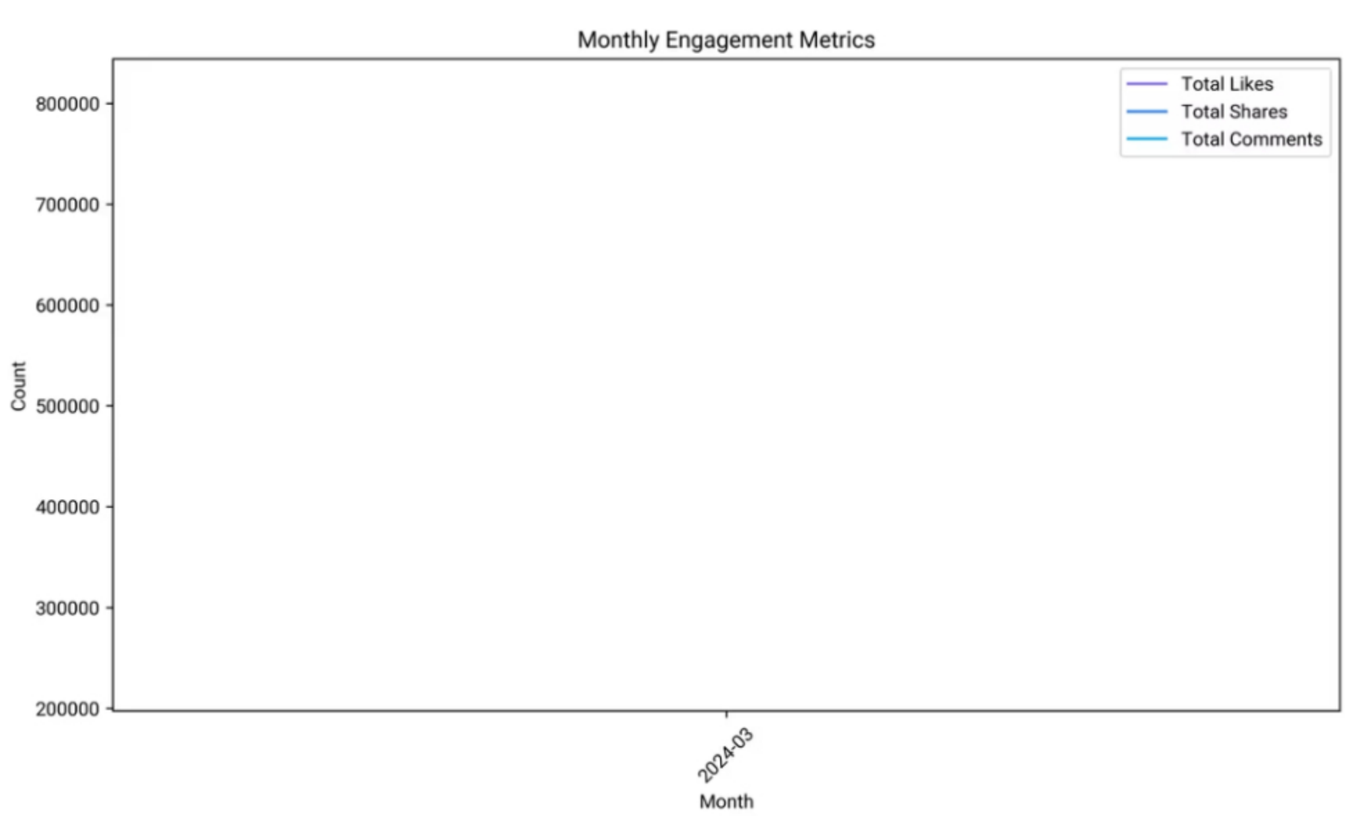
Analysis of Social Media Engagement and Sentiment

About the dataset

Analysis of Social Media Engagement and Sentiment

Relevant Inquiries

Q1.What are the monthly trends in user engagement metrics (Likes, Shares, Comments) over the past year?



Data Aggregation

- **Time Frame**: The data was filtered to include only records from the past year, specifically March 2024.
- **Metrics Aggregated**: The sum of Likes, Shares, and Comments was calculated for March 2024.
- **Values**:
  - **Total Likes**: 814,688
  - **Total Shares**: 425,597
  - **Total Comments**: 227,010

