# Report on Customer Feedback Data

1. ## Introduction
   The main objective of this report is to analyze customer feedback data to gain insights into customer experiences and identify potential areas for improvement. The dataset included information on customer demographics, engagement metrics, and feedback ratings.

2. ## Data Preprocessing
   Data preprocessing was essential to ensure the quality and reliability of the dataset for analysis. The preprocessing steps involved:

   ### Data Cleaning:
   - **Removal of Rows with Zero Values:** Rows where columns had zero values, specifically in the tenure column, were identified and removed as all customers have customer ID, that means they have spent some time with company, so there tenure cannot be zero. This step was crucial to eliminate irrelevant data as this type of data can make the process of data analysis less efficient.

   - **Removal of Blank Rows:** Rows with missing values were removed to ensure completeness of the dataset. This was achieved using Excel techniques:

   ### Data Understanding: The relationships between columns were analyzed to understand their interdependencies and relevance to customer feedback. This involved:
   - Reviewing how tenure, completed tutorial, and other columns correlate with customer feedback ratings and scores.
   - Identifying any patterns or anomalies that could inform further analysis.

3. ## Exploratory Data Analysis (EDA)
   Following data cleaning, the dataset is subjected to exploratory data analysis to visualize and interpret the data.

- **Data Visualization:**

- **Count Of Completed tutorial by Theme**



Count of Completed tutorial by Theme

Completed tutorial ● no ● yes

reliability
- no 1871
- yes 3034
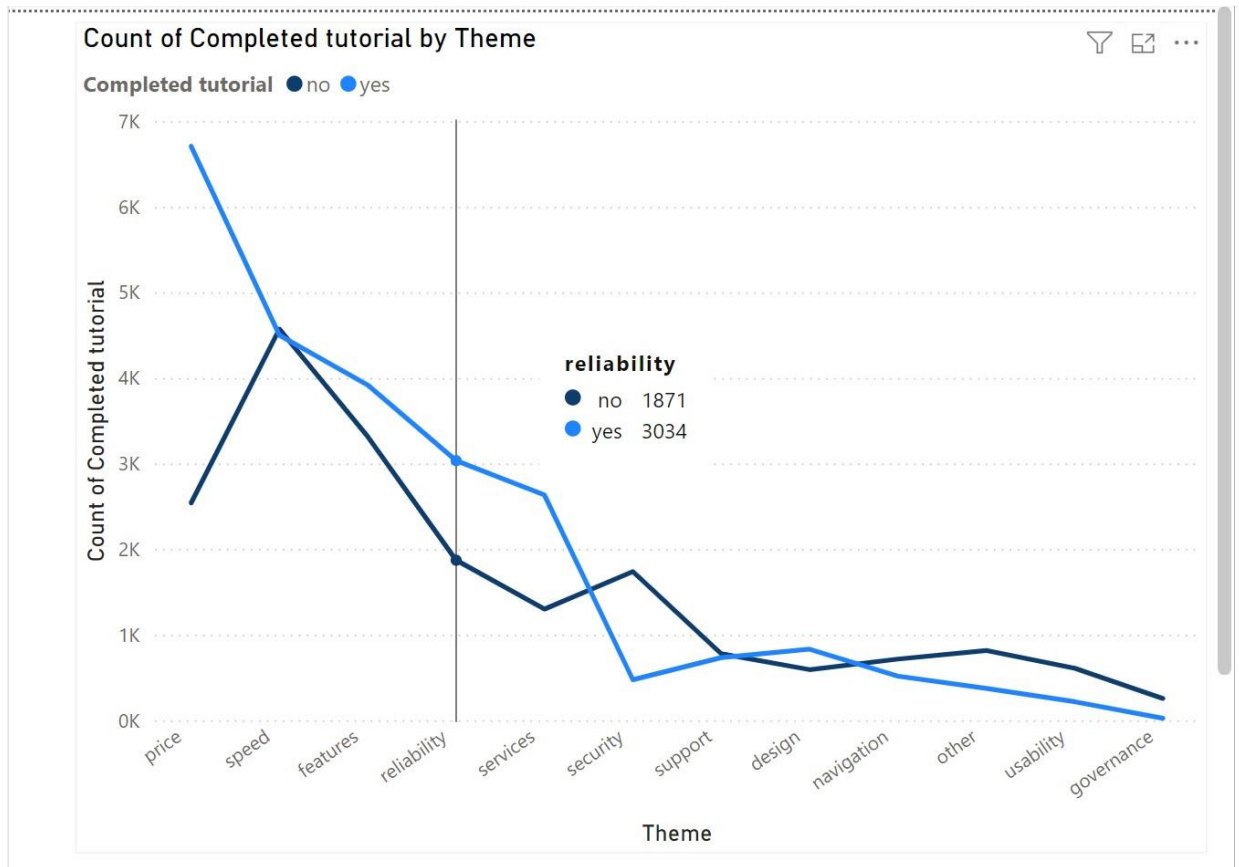
This Line chart compares the number of completed tutorials for various themes, categorized as "Reliability," "Services," "Security," and more. The x-axis represents the different themes, while the y-axis indicates the count of completed tutorials. The chart also includes a breakdown of tutorials by completion status ("yes" or "no").

