

# **STARBUCKS DATASET ANALYSIS**

**SHALU JAIN  
MABSPG24024  
ALPHA**

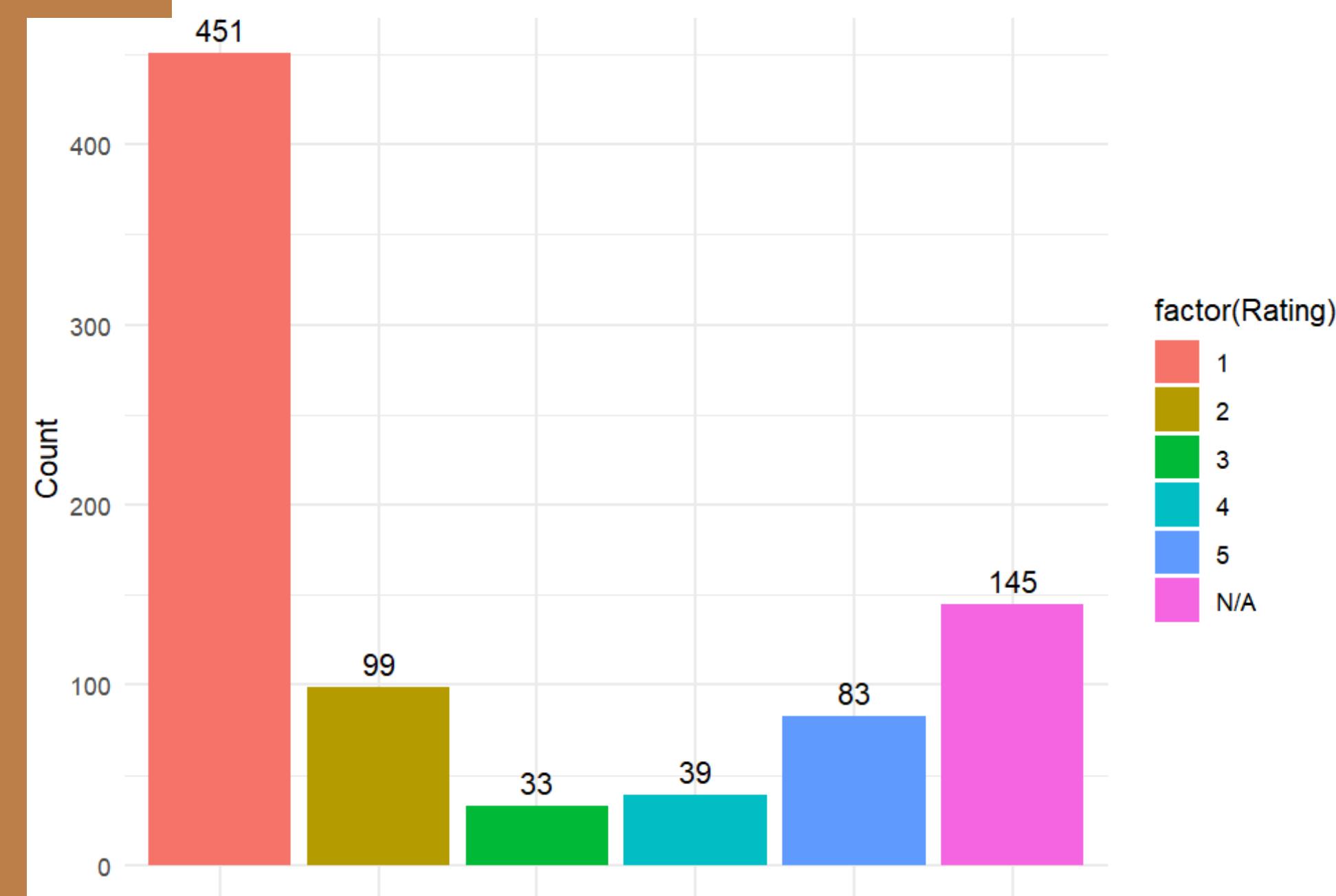


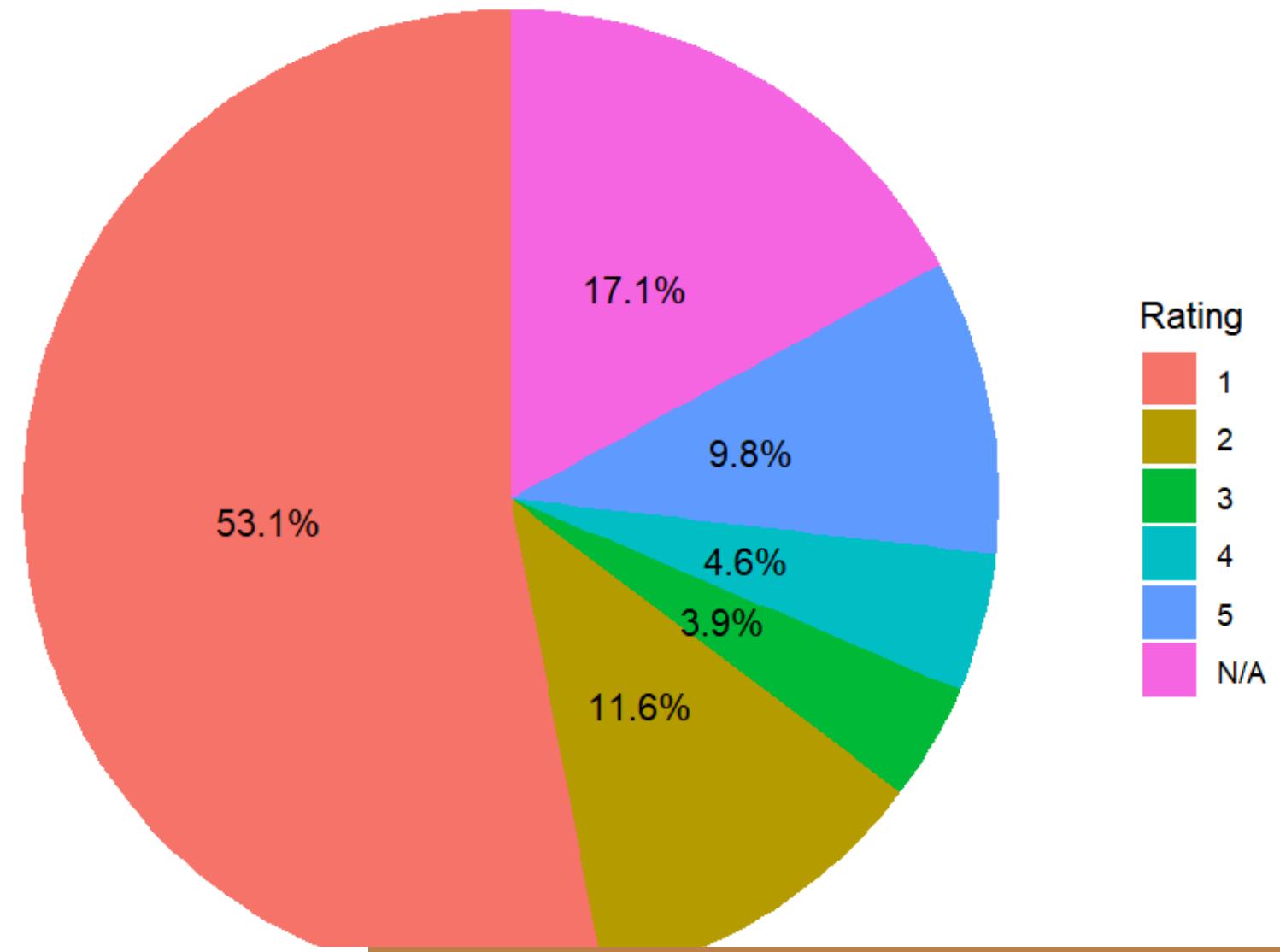
## MAIN OBJECTIVE

- Understanding the Data: Interpretation aims to understand what the data represents, how it was collected, and any potential limitations or biases.
- Identifying Trends and Patterns: By analyzing the data, we can identify trends, patterns, and relationships between different variables.
- Making Predictions: In some cases, interpretation can be used to make predictions or forecasts about future events based on historical data.
- Supporting Decision Making: The insights gained from interpretation can inform decision-making processes in various fields, such as business, research, and policy development.

# INTERPRETATION 1

The bar chart illustrates the distribution of review ratings. The highest number of reviews falls into the "1" category with 451 reviews. This is followed by "2" with 99 reviews, and then "5" with 145 reviews. The remaining categories ("3," "4," and "N/A") have significantly lower counts.



