SENTIMENT ANALYSIS

DERRICK WEKESA

CONTENTS

INTRODUCTION

In today's digital age, Twitter and similar social media platforms wield significant influence for gathering rapid customer feedback. It's essential for businesses to grasp customer sentiment about brands and products to make informed decisions, enhance satisfaction, and maintain a positive brand image. This project's goal is to develop a tailored sentiment analysis model for Twitter data related to Google, Apple, and other products.

PROBLEM STATEMENT

As a consulting firm, Twitter has entrusted us with the responsibility of constructing a model capable of assessing the sentiment expressed in a Tweet by analyzing its content. This model should accurately classify Twitter sentiments regarding Apple and Google products into categories of positivity, negativity, or neutrality. The primary objective is to extract valuable insights from public sentiment, enabling businesses to make well-informed decisions in shaping their strategic approaches and enhancing overall customer satisfaction.

OBJECTIVES

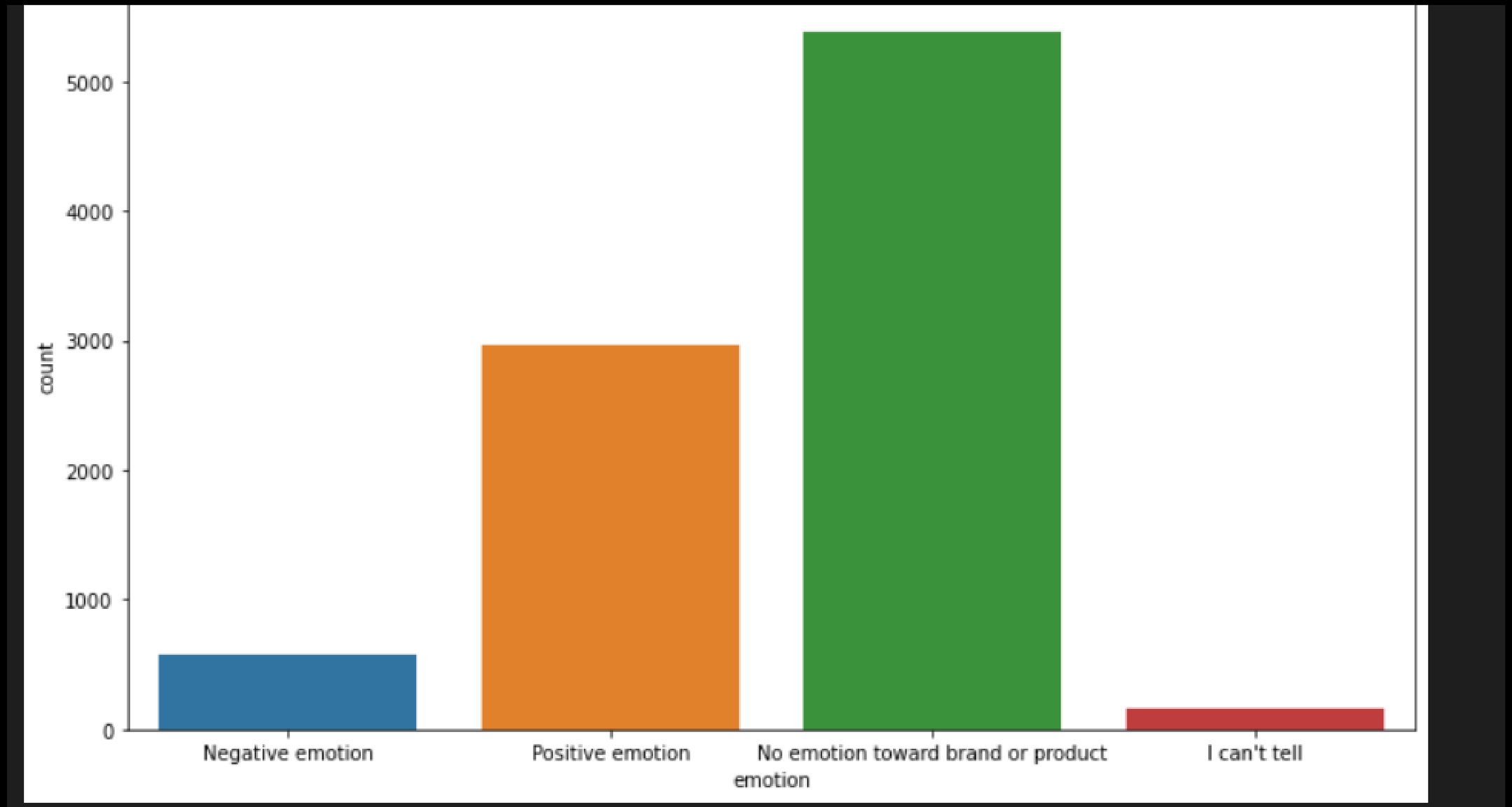
- To develop a binary classifier that can classify tweets into positive or negative sentiment categories. This will serve as a proof of concept and provide a foundation for further analysis.
- 2. To expand to a multiclass classifier, thereby incorporating the neutral tweets to create a multiclass classifier that can accurately classify tweets as positive, negative, or neutral. This will provide a more comprehensive sentiment analysis of the tweets.
- 3. To compare sentiment between Apple and Google products by analyzing the sentiment distribution of tweets mentioning Apple, Google, and other products

