

# PROJECT RAAM



Anita Sun, Amrita, Pathak, Mia Sowder, and River Wang

# RESEARCH QUESTION

How did anxiety increase during the COVID-19 pandemic, and how did the increase vary across different groups?

## AIM

Explore how the sentiments of users in the subreddits R/Teachers, R/Nurses, and R/HealthAnxiety compares to the global anxiety rates reported before, during, and after the pandemic.

# DATA COLLECTION

## Reddit: (Feb 2019–Dec 2022)

- r/Nursing
- r/Teachers
- r/HealthAnxiety

## WHO: (Apr 2020–Apr 2022)

- 109 countries

## CDC: (May 2020–September 2024)

- State and nationwide



# DATA COLLECTION

## Workflow

- **For each subreddit, retrieve links to posts from 2019 to 2022**
- **Skip posts made of images or videos**
- **Access the links and extract texts from posts and their comments**

## Error Handling

- **Use `time.sleep()` to ensure that we send requests to less than 100 posts per minute**
- **Temporarily increase the wait time whenever the status code is 429**

```
def find_posts(n, subreddit):  
    '''  
    Find posts within a subreddit posted between January 2019 and  
    January 2022 under the "top" section.  
  
    Args:  
        n (int): total number of posts to be searched (increments by 25)  
        subreddit (str): name of the subreddit  
  
    Returns: a list of all URLs to the posts within a subreddit  
    '''
```

```
def scrapeReddit_posts(post_urls, csv_filename, max_comments=50):  
    '''  
    Scrapes details from Reddit posts and their comments.  
  
    Args:  
        post_urls (list): List of Reddit post URLs to scrape.  
        csv_filename (str): Name of the output CSV file.  
        max_comments (int): Maximum number of comments to extract per post.  
    '''
```

```
if response.status_code == 429:  
    print(f"Failed to fetch {post_url}: {response.status_code}")  
    time.sleep(30.0)  
    continue
```

## 5

	post_id	timestamp	username	post_title	post_body	score	comments_count	post_url
0	1	2021-09-24T21:54:20+00:00	aroc91	Nurses, what's your best/worst vagina story?	Mine: An immobile and incontinent but otherwis...	11.9k	458 comments	<a href="https://old.reddit.com/r/nursing/comments/putf...">https://old.reddit.com/r/nursing/comments/putf...</a>
1	2	2022-01-07T13:25:23+00:00	erinpdx7777xdpnr	Who is going to be the one to tell everyone el...	I've read the threads, spoken with friends and...	11.8k	458 comments	<a href="https://old.reddit.com/r/nursing/comments/ry7b...">https://old.reddit.com/r/nursing/comments/ry7b...</a>
2	3	2021-10-25T11:42:50+00:00	nomad_9988	Patient made me reconsider being an organ donor	Not really, but almost.\nI had an (unvaccinate...	11.5k	458 comments	<a href="https://old.reddit.com/r/nursing/comments/qfev...">https://old.reddit.com/r/nursing/comments/qfev...</a>
3	4	2021-10-02T11:26:57+00:00	igordogsockpuppet	To all you eat-your-young nurses out there, ju...	Feeling overstaffed and overworked? You've j...	10.6k	458 comments	<a href="https://old.reddit.com/r/nursing/comments/pzte...">https://old.reddit.com/r/nursing/comments/pzte...</a>
4	5	2021-11-23T19:16:50+00:00	Aloha_Snackbar357	Struggling to Not Care about my Antivax	On paper, it's not a problem. Fuck Around	10.2k	458 comments	<a href="https://old.reddit.com/r/nursing/comments/0Kt...">https://old.reddit.com/r/nursing/comments/0Kt...</a>



# DATA CLEANING

6

## Delete empty posts:

- Many comments were empty, deleted, or from community bots
- Delete all records that do not reflect the sentiments of the communities

Raw Data Table (Teaching)

