Nouman Ali

Modult project Data Visualization

**Link to the dashboard:**

<https://public.tableau.com/views/ModuleProjectNoumanAli/Dashboard1?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link>

<https://public.tableau.com/views/ModuleProjectNoumanAli1/Dashboard2?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link>

**Introduction:**

Here in this assignment, I have found a dataset from Kaggle which I thought would be useful for this task. This sales data represents the sales of products for the year 2003-2005. This Data set has 10414 rows and 25 columns, and it was clean and ready to be used. It has customers' names and product sales for each customer which is always what we like to analyze to examine the sales trends. Also, it has country names, the sales quarterly, and monthly as well which will be helpful in finding trends using Tableau.

In the prospect of AI/ML, we can do sales forecasting later in this course where we will be using the historical data to make predictions about future sales and trends. Also, by analyzing customer behaviors over the year we can predict what can be done to improve sales and give recommendations to the store/seller to make improvements.

**Visualization Process:**

Since the data had many columns and it had many details. However, covering all these details would take much time and they were not that necessary. I have made 8 plots/ charts which I thought were delivering the details.

1. Sales by order date: This is a very basic thing to observe in sales data. This will show a trend of overall sales that the company have made during a specific time. So in the columns, I have added “order date” and we can choose month/year to see the particular details and in rows the sum(sales) which gives the total sales made at that particular time. It is a time series data and the line chart is the best to visualize such data. We can easily keep track of the changes over time and see the trends like sudden spikes and dips.
2. Sales by product line: It is also significant for the company to see the sales of each product. This will give the idea of which product is doing well and contributing more to sales and which product might need more attention. For this, I have used bar charts as they are the best to compare objects. You put the product in ascending or descending order to have a clear picture of the performance of the products.
3. Top N customers by sales: It is similar to the previous graph but here we are considering the customer who had the most sales. On rows, we set the names of customers and on columns we gave the sales they did. Again, we are comparing the top customers using a bar chart is ideal as it is easy to compare the customers based on an order. Here I used horizontal bar charts to make the visualization much diverse.
4. Sales by country: The customer belongs to diverse countries and to see which country is responsible for what number of sales we decided to show it in a map chart. Since there were many countries using a pie chart would mess things up, so I used a map chart and visually it looks appealing to the viewers. To create this, drag the country to the rows and Tableau will automatically select the map if it didn’t select the map chart yourself. In the marks section ass Sales for text, colour, size and country are also in the text. This will make the desired plot.
5. Sales vs Quantity order: To see the trend between the sales and order size we created a scatter plot using quantity ordered on the X-axis and Sales(sum) on the Y-axis. Since there were too many points in the plot it didn’t look visually appealing, so we added order date as a filter to check this trend over small intervals of date.
6. Deal Size Distribution: This data set contained three deal sizes small, medium, and large. To visualize this a pie chart would be ideal. To create a pie chart, we placed the sum(sales) in text, size and angle and the deal size in colour. This will show which deal size brought the most sales. Also, we added a country filter to this trend for every country in the dataset.
7. Sales by shipping status: This is also an important aspect of sales data to see the order status. This will tell the seller about the status of shipping and what needs to be done to improve the shipping process see the orders in process and improve the shipping service. For this, I used a bubble chart which clearly serves the purpose. To make this graph in Tableau drag the sum(sales) to marks for size and text and drag the status in the marks for text and colour.
8. Sales ordered by quarter: Lastly, we want to see the sales quarterly and for this, we used a heatmap. the intensity of the heat map shows the sales, and the boxes represent each quarter of the year. The reason for using a heatmap was to easily find the trend based on the intensity of the colour for each quarter. To create a heatmap we place qtr id in columns and Year id in rows. And used sum(sales) in the marks box.

**Decision-Making Justification:**

I have created 8 different graphs all of which tell a different stat about the sales. I tried to use different sets of colours to make the designs more user-friendly and easy to understand. The main reason for making charts or graphs is to ease the cognitive load and deliver a summary in just one visualization. I made sure the title tells what is being displayed, properly labelled the axis so there won’t be any confusion and added the texts where I thought more detail would be better. Since there were 8 graphs/charts in total putting them in one didn’t make any sense so I decided to make 2 dashboards for that. The filter I used in the graphs was also present in the dashboard next to the respective charts so users can easily adjust the filter and get instant results on the dashboard without having to toggle between each sheet.

**Challenges and Solutions:**

The dataset had 25 columns, so the main challenge was to analyze the data and make a decision about which of those 25 rows was of the most value for the data visualization part. By carefully analyzing the data it made sense which columns told about the trend, and which were there to provide additional information. Secondly, to find the top n customer we had to parameter and use this parameter to make a calculated field and sort the top n customers. This was a bit difficult but found details on the internet and it work as required.

**Conclusion:**

In a nutshell, it was a great project to have some practical of the concepts learned in this class. It was a challenge at each step and this also made it more interesting. Repeating the same mistakes over and over again and then finally finding a solution to those was the most repetitive task of this assignment.