MENTAL HEALTH SUPPORT FOR AUTISM

2021-IMT-015 Apoorv Jain
2021-IMT-048 Astitva kamble
2021-IMT-116 Yashveer

INTRODUCTION

- Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that affects communication, social interaction, and behavior. It varies widely in how it presents in individuals.
- While clinical definitions and research offer essential frameworks, understanding how autism impacts day-to-day life requires hearing from those who live with it. Real-life experiences help uncover emotional, social, and practical aspects that often go unnoticed in academic literature. These perspectives are crucial for building empathy, designing better support systems, and spreading awareness.
- For this project, we used Quora as a source of real-world insights, collecting posts where people shared their experiences, questions, and advice related to autism.

DATA COLLECTION

1. Identified the search keywords related to the autism.
2. webscraper.io is used to scrap the data from the Quora.
3. corresponding to each keywords we collected around 100-150 posts.
4. keywords: Autism spectrum disorder, What is autism, Signs of autism, Symptoms of autism in adults, Autism diagnosis, High functioning autism, Asperger's syndrome, Autism in girls, Autistic meltdowns, Stimming behavior, Speech delay autism, Autism and sensory issues, Parenting autistic child, Autism and social anxiety, Non-verbal autism, Early signs of autism, Autism in toddlers, Neurodivergent meaning, How to support autistic people, Masking autism, ABA therapy autism, Autism and empathy, How to communicate with autistic adults, Autism and school, Autism relationships, Autism vs ADHD, Autism diagnosis in women, Living with autism, Autism burnout, Neurodiversity vs neurotypical

INITIAL SCHEMMA

author_name	Name of the person who asked the main question
question	The main question related to Autism
updated_year_main	Time when the top answer was posted .
author_main	Name of the author who provided the top answer
answer_main	Full answer text provided by the main author
upvotes_main	Number of upvotes the main answer received
views_main	Number of views on the main answer

shares_main	Number of times the main answer was shared
comments	Total number of comments under the main answer
comment_author_name_l	Name of the top comment's author
date_l	Time since the top comment was posted
comment_answer_l	Text of the top comment
upvotes_l	Number of upvotes the top comment received

DATA PREPROCESSING

1. Removed entries lacking either a question or an answer.
2. Extracted engagement metrics: upvotes, views, shares, comments, and comment upvotes from raw column data.
3. Anonymized data by removing author names.
4. Standardized time-related fields using Python's datetime library.