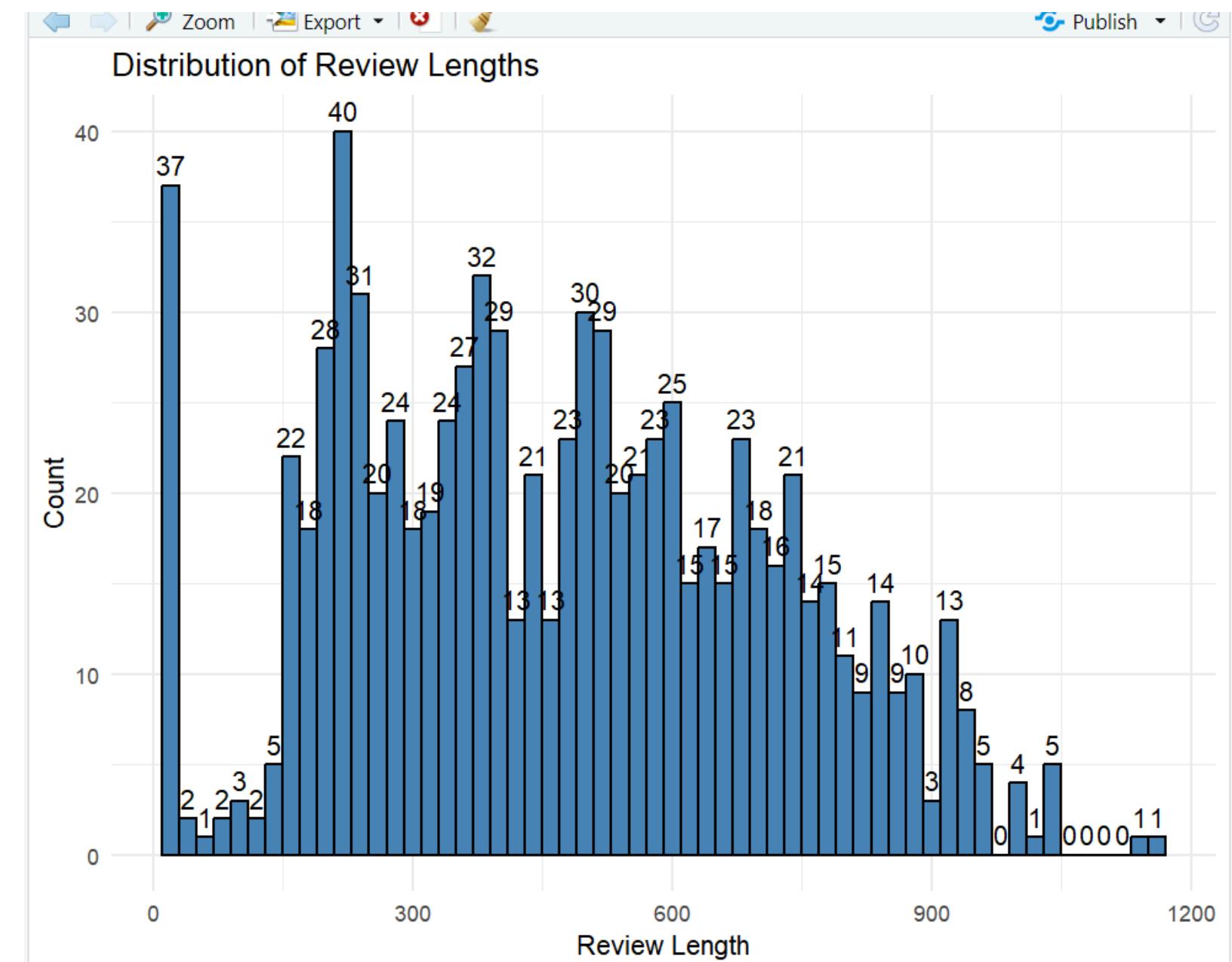


## INTERPRETATION 2

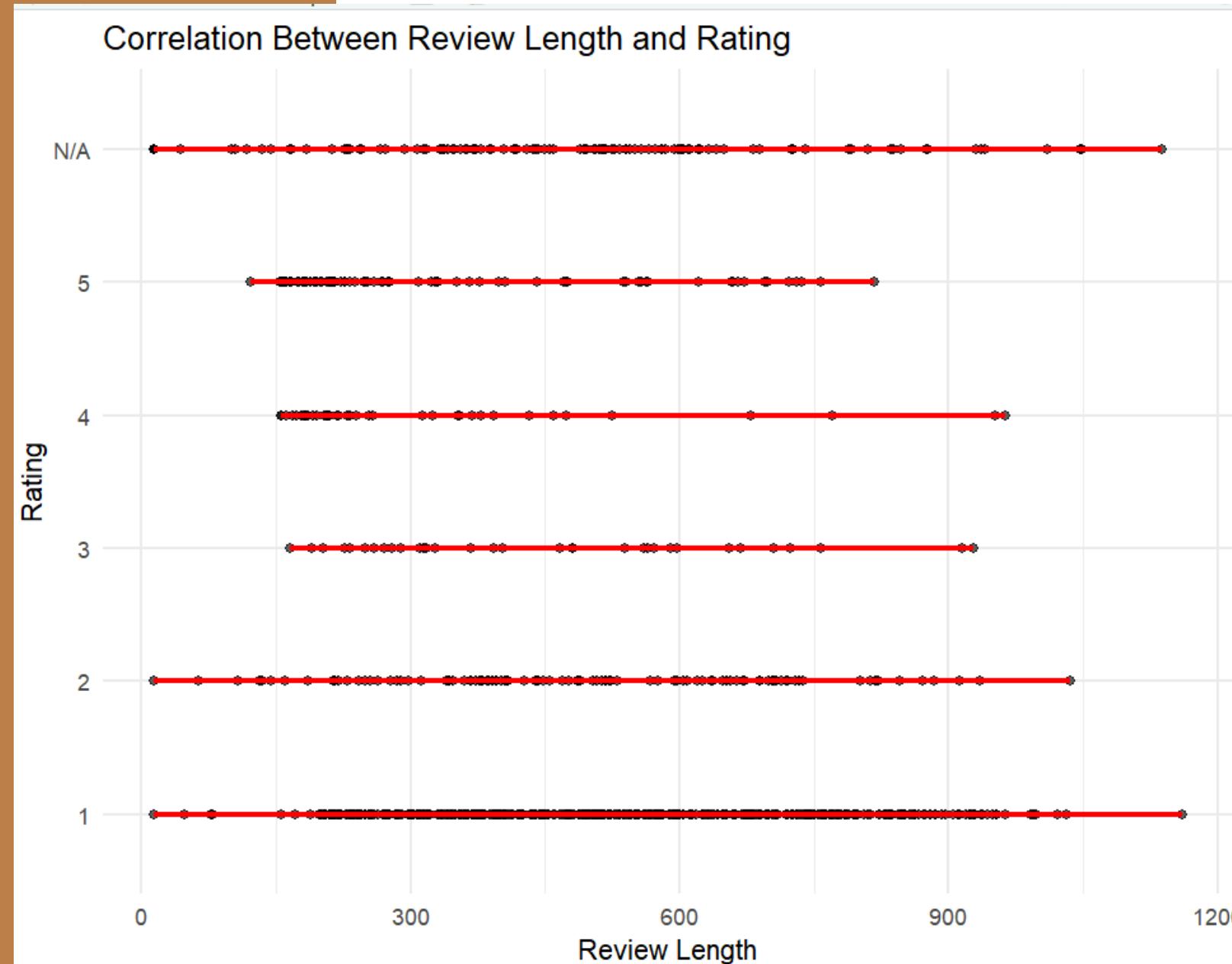
The pie chart illustrates the distribution of review ratings. "Rating 1" dominates with 53.1%, indicating a high proportion of negative or dissatisfied reviews. Ratings 4 and 5 have significantly smaller shares, highlighting the lower frequency of positive or neutral reviews. A notable 9.8% of reviews fall into the "N/A" category.

# INTERPRETATION 3

The histogram illustrates the distribution of review lengths. Most reviews fall within the range of 0 to 300 characters with a peak around 150 characters. A smaller number of reviews are longer, extending up to around 1200 characters. This suggests that a majority of reviews are relatively concise.