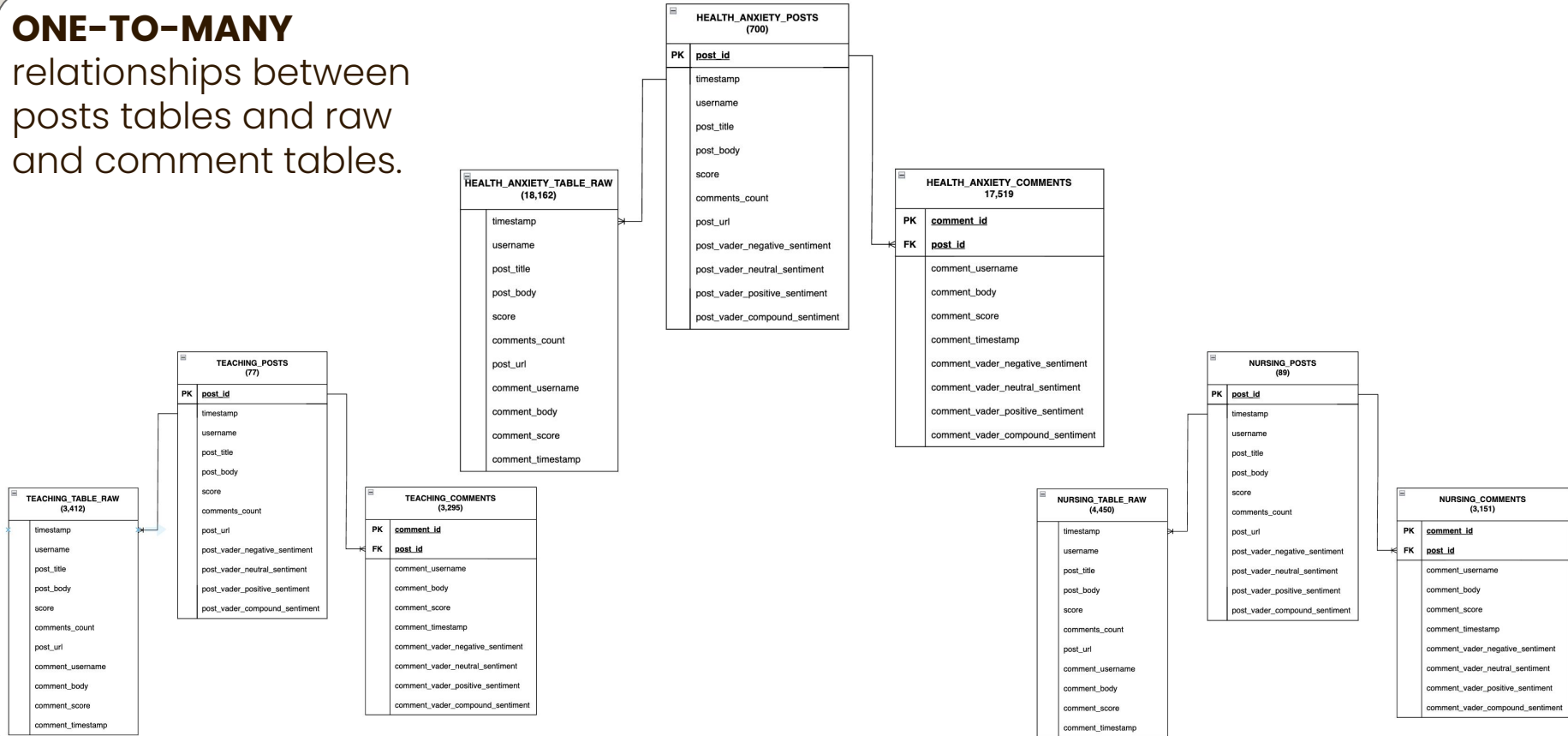
	post_title	comment_username	comment_body	comment_score
0	My school has paid for a huge abstract art ins...	AutoModerator	Welcome to /r/teaching. Please remember the ru...	Unknown
1	Why is it the teacher's responsibility to call...	AutoModerator	Welcome to /r/teaching. Please remember the ru...	Unknown
2	Dear students	AutoModerator	Welcome to /r/teaching. Please remember the ru...	Unknown
3	Science class: 'What can we infer about Mrs. T...	AutoModerator	Welcome to /r/teaching. Please remember the ru...	Unknown
4	"These 13 year olds were their own bosses for ...	AutoModerator	Welcome to /r/teaching. Please remember the ru...	Unknown
5	It begins	Anonymous	[deleted]	Unknown
6	It begins	Anonymous	[deleted]	Unknown
7	Some News!	Anonymous	[deleted]	Unknown
8	Some News!	AutoModerator	Welcome to /r/teaching. Please remember the ru...	Unknown
9	Today I was assaulted	AutoModerator	Welcome to /r/teaching. Please remember the ru...	Unknown
10	Today I was assaulted	Anonymous	[deleted]	Unknown
11	Today I was assaulted	Anonymous	[deleted]	Unknown
12	Today I was assaulted	Anonymous	No Comment	Unknown
13	Thank you to all the sweet older brothers and ...	AutoModerator	Welcome to /r/teaching. Please remember the ru...	Unknown
14	TIME TO STAND WITH TEACHERS	Anonymous	[deleted]	Unknown
15	TIME TO STAND WITH TEACHERS	AutoModerator	Welcome to /r/teaching. Please remember the ru...	Unknown
16	I saw one of my favorite kindergarteners (now ...	Anonymous	[deleted]	Unknown
17	I saw one of my favorite kindergarteners (now ...	Anonymous	[deleted]	Unknown
18	I saw one of my favorite kindergarteners (now ...	AutoModerator	Welcome to /r/teaching. Please remember the ru...	Unknown

# SQLite Architecture

7

## ONE-TO-MANY

relationships between posts tables and raw and comment tables.

