

# EXPLORING PARENTING ADVICE

---

Unveiling Trends, Sentiments, and  
Connections in Online Parenting Discussions



# UNLOCKING INSIGHTS FROM ONLINE PARENTING

- In an era where digital forums have become a cornerstone for community support and knowledge exchange, understanding the nature of these interactions, especially in the context of parenting, is more crucial than ever. This project employs advanced text mining techniques to delve into the vast and varied landscape of parenting advice shared on Reddit.

# DATASET OVERVIEW

- The dataset is a compilation of text-based posts and comments where individuals seek guidance or provide advice related to parenting. Each entry is enriched with metadata such as post content, comment content, comment score, comment date, and provider. By mining this rich unstructured data source, we aim to interpret the collective narrative of parenting in the digital age.

	Post_ID	Post_Content	Comment_Content	Comment_Score	Comment_Date	Provider
0	1bk58ru	My first baby was born after a long old 36 hou...	My friend gave birth in the car in the hospita...	359	2024-03-21	r/beyondthebump
1	1bm36ak	Went to bed around 11 after bub had his bottle...	I mean, you did it. You hit the partner teamwo...	1629	2024-03-24	r/beyondthebump
2	1c2aazi	Just a scenario i thought was sad but also a l...	Sorry I laughed out loud at the title alone. B...	719	2024-04-12	r/beyondthebump
3	1brvfep	I was angry at you today. You asked me to go p...	This is.. rough. In your writing I can feel th...	799	2024-03-30	r/beyondthebump
4	1bnfst8	Mine is that part of the reason newborns cry i...	That colic is a lazy diagnosis and synonymous ...	1252	2024-03-25	r/beyondthebump



# METHODOLOGY

- **NLP Pipeline:**

**Tokenization:** Break down complex text into singular words, or 'tokens,' to facilitate deeper analysis.

**Removing Stopwords:** Common words add more noise than meaning, and they are filtered out.

**Lemmatization:** Words are reduced to their base, aiding in the standardization of the dataset.

**Removing Special Characters:** Non-alphanumeric characters are removed to clean the data further.

- **Sentiment Analysis:**

This process involves assigning a sentiment score to each comment, classifying them as positive, negative, or neutral. This sentiment score is crucial to gauge the overall emotional tone of the discourse within the parenting community on Reddit.

- **Text Similarity Analysis:**

By comparing documents based on their content, we identify posts that discuss similar issues or share advice. This approach is vital for recognizing common patterns and themes in parenting discussions.

- **Question-Answering System:**

We utilize an interactive system that allows users to query the dataset in natural language. The system then retrieves the most relevant answers, enhancing user engagement and making the insights more accessible.

# NLP PIPELINE

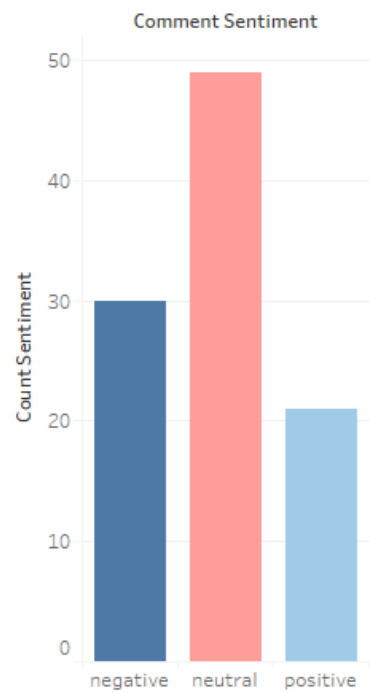
	Post_ID	Post_Content	Comment_Content	Comment_Score	Comment_Date	Provider	Cleaned_Post_Content	Cleaned_Comment_Content
0	1bk58ru	My first baby was born after a long old 36 hou...	My friend gave birth in the car in the hospita...	359	2024-03-21	r/beyondthebump	first baby bear long old hour labour drag lot ...	friend give birth car hospital parking lot rd
1	1bm36ak	Went to bed around 11 after bub had his bottle...	I mean, you did it. You hit the partner teamwo...	1629	2024-03-24	r/beyondthebump	go bed around bub bottle diaper change swaddle...	mean hit partner teamwork jackpot sleep night ...
2	1c2aazi	Just a scenario i thought was sad but also a l...	Sorry I laughed out loud at the title alone. B...	719	2024-04-12	r/beyondthebump	scenario think sad also little funny embarrass...	sorry laugh loud title alone bless
3	1brvfep	I was angry at you today. You asked me to go p...	This is.. rough. In your writing I can feel th...	799	2024-03-30	r/beyondthebump	angry today ask go pick old grandparent meet f...	rough writing feel love support husband damn f...
4	1bnfst8	Mine is that part of the reason newborns cry i...	That colic is a lazy diagnosis and synonymous ...	1252	2024-03-25	r/beyondthebump	mine part reason newborn cry hormonal one talk...	colic lazy diagnosis synonymous know probably ...



# SENTIMENT ANALYSIS

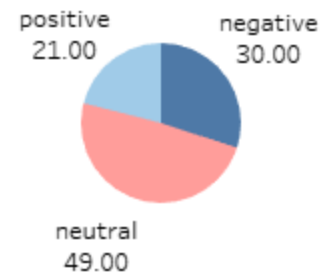
Count by Sentiment Categories

Neutral surpasses any other category of sentiment.



% of Sentiment Category

Neutral covers nearly half of the sentiment.



# TEXT SIMILARITY

```
def advice_recommender(advice_title, advice=advice_list, doc_sims=doc_sim_df):  
    advice_idx = np.where(advice == advice_title)[0][0] # find advice indx of the advice you want  
    advice_similarities = doc_sims.iloc[advice_idx].values # produce cosine similarity values  
    similar_advice_idxs = np.argsort(advice_similarities)[1:6] # output top 5 advice in descending order that are similar to chosen advice  
    similar_advice = advice[similar_advice_idxs] # produce advice titles  
    print("Based on your interest in \"{}\", I'd recommend checking out:".format(advice_title))  
    for advice in similar_advice:  
        print("- {}".format(advice))
```

```
advice_recommender('If you really really feel something is wrong, call someone.') # input advice
```

Based on your interest in "If you really really feel something is wrong, call someone.", I'd recommend checking out:

- So hard to loose it! No advice but I fell your pain. I never lost my baby weight 2nd time around
- Unrelated but holy cow! That baby has so much hair! Mine came out bald but with sideburns 🤔
- I'd say that's a strong positive.
- Sounds like your mom has an awesome sense of humour.
- I would be blasting them so hard on their FB right now.

How rude!



# QUESTION-ANSWERING SYSTEM





# NEXT STEPS

- **Enhance Analysis Techniques:**

**Sentiment Analysis:** consider using more sophisticated sentiment analysis techniques or tools for deeper insight.

**Text Similarity:** explore advanced algorithms for better accuracy in finding similar texts.

- **Data Visualization:**

Refine the visualizations in Tableau to make them more intuitive and informative.

- **Testing and Feedback:**

Conduct testing sessions with potential users to gather feedback on usability and functionality. Use the feedback to make necessary adjustments and improvements.

- **Publish and Share Findings:**

Share findings on platforms like GitHub and LinkedIn.