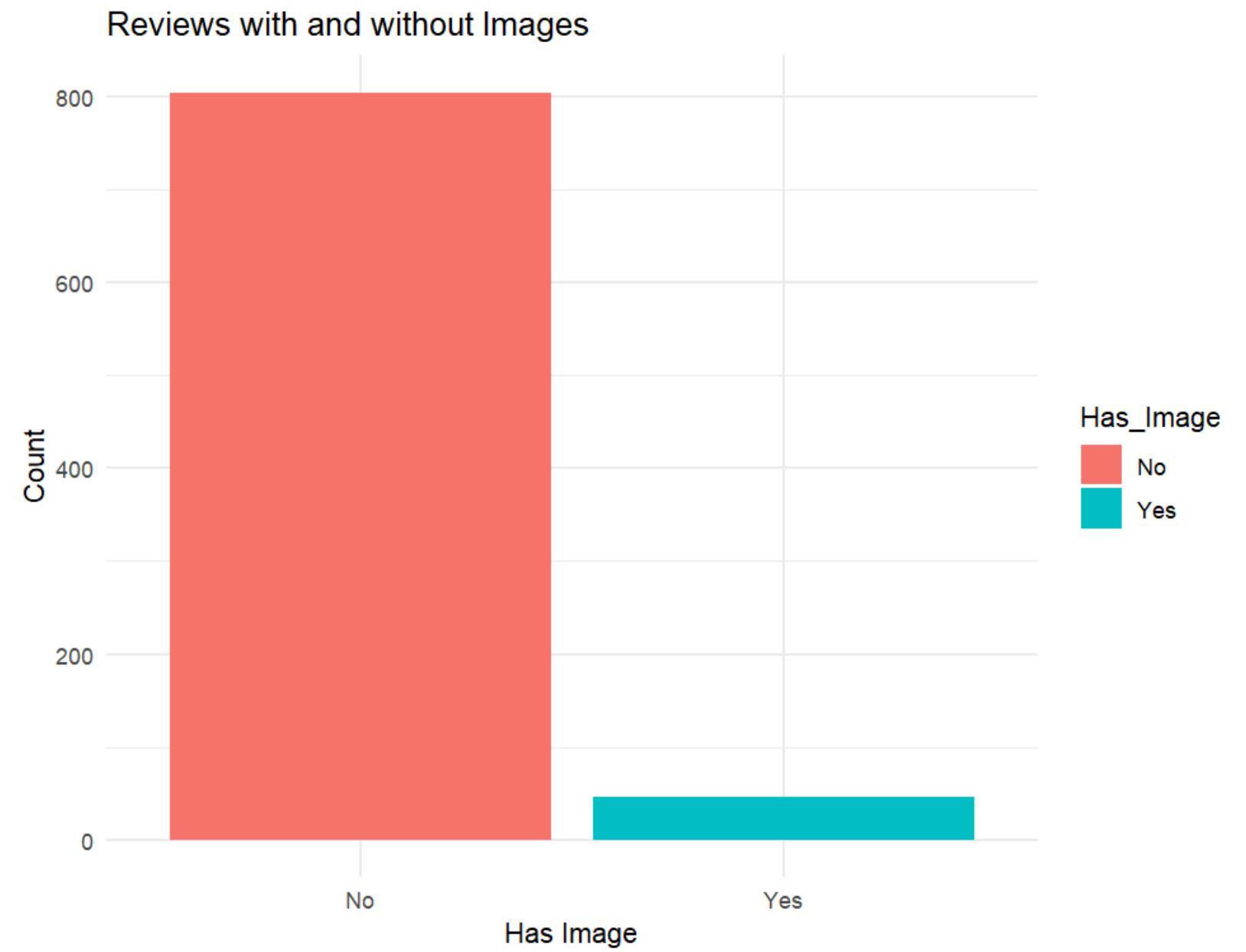


# INTERPRETATION 4

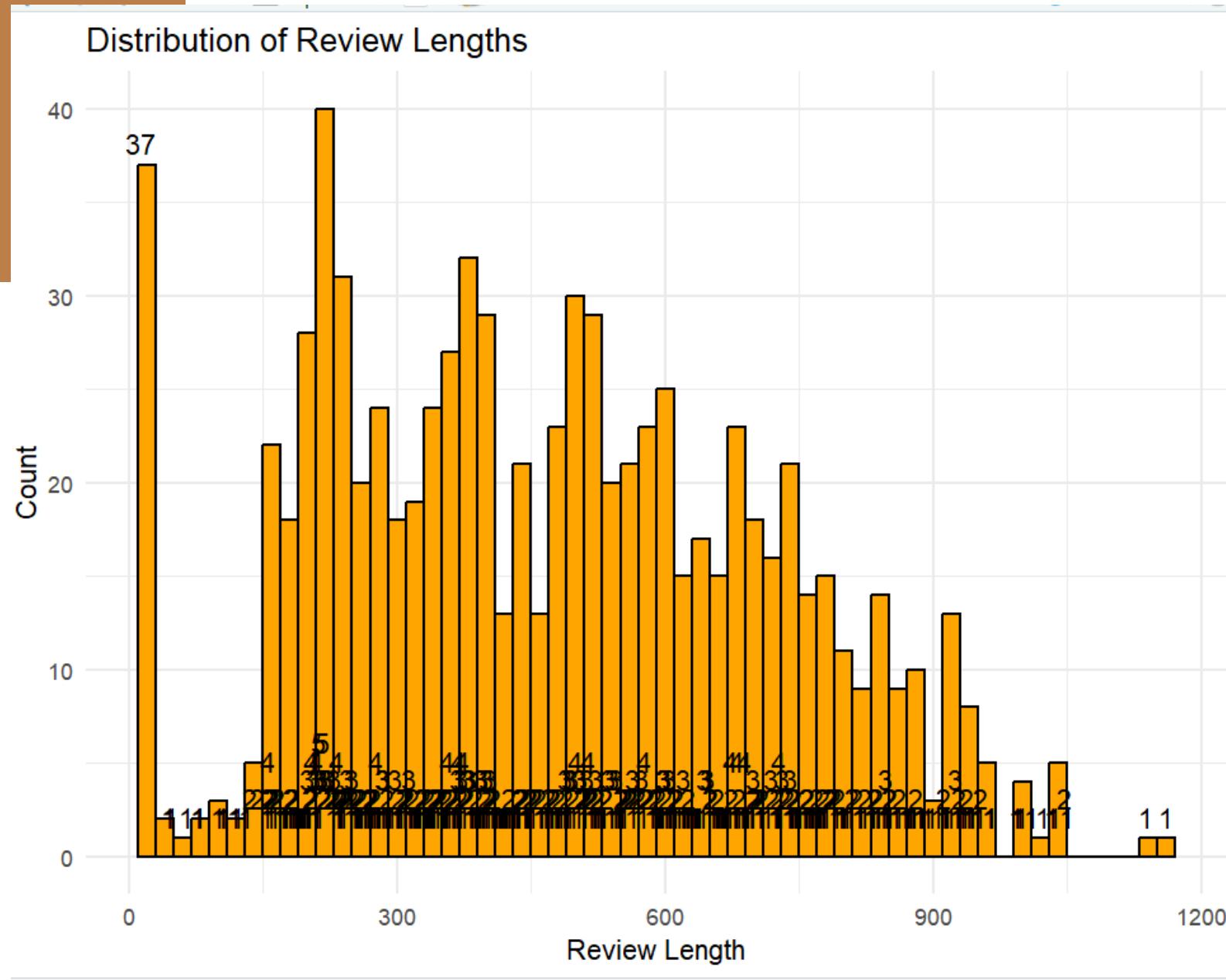


# INTERPRETATION 5

The bar chart illustrates the distribution of reviews with and without images. A significantly larger number of reviews (around 800) do not have images associated with them, while a much smaller number (around 50) do. This suggests that the majority of reviews are submitted without accompanying images.



# INTERPRETATION 6



The histogram illustrates the distribution of review lengths. Most reviews fall within the range of 0 to 300 characters with a peak around 150 characters. A smaller number of reviews are longer, extending up to around 1200 characters. This suggests that a majority of reviews are relatively concise.

The bar chart illustrates the top 10 reviewers based on the number of reviews they have submitted. Linda is the most prolific reviewer with 13 reviews, followed by John with 10 reviews. The remaining reviewers have submitted between 5 and 9 reviews. This suggests that a small group of reviewers contributes a significant portion of the total reviews.

