SENTIMENT ANALYSIS

BY:SUSAN WANJIRU

Title and Content Layout with List

This project aims to develop a Natural Language Processing (NLP) model to analyze sentiment in Tweets related to Apple and Google products. By classifying the sentiment of these Tweets as positive, negative, or neutral, the model will provide valuable insights into public perception, aiding businesses in marketing strategies and product development.

BUSINESS UNDERSTANDING

-
-
-

BUSINESS PROBLEM.

-

MAIN OBJECTIVES

gle an

1. Data Understanding Analyze the dataset, which includes columns such as tweet_text, emotion_in_tweet_is_directed_at, and is_there_an_emotion_directed_at_a_brand_or_product. Address any anomalies like missing values and duplicates.

2. 2. Data Preparation Remove duplicate entries. Populate missing values in th emotion_in_tweet_is_directed_at column with "none.“ Apply text preprocessing techniques: tokenization, lowercasing, stopwordremoval.

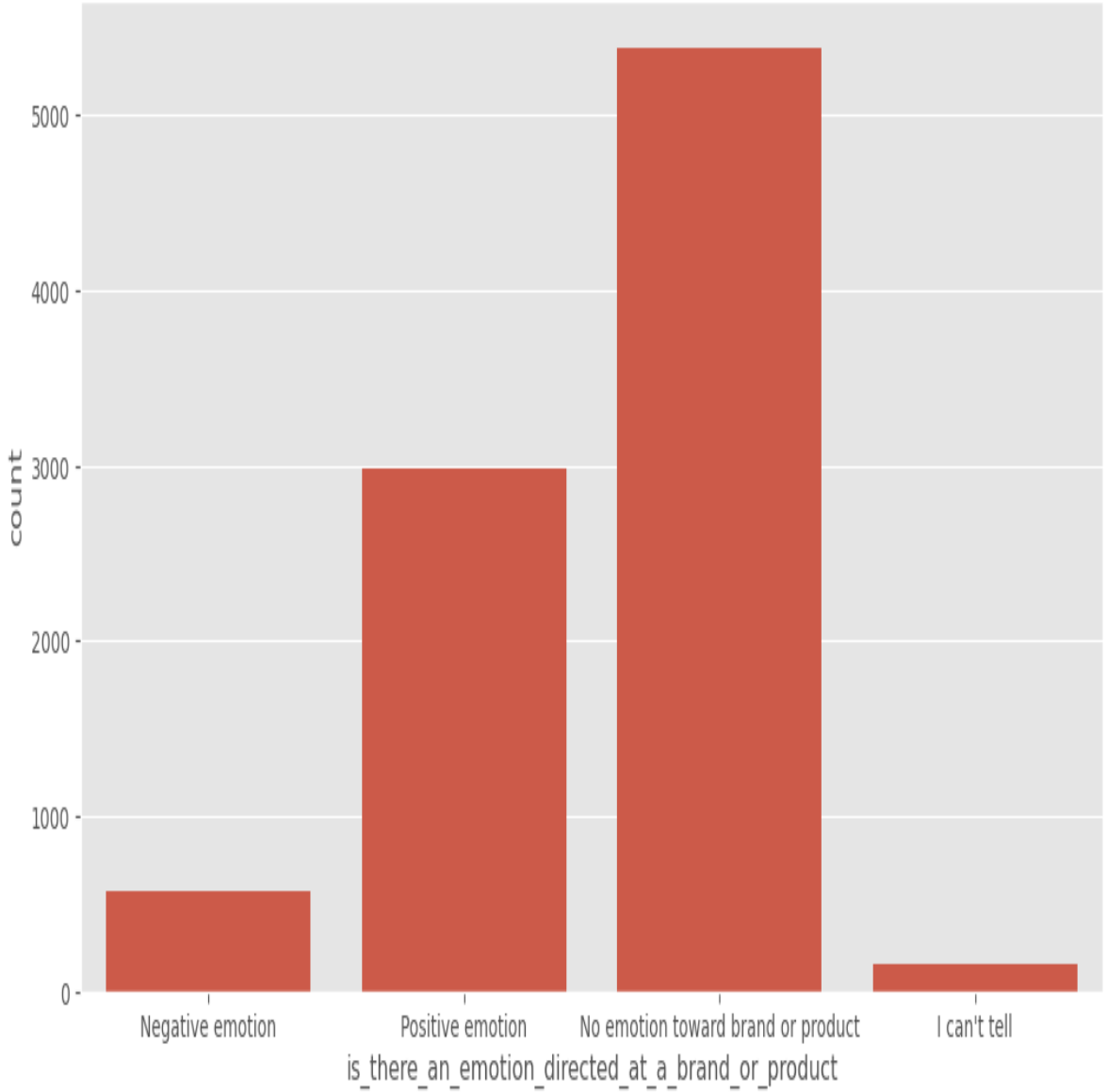
3. 3. Modeling Utilize libraries like bayes naves), sklearn'sCountVectorizer(for vectorization), and pandas (for data handling). Build a logistic regression model for binary classification (positive/negative sentiment), aiming for 70% accuracy. Expand the model to a multiclass classifier to include neutral sentiments.