DATA UNDERSTANDING

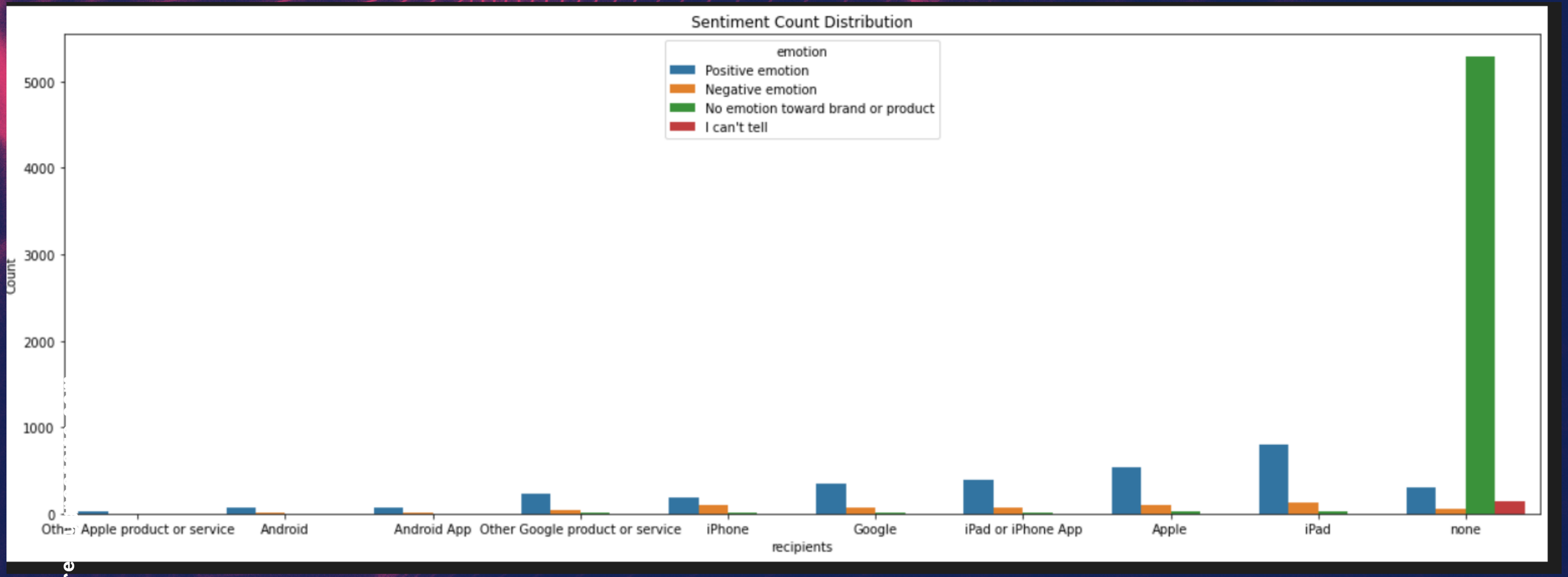
- I collected data from data world of customer tweets related to different products, including different brands like Apple, Google
- I chose Data World as the primary platform for data collection twitter provides a real-time customer opinions and feedbacks making it ideal for sentiment analysis

