

# Sound Harmony: A Review and Technical Data Approach to Headphones Selection

## i. Abstract

*How difficult and tiring has it been the last time you wanted to invest in a good pair of headphones?*

Headphones have become an essential part of our daily lives, whether we're listening to music, watching videos, or making calls, and choosing the perfect headphones can feel tricky with so many options available.

In this project, I analyzed how people rated eleven specific headphones through reviews and comments on Reddit platform. Moreover, since many headphone brand websites lack reviewer sections, Reddit offers a platform for users to share genuine experiences and opinions. Given that, I want to see if the reviews/comments from Reddit align with the article rating for those specific headphones. Additionally, I want to see which rating from these two has a stronger link with the technical specifications of the headphones, and which rating is more reliable for the potential customers.

As for the sentiment analysis, I will be analyzing three different models, and look deeply into which model is performing best for this project's topic. I will also perform LDA modelling to check which topics are mostly discussed when it comes to the headphones I chose.

Overall, this project shows the importance of balancing technical specifications with user experiences when choosing headphones. By adding real user feedback, individuals can make better informed decisions for their needs.

## ii. Introduction

In today's world, we see new headphones hitting the market all the time. With so many brands and features to choose from, it's not easy to find the headphones that are just right. Some people might just look at the price and hope for the best, but a higher price doesn't always mean better sound or quality.

On the other hand, some people turn to online articles to find the trendiest headphones. While these articles can be helpful, they might not always reflect what actual users think or feel about those headphones.

That's why I've decided to take a closer look at reviews and comments from Reddit for eleven different headphones. Alongside this, I'll also be analyzing articles from various online sources to see how they rate these headphones.

### iii. Data

For this project, I will use 2 different data sets. The first one is the data set which consists of 11 headphones found from different websites/articles and the second data set has the reddit reviews of each headphone. Some headphones have more reviews such as around 500 reviews, whereas some have around 90 reviews, where I still hope to achieve a good sentiment analysis for those products by doing weighted formula on the results.

#### Headphone\_ArticleList Data set

From the articles I have been able to gather technical data for eleven headphones, where eight are ranked as "good" according to the articles and three of them are ranked as "not good" headphones. The reason why I got "good" and "bad" headphones, is because I wanted to have a more balanced dataset which consisted of different rated headphones, brands, types, connectivity and so on. Furthermore, six of them are from articles published in 2024, four of them are from 2023 and one of the headphones in the list is from the article published in 2022. I was only able to gather the data manually since the websites of the headphone brands don't allow to scrape the data.

Description of the variables in headphones data set:

Field	Description
Headphone Name	The specific name of the headphone model.
Brand	The brand name of the headphone.
Article/Website	The name of the article or website that published the headphone.
Year	The publication year of the article source.
Type	The headphone design type: over-ear, in-ear, or on-ear.
Connectivity	The connectivity options of the headphone: wireless or hybrid (both wired and wireless).
Battery Life (hours)	The duration the headphone battery can last on a single charge, measured in hours.
Battery Capacity with Case (mAh)	The battery capacity of the headphone and its charging case, measured in milliampere-hours (mAh).
Active Noise Cancellation	Indicates if the headphone features active noise cancellation technology.
Bluetooth Version	The version of Bluetooth technology supported by the headphone.
Driver Sensitivity (dB)	The sensitivity of the headphone's driver, measured in decibels (dB).
Driver Size (mm)	The diameter size of the headphone driver, measured in millimeters (mm).
Tag/Comment	Additional labels or comments provided by the article or publication about the headphone.
Article Rate	The rating or evaluation given to the headphone by the article or publication (good or not good).

Initially, I got the Price of each headphone as well, but since the prices vary between different websites and geographical locations, I removed it. I also removed the Frequency Response since almost all of the headphones were in the range of 20Hz~40KHz (wouldn't be a helpful variable).

### **RedditComments Data Set**

For this data set I have been able to gather in total 6100 reviews/comments for the 11 headphones, by using the help of Reddit Api. Initially, I had more reviews/comments, but some posts were not related to the headphones, due to the fact that reddit also shows posts that have a similar word in the search query, therefore I removed those posts manually, and it resulted in removing around 500 hundred reviews/comments. I have gathered some reviews in the r/headphones community(subreddit), since it has 1.1 Million members and that seems like the right place to find some decent discussions around the headphones. However, for some of the headphones in my data set, there wasn't much discussions in r/headphones community, therefore I tried to gather some comments/reviews using the "all of Reddit" (different subreddit - not a specific one) in the search bar.

Description of the variables in reddit data set:

Field	Description
Headphone Name	The specific name of the headphone model.
Subreddit	The name of the subreddit where the post was published.
Post_Title	The title of the post.
Post_ID	The unique identifier for the post.
Post_Author	The username of the author who created the post.
Post_URL	The URL link to the post.
Post_Score	The net upvotes received by the post, indicating its popularity among users.
Num_Comments	The total number of comments on the post.
Created	The Coordinated Universal Time (UTC) timestamp when the post was created on Reddit.
Comment	The comment or review related to the headphone in the specific post.
Comment_Author	The username of the author who posted the comment.
Comment_Score	The net upvotes received by the comment, indicating its popularity among users.

#### **iv. Preprocessing Steps**

As for the preprocessing steps, I decided to do data transformation and also data cleaning, therefore in this part, I transform categorical variables in the headphones dataset into binary columns.