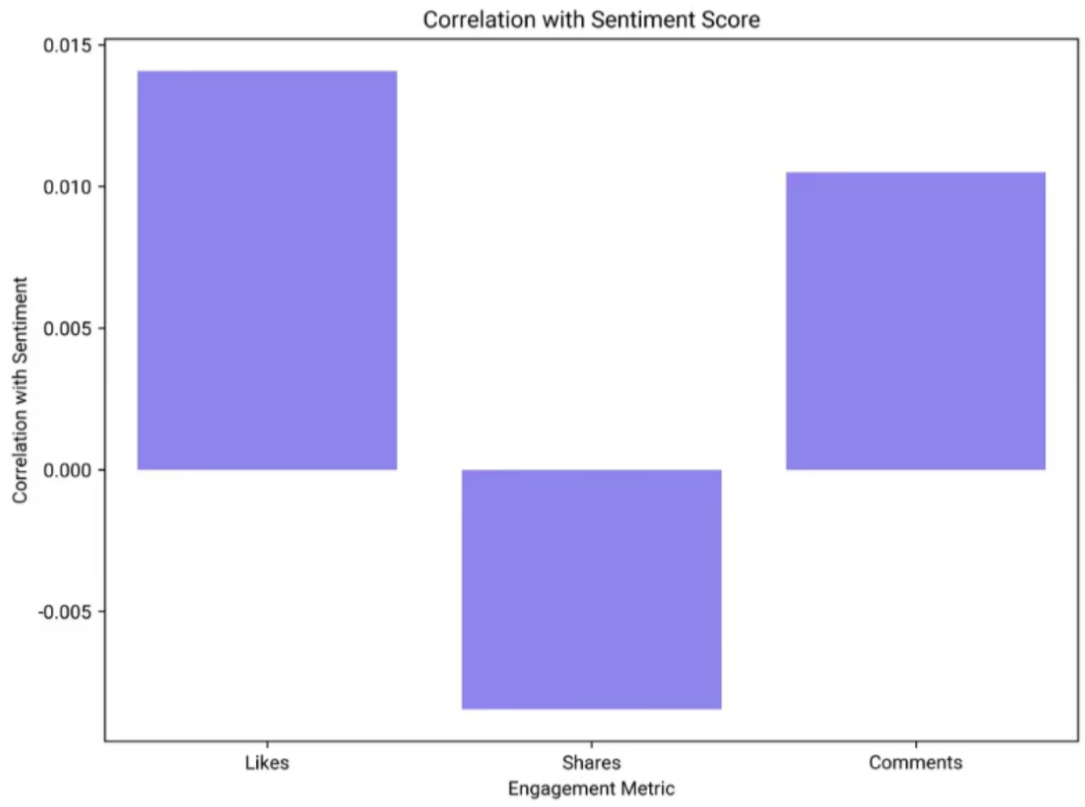
Visualization

- **Chart Type**: A line chart was used to display the monthly engagement metrics.
- **Observation**: The chart shows the engagement metrics for March 2024, with Likes, Shares, and Comments plotted.

Conclusion and Insights

- **Single Month Data**: The data and visualization only cover March 2024, indicating a need for more monthly data to observe trends over the entire year.
- **High Engagement**: March 2024 shows significant user engagement, with Likes being the highest metric, followed by Shares and Comments.

Q2.What is the correlation between Sentiment_Score and user engagement metrics (Likes, Shares, Comments)?



Correlation Coefficients

- **Likes**: The correlation with Sentiment_Score is **0.0141**, indicating a very weak positive relationship.
- **Shares**: The correlation with Sentiment_Score is **-0.0084**, showing a very weak negative relationship.
- **Comments**: The correlation with Sentiment_Score is **0.0105**, suggesting a very weak positive relationship.