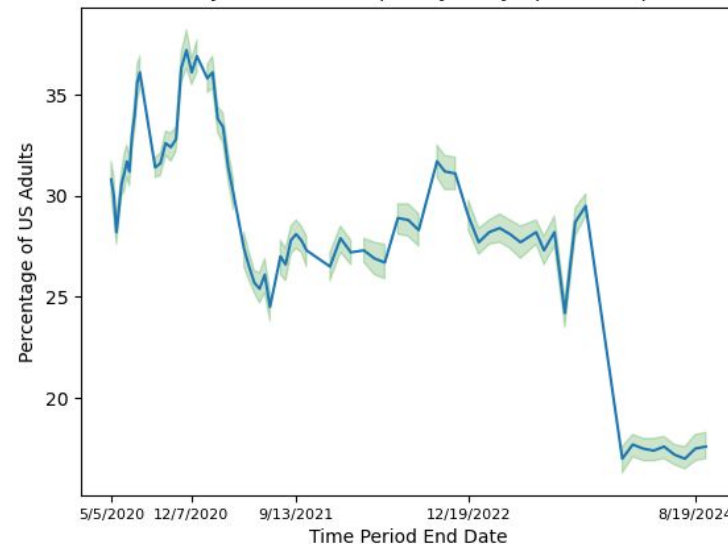


# EDA: CDC Survey Data

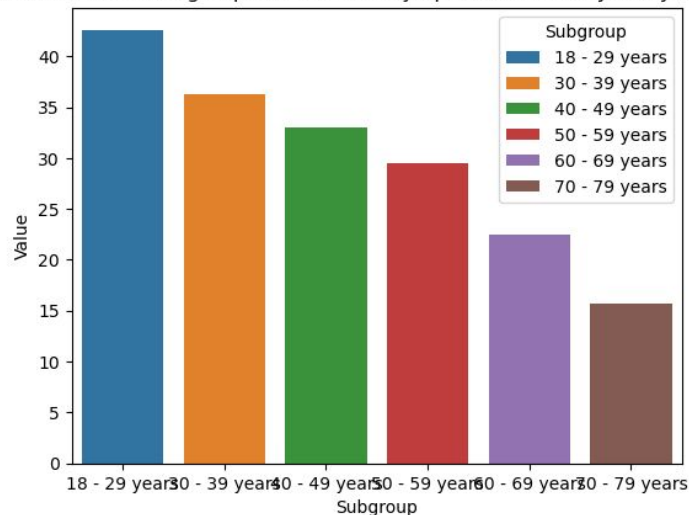
8

- Decreasing survey frequency over time
- Spike at December 2020
- Spike at December 2022
- Drop at July 2024

CDC Indicators of Anxiety Based on Frequency of Symptoms - Apr 2020 to Sept 2024



Average Percent of Each Subgroup that Exhibits Symptoms of Anxiety - May 2020 to Dec 2022