Specifically, 'Article Rate' is mapped to binary values where 'good' is represented as 1 and 'not good' as 0. Similarly, 'Active Noise Cancellation' is converted to binary, with 'yes' as 1 and 'no' as 0. Lastly, 'Connectivity'

is transformed into binary form, where 'both' corresponds to 1 and 'wireless' to 0. These transformations were done for effective use in data exploration and modeling steps.

After this step, I define a function named `clean_text` to preprocess textual data. Firstly, it removes URLs from the text using regular expressions. Next, it eliminates special characters and numbers, ensuring that only alphabetic characters and spaces remain. Finally, it removes any extra spaces and trims the text to produce a clean and normalized version. Therefore, I apply this `clean_text` function to the 'Comment' column in the reddit data set, ensuring that the comments are appropriately cleaned for next text analysis tasks.

## v. Data Visualization

In this section, I want to visualize the distribution and correlations of various features within the headphones dataset.

### Count Distribution Plots

I defined a custom function, `plot_count_distribution_with_annotations`, to plot the count distribution of categorical variables. I apply this function to visualize the distribution of 'Article Rate', 'Brand', 'Type', and 'Connectivity' within the dataset, providing insights into the frequency of each category.

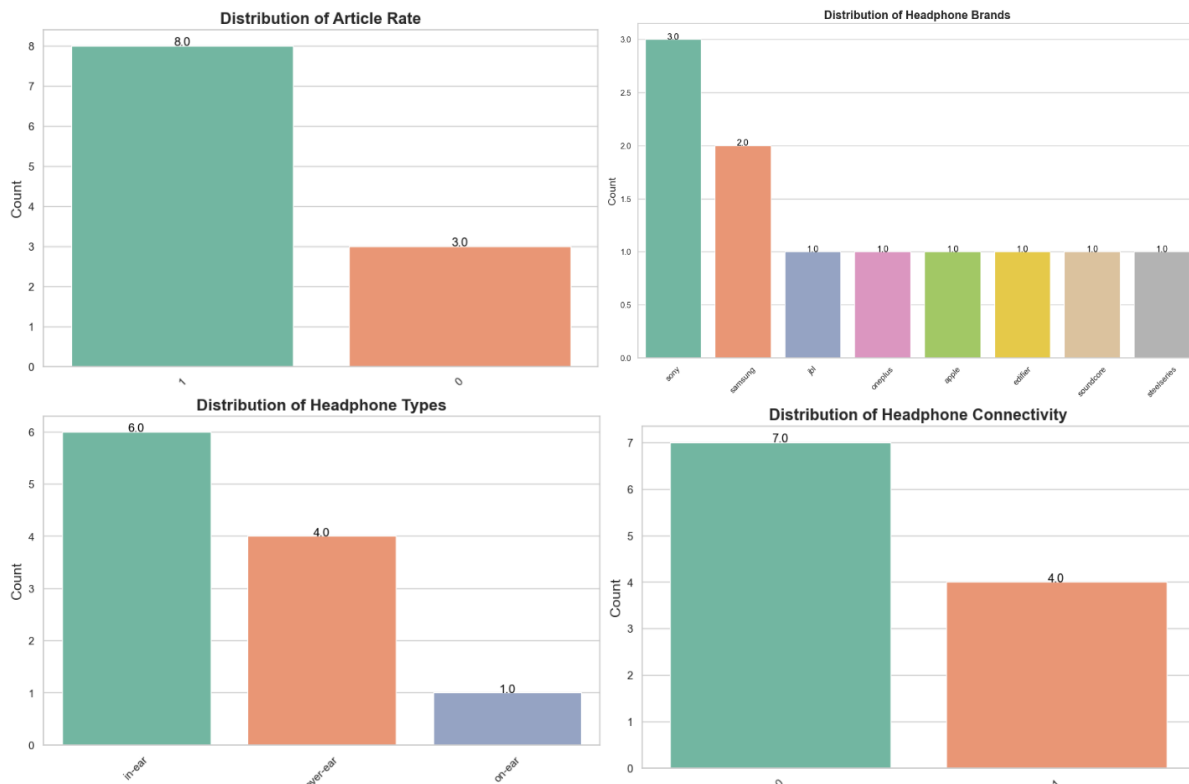


Fig. 1 Barplots of the distribution of categorical variables in the `Headphones_ArticleList` data set.

These visualizations show the information about the headphones I have in the headphones data set, in terms of the how many headphones are rate as “good” or “not good”, the distribution of brands in the data set, how many of them are in-ear, over-ear and on-ear, and how many of them are wireless or have hybrid

connectivity. I wanted my data set to not be biased in terms of only “good” headphones. That’s why I chose some of them that were rated as “bad” headphones in the articles. Initially, I had over 50 headphones in my data set, but it was very difficult to find all the necessary technical data on the internet for all of them (since every headphone brand puts the information differently on their official websites). Therefore, there were too many missing values to keep all of the headphones, that’s why I chose to stick with all these 11 headphones, where eight of them are rated as “good” and three of them are rated as “not good”. As for connectivity, four of them are hybrid (meaning wireless and wired) and seven of them are wireless.

## Correlation Matrix Plot

Additionally, I defined another function, `plot_correlation_matrix`, to visualize the correlation matrix of numerical features in the dataset, which shows the relationships between them.

