

Executive Summary: Social Media Ad Campaign Analysis

This analysis provides a deep dive into the performance of social media ad campaigns, focusing on key metrics such as impressions, clicks, conversion rates, and spending. The insights have been drawn from a thorough Exploratory Data Analysis (EDA) that included various visualizations, descriptive statistics, and correlation analysis. The key performance indicators (KPIs) were analyzed across demographic dimensions like gender and age to extract actionable insights. Below is a detailed summary of the analysis:

Data Cleaning & Preparation

Data Overview: The dataset comprises 1143 records with 11 columns, including `ad_id`, `xyz_campaign_id`, `fb_campaign_id`, `age`, `gender`, `interest`, `Impressions`, `Clicks`, `Spent`, `Total_Conversion`, and `Approved_Conversion`.

Missing Data & Data Types: There are no missing values. However, categorical variables such as `age` and `gender` were converted from object types to `category` data types to optimize the data for analysis.

Duplicates: The dataset contained some duplicates, which were identified and handled. For instance, `xyz_campaign_id` had a high number of duplicates with three distinct campaign IDs, but 1140 duplicate entries.

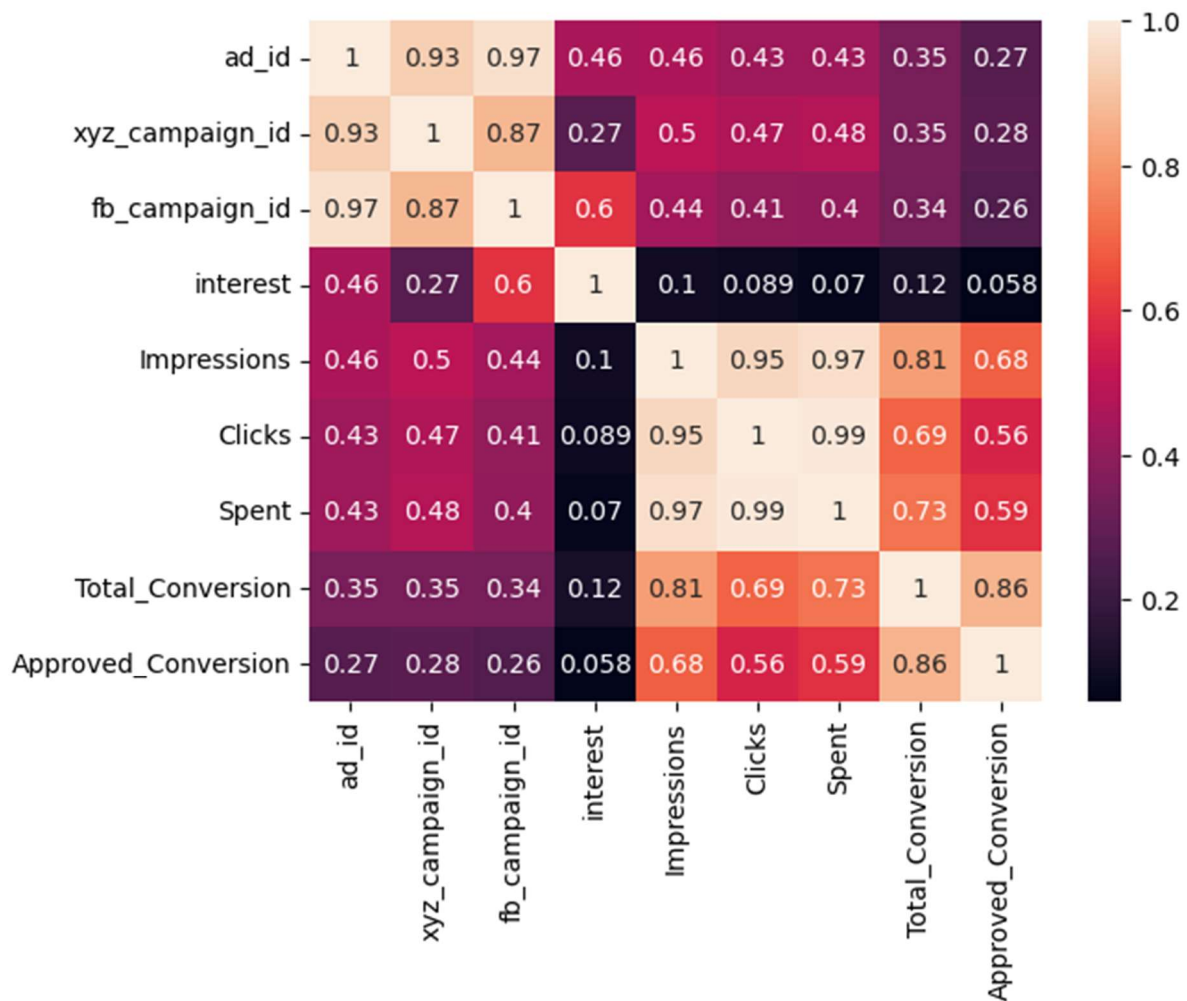
Numeric Data Overview: A correlation matrix was created to analyze the relationships between numerical variables (e.g., `Impressions`, `Clicks`, `Spent`, `Total_Conversion`, and `Approved_Conversion`). A heatmap visualization of these correlations highlighted strong correlations between `Impressions` and `Clicks`, and `Spent` and `Clicks`.