**Key Insights:**

**"Price" boasts the highest completion rate , while "Speed" has the most unfinished tutorials. This suggests customers are more interested in learning about pricing compared to speed optimization.**

- ## **Count of customers by Theme**
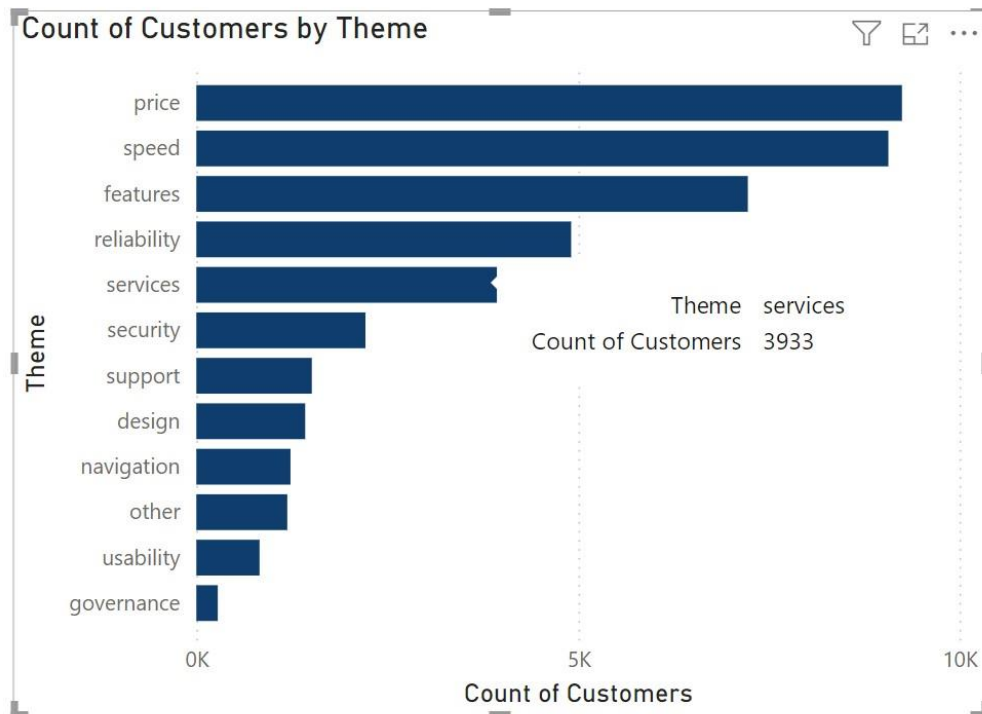


This clustered bar chart compares the number of customers interested in various themes, categorized as "Price," "Speed," "Features," and more. The y-axis represents the different themes, while the x-axis indicates the count of customers.