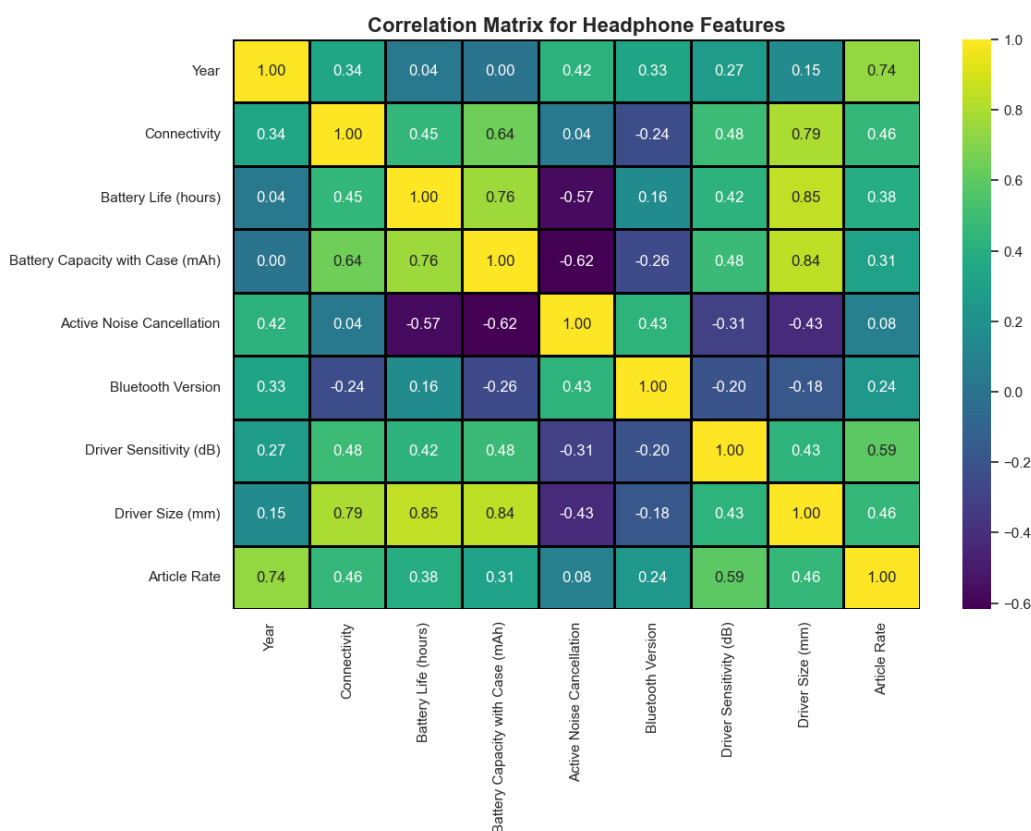


Fig. 2 Correlation heatmap of the numerical variables in the `Headphones_ArticleList` data set.

From this visualization, I can see that there’s some correlation between Article Rate and the other variables in the headphones data set. For example, the variables Connectivity, Driver Sensitivity and Driver Size are moderately correlated with Article Rate, whereas the other variables are weakly correlated with it. I can also see that Driver Size is strongly correlated with Connectivity, Battery Life and Battery Capacity.

## Text Length Analysis

Here, I create a new feature, `Comment_Length`, in the reddit data set to represent the length (the number of characters) of each comment across different headphones.

Next, I define the function `plot_text_length_histogram`, to plot histograms for the comment length distribution. Each subplot represents a unique headphone, and its color is determined using the 'husl' color palette from Seaborn, ensuring distinct colors for each headphone.