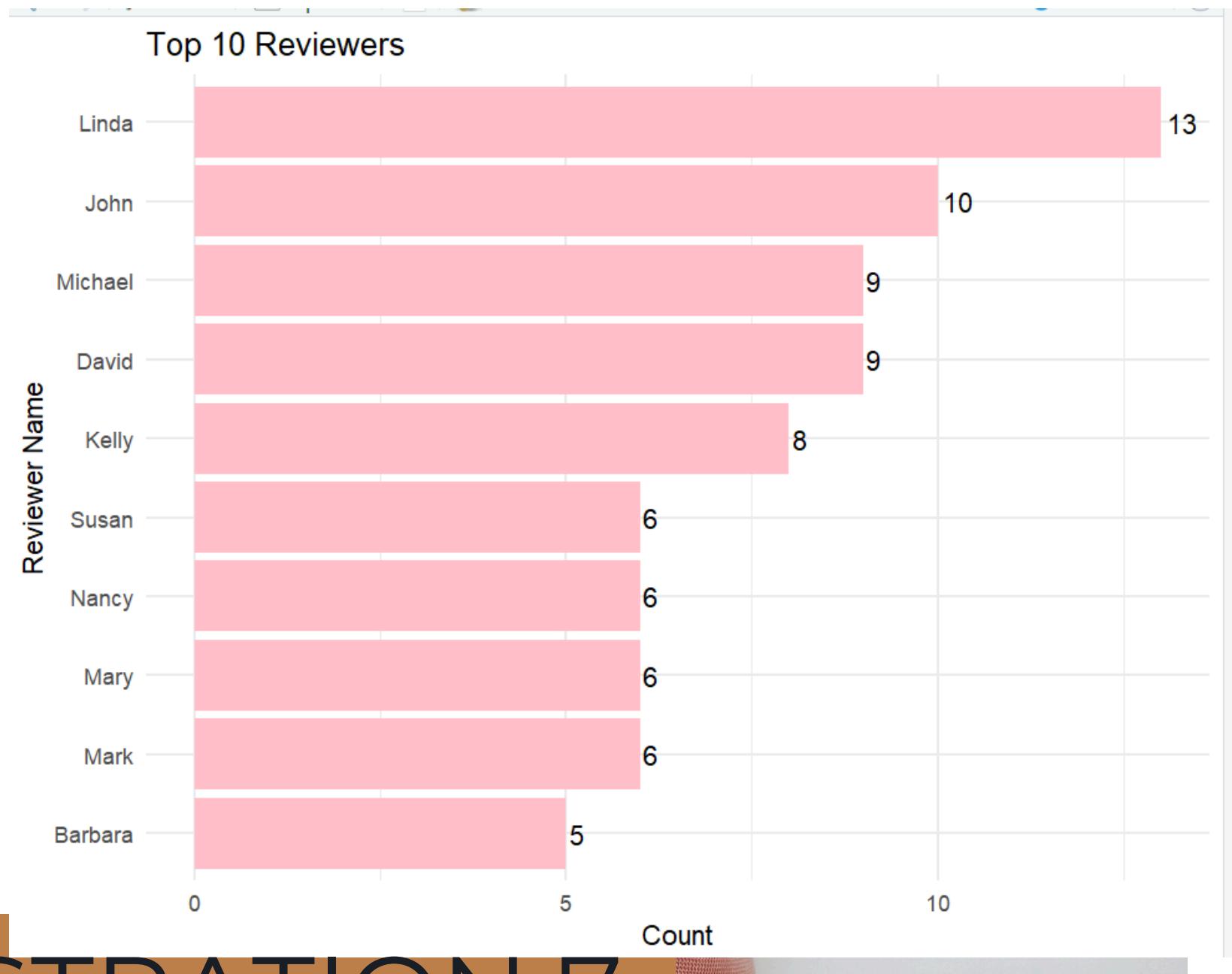
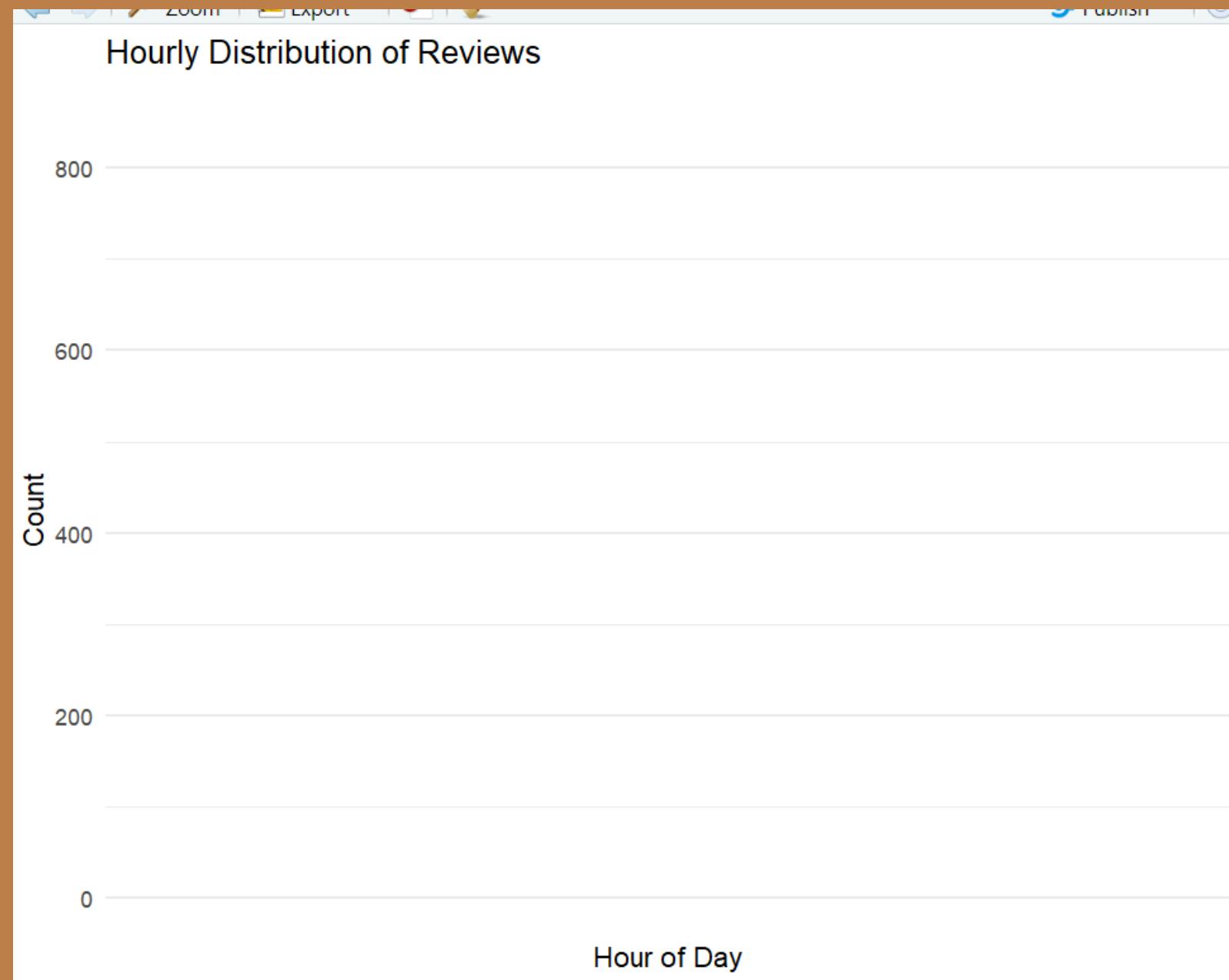