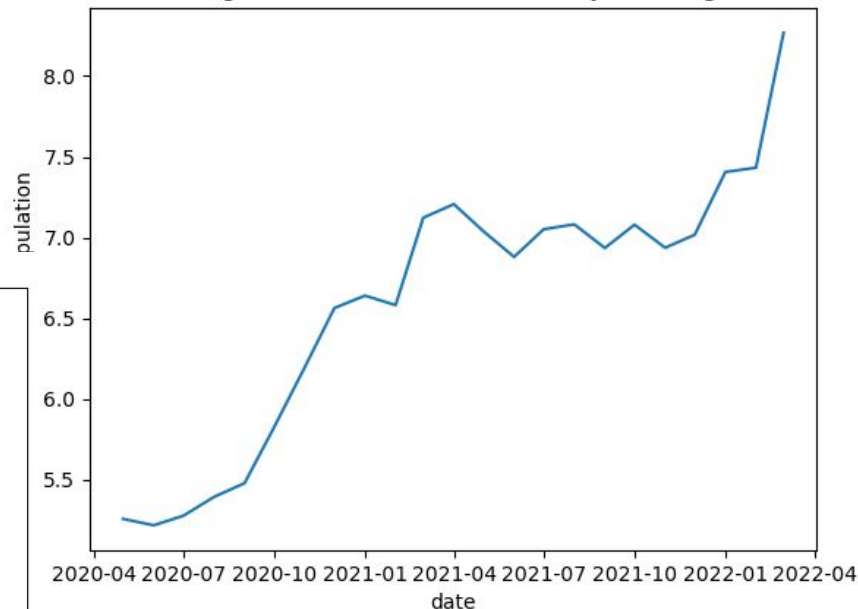


# EDA: WHO Survey Data

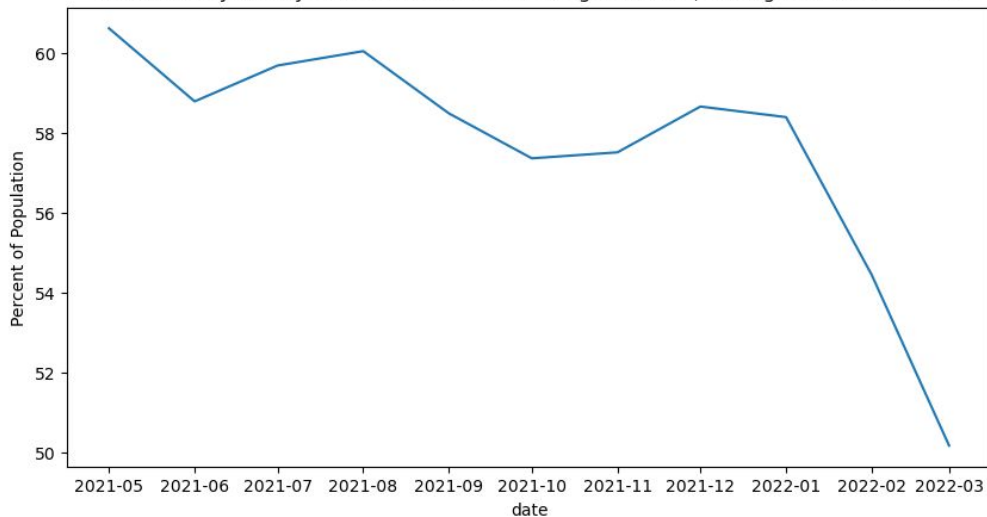
9

- Gradual increase over time
- Most Anxious Countries: Lebanon, Tunisia, Türkiye, Algeria, Jordan, Yemen, Egypt, Iraq, Azerbaijan, United Republic of Tanzania

WHO Worldwide Average - Felt Anxious for Past 7 Days, Averaged across Countries



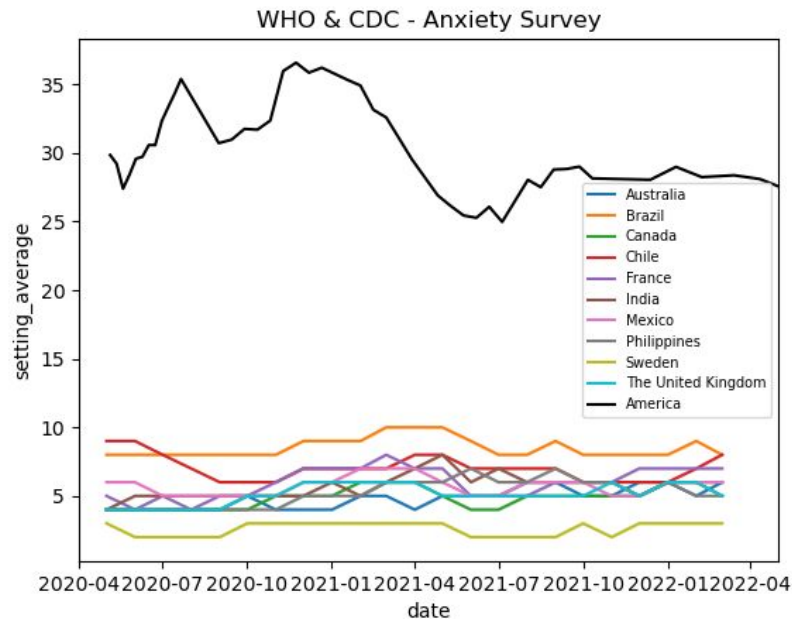
WHO Anxiety Survey - Anxious about Contracting COVID-19, Average across Countries



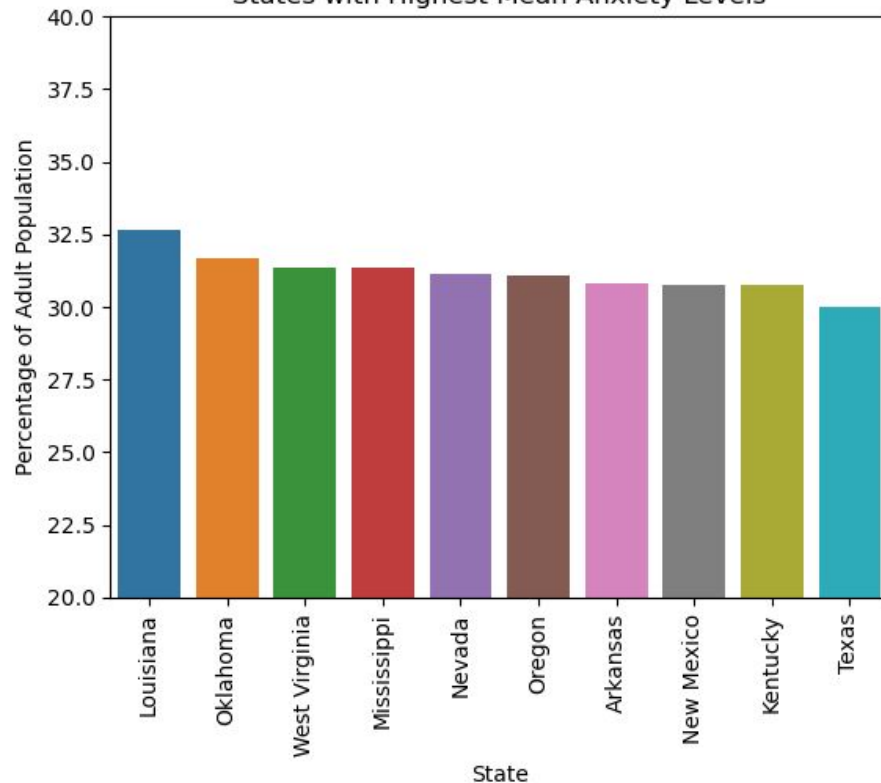
# EDA: Cont.