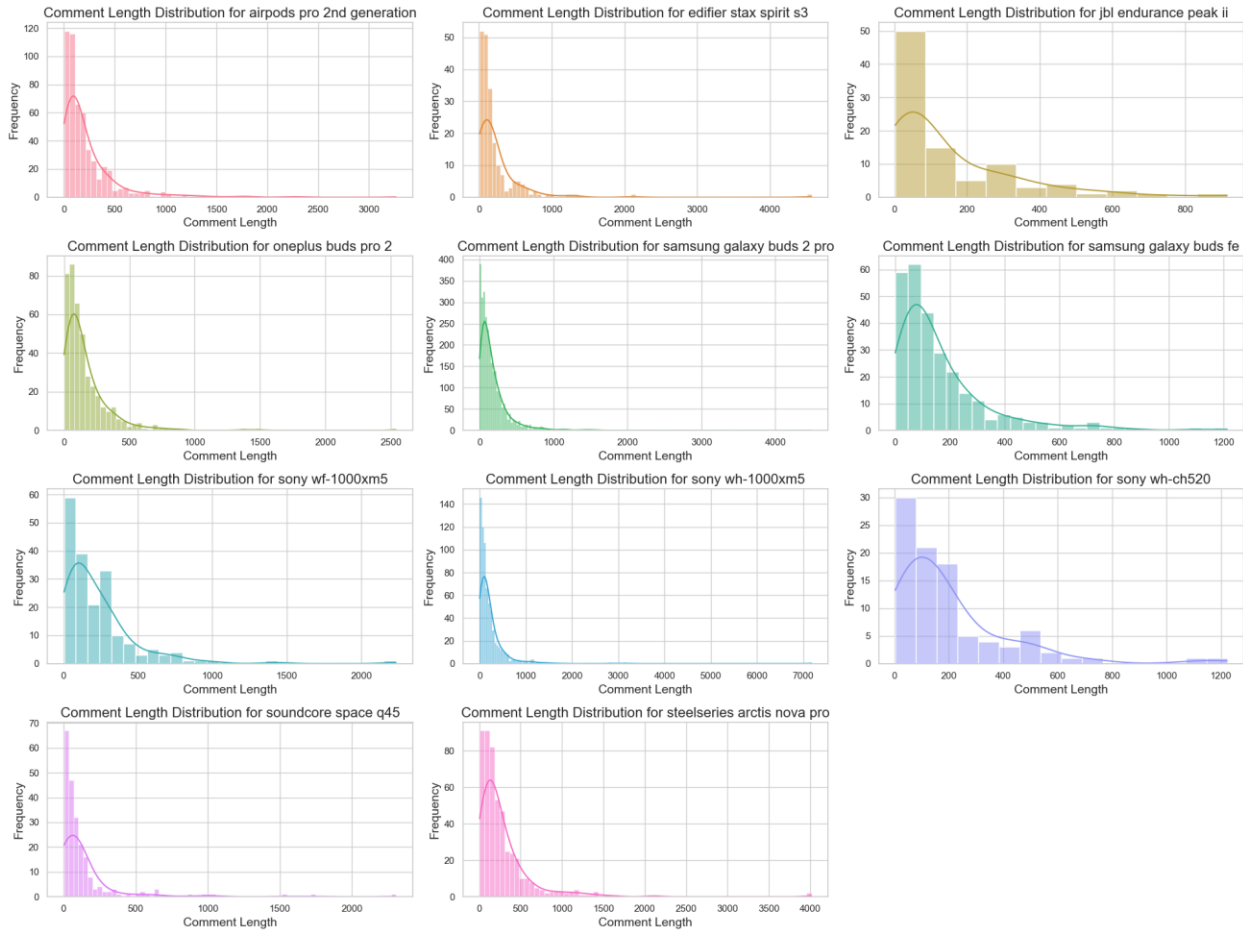


Fig. 3 The Reddit Comment Length Distribution for each headphone in the `Headphones_ArticleList` data set.

From this visualization, I can see that none of the headphones has the same comment length distribution as the others. Some of them have a lot more comments than others by just looking at the histogram bin widths and the frequency limits on the y-axis, for example, `samsung galaxy buds 2 pro` and `sony wh-1000xm5` have a lot more comments than `sony wh-ch520` and `jbl endurance peak ii`. Due to this fact, the further analysis with these headphones could be considered as “biased” if I don’t properly interpret the sentiments results by using the weighted formula to calculate the overall score.

## Word Cloud

For this part, I visualize the most frequent words present in the Reddit comments using a word cloud. A word cloud provides a visual representation of text data, where the size of each word indicates its frequency in the text (Wordcloud: What's it all about?, 2023).

First, I define the function `generate_word_cloud`, to generate a word cloud from the “Comments” variable. The resulting word cloud is saved as ‘word\_cloud.png’ then loaded and displayed as below:

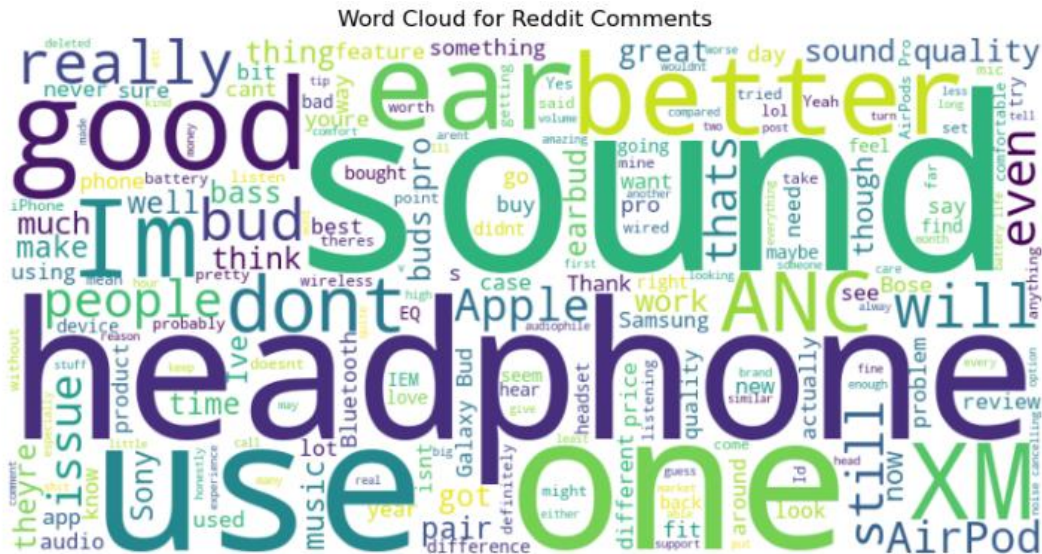


Fig. 4 The word cloud visualization on Reddit Comments for the headphones in the Headphones\_ArticleList data set.

From this visualization, I can see that the most terms used in the comments are Sound Quality, Ear, Buds, ANC, Audio, Music, Bass, Battery, Mic and Bluetooth which all are related to this topic. Additionally, I can also see some of the headphone brands, which had the highest number of comments as I stated before, such as Samsung Galaxy Buds and AirPods.

## Topic Modeling using Latent Dirichlet Allocation (LDA)

Now, I perform topic modeling on the Reddit comments using Latent Dirichlet Allocation (LDA). Topic modeling is a technique used to discover topics present in a collection of documents/comments. LDA is a probabilistic model that assumes each document is a mixture of a small number of topics, and each word in the document is attributable to one of the document's topics (Kulshrestha, 2019).

I set the maximum word frequency (max\_df) to 0.95 (95%) and the minimum word frequency (min\_df) to 2 to filter out words that are too frequent or too rare. Additionally, I remove English stop words to focus on meaningful words.