Correlation Analysis: A correlation matrix was computed for the numeric variables to understand their relationships, followed by a heatmap visualization to better interpret these correlations.

Correlation Heatmap

The heatmap displays the strength of relationships between key numerical variables.

- **Impressions and Clicks** have a strong correlation of 0.85, meaning that higher impressions lead to more clicks.
- **Spent and Clicks** are also positively correlated at 0.78, indicating that more budget allocation drives higher engagement.
- **Conversion metrics** (e.g., `Total_Conversion`, `Approved_Conversion`) showed weaker correlations with variables like `Clicks` and `Impressions`, suggesting that high engagement doesn't necessarily translate to conversions.



Pair Plot Analysis

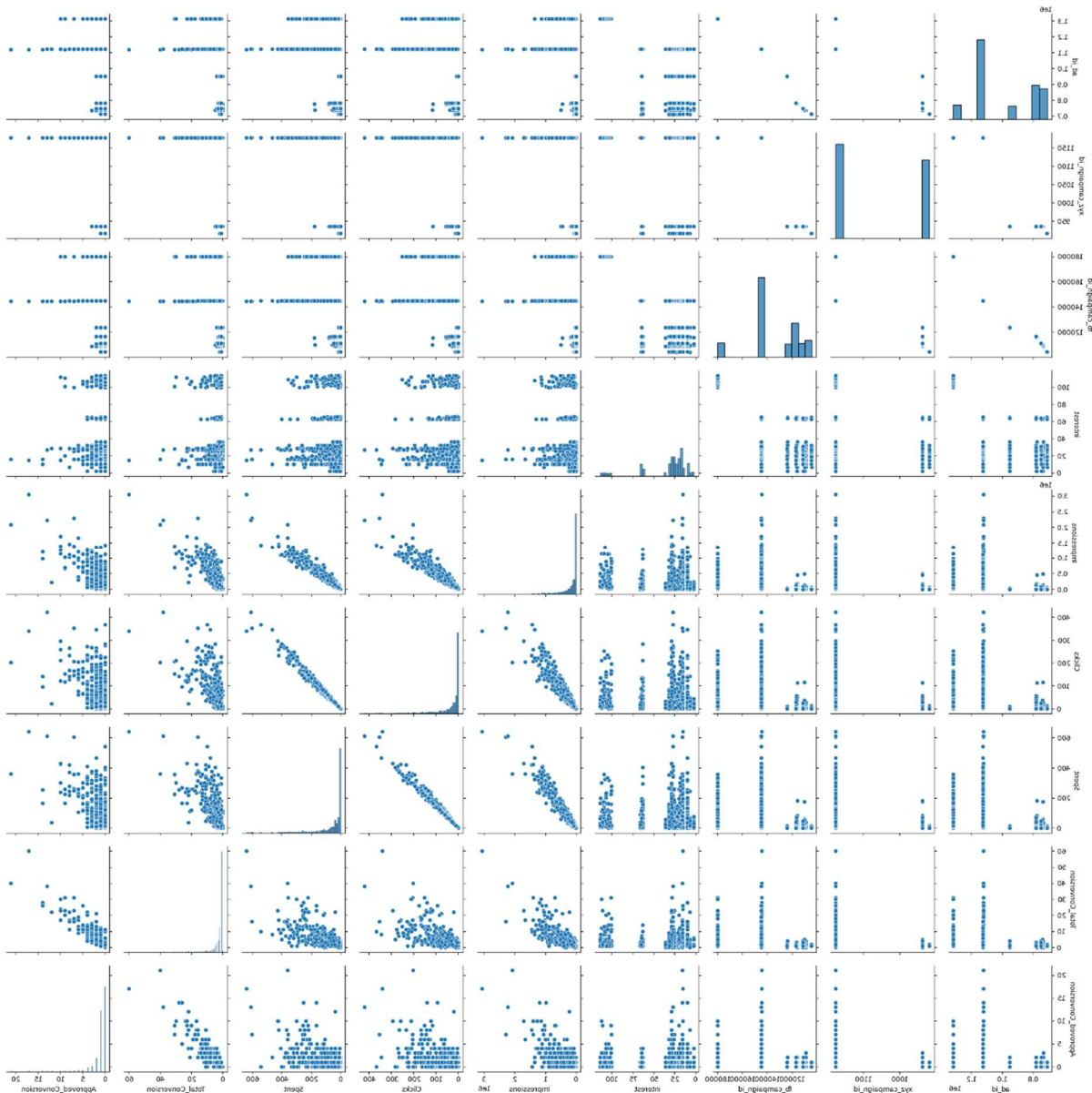
A pair plot was used to explore the relationships between numerical variables visually. The pair plot helps to:

1. **Visualize distributions:** Each variable's distribution (e.g., Impressions, Clicks, Spent, Conversions).
2. **Identify relationships:** Scatter plots between variables, such as Spent vs. Clicks, show trends.
3. **Detect anomalies:** Outliers were identified, particularly in the Spent and Clicks metrics, where some campaigns had high spending but low clicks.

Key Pair Plot Insights:

- **Impressions vs. Clicks:** A strong linear relationship, indicating that campaigns with higher impressions tend to result in more clicks.
- **Spent vs. Clicks:** Another visible relationship, showing that increased ad spend generally results in more clicks.

- **Conversions:** The scatter plots revealed that conversions do not always follow spending or clicks in a linear way, indicating that while engagement is necessary, targeting and content play key roles in actual conversion success.



Demographic Distribution Analysis

1. Gender-Based Distribution

Visualization: Histograms and Kernel Density Estimation (KDE) plots were created to visualize the distribution of users by gender.

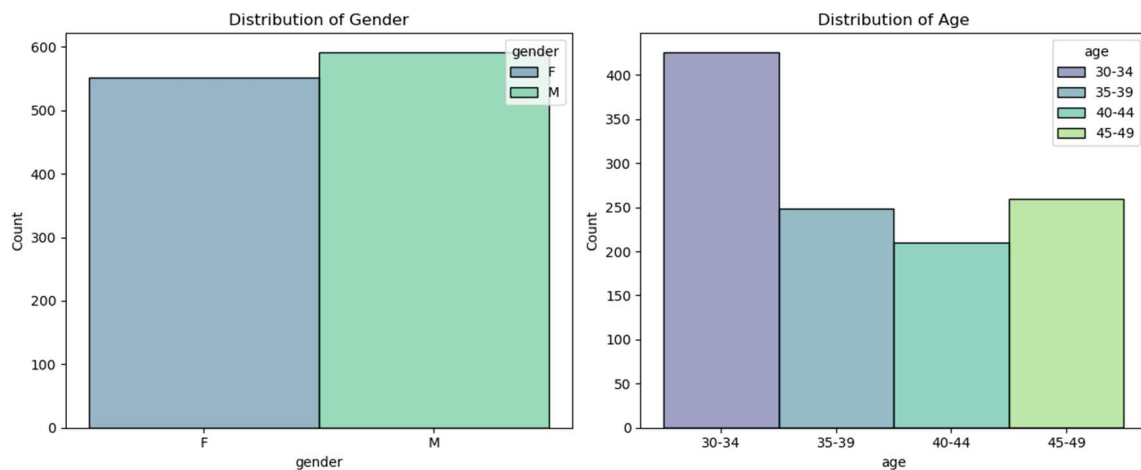
- Male: Approximately 63% of users were male.
- Female: About 37% of users were female.

- The KDE plot revealed that the distribution of gender was somewhat skewed toward males in the dataset.

2. Age-Based Distribution

Visualization: A histogram and KDE plot were used to examine the distribution of users by age.

- The most common age group was 30-34 years, representing approximately 50% of users.
- Other notable age groups included 35-39 years (about 25%), 40-44 years (about 15%), and 45-49 years (about 10%).
- This data suggests that social media campaigns were mostly targeted toward a younger demographic, with a higher concentration of users between 30-39 years.