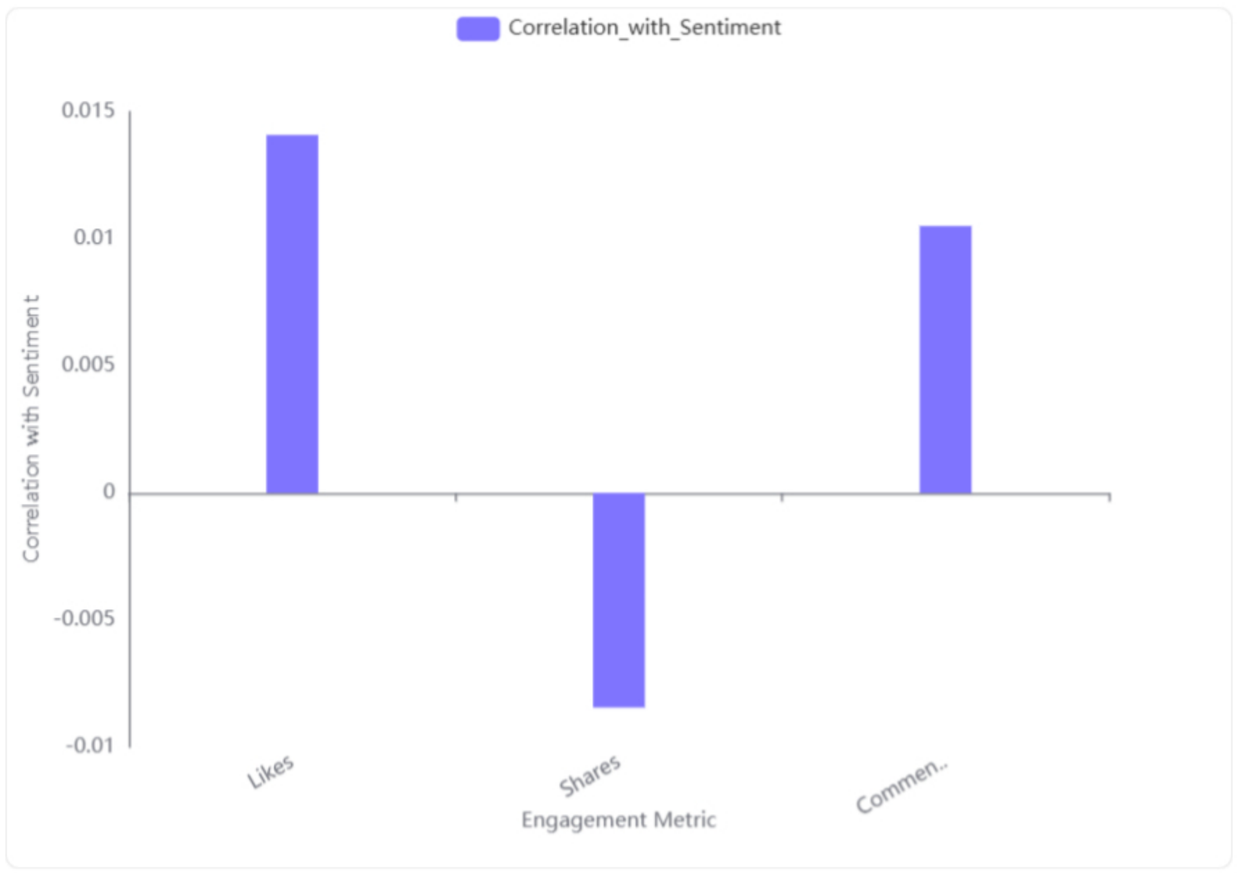
Visualization of Correlations

- The bar chart illustrates the correlation values, confirming the weak relationships between Sentiment_Score and each engagement metric.

Conclusion and Insights

- **Overall Weak Correlations**: The correlations between Sentiment_Score and the engagement metrics (Likes, Shares, Comments) are very weak, indicating that changes in sentiment do not strongly influence these metrics.
- **Implications for Strategy**: This suggests that other factors might be more influential in driving user engagement, and sentiment alone may not be a strong predictor of engagement levels.

Q3.Analyze the correlation between the Sentiment_Score and the user engagement metrics (Likes, Shares, Comments) to determine if sentiment has a significant impact on post virality.



Correlation Coefficients

- **Likes**: The correlation with Sentiment_Score is **0.014**, indicating a very weak positive relationship.
- **Shares**: The correlation with Sentiment_Score is **-0.008**, suggesting a very weak negative relationship.
- **Comments**: The correlation with Sentiment_Score is **0.011**, showing a very weak positive relationship.

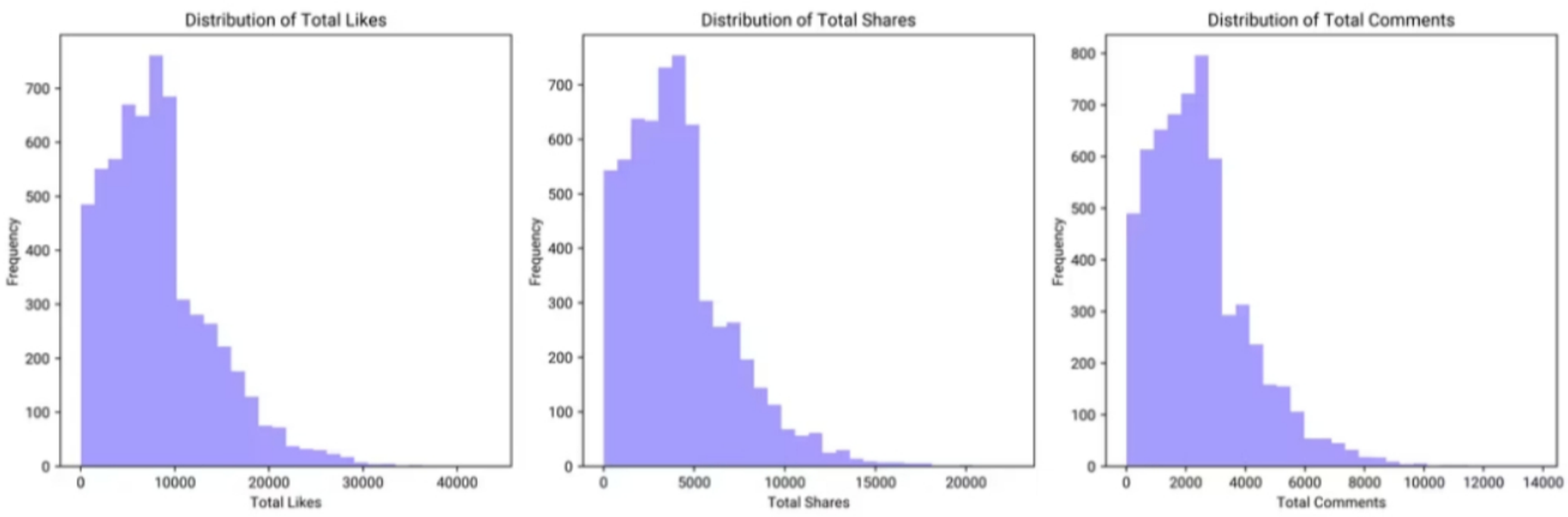
Visual Assessment

- **Scatter Plots and Heatmap**: The visual representation confirms the weak correlations between Sentiment_Score and the engagement metrics (Likes, Shares, Comments).