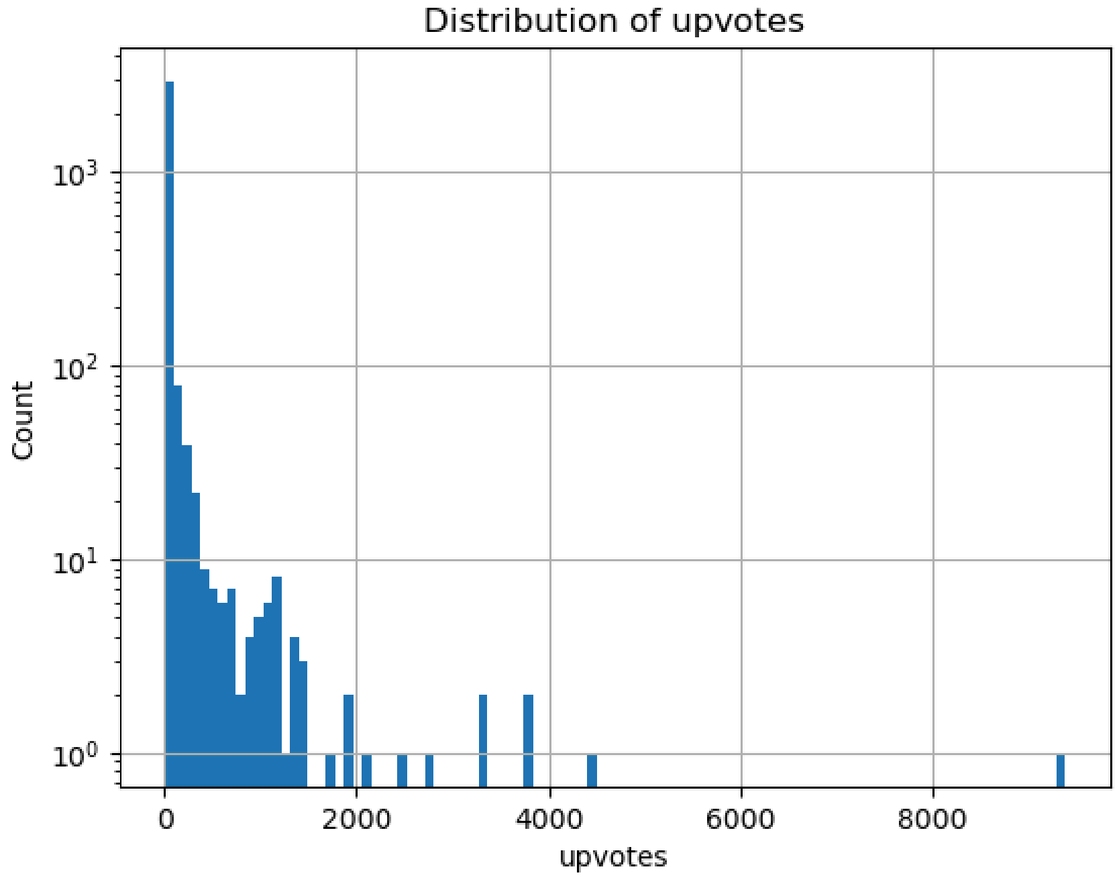
FINAL SCHEMMA

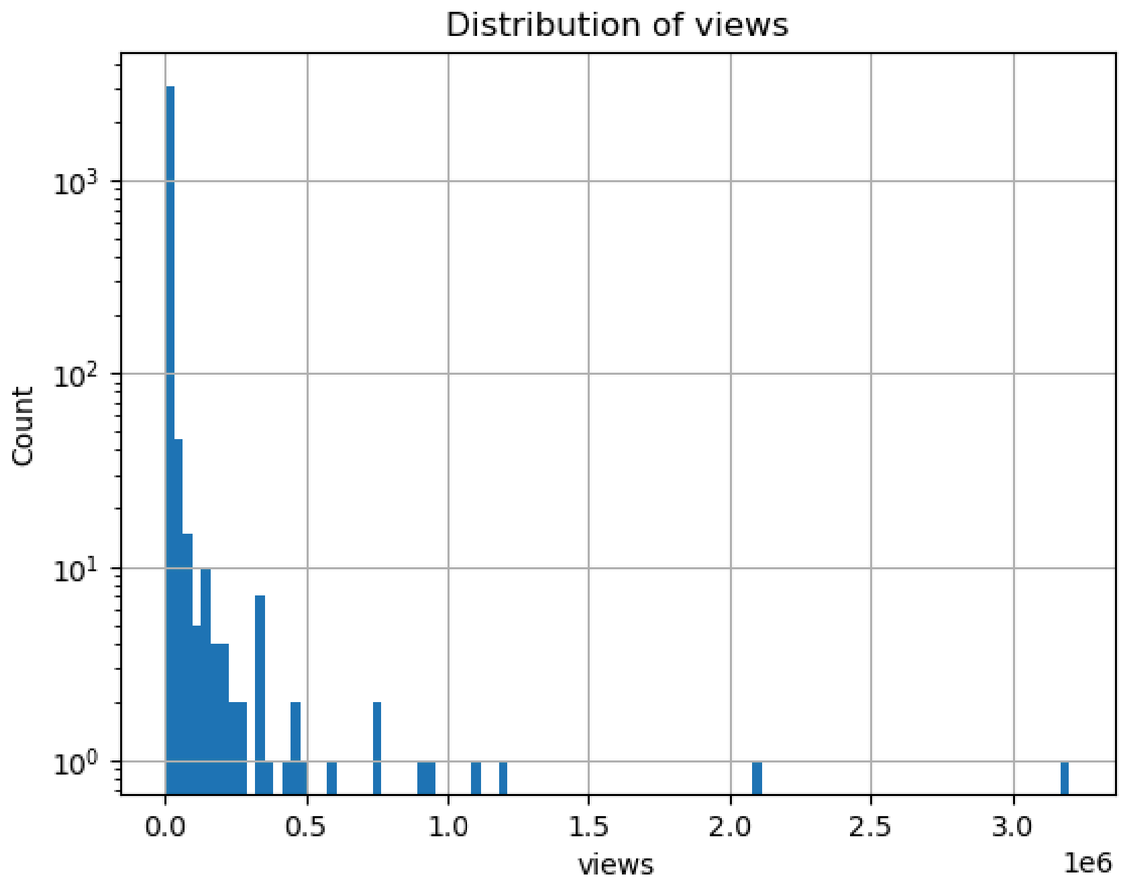
question	The main question asked
year_ago	number of years ago the question was posted
answer	The top answer to the question
upvotes	Number of upvotes to the answer
views	Number of views the answer received
shares	Number of times the answer was shared
comments	Total number of comments

