

Amazon Reviews Analysis with NLP

Group 2

**Wambui Githinji, Lynette Mwiti, Felix
Njoroge, Wilfred Lekishorumongi, Monica
Mwangi, Joan Maina**



CONTENTS



01
The Problem Statement

02
Objectives

03
Exploratory Data Analysis (EDA)

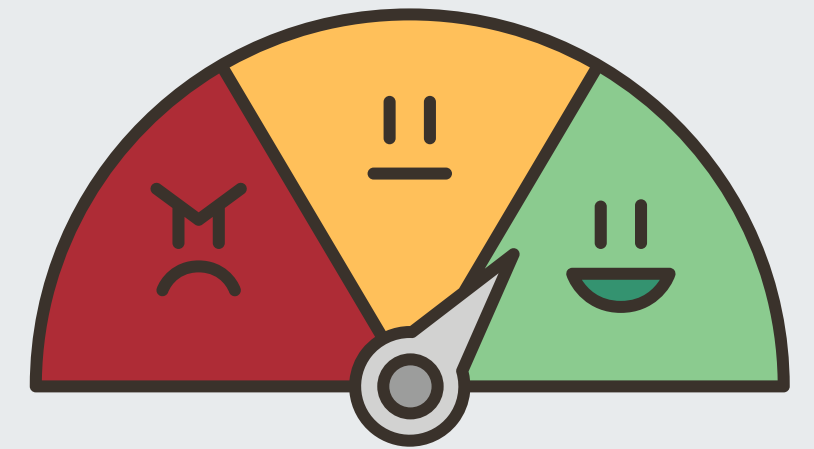
04
Sentiment Analysis

05
Modeling

06
Reccomendations

PROBLEM STATEMENT