Conclusion and Insights

- **Weak Correlation**: The correlations between Sentiment_Score and engagement metrics are very weak, indicating that sentiment does not have a significant impact on post virality.
- **Engagement Factors**: Other factors beyond sentiment may play a more substantial role in influencing user engagement metrics such as Likes, Shares, and Comments.

Q4.What is the distribution of user engagement (Likes, Shares, Comments) across different user segments?



User Engagement Metrics

- **Total Likes**: The average number of likes per user is 8,280.36, with a standard deviation of 5,672.18. The range of likes spans from 4 to 43,581.
- **Total Shares**: Users have an average of 4,132.04 shares, with a standard deviation of 2,915.22. Shares range from 1 to 22,633.
- **Total Comments**: The average number of comments is 2,458.07, with a standard deviation of 1,698.42. Comments range from 1 to 13,780.

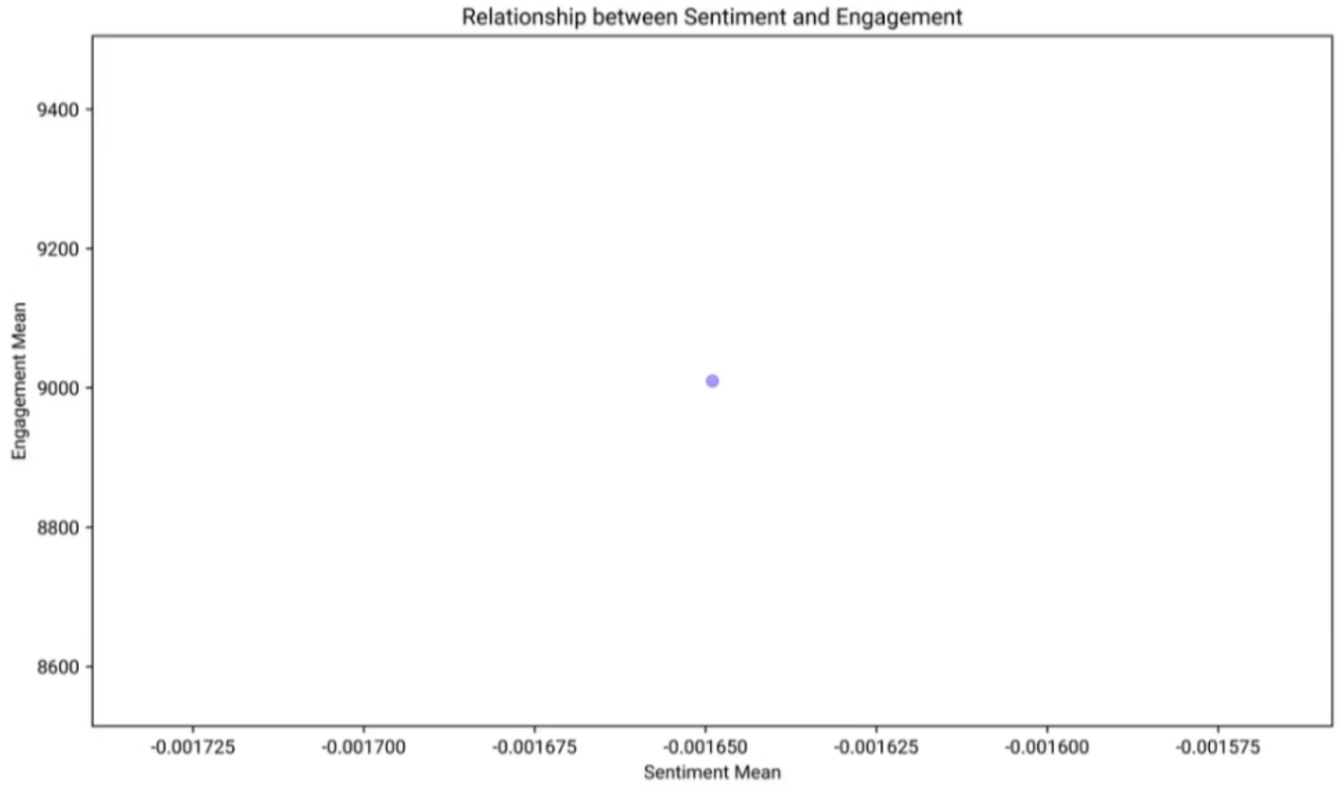
Visual Distribution Analysis

- **Likes Distribution**: The histogram shows a right-skewed distribution, indicating that most users receive fewer likes, with a few users receiving a significantly higher number.
- **Shares Distribution**: Similar to likes, the distribution of shares is right-skewed, with most users having fewer shares.
- **Comments Distribution**: The distribution of comments also follows a right-skewed pattern, with most users having fewer comments.