I specify `n_components=5` to indicate that I want to identify 5 topics in the dataset.

Then I define the function `display_topics` to display the top words associated with each topic. This function sorts the words in each topic by their contribution to the topic and displays the top `no_top_words` words for each topic.

Lastly I group the comments by 'Headphone Name' and perform LDA topic modeling separately for each group. This allows me to identify and display the dominant topics specific to each headphone brand.

Headphone Name	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
airpods pro 2nd generation	like sound apple airpods quality	anc gen sound noise airpods	airpods real got apple pro	sound real refurbished like pair	good sound better just airpods
edifier stax spirit s3	headphones like sound wireless stax	stax ms really think wireless	audeze products bit price look	closed lcd mobius audeze like	stax edifier audeze sound headphones
jbl endurance peak ii	jbl thank doesnt tried ive	thanks solution works alot entering	worked case right bluetooth times	thanks worked im good right	thank gave lol works like
oneplus buds pro 2	buds pro sound oneplus good	im tips foam noise using	oneplus pro buds earbuds phone	buds pro battery great life	sound ive did noise just
samsung galaxy buds 2 pro	sound quality buds headphones like	ear noise like buds airpods	buds ear just ive im	buds deleted pro apple galaxy	good sound bass consumer just
samsung galaxy buds fe	buds sound like fe fit	buds just tips anc pair	buds ear use pro just	buds fe pro fit galaxy	good touch yes foam samsung
sony wf-1000xm5	xm like im sound dont	anc noise xms ears just	pro ear im better buds	years great xms better sound	xm xms just sound bass
sony wh-1000xm5	headphones sound xm like im	xms max airpods xm sony	got use time yes definitely	welcome deleted xms youre dont	xm sound xms anc better
sony wh-ch520	like sound headphones battery im	thanks headphones ear bot automatically	ear pads im working replacement	im ch sidetone headphones like	good quite ive theyre clear
soundcore space q45	battery headphones sound bass quality	listening qc battery issues happened	send just volume space got	wired just qs headphones batteries	battery did good brand batteries
steelseries arctis nova pro	sound like pro im headset	mic just headset like ear	wireless just pro headset nova	gaming sound wireless guide headsets	good fulla impressed offer yesterday

Fig. 5 The most discussed Topics on Reddit for the headphones in the Headphones\_ArticleList data set.

From this visualization I can look into the most common discussions or interests across all the headphones. As for example for the sony wh-ch520, I can see the word “clear” therefore I might suppose that they have clear sounds, the word “replacement” together with “pads” can mean that the ear pads may need replacement after a while. For samsung galaxy buds 2 pro, “good sound bass” can mean that they have a good bass. For jbl endurance peak ii, I can see a lot of “thanks, worked” which can mean that a lot of the comments were about the issues customers were having and they were trying to get some help at reddit. For soundcore space q45, I can see a lot of “good battery” which can mean that the battery capacity/life is good from the experience of the commenters. Moreover, for the steelseries arctis nove pro, “gaming” can mean that these headphones are recommended for gaming enthusiasts.

## vi. Sentiment Analysis – Roberta Model

For this part of the project, I tried to compare three different sentiment analysis models and looked at how each model rated some randomly chosen comments, and I tried to chose which one was the best in this scenario. The models I used were:

- **TextBlob** (Uses the rule-based approach, and if a comment contains slang, emojis, or non-standard language, TextBlob might struggle to accurately interpret the sentiment (Gupta, 2023).)
- **Vader** (Uses a lexicon and rule-based approach specifically designed for sentiment analysis of social media text (Berl, 2020).)
- **Roberta** (Uses training data approach, where the training data can influence the model's understanding and interpretation of sentiment. Its complexity might allow it to capture nuances in sentiment that simpler models may miss (Efimov, 2023).)



Now I am going to show three of some random comments, which got these results:

---

**Comment:** *Interesting I guess you have bad luck with them then Have you tried many other brands*

TextBlob Sentiment: Positive (0.04)  
Roberta Sentiment: Negative (Pos: 0.03, Neu: 0.38, Neg: 0.59)  
Vader Sentiment: Positive (0.30)

Here the user is responding to someone's experience, stating that the person has had bad luck with those headphones and also asking if the person has tried other headphones, which looks like a negative comment for those specific headphones.

---

**Comment:** *Which speaker setup are you using and is it capable of playing dolby atmos codec*

TextBlob Sentiment: Positive (0.20)|  
Roberta Sentiment: Neutral (Pos: 0.04, Neu: 0.93, Neg: 0.03)  
Vader Sentiment: Positive (0.53)

Here the user is only asking a question, which can only be labeled as neutral comment.

---

**Comment:** *I guess its just luck really Some people are allergic to certain things others arent I have Galaxy Buds Pro and zero issues However latex gloves make my hands very very rough in a matter of minutes like grit sandpaper I have a latex allergy Most people dont xB If you bought the Galaxy Buds Pro and have an allergic reaction then I hope youll be able to return them For me theyre all good*

TextBlob Sentiment: Positive (0.33)  
Roberta Sentiment: Neutral (Pos: 0.40, Neu: 0.43, Neg: 0.17)  
Vader Sentiment: Positive (0.84)

This comment can be more difficult to understand since this user is stating that they have the Galaxy Buds Pro and they have no issues with them, but there are many people who could be allergic to the latex material, hence could have issues with the headphones. That's why this comment is considered as neutral comment.

From all these examples, I can clearly see that the Roberta Model rated the comments most accurately, that's why I decided to use this model for my further sentiment analysis.