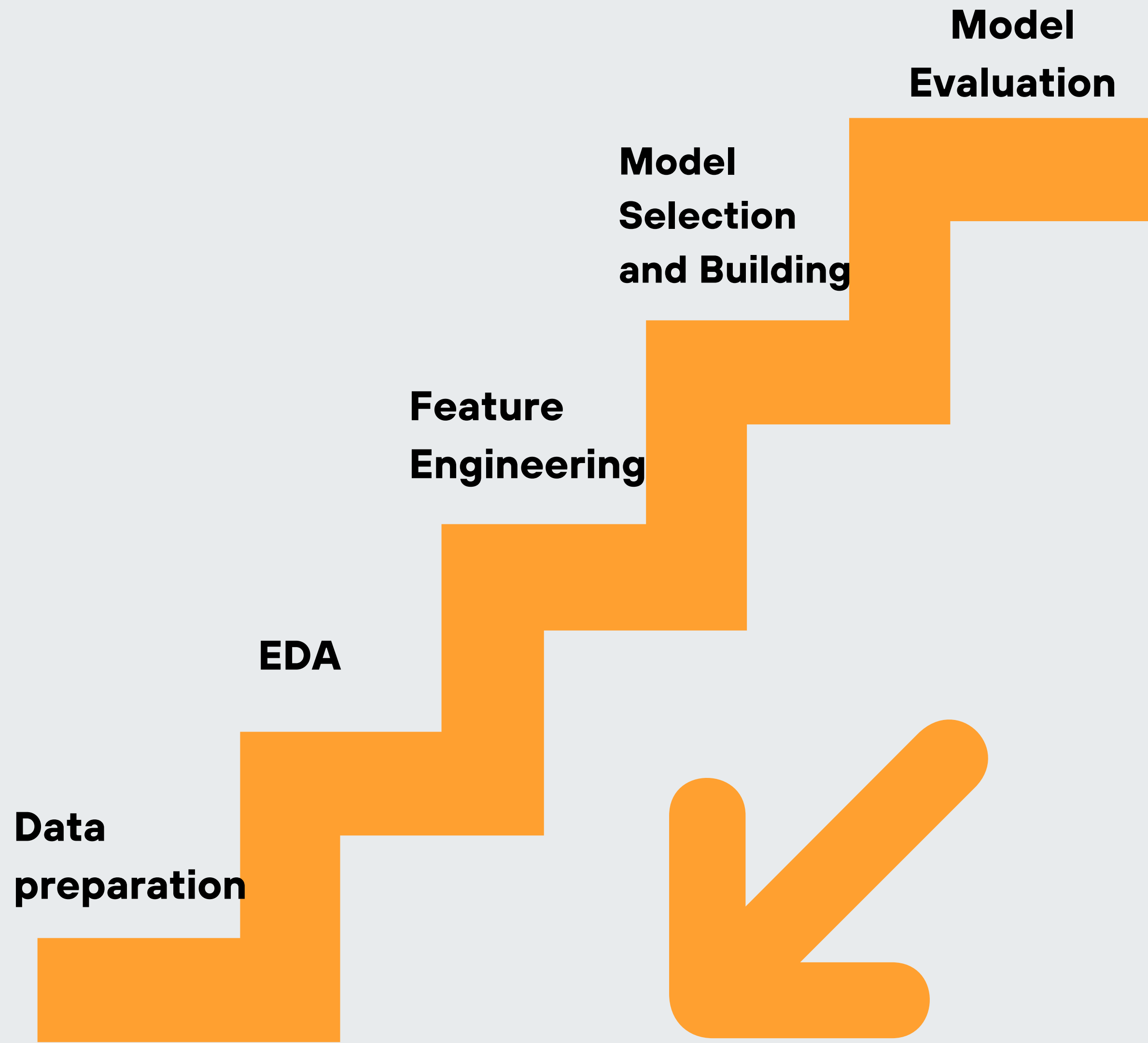
- Reviews are critical to businesses as they offer insights into customer satisfaction, preferences and areas of improvement.
- Businesses need to understand and interpret these reviews in order to cut through the competition. Lots of reviews are generated daily and manually analyzing them is impractical.