**Key Insights:**

**"Price" theme has the most customers i.e 9248 and "governance" theme has the least amount of customers i.e 280. This suggests more customers are inclined towards price theme rather than governance.**

- **Count of Geography by Theme**
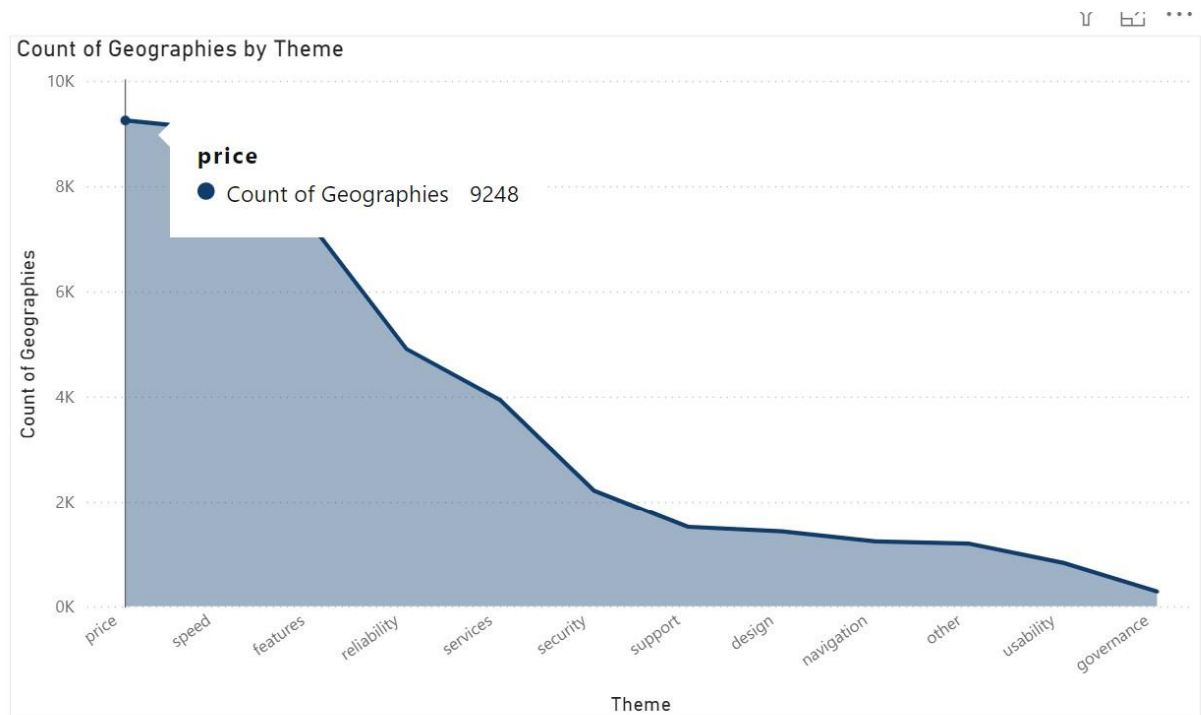


This Area chart compares the number of geographies associated with various themes. The y-axis represents the count of geographies, while the x-axis indicates the different themes.

**Key Insights:**

**"Price" has the highest number of associated geographies, followed by "Speed" and "Features." Themes like "Reliability" and "Services" also have a significant number of geographies. "Governance" stands out with the lowest count of geographies.**

- **Average Of Original Score by Role**



This line chart displays the average original score for three different roles, labeled 1, 2, and 3 on the x-axis. The y-axis represents the average original score.