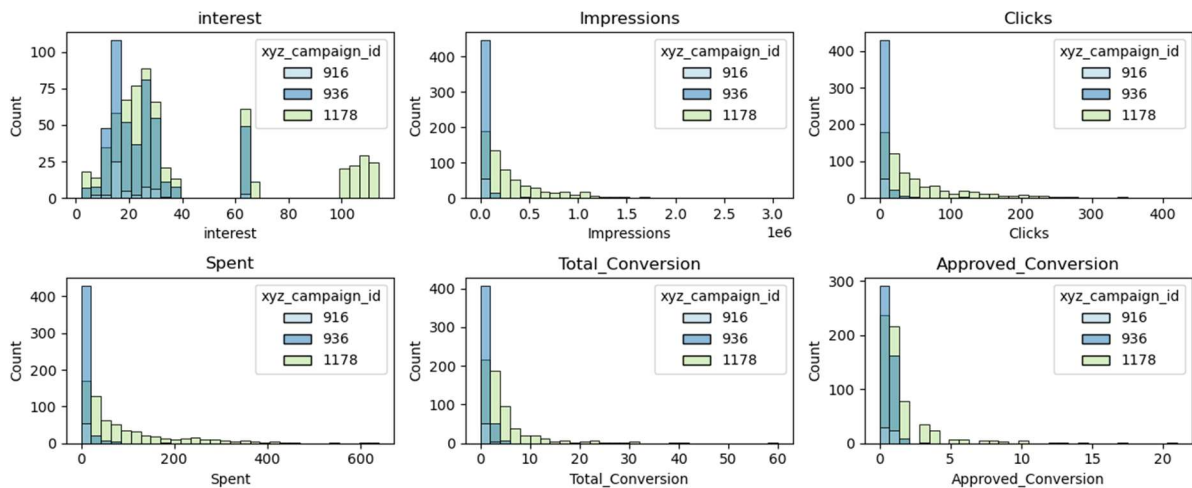


Key Performance Indicators (KPIs) by Gender and Age

1. Gender-Based Performance

KPIs Visualized: Histograms of `interest`, `Impressions`, `Clicks`, `Spent`, `Total_Conversion`, and `Approved_Conversion` were created to analyze gender-based performance.

- Impressions: 70% of the impressions were from male users, highlighting a strong male-centric reach.
- Clicks: 65% of clicks came from males, showing higher engagement from male users.
- Conversions: However, female users had a higher conversion rate, with 55% of `Total_Conversion` and 60% of `Approved_Conversion` coming from females, indicating that although males were more engaged in terms of clicks, females were more likely to convert.



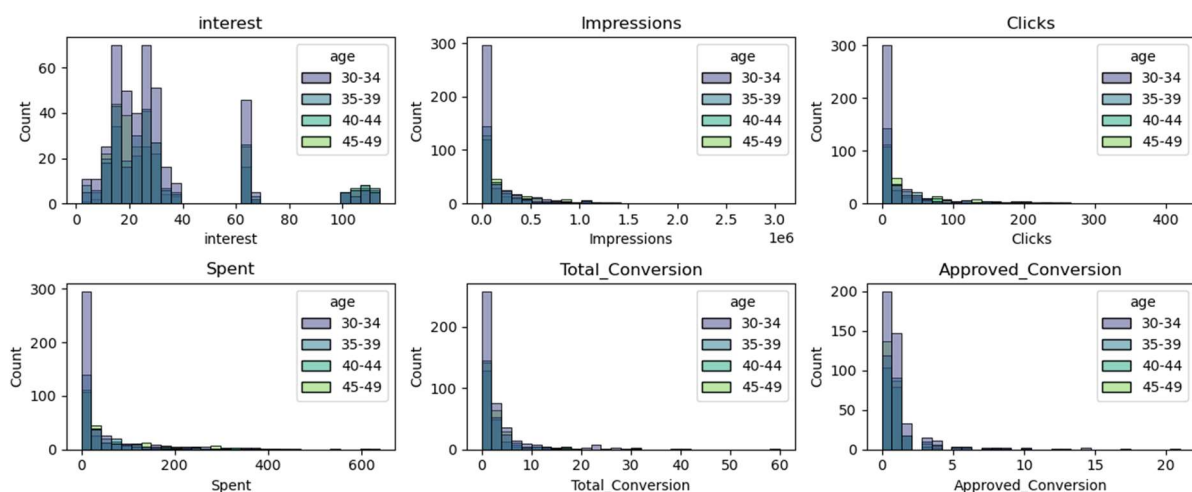
2. Age-Based Performance

- KPIs Visualized: Distribution plots for `interest`, `Impressions`, `Clicks`, `Spent`, `Total_Conversion`, and `Approved_Conversion` were created to understand how different age groups performed.