OBJECTIVES

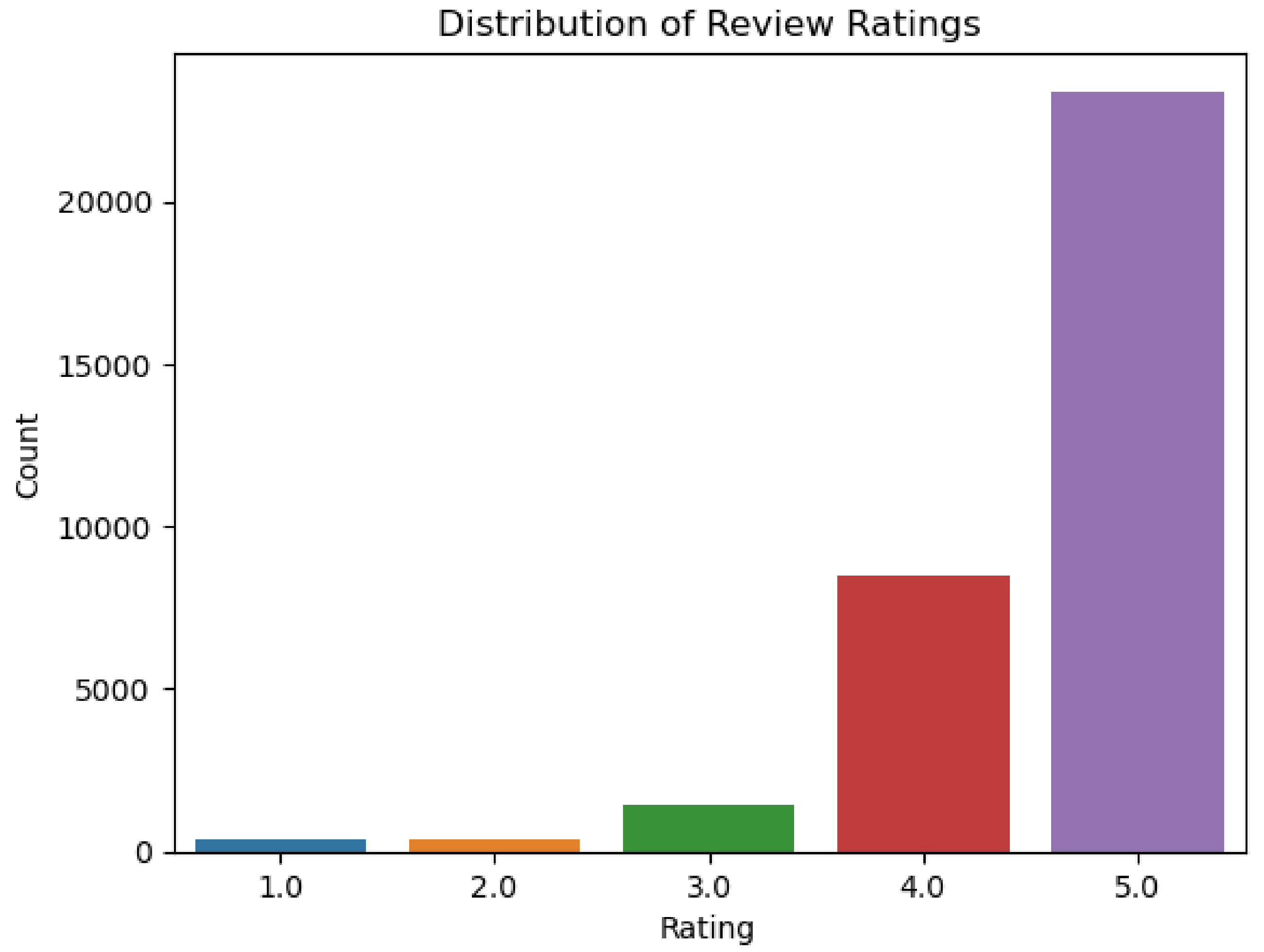
- **Broad Objective:** Use Sentiment analysis to help the businesses get actionable insights from the feedback received from customers.
- **Specific objectives:**
 - Conduct exploratory data analysis to understand the distribution of reviews over time, across brands and products.
 - Determine the sentiment of the reviews (positive or negative) to understand overall customer satisfaction and feedback.
 - Utilize sentiment analysis to help our stakeholders understand customer preferences across various products.
 - Leverage customer reviews to identify areas for improvement in products based on user experience.
 - Build a classifier model to help predict reviews as positive or negative