PATTERNS AND INSIGHTS

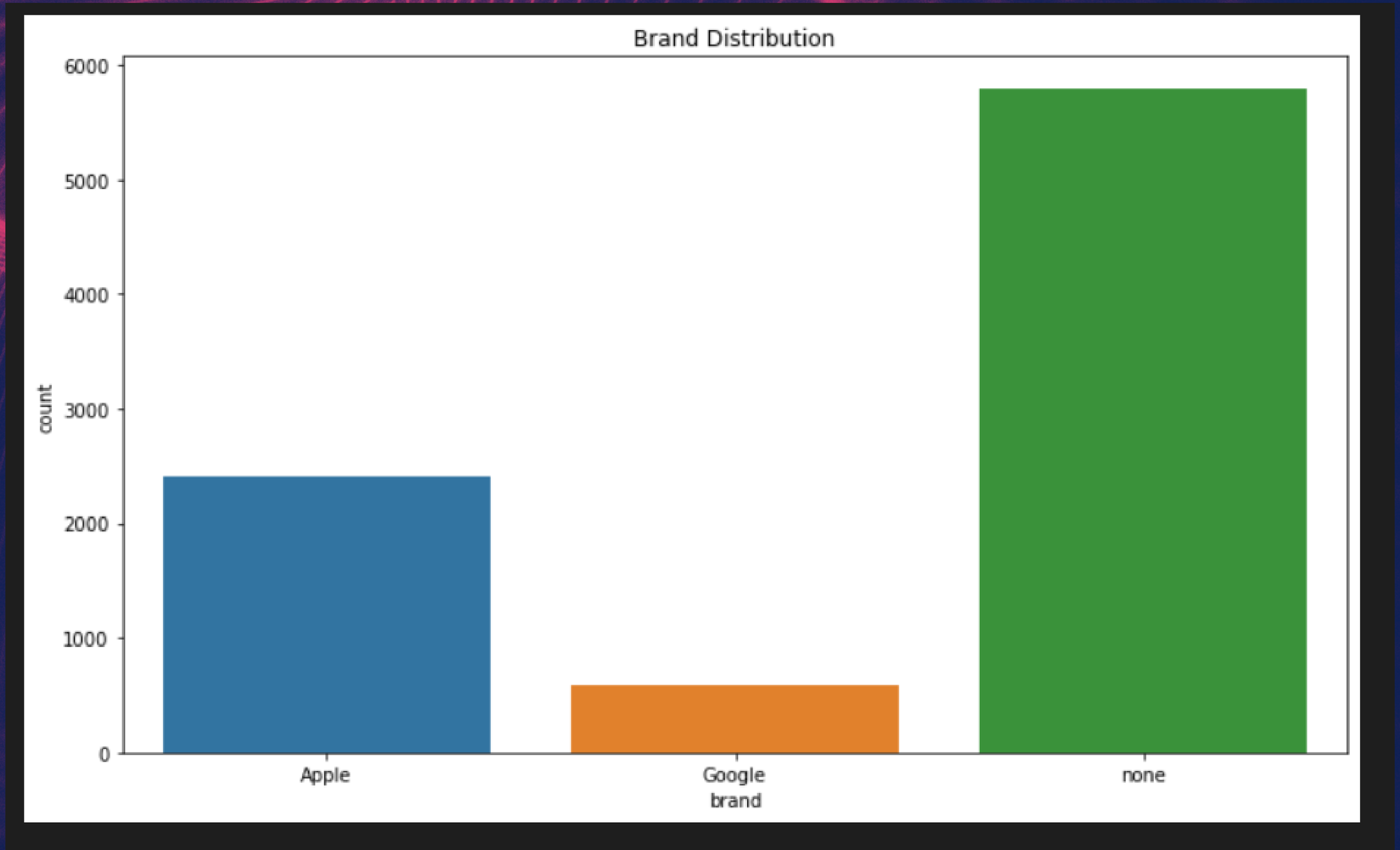
EMOTION DISTRIBUTION COUNT



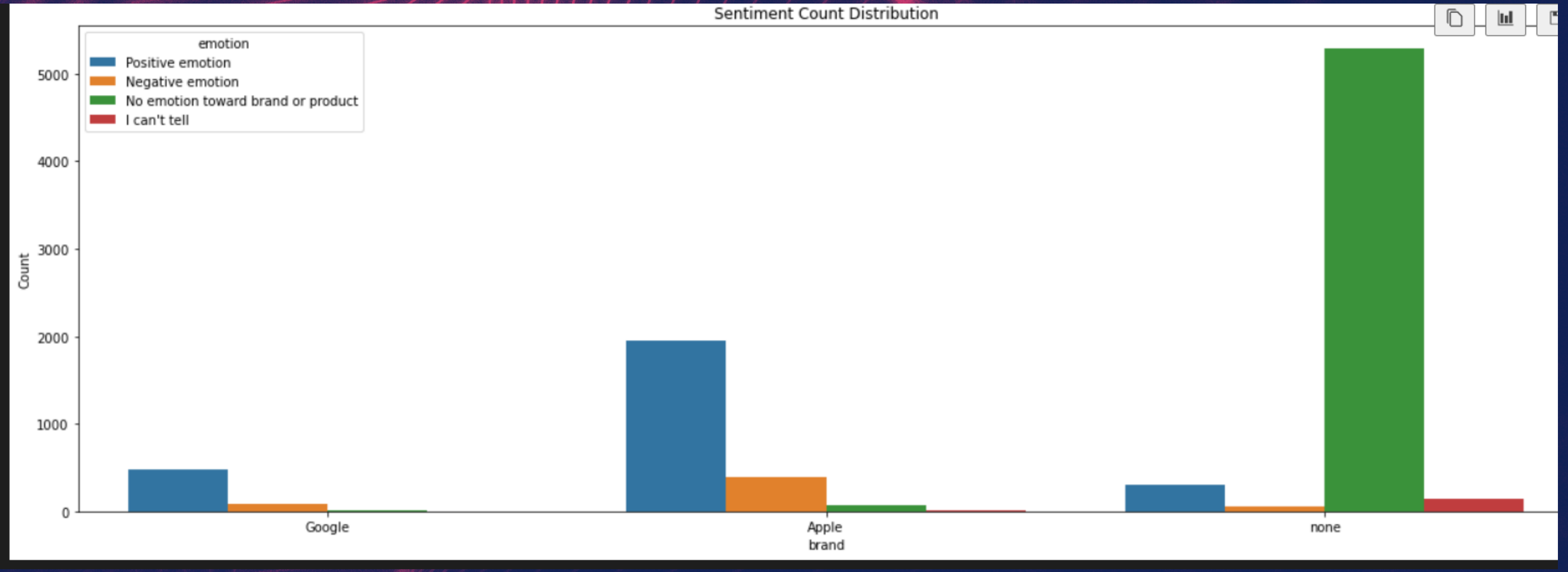
SENTIMENT DISTRIBUTION COUNT

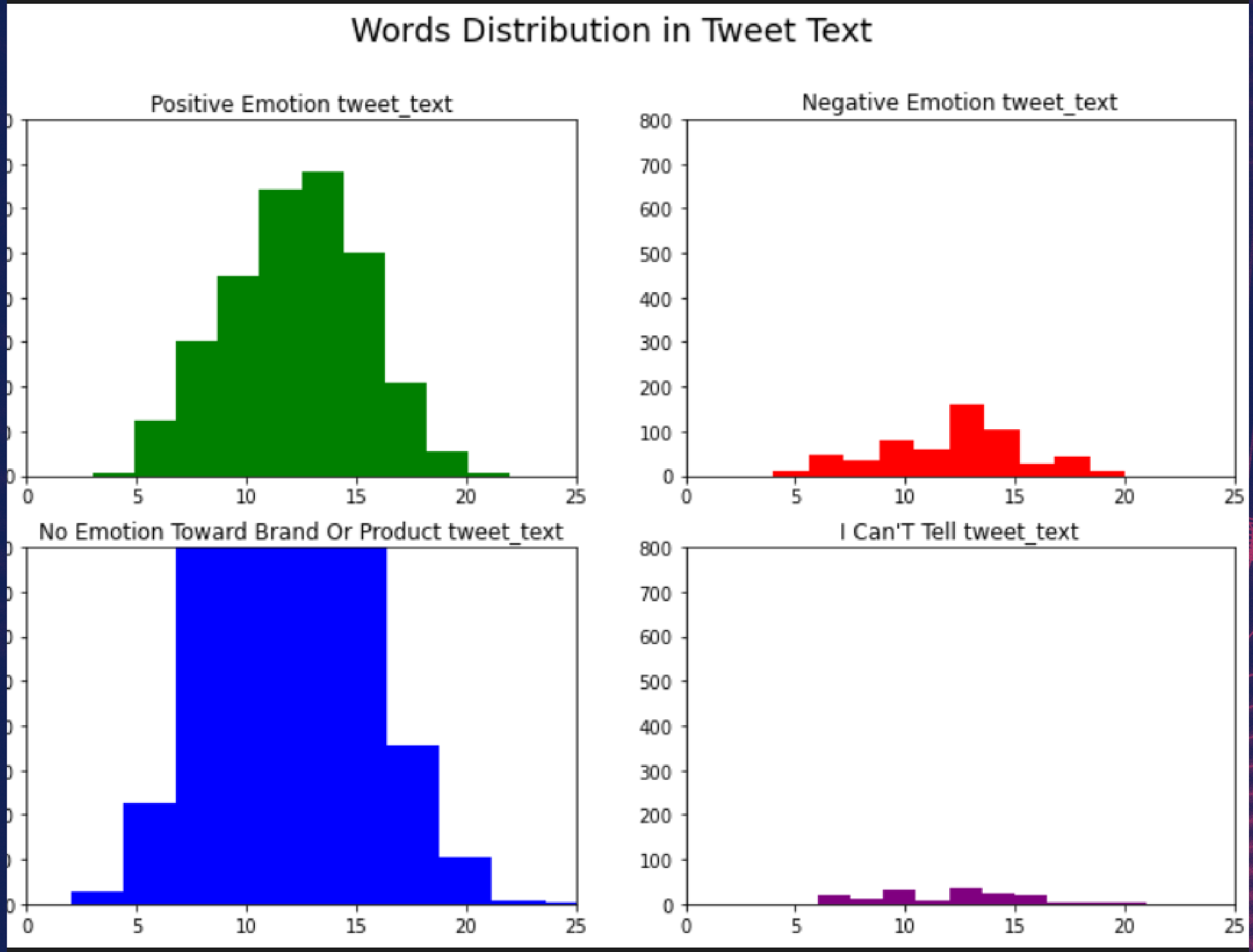


BRAND DISTRIBUTION

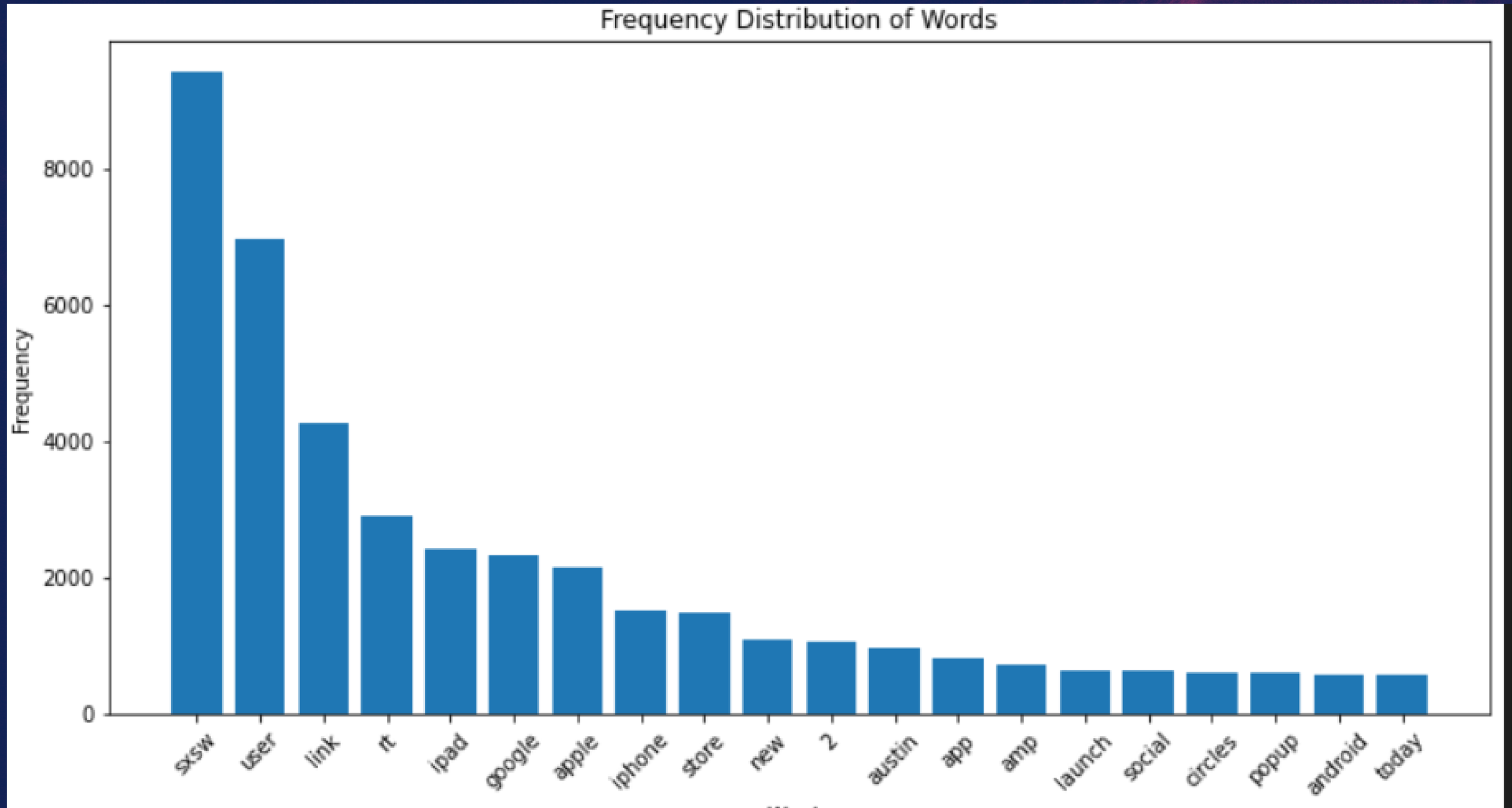


SENTIMENT COUNT DISTRIBUTION




Words Distribution in Tweet Text
Positive Emotion tweet_text
Negative Emotion tweet_text
No Emotion Toward Brand Or Product tweet_text
I Can'T Tell tweet_text
0
100
200
300
400
500
600
700
800
0
5
10
15
20
25

FREQUENCY DISTRIBUTION OF WORDS



FINDINGS

- Most tweets were directed to None brand category. This may indicate that customers were not engaging with the brand.
- Positive sentiments had the highest count compared to Negative sentiments, indicating that most people in general liked respective brands(Google and Apple)
- Most of the positive tweets were directed to Apple brands
- In the field of sentiment analysis, one of the significant challenges is dealing with language ambiguity and sarcasm detection. Natural language is complex and often subjective, making it difficult to accurately interpret sentiments from text.
- On average most of the tweets were 10-15 words long - more words increase ambiguity.

RECOMMENDATIONS

From the analysis I recommend that there be more customer engagement.

Probably check on this areas;

* Churn ratio - rate at which customers discontinue their relationship with a product company within a given time period
* Social media influencers through brand or product endorsement
* Customer feedback - The brands can introduce a rating system to accurately capture the sentiments of their customers