19

Note: the WHO survey asks if the respondent felt anxious for most or all of the past 7 days.  
The CDC survey is more broad and asks about various symptoms of anxiety



States with Highest Mean Anxiety Levels



# Data Analysis: VADER

11

**Mean Vader Compound Score (-1.00 to 1.00)**

Teaching Posts

0.217

Nursing Posts

-0.122

Health Anxiety Posts

-0.095

Teaching Comments

0.198

Nursing Comments

0.033

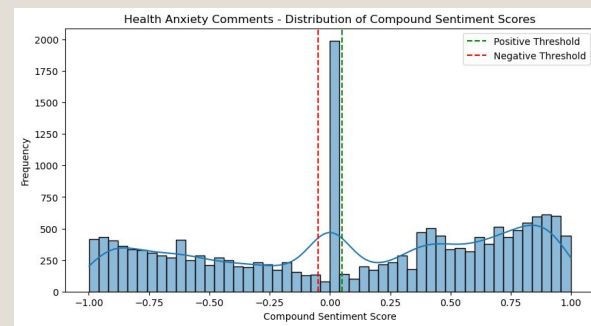
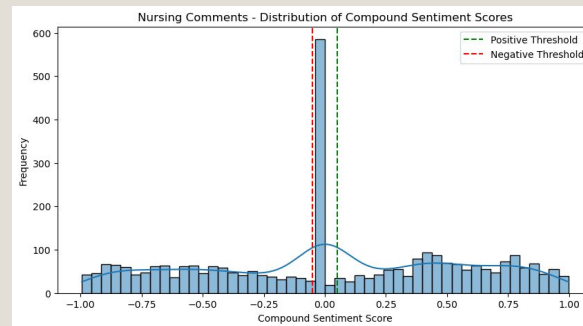
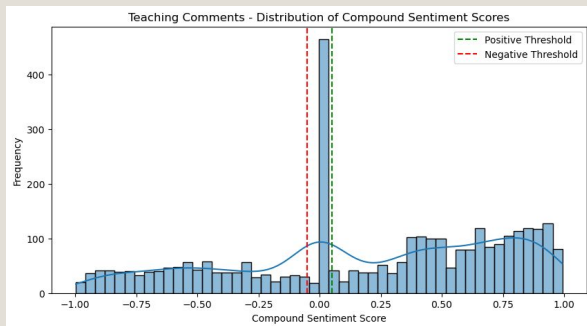
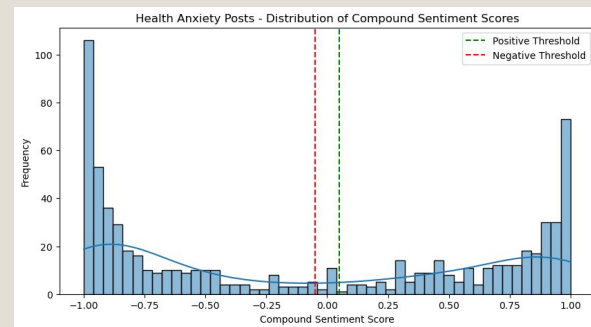
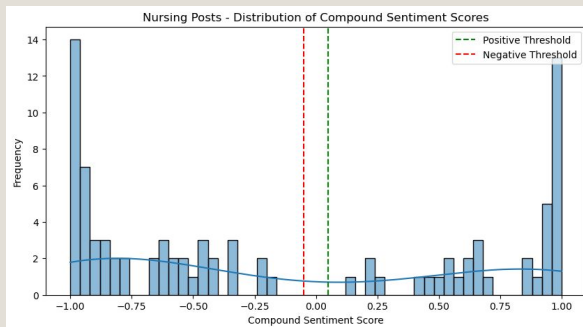
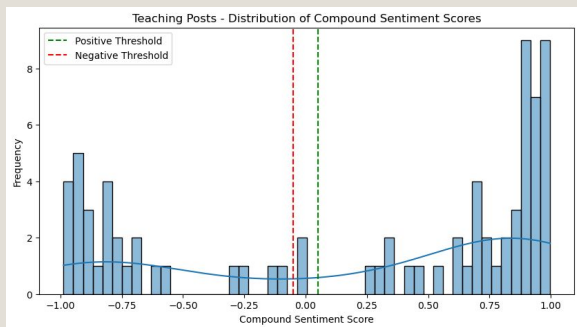
Health Anxiety Comments