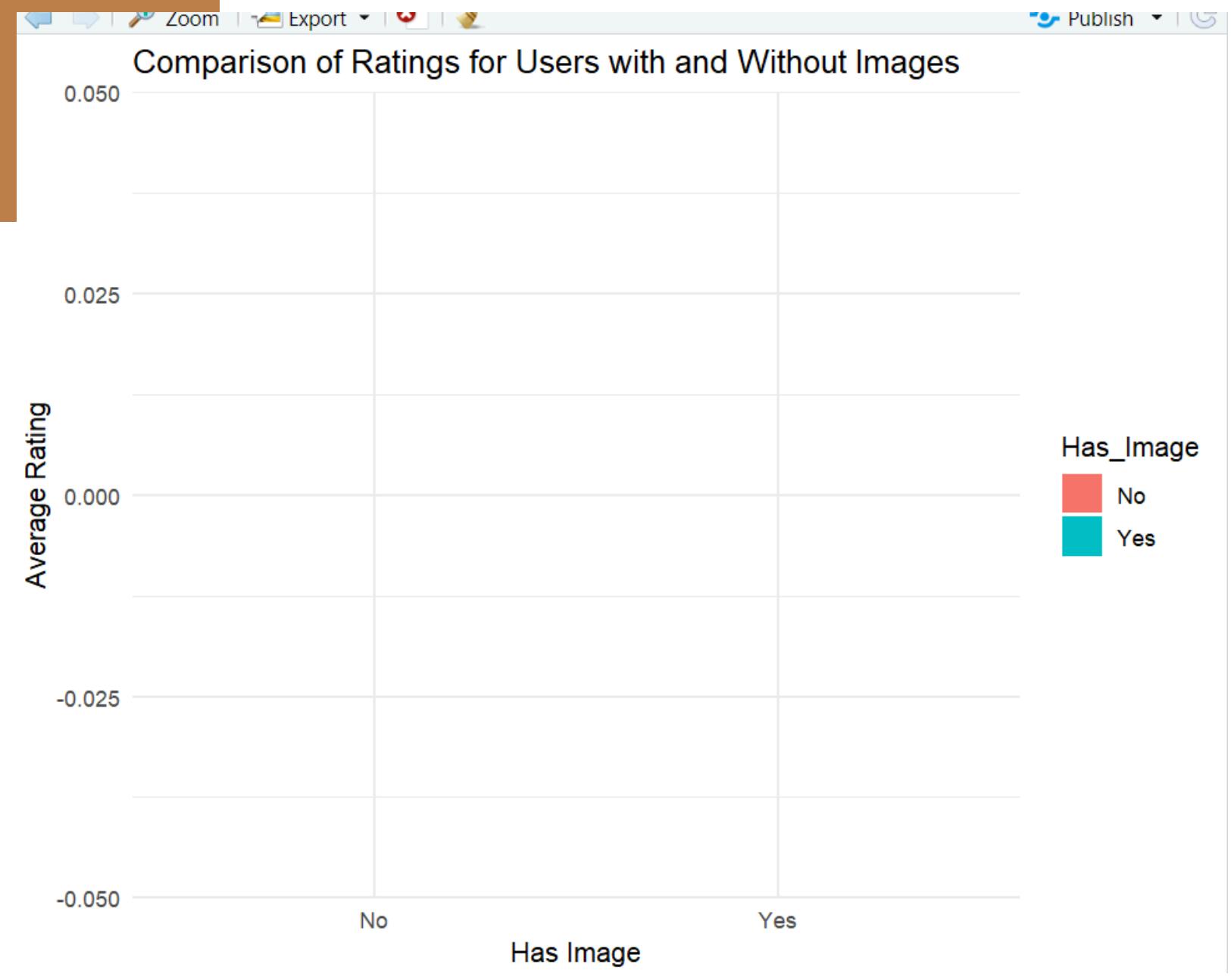


