**Customer Sales Pattern Analysis**

**Link:** https://github.com/Arvin6995/Projects/blob/9a6d053c369496e8580c218707a286405833f70a/KAGGLEARVIN1.ipynb

**Overview**

Over the past years, customers from all over the world have purchased products and placed orders on a day-to-day basis. Vendors from all over the world have provided their services satisfying customers as that is their main goal. Over the years, the sales drive has been varying from year to year where in certain years the sales go up and in certain years sales have gone down.

**Objective**

My objective is to analyze the sales pattern to identify the trends, on what basis we observe the rise in sales and what factors cause the sales to depreciate. The dataset provides raw data consisting of customer details, order details, sales, quantity ordered, product line, etc. Once, I had a understanding of the situation, I began my analysis by performing the data wrangling techniques on the dataset that I gathered from Kaggle.

**Analysis**

After performing the data cleaning followed by data analysis, I have hereby furnished my observations. I have used Python3 to perform my analysis and determine my observations. Considering Sales as the target dependant variable:

I first began by determining the total sales by region to observe which country has dominated with the most sales. As shown in the image, we can see USA being the top-most provider of products and product lines placed by customers over the years, followed by Spain and France being the 2nd and 3rd providers. We can also see Ireland and Philippines being the least providers to maintain good sales ratio.



Here is the respective graph that depicts the total sales by country:

A screenshot of a computer

Description automatically generated

Going further, I also determined the total sales by Product Line to see which product line is the most frequently ordered, we can see that Customers from all over the world order Classic Cars, Vintage Cars and Motorcycles more often compared to trucks and Buses, Planes or other means of transport. The table is shown below:



We can see that from the observation, classic cars, vintage cars and motorcycles have been the most frequently ordered product line. The observation in the graph is shown:

A graph with blue bars

Description automatically generated

I was also able to determine the count of the orders by product line to strengthen my observation in this case:

A screenshot of a pie chart

Description automatically generated

We can see that Classic Cars hold most of the proportion by 34.25%, followed by Vintage Cars that holds the proportion to 21.50% and Motorcycles holding the proportion to 11.73%. We can verify and confirm that these 3 product lines are the most placed orders by customers from all over the world.

I proceeded further with identifying the major and the most important factor which is to determine where exactly the sales began to rise and where exactly the sales began to drop. The Graph is shown below:

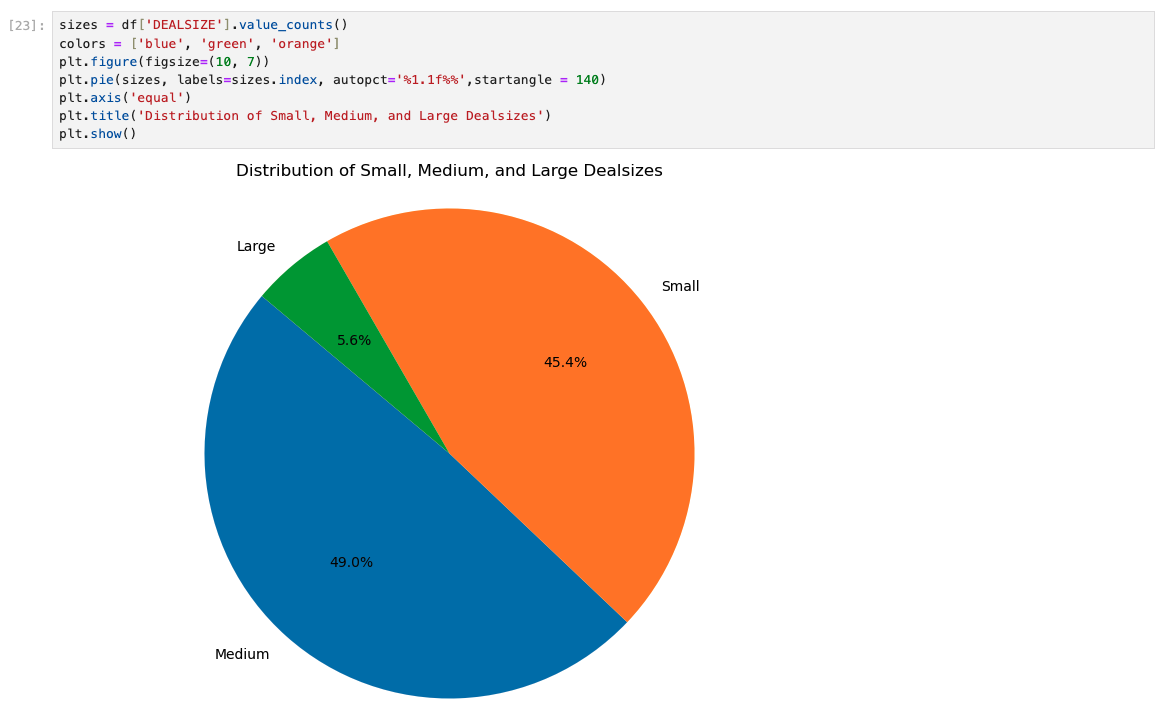
A graph with a red line

Description automatically generated

From this, we can observe that the sales had been progressing well from 2003 to 2004, then we can see that there is a drop in sales after 2004. The sales had been increasing up to more than $4.5M and then dropped to less than $2M in 2005.

I also determined the sales proportion with respect to the deal size distributions, which is the sales proportion with respect to small, medium and large deal size distributions.

The graphical representation is shown:



As shown in the graph, we can see that the sales proportion is heavily relied on medium size deal distributions taking half of the proportion which is approx. 50%, while the rest of the deal size proportions are occupied by small deal sizes having 45% and only 5 to 6% occupied by large deal sizes.

We can also see the sales by deal size and by country, as we already confirmed that USA is the leading provider with the top sales, we can see the deal size proportions by country wise which is shown in the graph:

A graph with different colored bars

Description automatically generated

**Conclusion and Recommendations:**

From the analysis as shown above, we can provide recommendations by focusing on customers in countries such as Ireland, Philippines, Belgium, Japan, etc by providing more promotions and deal packages that could bring the possibility to regain customers and increase sales. For instance, Europe has been the top leader in producing automobiles and Japan has been the top leader in Japanese variant based vehicles. The possibility to bring exclusive deals and discounts could help customers’ demand to rise and eventually anticipating a rise in sales as well. Other countries should be recommended to follow the same by focusing on what the customers prefer purchasing orders and elevate those product lines and deal sizes with promotions, discount deals and incentives. In this way, we can expect a rise in sales from these countries thereby bringing competition to the sales performance done by USA.