0.094

# Data Analysis: VADER

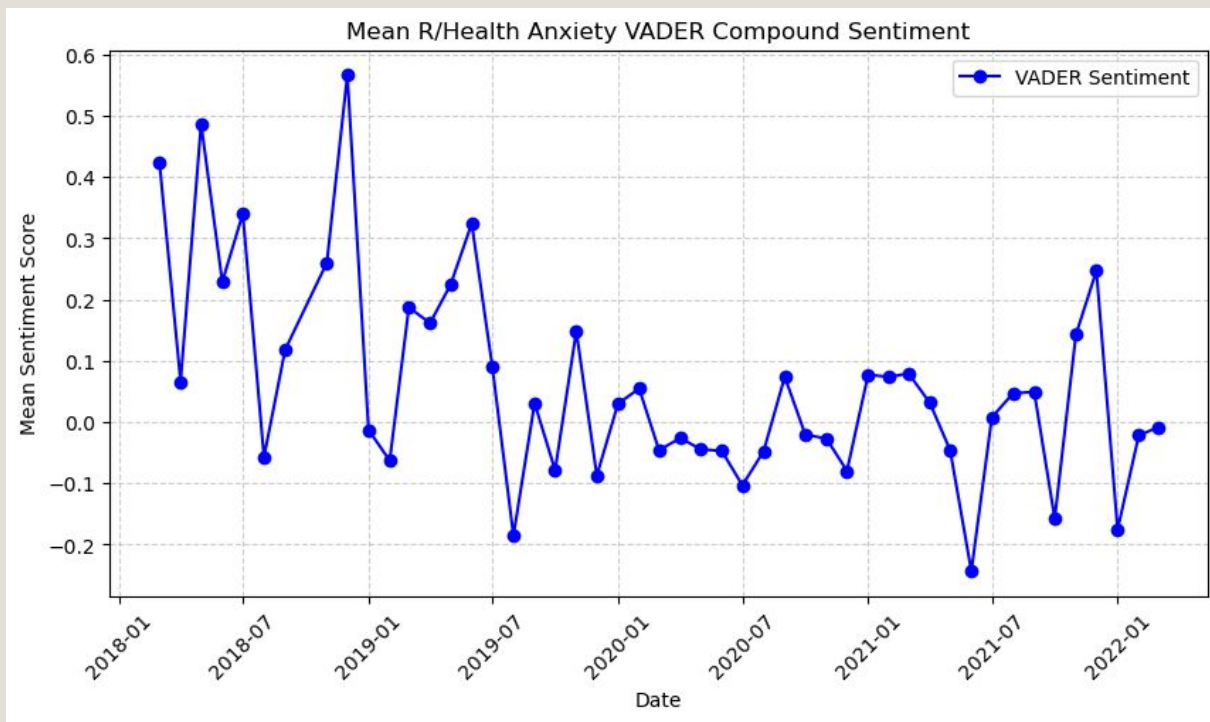
12

## Graphs - Vader Histograms



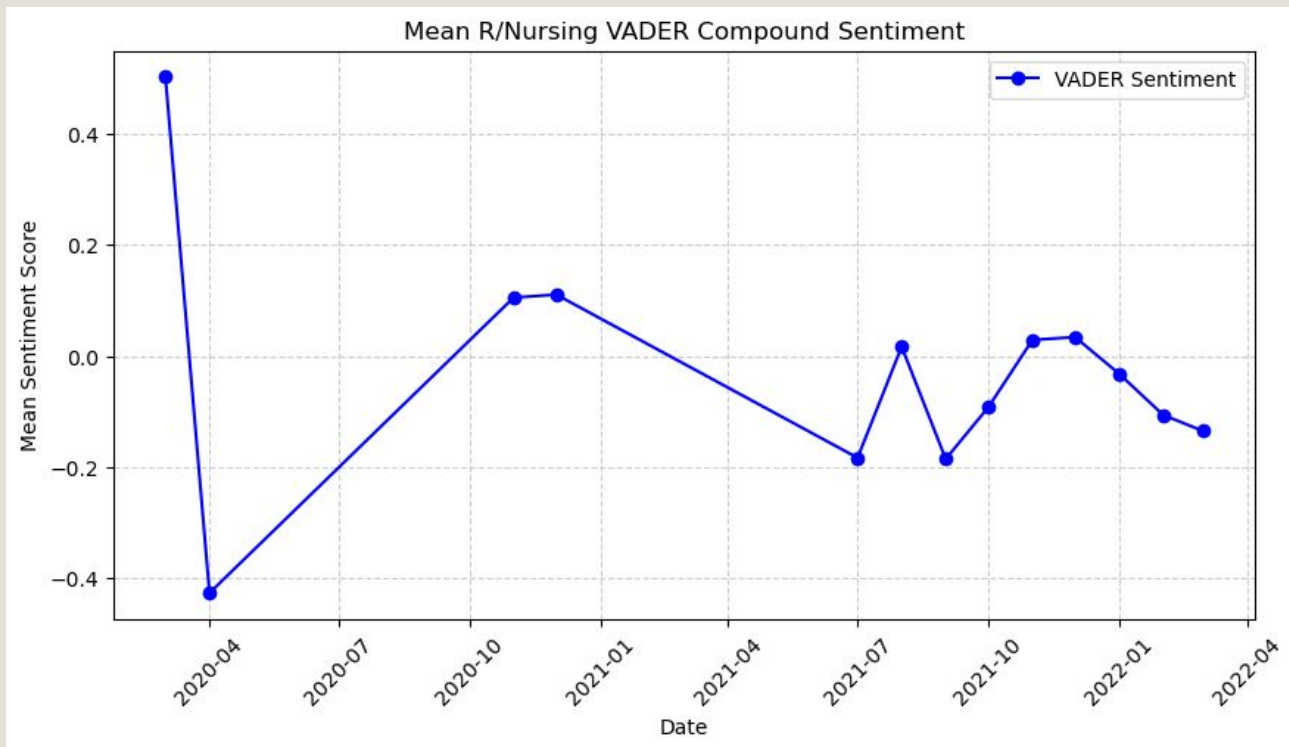
# Data Analysis: VADER

## Graphs – Vader