ILLUSTRATION 7

# INTERPRETATION 8





The chart compares the average ratings for users with and without images. There appears to be a very slight difference in average ratings between the two groups. Users with images have a slightly higher average rating than those without, but the difference is minimal. This suggests that the presence of images may have a very small or negligible impact on the average ratings received by users.

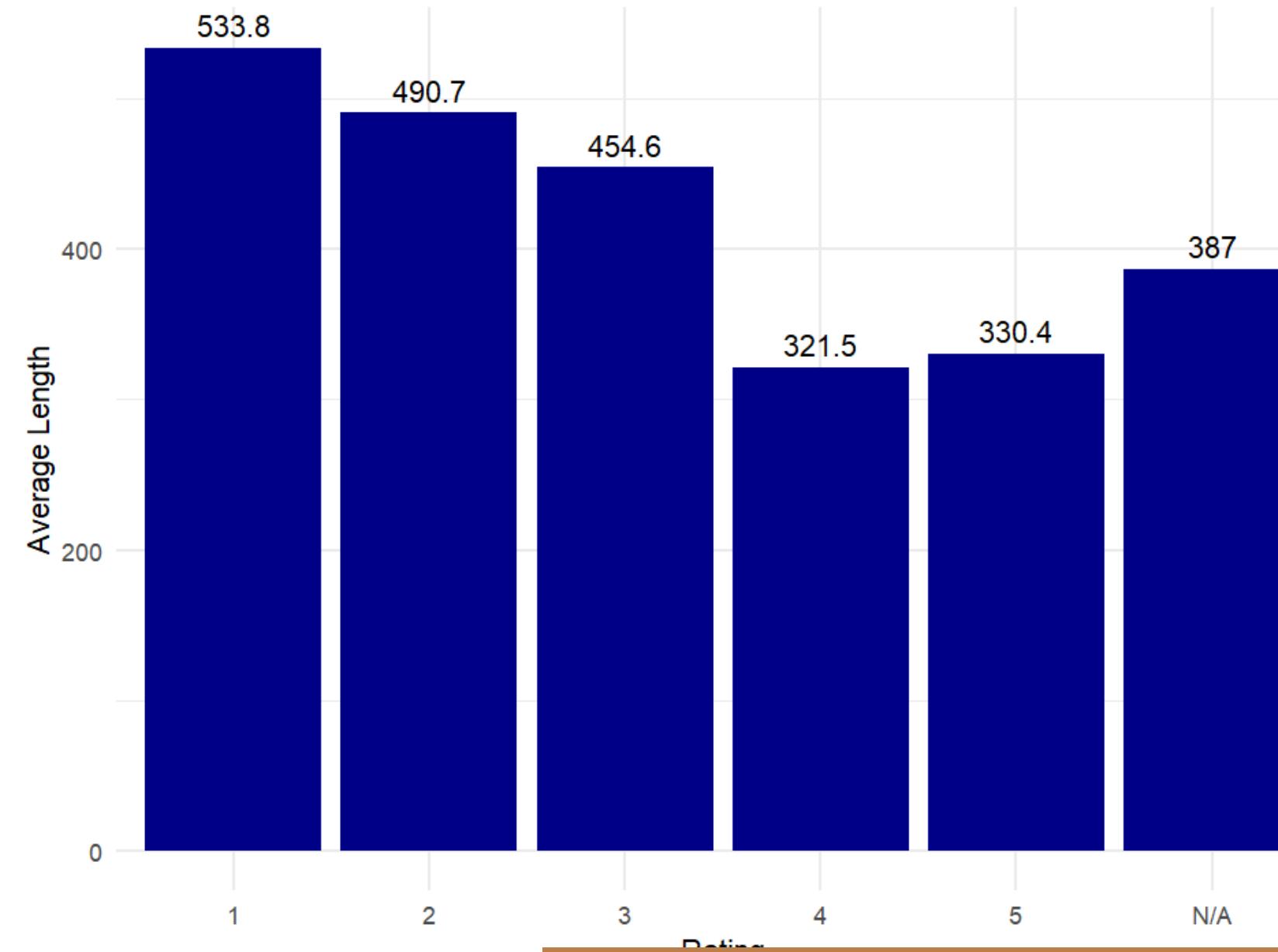
## INTERPRETATION 9

# INTERPRETATION 10

The box plot displays the relationship between review length categories (Long, Medium, Short) and ratings. For each rating category (1-5 and N/A), the boxes show the median rating and the interquartile range, indicating the spread of ratings within each review length category. There appears to be some variation in the distribution of ratings across different review lengths, but the overall differences seem relatively small.