**As a first step,** I'm initializing a pre-trained sentiment analysis model called 'cardiffnlp/twitter-roberta-base-sentiment'. This is a Roberta-base model which is trained on ~58M tweets and finetuned for sentiment analysis with the TweetEval benchmark and is suitable for English

(Francesco Barbieri, 2020). Then, I'm using the Hugging Face Transformers library to load both the tokenizer and the model for sequence classification from this pre-trained model.

**As a second step**, I've defined a function named *polarity\_scores\_roberta* to perform sentiment analysis using the ROBERTA model. The function takes a comment as input, preprocesses it to fit the model's token limit, and then calculates sentiment scores for negative, neutral, and positive sentiments. These scores are returned as a dictionary. Consequently, I'm calculating sentiment scores for each comment using this function.

**As a third step**, I've defined a function named *calculate\_overall\_rate* to determine the overall sentiment rate for each comment based on the sentiment scores obtained from the ROBERTA model. The function identifies the maximum sentiment score among negative, neutral, and positive sentiments and assigns an overall rate of 1 for positive sentiment, -1 for negative sentiment, and 0 for neutral sentiment. I then apply this function to the sentiment scores calculated earlier for each Reddit comment to create a new column named **Overall\_Rate** in the reddit data set.

Additionally, I want to show the distribution of the **Overall\_Rate** it's visualization across all the headphones.

Overall_Rate	-1	0	1
Headphone Name			
airpods pro 2nd generation	104	183	241
edifier stax spirit s3	44	98	62
jbl endurance peak ii	19	35	38
oneplus buds pro 2	94	207	116
samsung galaxy buds 2 pro	855	1214	818
samsung galaxy buds fe	61	109	104
sony wf-1000xm5	43	79	67
sony wh-1000xm5	119	295	278
sony wh-ch520	19	42	32
soundcore space q45	68	104	49
steelseries arctis nova pro	164	175	164

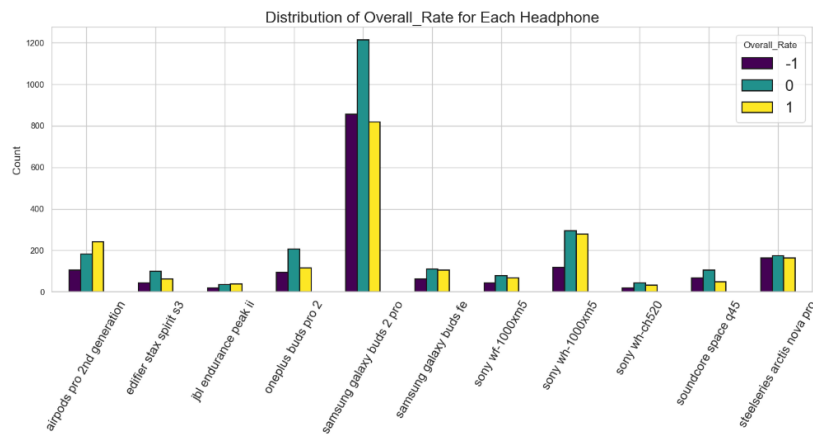


Fig. 6 Overall\_Rate distribution and its visualization for each headphone in the Headphones\_ArticleList data set.

From this table, and it's visualization, I can see how many positive, neutral, negative comments each of the headphones have. These don't tell that much about which headphones were the best/worst rated since all of them had different numbers of comments, but I will look into that further below.

**As a forth step**, I am firstly computing the mean of Overall\_Rate for each headphone, named **Mean\_Rate**. Then, I'm defining a **Weighted\_Rate** variable which consists of the formula:

$$\text{Weighted\_Rate} = \begin{cases} \text{Mean\_Rate} * \text{number of positive comments}, & \text{if } \text{Mean\_Rate} > 0 \\ \text{Mean\_Rate} * \text{number of negative comments}, & \text{if } \text{Mean\_Rate} < 0 \\ 0, & \text{if } \text{Mean\_Rate} = 0 \end{cases}$$

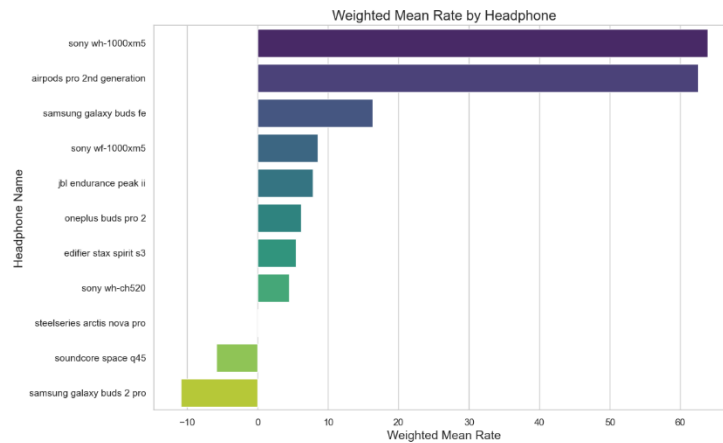
The reason why I chose this formula is because of different distributions of comment number across the headphones I have in my data set. If one headphone has the Mean\_Rate negative, I would want to multiply it by the number of negative comments it had, and the same goes for the positive scenario. In the last table, I also saw that one of the headphones(steelseries arctis nova pro) had the same number of positive and negative comments, therefore the Mean\_Rate of that headphone is 0(neutral).

The table below shows each headphone with the number of negative and positive comments, Mean\_Rate and Weighted\_Rate.