STEPS

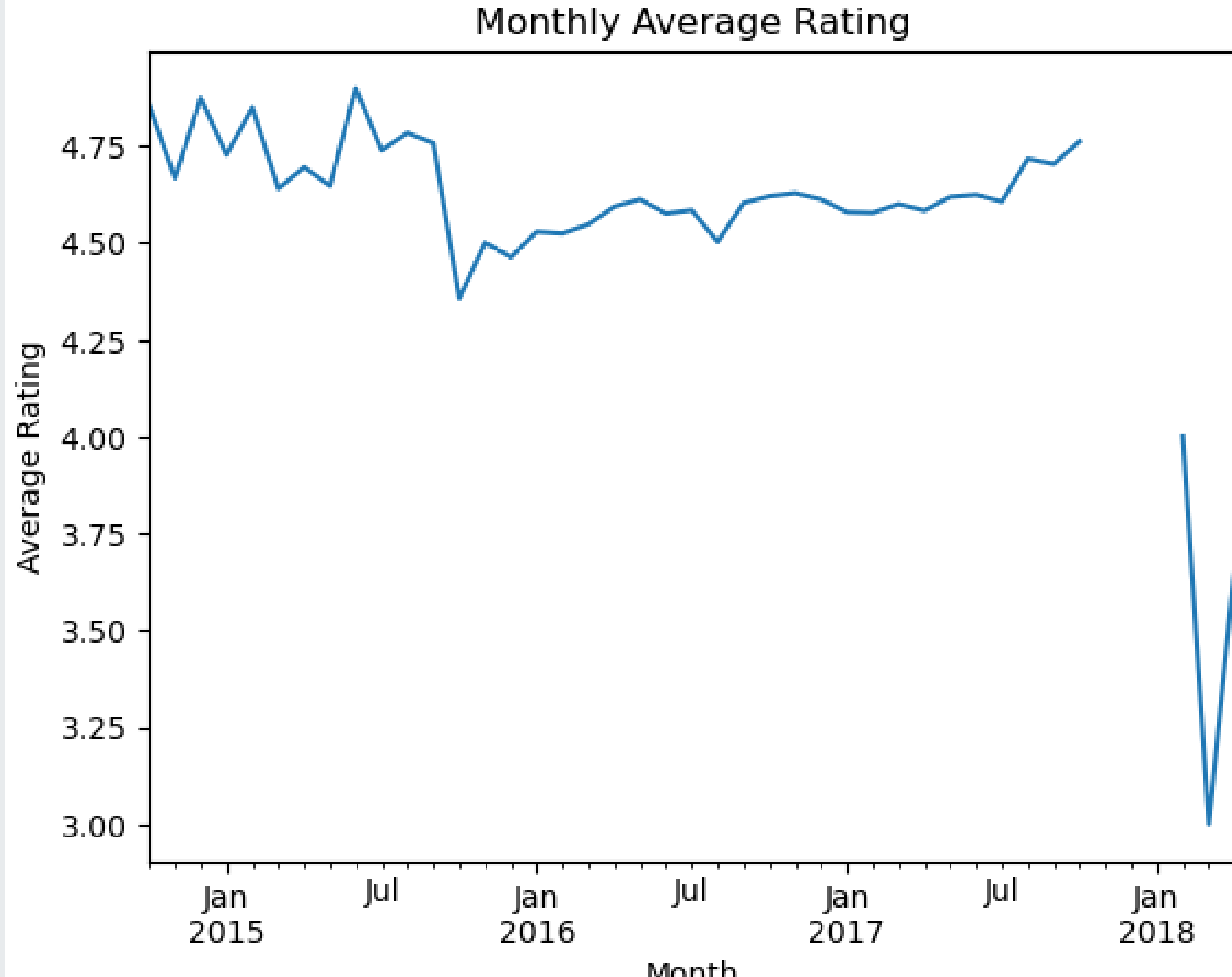


EDA

Reviews with higher ratings tend to have a higher count.