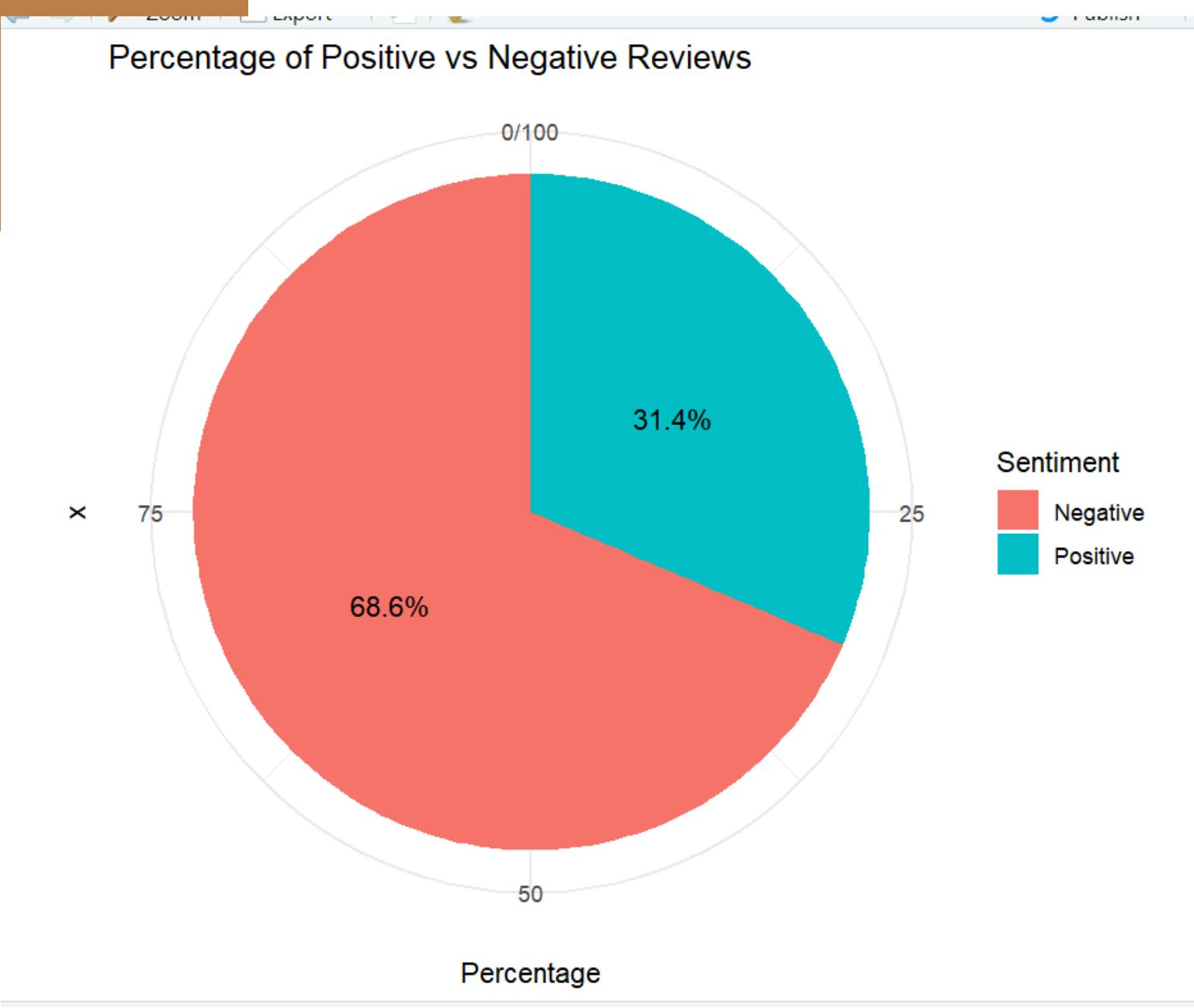
Average Review Length by Rating



## INTERPRETATION

11

The chart shows the average review length for each rating category. Reviews with a rating of "1" tend to be the longest, followed by ratings of "2" and "3." Reviews with ratings of "4," "5," and "N/A" have progressively shorter average lengths. This suggests a possible correlation between negative ratings (1 and 2) and longer review lengths, potentially indicating more detailed explanations for lower scores.



The pie chart shows the percentage of positive versus negative reviews. A significantly larger portion of the pie, representing 68.6%, is attributed to negative reviews. This indicates that a substantial majority of the reviews analyzed fall into the negative sentiment category. In contrast, positive reviews account for 31.4% of the total.

## INTERPRETATION 12

The bar chart shows the top 10

locations with the most reviews.

New York, NY leads with the highest number of reviews (14), followed by Los Angeles, CA with 12 reviews. The remaining locations have varying numbers of reviews ranging from 6 to 9. This indicates that these top 10 locations are popular destinations and have received a significant amount of feedback.