4. 4. Evaluation Use accuracy as the primary evaluation metric, assessing the model’s ability to classify sentiments correctly. Address potential limitations such as missing values and data quality issues.

VISUALIZATIONS

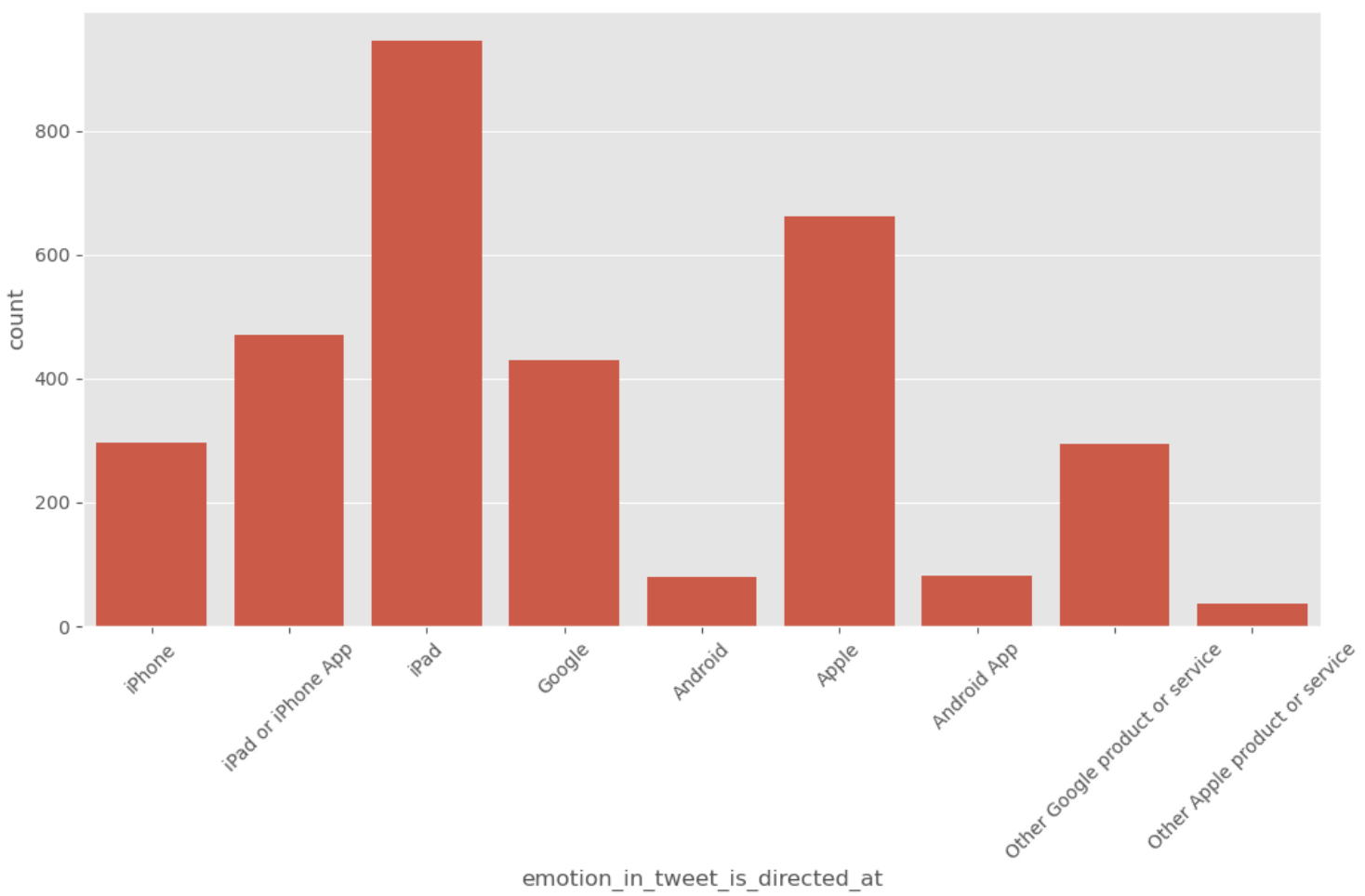
Sentiment Comparison Between Apple and Google Products

