product Page**



**Customer Page**



**Location Page**



**Identify Business Questions:**

1. Which products have the highest sales?

The products with the highest sales is the Classic Cars.

1. What is the correlation between Profit and deal-size?

Profit increase with Deal-size up till Medium and from there, it became a downward trend.

1. 3.How do different regions perform in terms of sales and deal-size?

The EMEA territory which has more countries than other territories has the highest number of sales but individually, the sales from USA dwarfs every other country.

1. 4.Which products contribute the most to the total sales?

The biggest contributors to sales are easily the Cars, The Classic Cars and The Vintage Cars

1. What are the top Seven countries contribute with high sales ?

USA & Spain contribute with highest sales

1. Identify Countries with high average of CLTV ?

Spain, USA , Switzerland

1. Identify Customer Segmentation percentage and its contribute with total sales?

Customer segmentation is segment depend on purchased frequency and monetary, Champions share with 39% , Loyal-customers share with 16.7%, Hibrnating share with 15.5%

**Insights& Recommendations:**

Summarize your findings and insights from the analysis.

* The bestselling products are the cars, the biggest sales come from the EMEA Territory with The USA contributing the highest sales per country.
* Identify potential areas for further exploration or analysis, such as studying the impact of promotions, seasonality, or customer demographics on sales.
* Quarter 4, most especially November and October had more sales. This may likely be in anticipation of the festive period December. So increasing my marketing effort in quarter 4 will be very good for sales.
* Provide recommendations based on your analysis that could help improve sales performance.
* Well the bestselling products are the cars and the most lucrative territory is the EMEA territory considering the number of countries to target , it will be best marketing strength is targeted towards the US which us far more lucrative. So targeting The USA market with our cars will be a good marketing strategy.
* BY customer segmentation and CLTV it easily to determine Marketing Target segment (I observe that SPAIN and USA highest country of CLTV so, the best location for launching a marketing strategy )