Conclusion and Insights

- **Engagement Skewness**: All engagement metrics (likes, shares, comments) exhibit right-skewed distributions, suggesting that a small number of users have very high engagement.
- **User Segmentation**: The data indicates potential for segmenting users into high and low engagement groups, which could be useful for targeted marketing or content strategies.

Q5.How does the distribution of Sentiment_Score relate to the engagement levels of posts?



Sentiment Score Distribution

- **Mean Sentiment Score**: The average sentiment score is approximately -0.0016, indicating a neutral sentiment across posts.
- **Sentiment Score Range**: The scores range from -1.0 to 1.0, showing a full spectrum from negative to positive sentiment.
- **Standard Deviation**: The standard deviation is 0.58, suggesting moderate variability in sentiment scores.

Engagement Levels

- **Mean Engagement**: The average engagement level is 9010.02, indicating high interaction with posts.
- **Engagement Range**: Engagement levels vary from a minimum of 178 to a maximum of 17851, showing significant differences in post popularity.
- **Standard Deviation**: The engagement standard deviation is 3327.97, reflecting considerable variation in engagement.

Relationship Visualization

- **Scatter Plot Observation**: The scatter plot shows a single point, indicating a lack of clear relationship between sentiment and engagement due to the narrow sentiment range.

Conclusion and Insights

- **Neutral Sentiment Dominance**: The sentiment scores are predominantly neutral, which may not significantly influence engagement levels.
- **High Engagement Variability**: Despite the neutral sentiment, engagement levels vary widely, suggesting other factors may drive engagement.
- **Further Analysis Needed**: Additional variables or a broader sentiment range might be necessary to understand the sentiment-engagement relationship fully.

Q6.Can we forecast future Likes based on historical data trends?

Historical Data Analysis

- **Data Aggregation**: The dataset was aggregated by month, summing up the 'Likes' to observe trends over time.
- **Time Period**: The data covers three months: January, February, and March of 2024.
- **Likes Trend**: The total likes decreased significantly from January (22,541,469 likes) to March (6,903,650 likes).