Time Series: Mean Compound Sentiment Score by Month



# Data Analysis: VADER

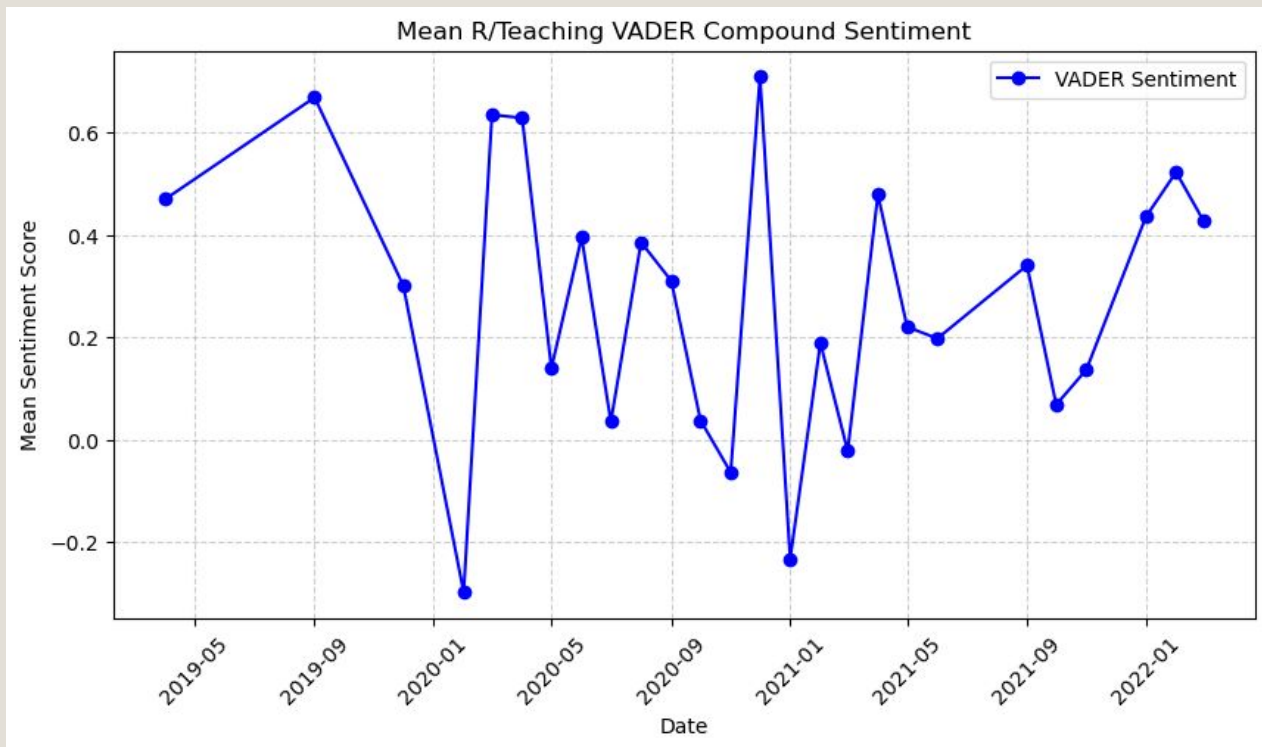
## Graphs – Vader Time Series: Mean Compound Sentiment Score by Month