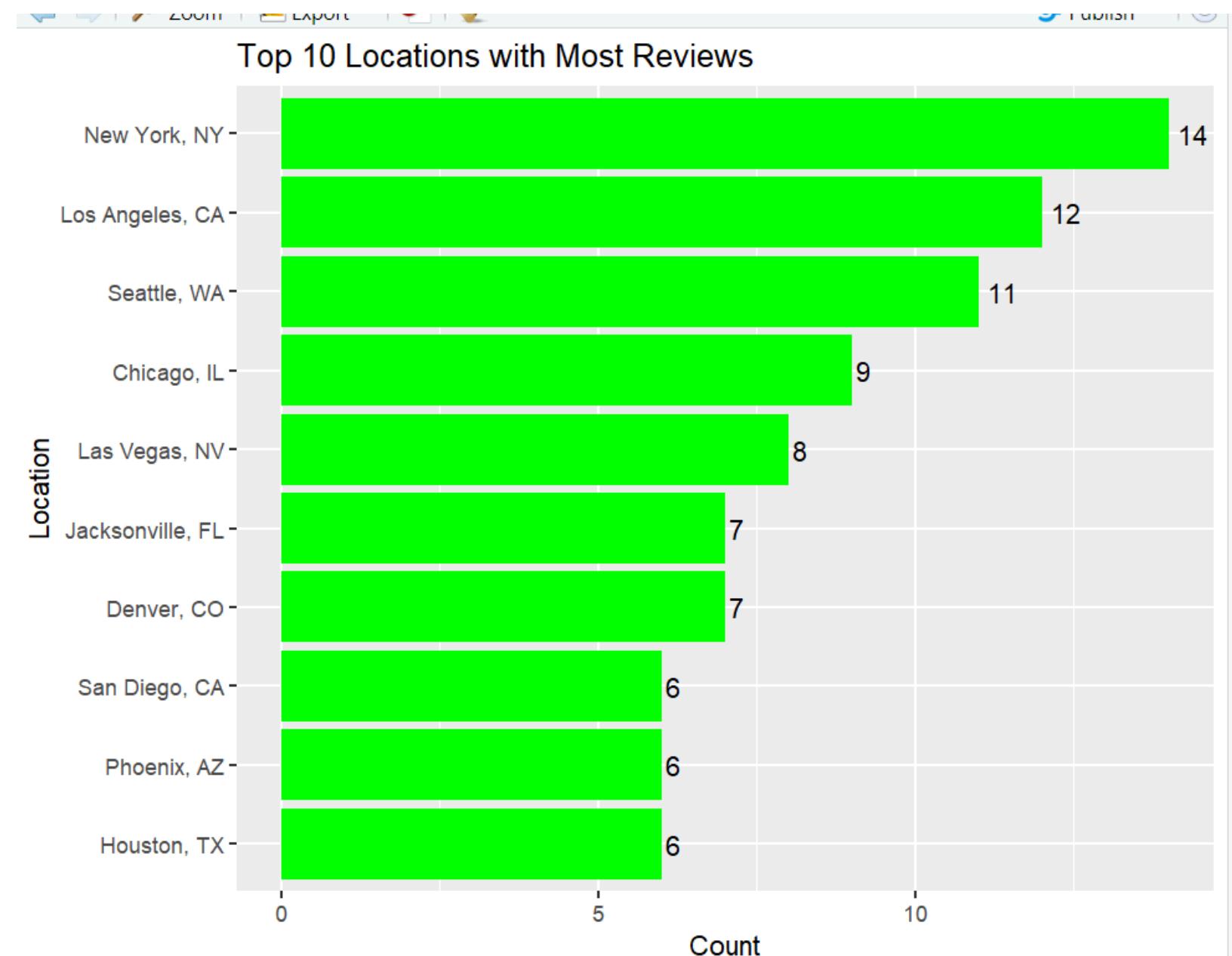
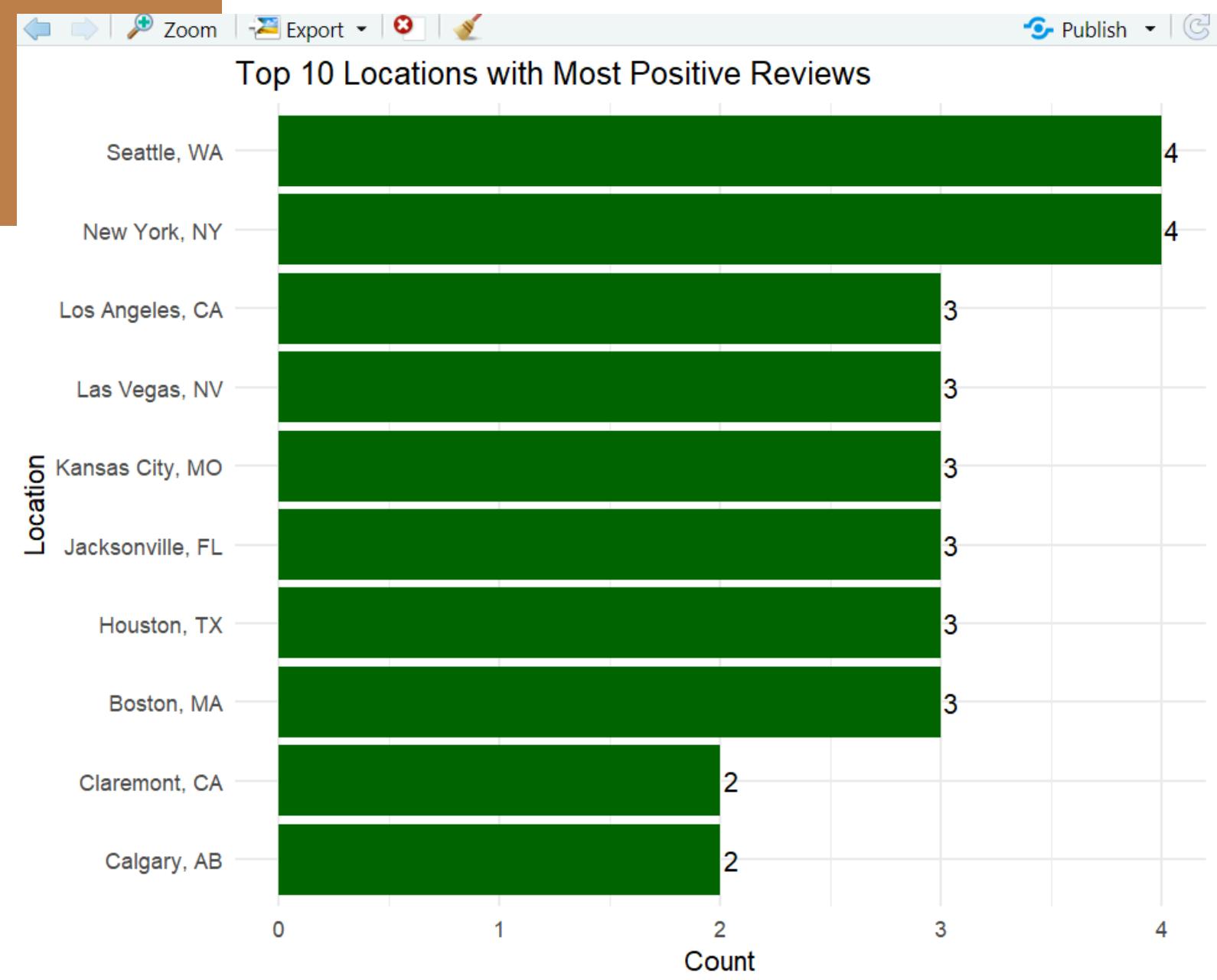


ILLUSTRATION 13



The chart shows the top 10 locations with the most positive reviews. Seattle, WA and New York, NY lead with the highest number of positive reviews (4 each). Several other cities, including Los Angeles, CA, Las Vegas, NV, and Kansas City, MO, also have a significant number of positive reviews (3 each). This indicates that these locations are highly regarded by reviewers and have received a considerable amount of positive feedback.

# INTERPRETATION 14



# THANKS

Coffee Business Presentation Template