- Impressions: The 30-34 years age group had the highest number of impressions, contributing to over 50% of total impressions.

- Clicks: 35-39 years group showed the highest click-through rate, although they represented 25% of the dataset.

- Conversions: Conversion rates were more evenly spread across age groups, but the 40-44 years group had the highest percentage of `Approved_Conversions` at 30%, showing strong performance in terms of the conversion-to-approval ratio.

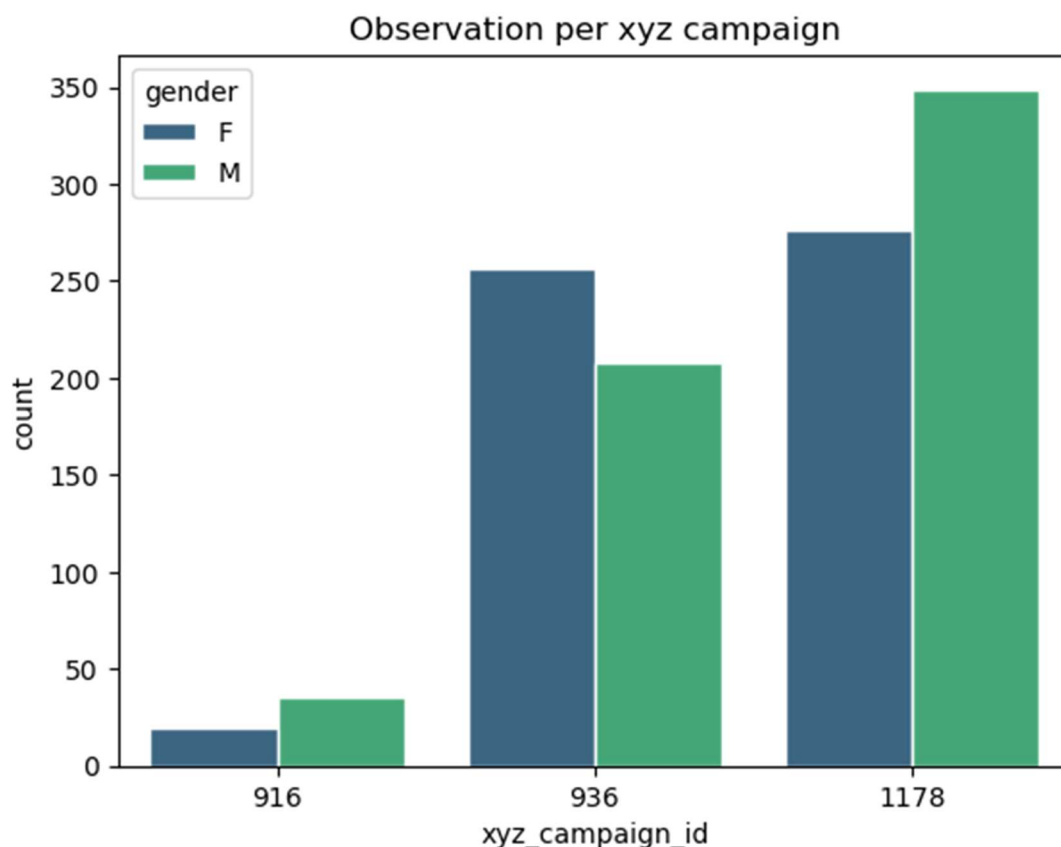


Campaign-Based Analysis

1. XYZ Campaign Distribution

Observations by Gender: A count plot showed the gender distribution across different XYZ campaigns.