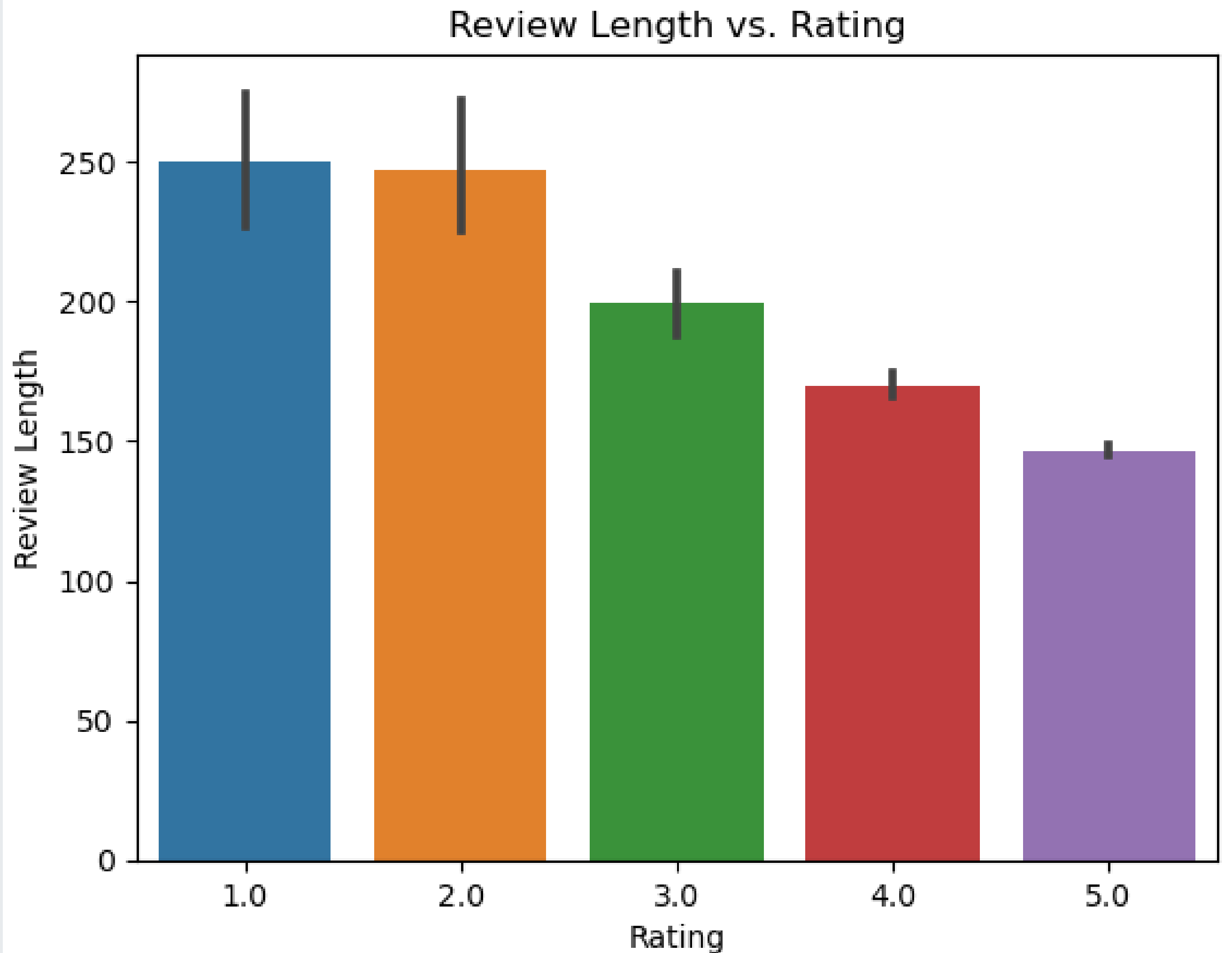
There is a slight fluctuation in average ratings over time, but no clear trend is evident from the monthly average ratings plot.