**Key Insights:**

**"Role 1" has the highest average original score, followed by "Role 2" and "Role 3". There's a significant drop in the average score between "Role 1" and "Role 2". "Role 3" shows a slight increase in the average score compared to "Role 2".**
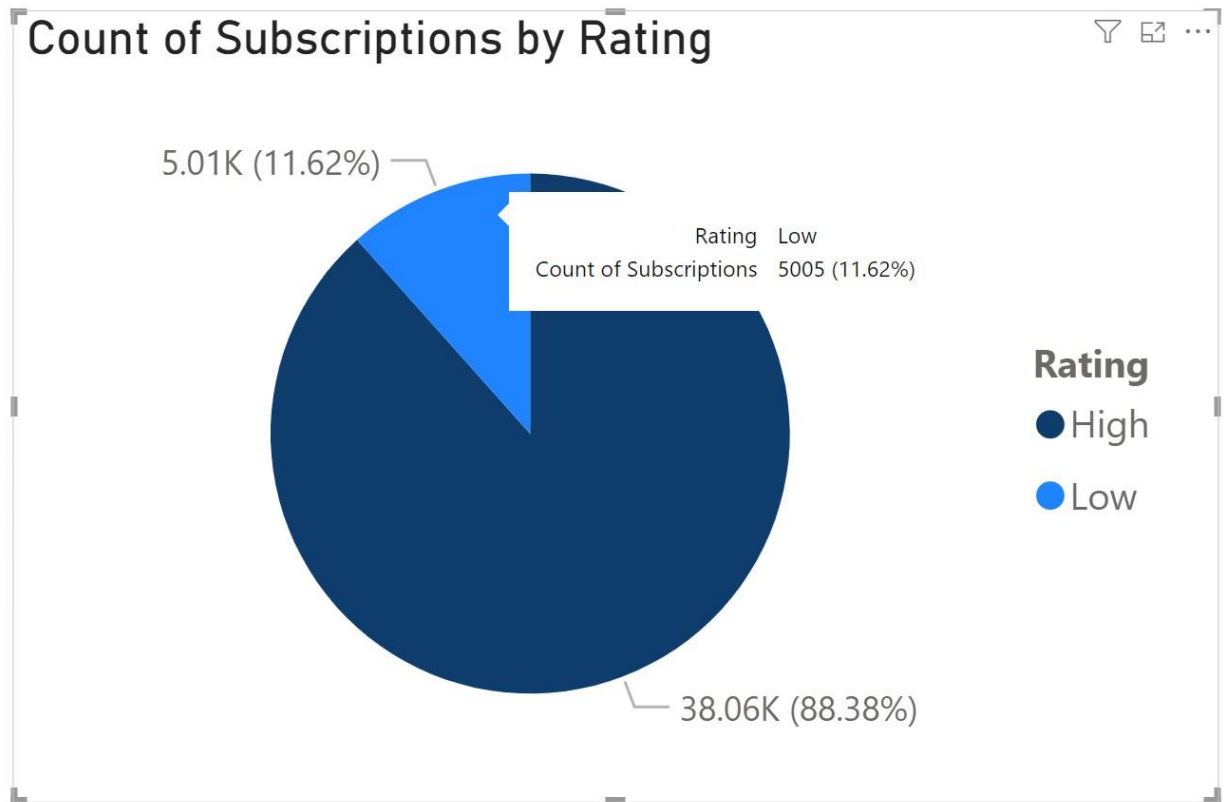
- **Count of Geographies by Original Score**
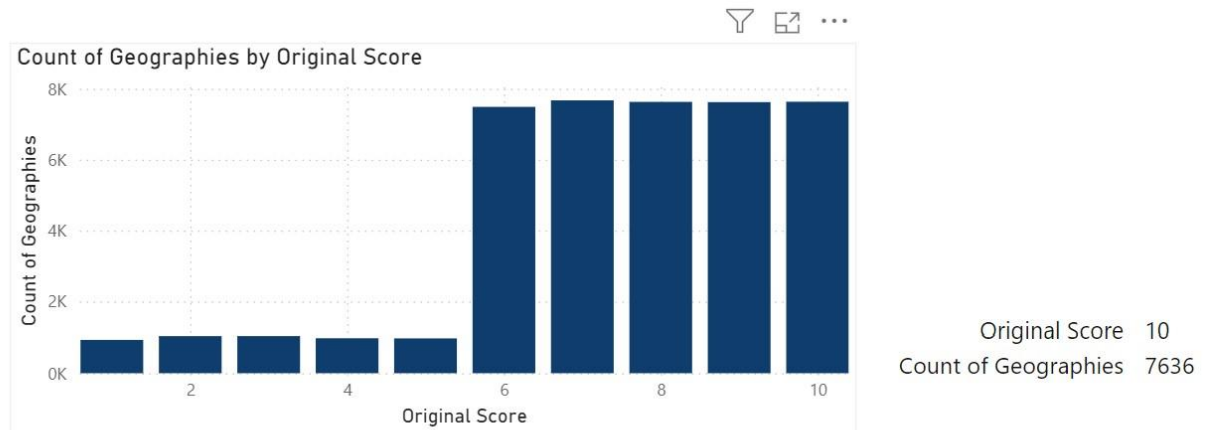


This stacked column chart illustrates the count of geographies across different original score values. The x-axis represents the original score, ranging from 1 to 10, while the y-axis indicates the count of geographies.