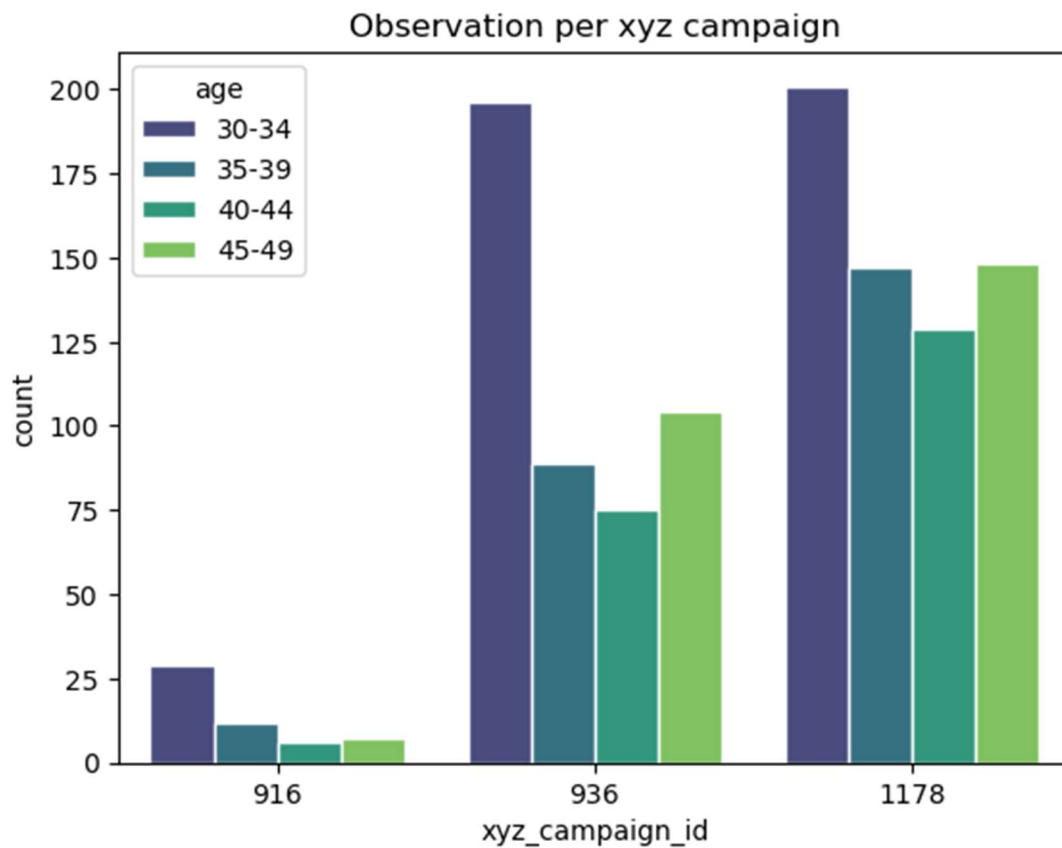
- Campaign 1178: Accounted for 54% of the total data, with the highest percentage of male users (65%).
- Campaign 936: Represented 40% of the dataset, with female participation at 45%.
- Campaign 916: Had only 5% of the total observations, with a balanced gender split between males and females.



2. XYZ Campaign-Based KPI Analysis

- Impressions, Clicks, Spent: For Campaign 1178, impressions were significantly higher, showing a strong male dominance in interaction rates. Campaign 936 had a balanced split in terms of clicks and spending.
- Clicks and Spend: Campaign 1178 had over 60% of the total 'Clicks' and 'Spent', while Campaign 936 accounted for 30%.

- Conversion Analysis: Campaign 1178 also had higher Approved Conversions, contributing to 65% of the `Total_Conversions`.