	Headphone Name	Number_of_Negative_Ratings	Number_of_Positive_Ratings	Mean_Rate	Weighted_Rate
0	sony wh-1000xm5	119	278	0.229769	63.875723
1	airpods pro 2nd generation	104	241	0.259470	62.532197
2	samsung galaxy buds fe	61	104	0.156934	16.321168
3	sony wf-1000xm5	43	67	0.126984	8.507937
4	jbl endurance peak ii	19	38	0.206522	7.847826
5	oneplus buds pro 2	94	116	0.052758	6.119904
6	edifier stax spirit s3	44	62	0.088235	5.470588
7	sony wh-ch520	19	32	0.139785	4.473118
8	steelseries arctis nova pro	164	164	0.000000	0.000000
9	soundcore space q45	68	49	-0.085973	-5.846154
10	samsung galaxy buds 2 pro	855	818	-0.012816	-10.957742

*Fig. 7 The general table of the rates for each headphone in the Headphones\_ArticleList data set.*

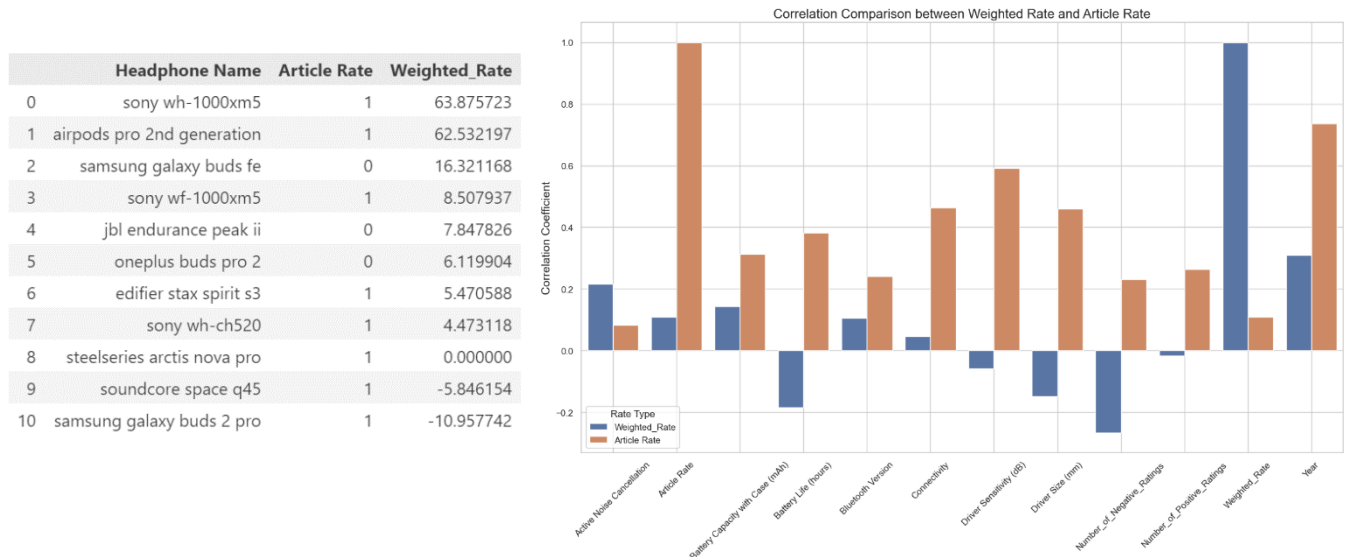
Now, I want to visualize the Weighted\_Rate by using bar plot, with the headphones sorted based on their weighted rates in descending order.



*Fig. 8 The Barplot visualization for Weighted\_Rate for each headphone in the Headphones\_ArticleList data set.*

From this visualization, I can see that only the Samsung Galaxy Buds 2 Pro and SoundCore Space q45 were rated as “not good” headphones from the reddit users/commenters. On the other hand, Sony wh-1000xm5 and AirPods Pro 2<sup>nd</sup> generation were rated the best from all these eleven headphones.

**As a last step,** I want to see how well the weighted rate aligns with the article's ratings for each headphone. Additionally, I want to also look at the correlations between the Weighted\_Rate and Article Rate with the technical data from the headphones dataset, which highlights potential alignment or discrepancies between the two rating systems.



*Fig. 9 The comparison table of the Article vs Reddit Rate results for each headphone in the Headphones\_ArticleList data set.*

From this visualization, I can see that the Reddit comments/reviews don't align with the website/articles ratings. A part of the reason is that articles may base their opinions on technical data, but the users on reddit share their experience with the products which in the end is more helpful to the ones who are looking to purchase those headphones. As for the correlation barplot visualization, I can see that there is a much weaker correlation between the technical data with Weighted Rate than with the Article Rate, which helps my reasoning that Article Rate may be based on technical data of the headphones.

## vii. Conclusion

In conclusion, in this project I wanted to analyze and compare user reviews on Reddit platform and online articles' ratings for eleven specific headphones. The project consists of sentiment analysis using three models, topic modeling with LDA, and correlation analysis with technical specifications of the headphones.

The Roberta model, which is trained on ~58M tweets, outperformed TextBlob and Vader in accurately assessing sentiment from Reddit comments, therefore I used this model to continue my sentiment analysis, with extracting the Overall Rate for each comment. On the other hand, the headphones I chose, had different number of comments (some of them being more discussed on Reddit), therefore in order to

perform a non-biased sentiment analysis I chose to create the Weighted\_Rate variable. This rate represents the mean of the Overall Rate for each headphone, multiplied by the number of comments (positive/negative/neutral) each headphone had.