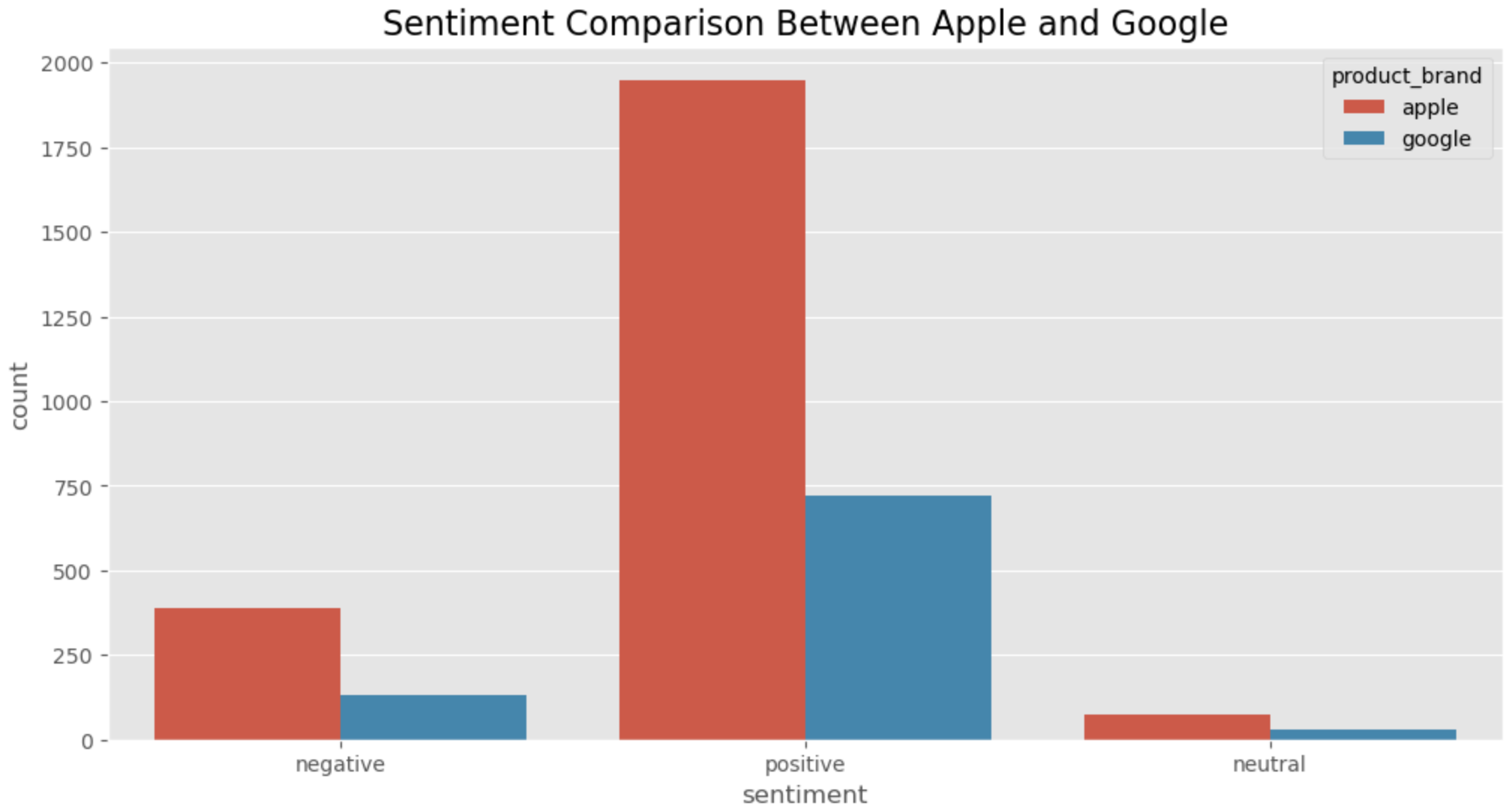
In this visualization there was more neutral reaction to the products .There was also a visualisation which couldn't be predicted and therefore we dropped it later.