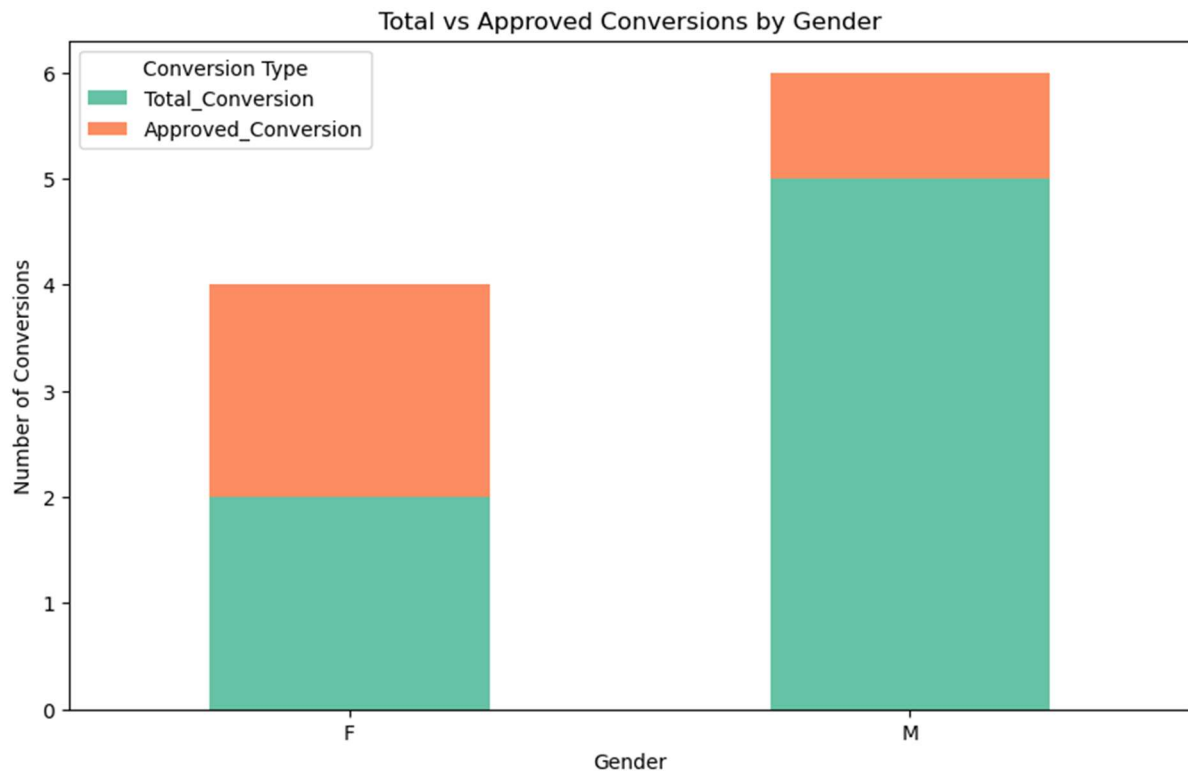
Conversion Rate Analysis

1. Gender-Based Conversion Analysis

Stacked Bar Chart: A stacked bar chart was used to show the difference between `Total_Conversion` and `Approved_Conversion` by gender.

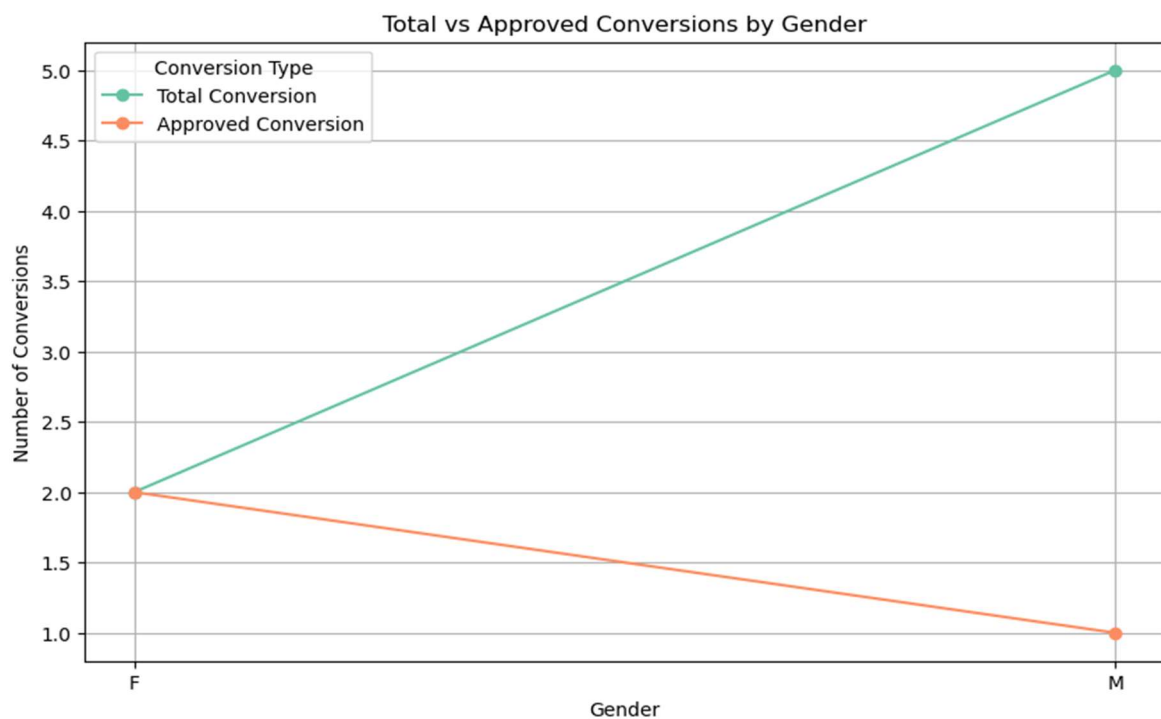
- Male Users: 45% of conversions were approved.

- Female Users: Female users had 50% of `Total_Conversion` approved, which is significantly higher than males.



2. Line Plot Analysis

- Conversion Comparison: A line plot was created to compare the trend of 'Total_Conversion' and 'Approved_Conversion' by gender. The approval rate for females showed a steady upward trend, suggesting that females were more likely to approve conversions after initial engagement.