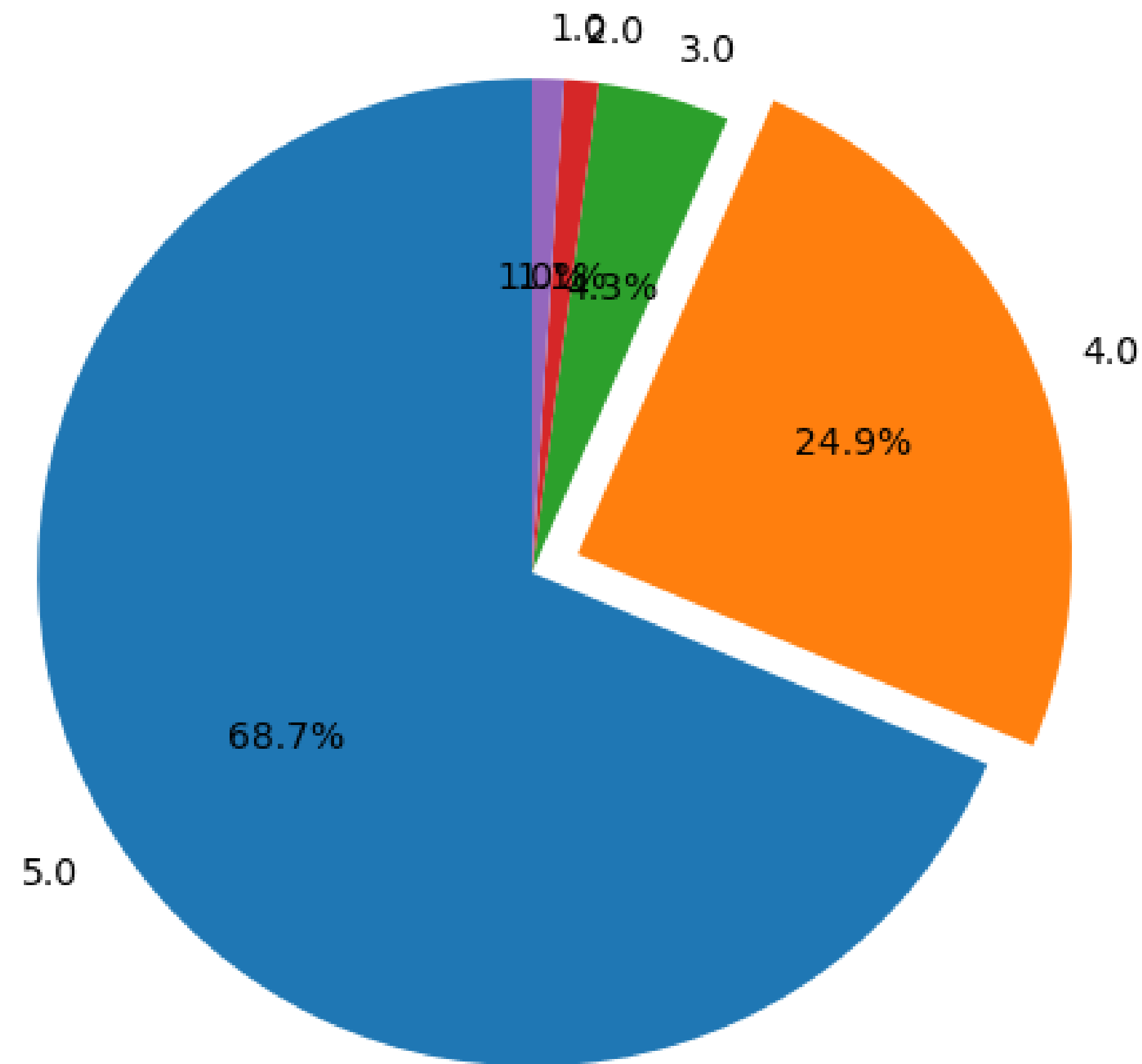
The analysis shows that reviews with a positive recommendation (review_do_recommend = True) generally have higher ratings compared to those without a recommendation.



It is evident that shorter reviews tend to receive higher ratings.



The analysis shows that reviews with a positive recommendation (review_do_recommend = True) generally have higher ratings compared to those without a recommendation.