Sentiment Breakdown and Visualization

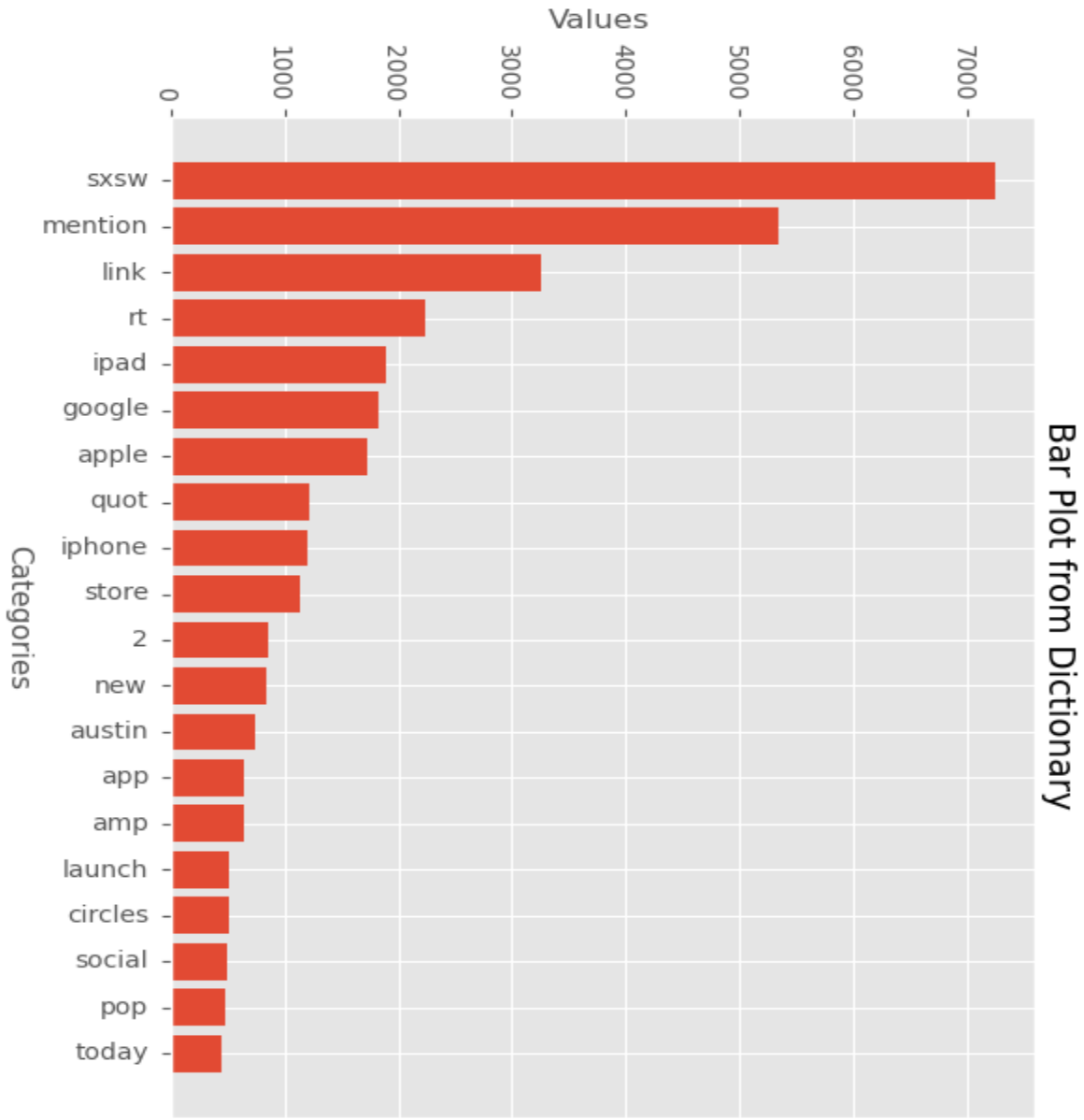
- The bar chart reveals that iPad is the most frequently mentioned product in the tweets, followed by other Apple products (iPad, iPhone, and Apple) and Google products. Android related products receive fewer mentions, highlighting the dominance of Apple products in user directed sentiments




Sentiment Comparison Between Apple and Google
product_brand
apple
google
2000
1750
1500
1250
1000
750
500
250
0
count
negative
positive
neutral
sentiment

The bar chart above compares the sentiment distribution between Apple and Google products. For positive sentiment, Apple has a significantly higher count compared to Google, indicating a strong positive reaction toward Apple products. Negative sentiment is more balanced but still higher for Apple than Google. Both brands have very low counts in the neutral sentiment category, with Apple showing slightly more mentions than Google. This comparison suggests that Apple products generate more engagement, particularly in positive sentiment, than Google products

From this visual, we notice that common terms like "link", "rt", "ipad", and "google" dominate the conversation, reflecting the general focus of discussions in the tweets.



RECOMMENDATIONS