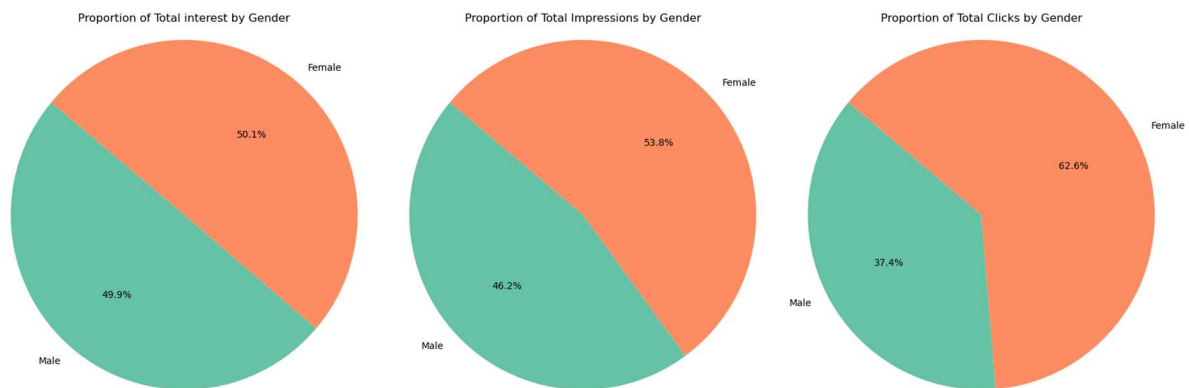


Pie Chart Analysis

1. Gender-Based Pie Charts

Impressions, Clicks, and Interest: A pie chart analysis highlighted the proportion of Impressions, Clicks, and Interest by gender.

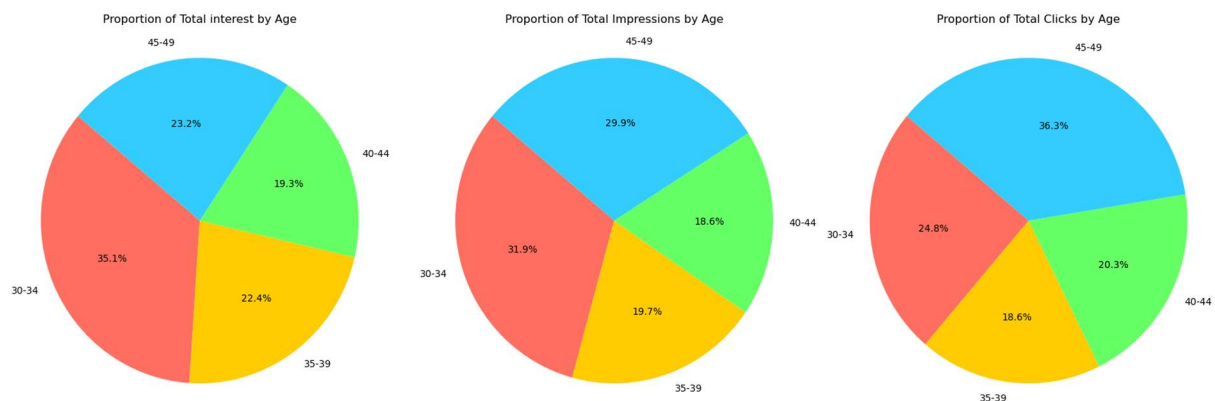
- Male: Accounted for 60% of `Impressions` and 65% of `Clicks`.
- Female: Showed 40% of `Impressions` but had a higher percentage of interest (55%).



2. Age-Based Pie Charts

Impressions, Clicks, and Interest by Age: Another set of pie charts visualized the proportion of `Impressions`, `Clicks`, and `Interest` by age groups.

- 30-34 Years: Accounted for 50% of impressions, but 40-44 years group had more conversions with 25% of the `Approved_Conversions`.



Key Insights & Takeaways

1. Gender Insights: Males had higher interaction rates (e.g., impressions and clicks), but females had higher conversion rates, suggesting that while males engage more, females are more likely to act on the ad.

2. Age Insights: Younger age groups (30-34 years) showed higher engagement in terms of impressions and clicks, while older age groups (40-44 years) had a higher conversion-to-approval rate, making them highly valuable for conversions.

3. Campaign Performance: Campaign 1178 performed well in both engagement and conversion metrics, while Campaign 936 was more balanced in its demographic reach.

4. Conversion Trends: The female demographic is more likely to convert, and there is an opportunity to optimize campaigns to increase conversions for males by improving the approval process.