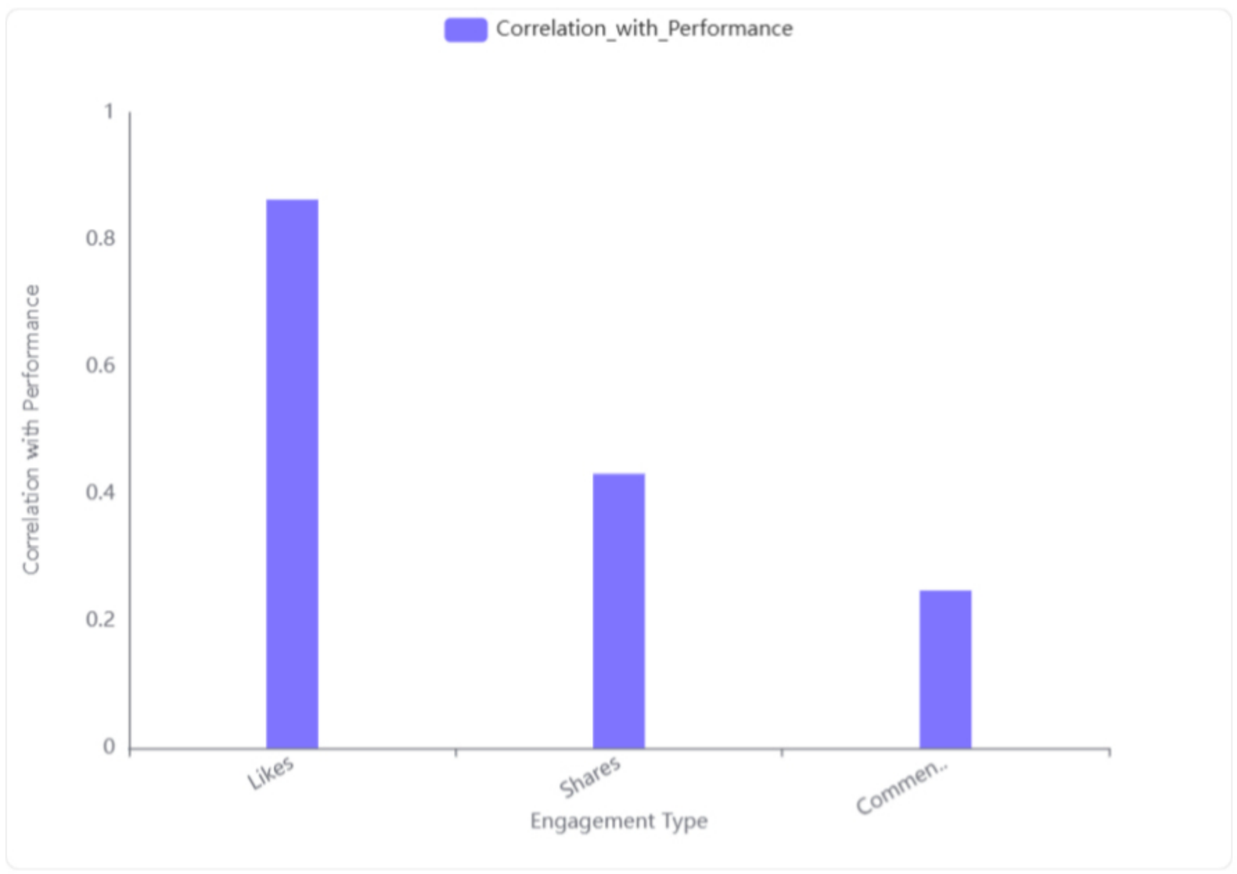
Statistical Insights

- **Mean and Variability**: The average number of likes per month is 16,723,568 with a standard deviation of 8,552,628.3, indicating high variability in monthly likes.
- **Trend Observation**: There is a noticeable downward trend in the number of likes over the three-month period.

Conclusion and Insights

- **Forecasting Feasibility**: Given the limited data spanning only three months, forecasting future likes with high accuracy is challenging. The observed downward trend suggests a potential decrease in future likes, but more data points are needed for a reliable forecast.
- **Recommendation**: To improve forecasting accuracy, it is recommended to gather more historical data over a longer period and consider external factors that might influence the number of likes, such as changes in content strategy or platform algorithms.

Q7.What is the comparative impact of different engagement types (Likes, Shares, Comments) on overall post performance?



Correlation Analysis

- **Likes**: The correlation with the performance metric is **0.86**, indicating a strong positive impact on overall post performance.
- **Shares**: The correlation is **0.43**, showing a moderate positive impact.
- **Comments**: The correlation is **0.25**, suggesting a weak positive impact.

