Good morning, everyone. My name is Moonmoon Samal and I’m learning skills to become a data analyst nd I'm here today to present my project on Online Retail Sale.

The first step in my project was to clean the data. I did this by checking for duplicate values, null values, and outliers. I also removed any rows where the quantity was below 1 unit or the unit price was below $0.

Adding a New Column

Once I had cleaned the data, I added a new column called "Revenue". This column calculates the total revenue for each order by multiplying the quantity by the unit price.

Visualization

Then in the next step i visualize the data. I did this by creating a number of charts and graphs. These charts and graphs show the distribution of the data, the trends over time, and the relationships between different variables.

I created a line chart to visualize the revenue data for each month in 2011. As you can see, October had the highest revenue of $52,681.27, which was 5,243.36% higher than August, which had the lowest revenue of $985.92. October accounted for 22.72% of total revenue for the year.

**Seasonal Trends**

As you can see from the line chart, there is a clear seasonal trend in the revenue data. Revenue is highest in the fourth quarter of the year, which is typically the holiday season. Revenue is lowest in the first quarter of the year.

As you can see on this clustered column chart, November had the highest revenue, at( 150 million 9 thousand 496.33 dollar )$1,509,496.33. This was 188.27% higher than February, which had the lowest revenue, at (5 million 23 thousand, 631 .89 dollar)$523,631.89. November accounted for 15.34% of revenue across all 12 months.

Revenue ranged from (5 million 23 thousand 631. 89 dollar )$523,631.89 to (150 million 9 thousand 496.33 dollar) $1,509,496.33. As you can see, there is a positive correlation between quantity and total revenue. This means that as the quantity of products sold increases, so does the total revenue.

The country with the biggest revenue and quantity divergence was EIRE. In EIRE, revenue was $132,018 higher than quantity. This could be due to a number of factors, such as higher prices in EIRE or a higher demand for our products in EIRE.

Nd now the top 10 customers by revenue. As you can see on this clustered column chart, customer ID 14646 has generated the most revenue, with (2 million 71 thousand 614.14 dollar) $2,71,614.14. The lowest revenue generated by a customer is (67 thousand 912.32 dollar) $67,912.32, by customer ID 16029.

As you can see, the top 10 customers generate a significant amount of revenue for our company. It is important that we target these customers and ensure that they remain satisfied with our products. We can do this by offering them discounts, loyalty programs, and other incentives.

As you know, we're always looking for new opportunities to expand our business. One way we do this is by identifying regions with high demand for our products.

We recently analyzed demand data for all countries, and we found that the Netherlands has the highest demand for our products. In fact, the Netherlands had 193,550 units sold in the last year.

This is a significant finding, as it suggests that there is a large market for our products in the Netherlands. We believe that by expanding our operations into the Netherlands, we can capture a significant share of this market.

Conclusion

Coming to the conclusion, I have successfully cleaned, analyzed, and visualized the data from the Online Retail Sale dataset. I have identified a number of interesting trends and insights, which I have presented to you today. I hope that this information will be helpful to you in making decisions about the future of of organigation.

Thank you for your time nd the opportunity

I hope this helps!

Have a good day