Visualising User Sentiment in Spotify Reviews

NLP & Data Visualisation Project

Marek Wisniewski
April 2025

Outline

- Project Objective
- Dataset Overview
- Data Cleaning
- Sentiment Analysis
- Word Frequency & Visualisation
- Key Insights
- Conclusion

Project Objective

- To explore user sentiment in Spotify reviews
- To identify which words dominate positive and negative sentiment
- To visualise the findings using Python and a Spotify-themed WordCloud

Dataset Overview

	0	1
0	Review	label
1	Great music service, the audio is high quality...	POSITIVE
2	Please ignore previous negative rating. This a...	POSITIVE
3	This pop-up "Get the best Spotify experience o...	NEGATIVE
4	Really buggy and terrible to use as of recently	NEGATIVE

Source: Kaggle

- This project uses a dataset of over 500,000 Spotify user reviews.
  Each review is labelled as either **POSITIVE** or **NEGATIVE**, enabling sentiment-based text analysis.

Below is a small sample of the raw data, showing:

- **Review** – the content of the user's opinion
- **Label** – sentiment category assigned to the review

Data Cleaning & Preprocessing

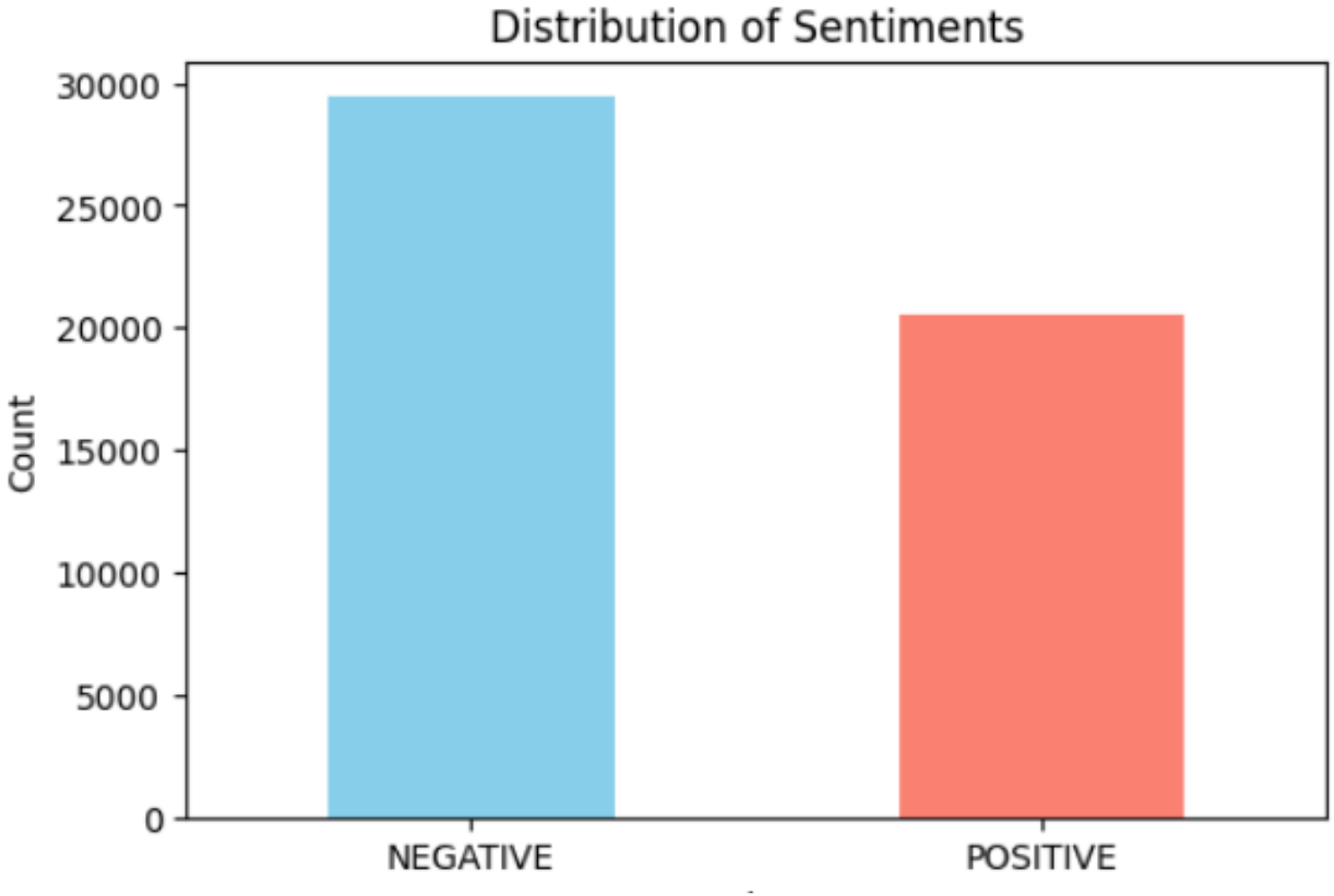
- Converted all text to lowercase
- Removed punctuation, special characters, and stopwords
- Used NLTK for stopword filtering
- Created cleaned text column

Sentiment Split

- Filtered reviews into positive and negative
- Extracted word frequency using *collections.Counter*
- Removed noise words

Top Frequent Words

Table of Top 5 Words in POSITIVE and NEGATIVE reviews



WordCloud Visualisation

Positive Reviews WordCloud
(with Spotify mask)

Negative Reviews WordCloud
(same style for comparison)

Positive Sentiment

Negative Sentiment

Insights

Top words in positive reviews:

- "music" (10,893), "app" (10,115), "spotify" (5,892), "love" (5,349), "songs" (4,291)
- Top words in negative reviews:
- "app" (18,387), "songs" (9,840), "music" (9,709), "song" (9,294), "spotify" (8,876)

Conclusion

- "App", "music", and "spotify" appear in both positive and negative reviews – they are central to user experience, and sentiment depends on context.
- "Love" is a highly frequent word in positive reviews, suggesting strong emotional approval.
- Negative reviews are more focused on functionality or usability issues – e.g., high mentions of "app", "songs", "song" may indicate dissatisfaction with performance or features.
- The fact that "app" is more common in negative sentiment may suggest users are more vocal when encountering issues.

Tools Used

- Python: Pandas, NLTK, Matplotlib, WordCloud
- Data source: Kaggle
- Environment: Google Colab

Thank you for your attention!

Feel free to connect or explore more:

- 📧 Email: marekwisniewskiuk@gmail.com
- 🔗 GitHub: https://github.com/M4R3K21
- 💼 LinkedIn: https://www.linkedin.com/in/marek-wisniewski-209930320/