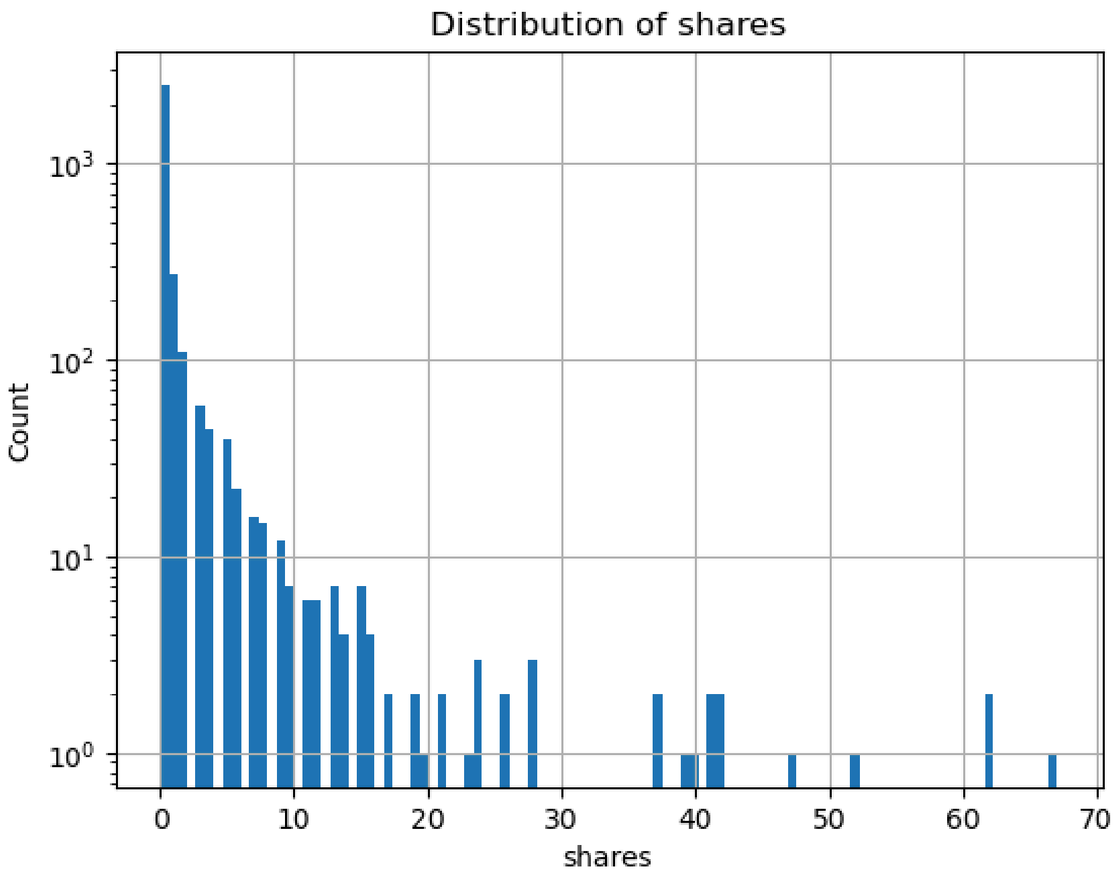
comment_years_ago	number of years ago the top comment was posted
comment_answer	Text of the top comment
comment_upvotes	Number of upvotes top comment received.