With the help of LDA topic modelling, by looking at the general topics that were discussed on Reddit I understood the user preferences and concerns around the headphones.

In the end, I observed that there was a weak correlation between Weighted\_Rate from Reddit comments and technical specifications, indicating that user experiences may diverge from technical assessments. Moreover, there was a stronger correlation between Article Rate and technical data, suggesting that article ratings may be influenced by technical specifications.

Overall, this project emphasizes the importance of considering both user experiences and technical specifications when evaluating headphones. While online articles may provide valuable technical insights, user reviews on platforms like Reddit offer authentic experiences that can guide potential customers in making informed decisions. On the other hand, accessing detailed technical information was quite challenging due to limited availability on brand websites therefore this project could be further enhanced if there was an easier access to technical data from official brand websites and by including a broader range of headphones for analysis. Such improvements would allow for a more detailed comparison between technical data and Reddit user comments.

## References

- AirPods Pro*. (n.d.). Retrieved from apple.com: <https://www.apple.com/de/airpods-pro/>
- Alice Richard, J. T. (2022, April 5). *4 earphones to avoid*. Retrieved from choice.com.au: <https://www.choice.com.au/electronics-and-technology/home-entertainment/personal-listening/articles/the-5-worst-in-ear-headphones>
- ARCTIS NOVA PRO WIRELESS FOR PC & PLAYSTATION*. (n.d.). Retrieved from steelseries.com: <https://steelseries.com/gaming-headsets/arctis-nova-pro-wireless-pc-playstation>
- Beri, A. (2020, May 27). *SENTIMENTAL ANALYSIS USING VADER*. Retrieved from Towards Data Science: <https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664>
- Caddy, B. (2023, December 30). *Headphone winners and losers of 2023*. Retrieved from techradar.com: <https://www.techradar.com/audio/headphones/headphone-winners-and-losers-of-2023-from-sony-to-samsung-to-beats>
- Editors, R. (2024, March 2). *The Best Headphones for 2024*. Retrieved from rollingstone.com: <https://www.rollingstone.com/product-recommendations/electronics/best-headphones-1234964365/>
- Efimov, V. (2023, September 25). *Large Language Models: RoBERTa — A Robustly Optimized BERT Approach*. Retrieved from Towards Data Science: <https://towardsdatascience.com/roberta-1ef07226c8d8>
- Francesco Barbieri, J. C.-C.-A. (2020). *TWEETEVAL: Unified Benchmark and Comparative Evaluation for Tweet Classification*. Santa Monica, CA 90405, USA: School of Computer Science and Informatics, Cardiff University, United Kingdom. Retrieved from <https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment>
- Galaxy Buds FE*. (n.d.). Retrieved from samsung.com: <https://www.samsung.com/de/audio-sound/galaxy-buds/galaxy-buds-fe-graphite-sm-r400nzaadbt/>
- Galaxy Buds2 Pro*. (n.d.). Retrieved from samsung.com: <https://www.samsung.com/de/audio-sound/galaxy-buds/galaxy-buds2-pro-graphite-sm-r510nzaadbt/>
- Gupta, S. S. (2023, May 4). *Introduction to TextBlob in NLP*. Retrieved from Scaler: <https://www.scaler.com/topics/nlp/nlp-textblob/>
- JBL Endurance Peak II*. (n.d.). Retrieved from de.jbl.com: <https://de.jbl.com/outlet/ENDURANCE+PEAK+II-.html>
- JBL Endurance Peak II True Wireless In-Ear Sports Headphones*. (n.d.). Retrieved from amazon.de: <https://www.amazon.de/JBL-Endurance-True-Wireless-Ear-Sportkopfh%C3%B6rer-Schwarz/dp/B08FB33K2R>
- Kulshrestha, R. (2019, Jul 19). *A Beginner's Guide to Latent Dirichlet Allocation(LDA)*. Retrieved from Towards Data Science: <https://towardsdatascience.com/latent-dirichlet-allocation-lda-9d1cd064ffa2>

*OnePlus Buds Pro 2*. (n.d.). Retrieved from oneplus.com:

<https://www.oneplus.com/global/oneplus-buds-pro-2>

*S3 Wireless Over-Ear Headphones*. (n.d.). Retrieved from edifier.com:

<https://www.edifier.com/global/p/over-ear-on-ear-headphones/s3>

*Space Q45 | Long-Lasting Noise Cancelling Headphones*. (n.d.). Retrieved from eu.soundcore.com:

<https://eu.soundcore.com/products/space-q45-a3040011>

Tim Gideon, B. M. (2024, March 23). *The Best Headphones for 2024*. Retrieved from uk.pcmag:

<https://uk.pcmag.com/headphones/63/the-best-headphones>

*WF-1000XM5*. (n.d.). Retrieved from sony.de: <https://www.sony.de/headphones/products/wf-1000xm5>

*WH-1000XM5 wireless noise canceling headphones*. (n.d.). Retrieved from sony.de:

<https://www.sony.de/store/product/wh1000xm5b.ce7/WH-1000XM5-kabellose-Kopfh%C3%B6rer-mit-Noise-Cancelling>

*WH-CH520*. (n.d.). Retrieved from sony.de: <https://www.sony.de/headphones/products/wh-ch520>

*Wordcloud: What's it all about?* (2023, October 30). Retrieved from DataScientest:

<https://datascientest.com/en/wordcloud-whats-it-all-about>