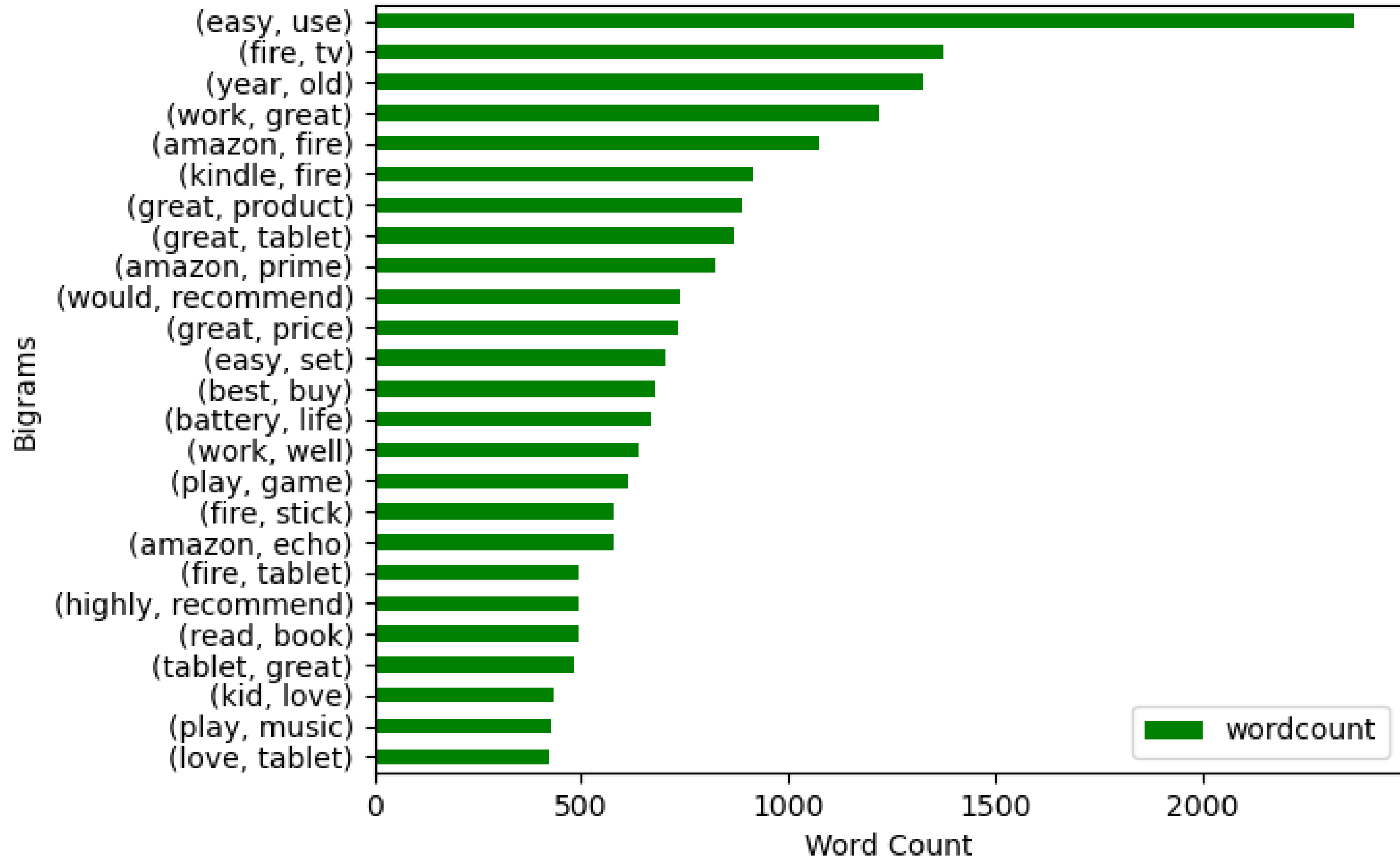


Sentiment Analysis

- Sentiment analysis was conducted to classify reviews as positive, negative, or neutral using NLP techniques



Top 25 Most Frequent Bigrams



Negative Reviews



[illegible]

Positive Reviews

love, easy, use, kindle, bought, great, amazon, good, product, book, device, fire, price, well, echo, speaker, play, music, apps, game, like, read, one, perfect, nice, recommend, find, picture, list, weather, wifi, reader, overall, value, news, definitely, easier, enjoy, feature, looking, everything

MODELLING

Two deep learning models are used;

- a Simple RNN
- an LSTM,

They are built to predict the the reviews as positive or negative.



Model Performance



These are the model performance results;

- SimpleRNN Model: With an AUC of 0.91, the SimpleRNN model performs much better than the LSTM model in distinguishing between the classes. It has a high true positive rate and a low false positive rate across various thresholds.
- LSTM Model: With an AUC of 0.53, the LSTM model is only marginally better than random guessing. This indicates that the LSTM model is ineffective for this particular task or Amazon reviews dataset.

Model Evaluation



- Hyperparameter tuning is performed to optimize the models for better performance.
- The Model tuning for simple RNN Accuracy score is 0.8694 which has no huge difference from the model before tuning which was 0.8658.
- The score is slightly higher than our objective of achieving 0.85 accuracy score. The model is therefore satisfactory.

RECOMMENDATIONS

- **Customer Feedback:**

Maintain product quality, enhance customer service, and solicit feedback to bolster positive reviews.

- **Product Preference:**

Expand preferred product categories, improve features based on feedback, and maintain competitive pricing.

- **Trend Analysis:**

Correlate sales spikes with launches, updates, or campaigns to adapt strategies.

- **User Experience:**

Incentivize detailed reviews to enhance credibility and trust.

thank you