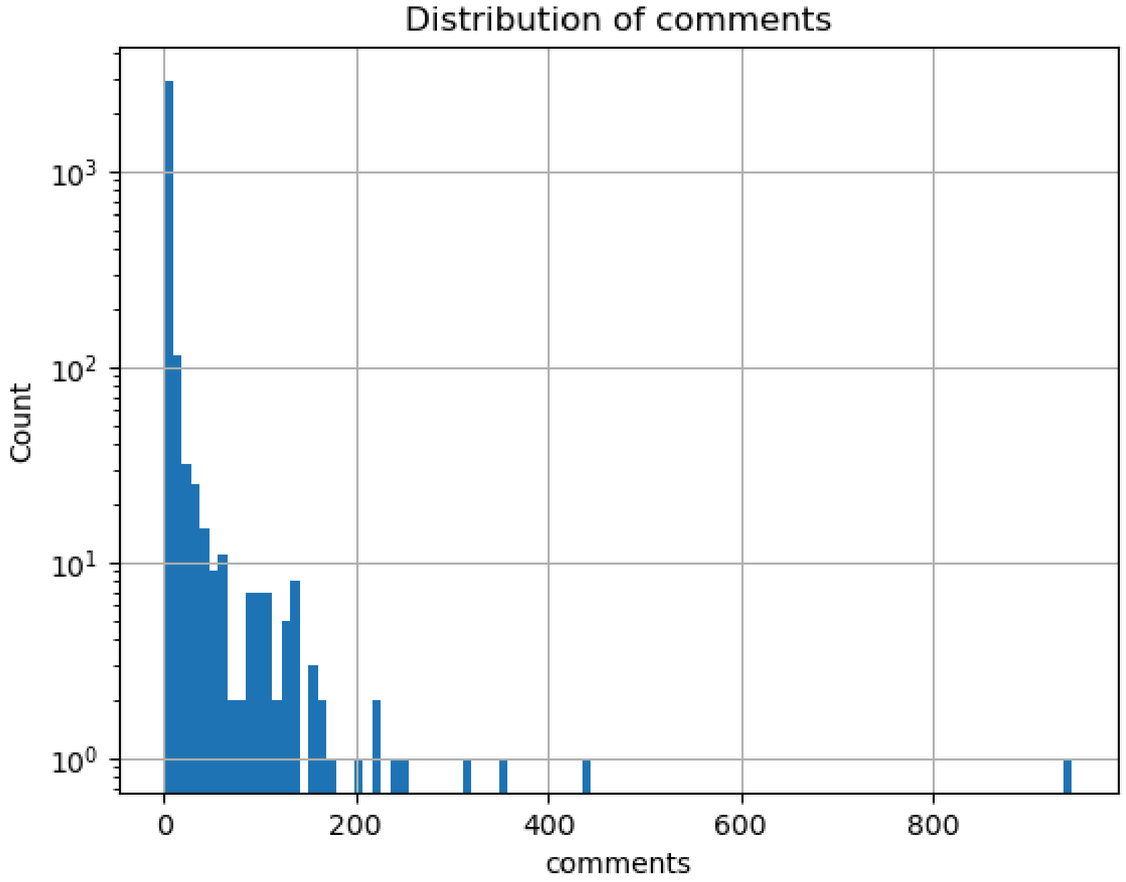
DATA STATISTICS

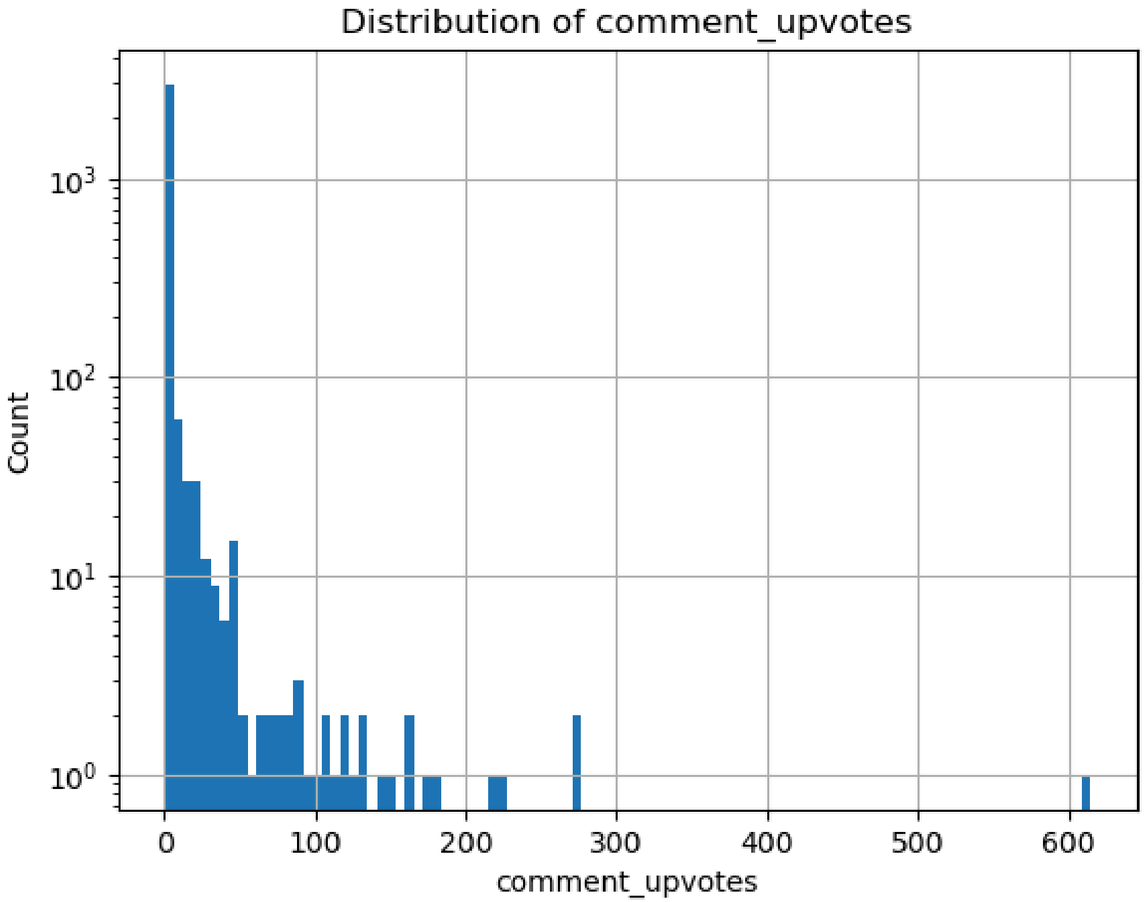
Statistic	Upvotes	Views	Shares	Comments	Comment Upvotes
Count	3164	3164	3164	3164	3164
Mean	45.63	9,595.09	0.95	5.15	2.88
Std Dev	275.22	85,469.61	3.93	26.97	18.26
Min	0	0	0	0	0
25th Percentile	0	176	0	0	0
Median (50%)	3	571.5	0	0	0
75th Percentile	12	2,200	0	2	0
Max	9369	3,200,000	67	944	614


Distribution of upvotes
Count
10^3
10^2
10^1
10^0
0
2000
4000
6000
8000
upvotes


Distribution of views
Count
10^3
10^2
10^1
10^0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
views
1e6


Distribution of shares
Count
10^3
10^2
10^1
10^0
0
10
20
30
40
50
60
70
shares


Distribution of comments
Count
10^3
10^2
10^1
10^0
0
200
400
600
800
comments



Metric	Variance
Upvotes	75,744.78
Views	7,305,054,000.00
Shares	15.43
Comments	727.21
Comment Upvotes	333.54

Topic Modeling with LDA

Input: “question” text

Vectorization

CountVectorizer with English stop-word removal

- min_df=20, max_df=0.9 to focus on terms appearing in at least 20 posts but not in >90%

Latent Dirichlet Allocation

Model: Scikit-learn’s LatentDirichletAllocation(n_components=5, learning_method=’batch’)

How it works:

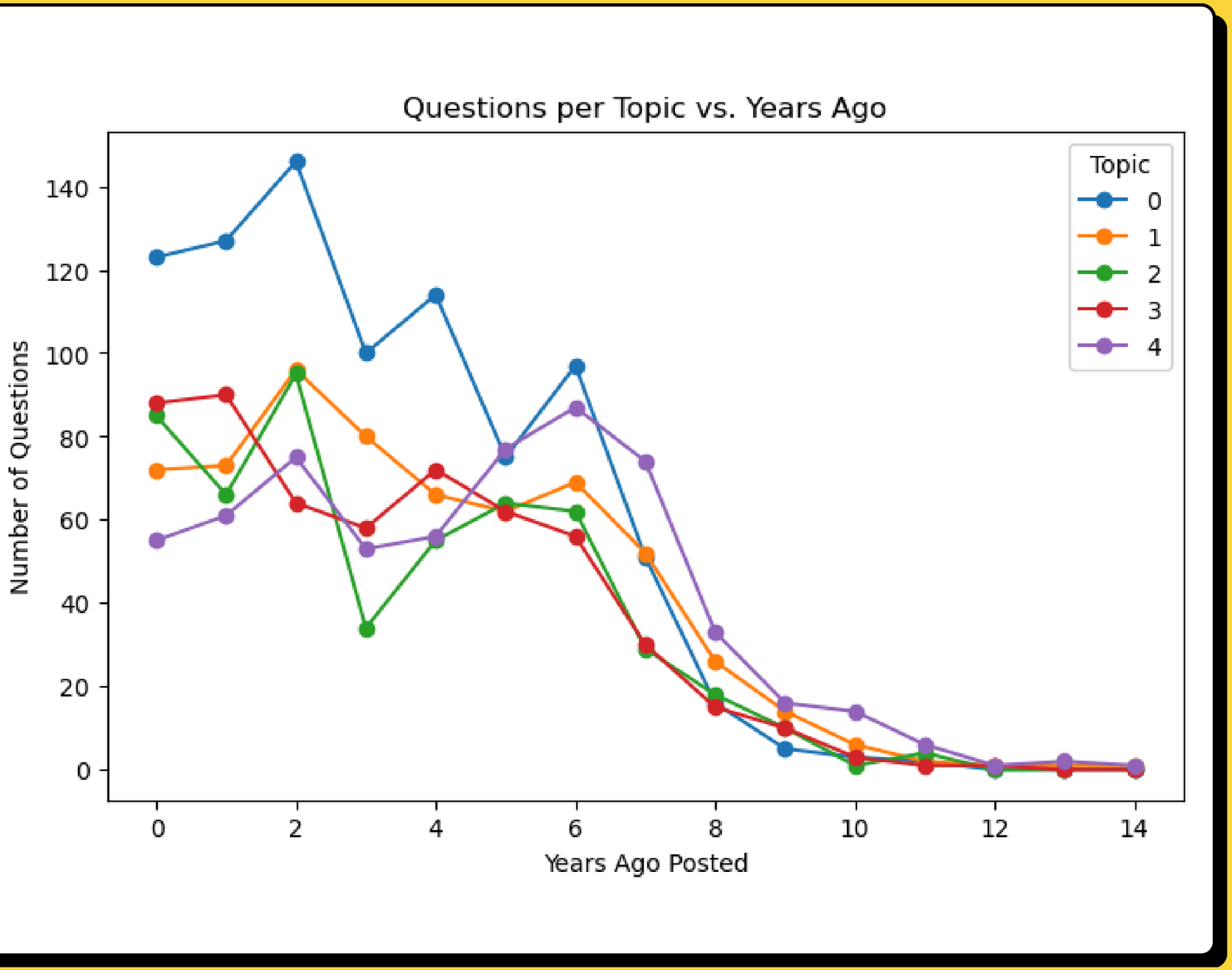
- Initialization: Randomly assign words in each document (question) to one of 5 topics.
- Expectation step: Re-estimate each document’s topic-distribution (θ) and each topic’s word-distribution (φ) given current assignments.
- Maximization step: Reassign each word to a topic based on updated θ & φ.
- Repeat until convergence.