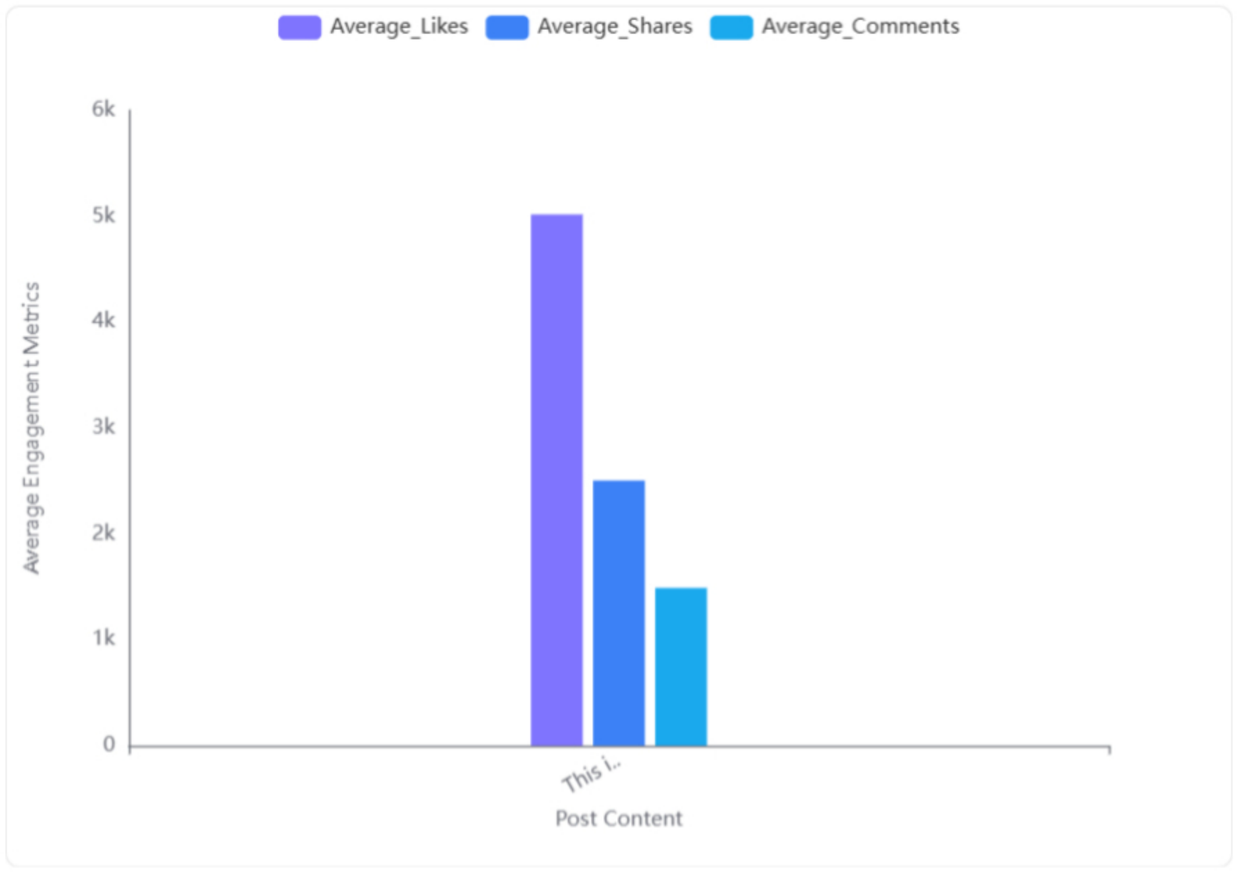
Visualization Insights

- **Likes**: Visually, Likes have the highest correlation with performance, reinforcing their strong impact.
- **Shares**: Shares have a moderate correlation, indicating a noticeable but lesser impact compared to Likes.
- **Comments**: Comments show the least correlation, highlighting their minimal impact on performance.

Conclusion and Insights

- **Likes are the most influential** engagement type on post performance, significantly more than Shares and Comments.
- **Shares have a moderate effect**, contributing positively but less than Likes.
- **Comments have the least impact**, suggesting they are less critical for enhancing overall performance.

Q8.How does the type of Post_Content affect the engagement metrics (Likes, Shares, Comments)?



Average Engagement Metrics

- **Average Likes**: The average number of likes for the post content is **5017.07**.
- **Average Shares**: The average number of shares is **2503.60**.
- **Average Comments**: The average number of comments is **1489.35**.

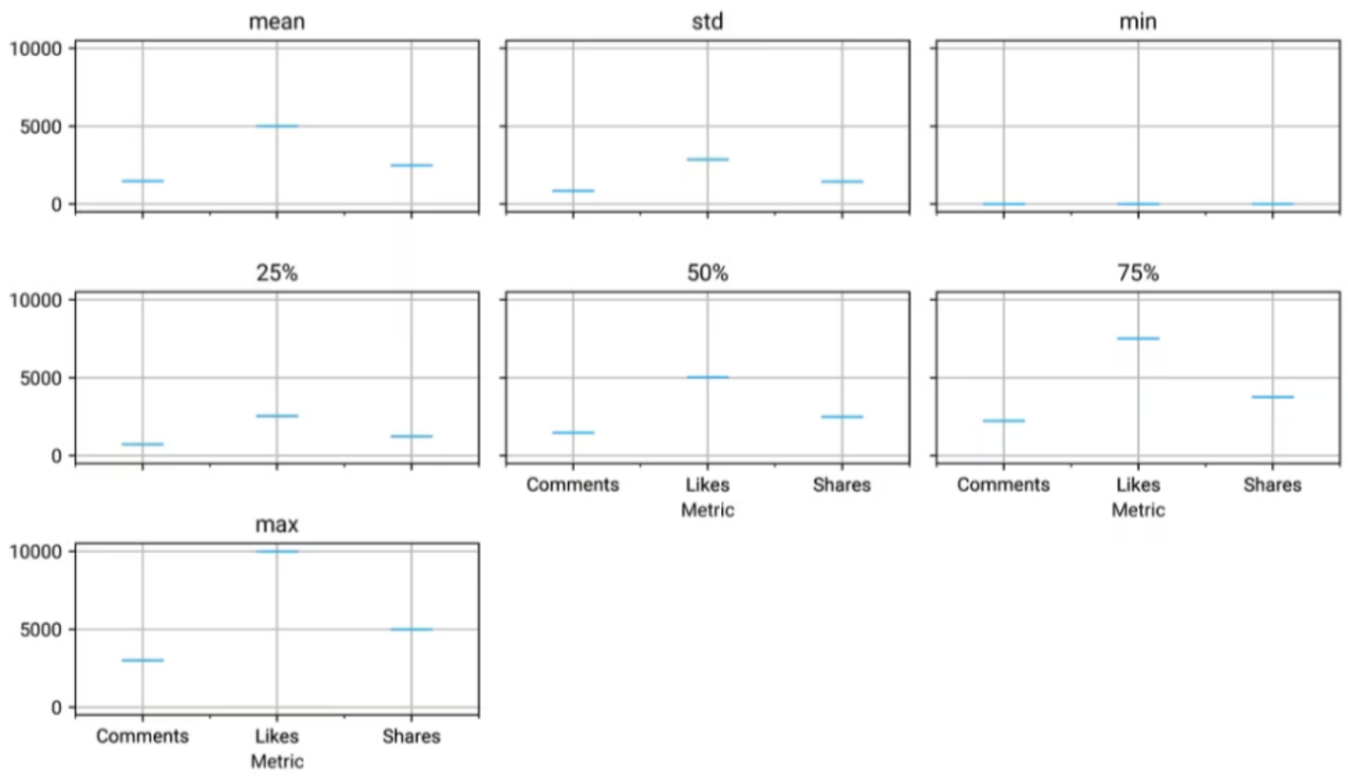
Visualization of Engagement Metrics

- **Bar Chart Insights**: The bar chart visually represents the engagement metrics, showing that likes are the highest, followed by shares and comments.