**Key Insights:**

**"Original score 7" has the highest number of geographies, with a count of 7675. There is a slightly downward trend in the number of geographies as the original score increases. "Original score 1" has the lowest count of geographies, with a count of 940 suggesting a lower representation in that category.**

- **Count of Subscriptions by Ratings**



This pie chart illustrates the distribution of subscriptions based on their rating.

**Key Insights:**

**The chart reveals that a significantly larger portion of subscriptions, representing 88.38%, fall under the "High" rating category. Conversely, only 11.62% of subscriptions received a "Low" rating. This indicates a strong preference towards high ratings among subscribers.**

- **Count of Customers by Original Score**
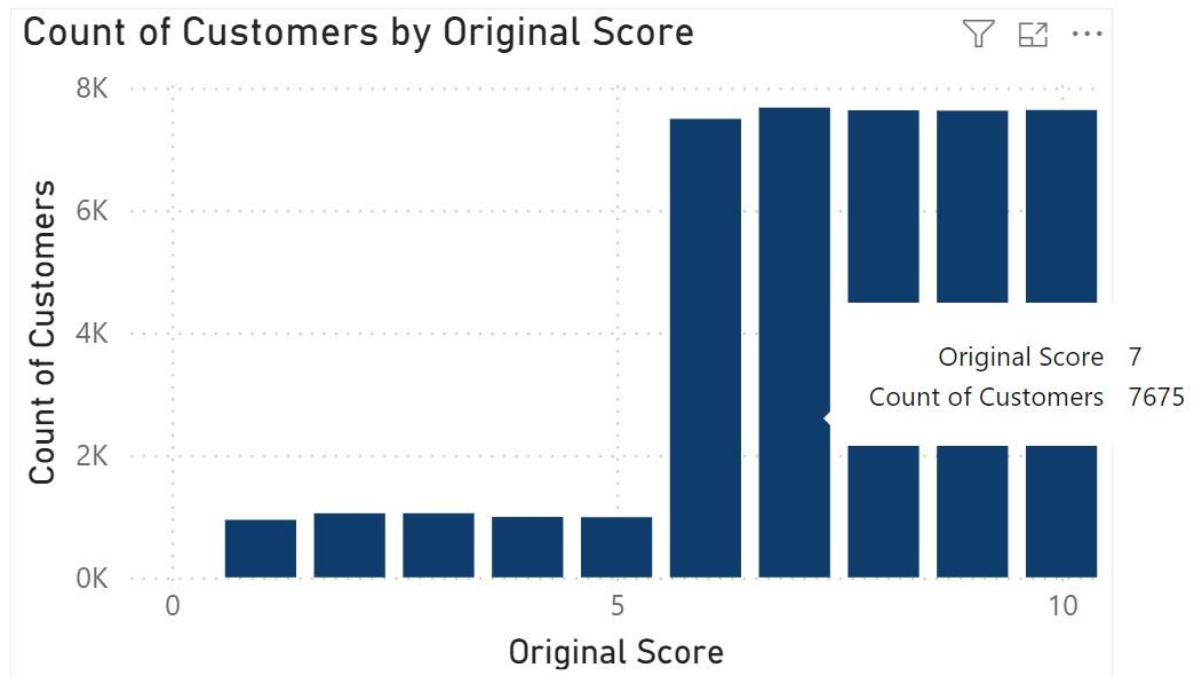


This clustered column chart illustrates the distribution of customers based on their original score. The x-axis represents the original score, ranging from 1 to 10, while the y-axis indicates the count of customers.

**Key insights:**

**"Original score 7" has the highest number of customers, with a count of 7675. There is a noticeable peak at original score 7, followed by a gradual decrease in customer count for higher scores. "Original score 1" has the lowest count of customers, suggesting a smaller segment with this score.**

## Conclusion

The data preprocessing and exploratory data analysis provided critical insights into customer feedback patterns. By cleaning the dataset and visualizing key metrics, significant trends and correlations were identified, highlighting areas for potential improvement in customer experience. The findings in this report underscore the importance of thorough data preparation and visualization in deriving actionable insights, guiding future analysis and decision-making.