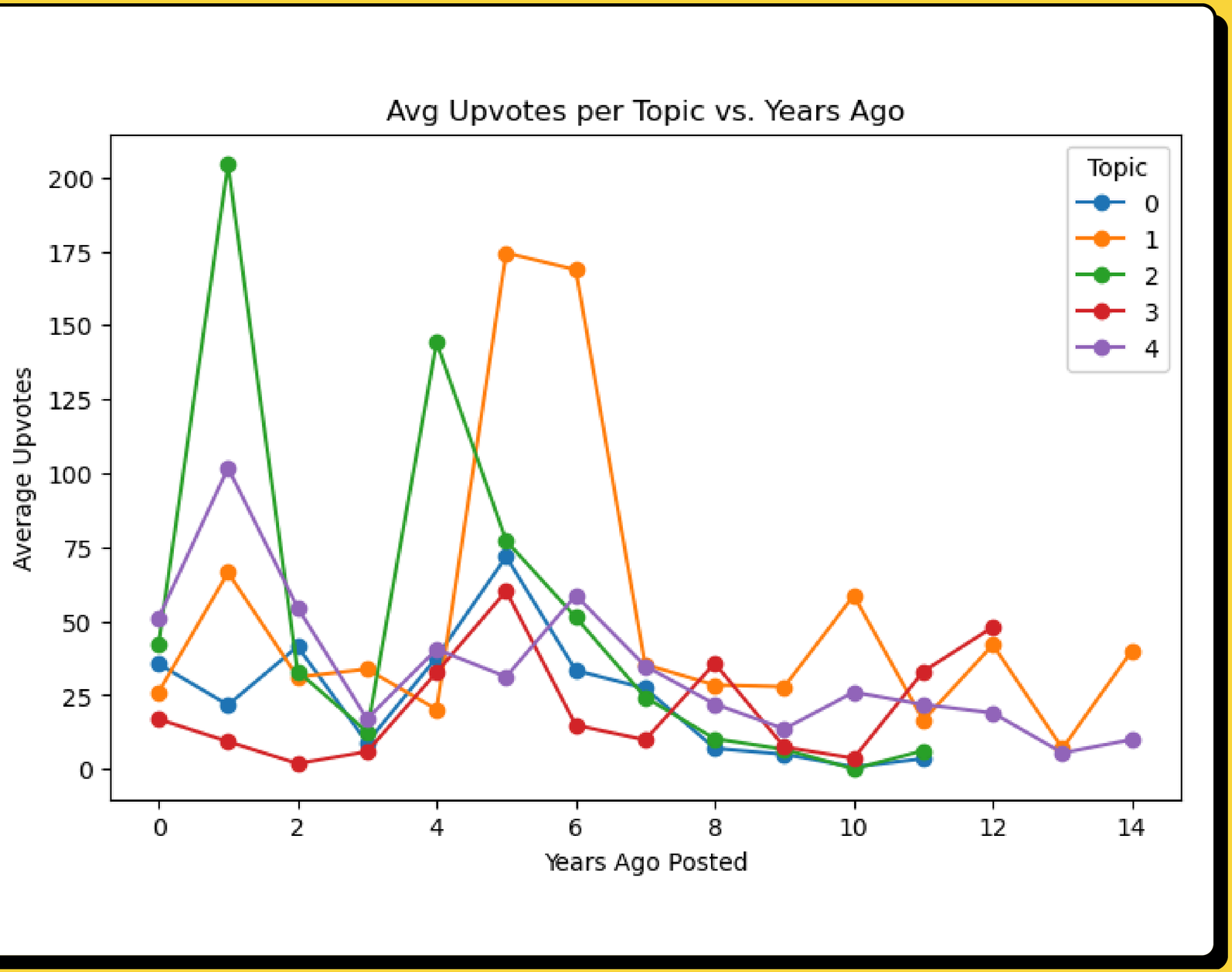
Extracted Topics using LDA

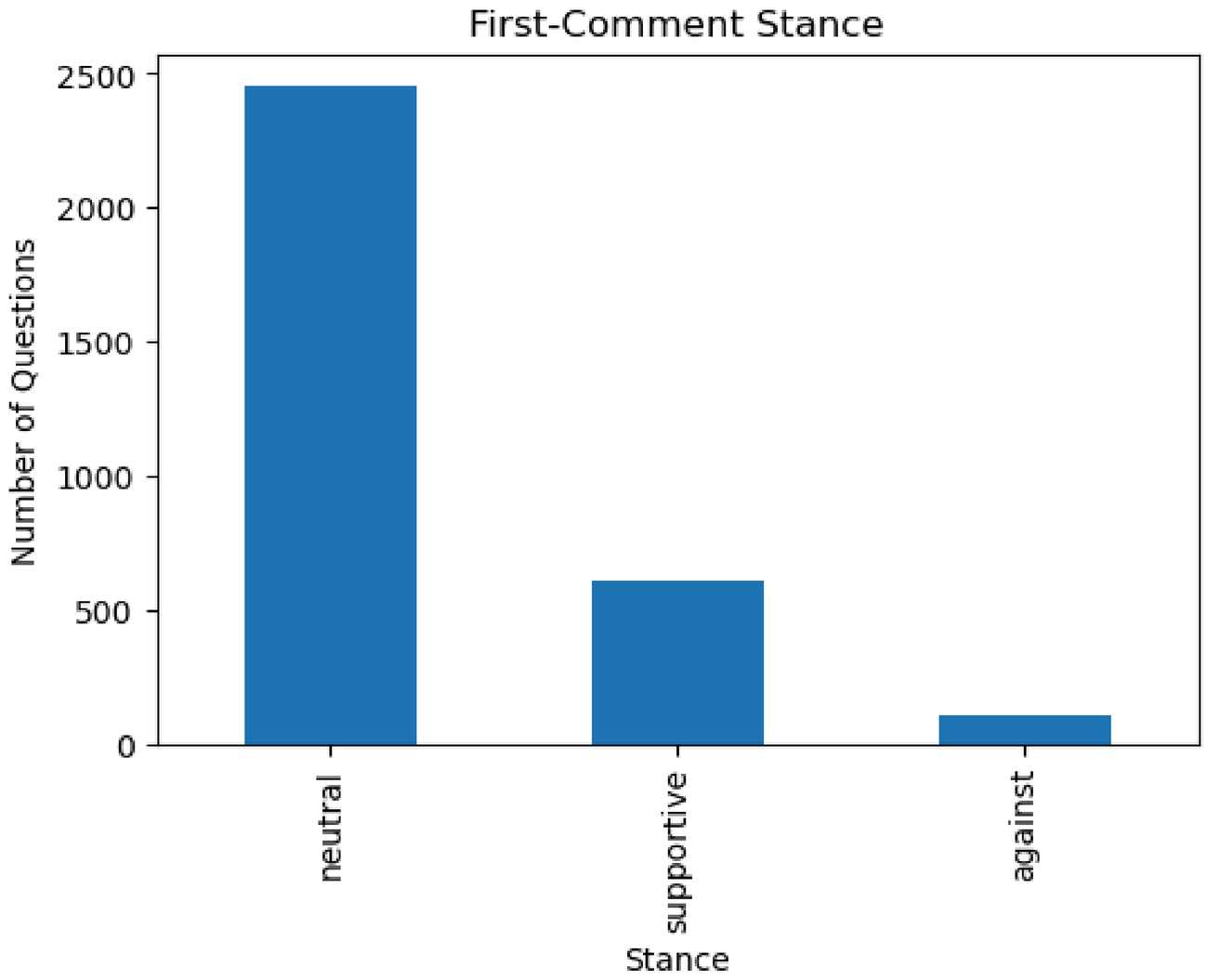
Topic 0: autism, signs, diagnosis, social, anxiety, old, speech, early, toddler, diagnosed
Topic 1: autism, people, like, adhd, high, asperger, school, functioning, syndrome, living
Topic 2: autistic, child, parents, therapy, aba, autism, does, meltdown, parent, children
Topic 3: autism, empathy, sign, sensory, issues, relationship, toddlers, individuals, lack, experience
Topic 4: autism, spectrum, disorder, functioning, high, asperger, non, syndrome, verbal, asd

By Observation, we can give unique titles to the above titles as below:

Topic 0 : Early Signs & Diagnosis
Topic 1 : Adulthood & ADHD/Asperger Functioning
Topic 2 : Parent-Led Therapy & ABA
Topic 3 : Sensory & Social Relationships
Topic 4 : High-Functioning Spectrum (ASD/Asperger)


Questions per Topic vs. Years Ago
Number of Questions
Years Ago Posted
Topic
0
1
2
3
4
0
20
40
60
80
100
120
140
0
2
4
6
8
10
12
14


Avg Upvotes per Topic vs. Years Ago
Average Upvotes
Years Ago Posted
Topic
0
1
2
3
4
200
175
150
125
100
75
50
25
0
0
2
4
6
8
10
12
14



Sentiment & Stance Analysis of First Comment

Sentiment Scoring

Tool: TextBlob's TextBlob(text).sentiment → returns

- Polarity (-1.0 to +1.0; negative vs. positive)
- Subjectivity (0.0 to 1.0; objective vs. opinionated)

Stance Classification by putting thresholds on polarity:

- Supportive if polarity > 0.1
- Against if polarity < -0.1
- Neutral otherwise

Comment stance ratios:

neutral	0.774336
supportive	0.193110
against	0.032554

FUTURE SCOPE

- ANOVA
- Tukey's HSD (Honestly Significant Difference)
- Apply dynamic and transformer-based topic models (e.g. DTM, BERTopic) for richer, time-aware themes.
- Develop supervised stance/emotion classifiers (beyond polarity) to capture nuance.
- etc....

THANK YOU