Conclusion and Insights

- **Likes Lead Engagement**: The type of post content significantly impacts likes, which are the highest among the engagement metrics.
- **Shares and Comments**: Shares and comments are also influenced by the post content, with shares being higher than comments. This suggests that the content is more likely to be shared than commented on.

Q9.What is the distribution pattern of engagement metrics (Likes, Shares, Comments) across all posts?



Statistical Analysis

- **Likes**: The mean is 5017.07 with a standard deviation of 2879.28. The distribution ranges from 1 to 9999, indicating a wide spread.
- **Shares**: The mean is 2503.60 with a standard deviation of 1452.88. The values range from 1 to 4999, showing moderate variability.
- **Comments**: The mean is 1489.35 with a standard deviation of 867.01. The range is from 0 to 2999, suggesting a narrower distribution compared to Likes and Shares.

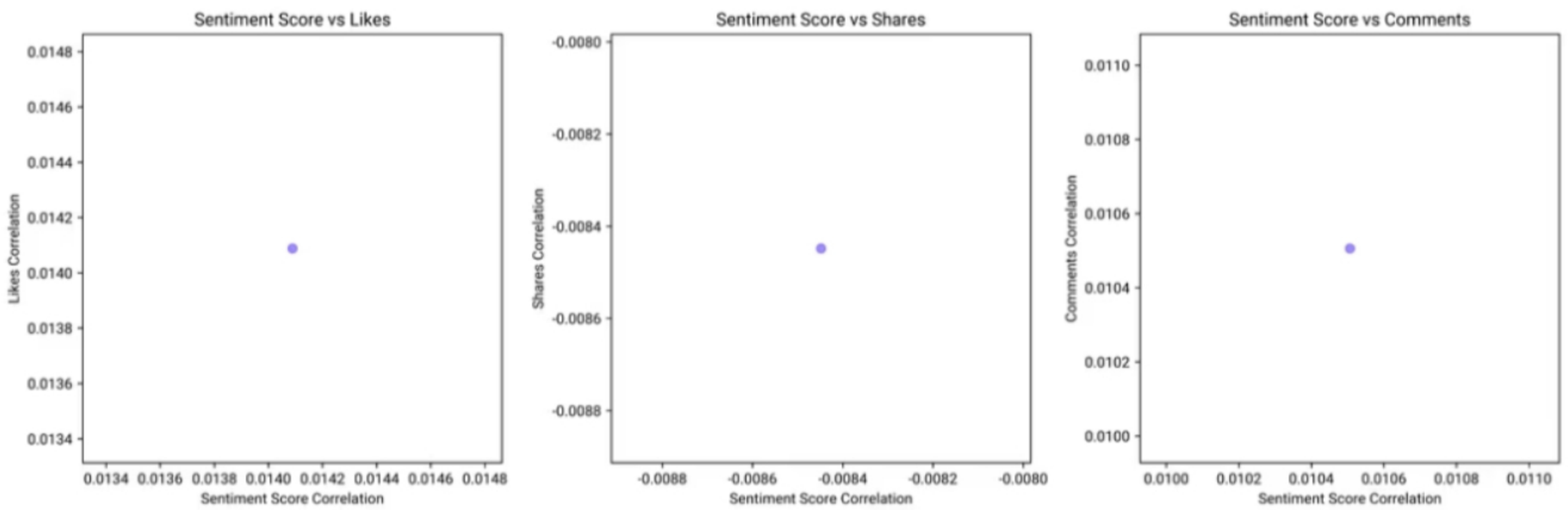
Visualization Insights

- **Mean and Median**: Likes have the highest mean and median, followed by Shares and Comments.
- **Spread and Outliers**: Likes show the widest spread, indicating potential outliers. Comments have the smallest spread.
- **Quartiles**: The 25th, 50th, and 75th percentiles for Likes are higher than those for Shares and Comments, reflecting a higher engagement level.

Conclusion and Insights

- **Engagement Variability**: Likes exhibit the most variability and potential outliers, suggesting some posts receive significantly more likes.
- **Relative Engagement**: Shares and Comments have lower means and narrower distributions, indicating more consistent engagement levels across posts.

Q10.How does Sentiment_Score impact the level of engagement (Likes, Shares, Comments) across posts?



Correlation Analysis

- **Sentiment_Score and Likes**: The correlation coefficient is **0.014**, indicating a very weak positive relationship.
- **Sentiment_Score and Shares**: The correlation coefficient is **-0.008**, suggesting a very weak negative relationship.
- **Sentiment_Score and Comments**: The correlation coefficient is **0.011**, showing a very weak positive relationship.

Visualization Insights

- **Scatter Plots**: The visualizations show minimal clustering or trend, reinforcing the weak correlations observed between Sentiment_Score and engagement metrics (Likes, Shares, Comments).

Conclusion and Insights

- **Minimal Impact**: Sentiment_Score has a negligible impact on engagement metrics, with all correlation values close to zero.
- **Further Analysis Needed**: Other factors might be influencing engagement more significantly, suggesting a need for additional analysis beyond sentiment.