# Data Analysis: VADER

## Graphs - Vader Time Series: Mean Compound Sentiment Score by Month



# Data Analysis: LIWC

16

## Mean LIWC Anxiety Score

Health Anxiety

1.980

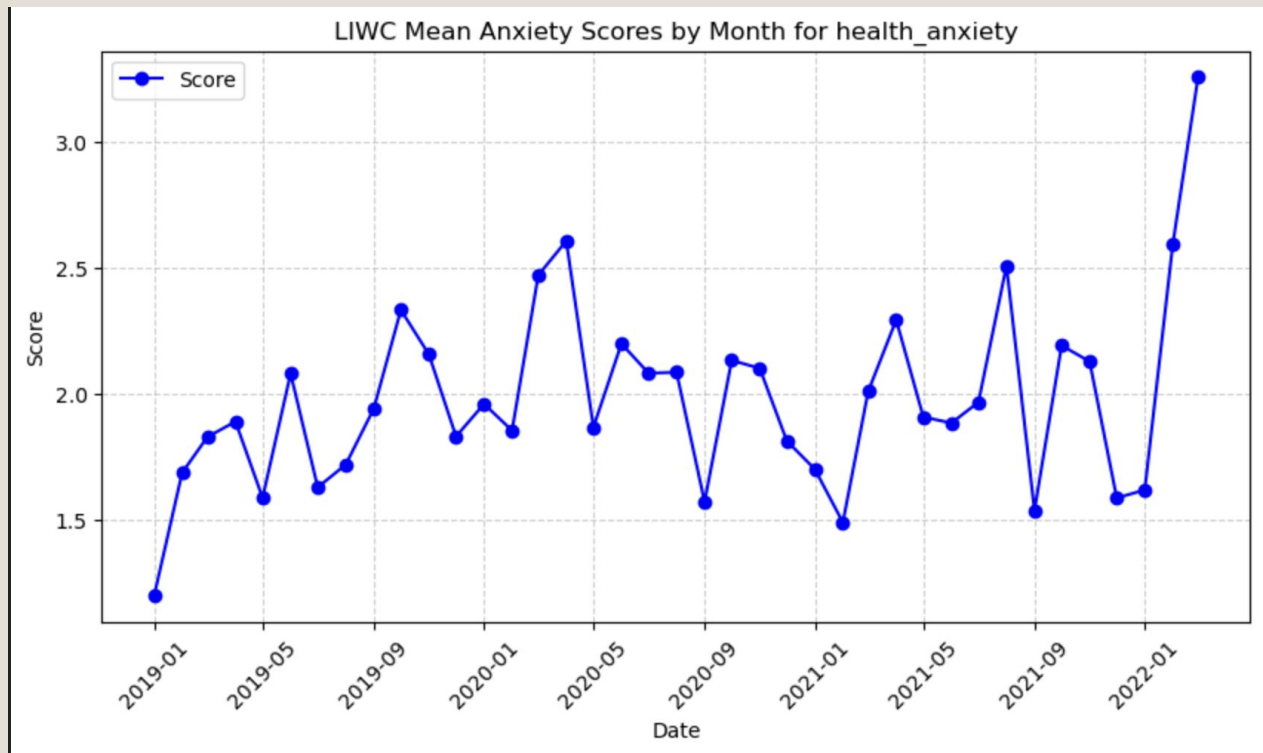
Nursing

0.222

Teaching

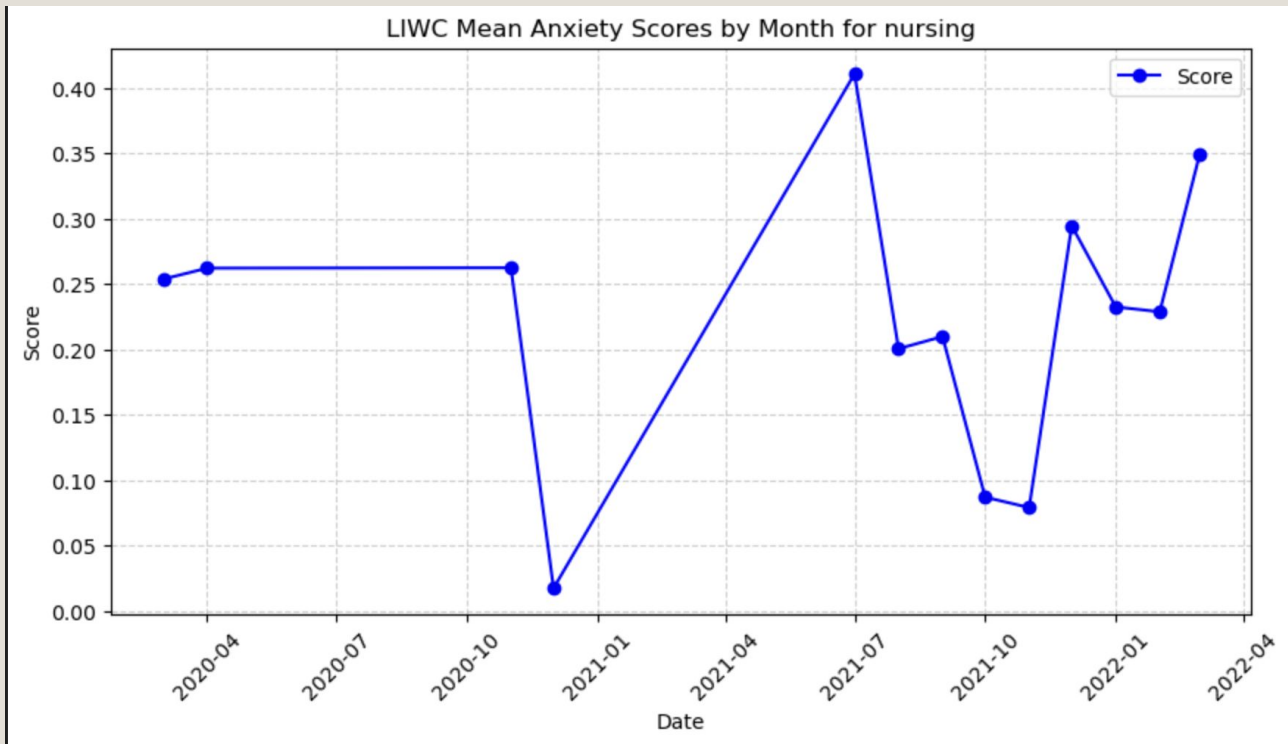
0.112

# Data Analysis: LIWC



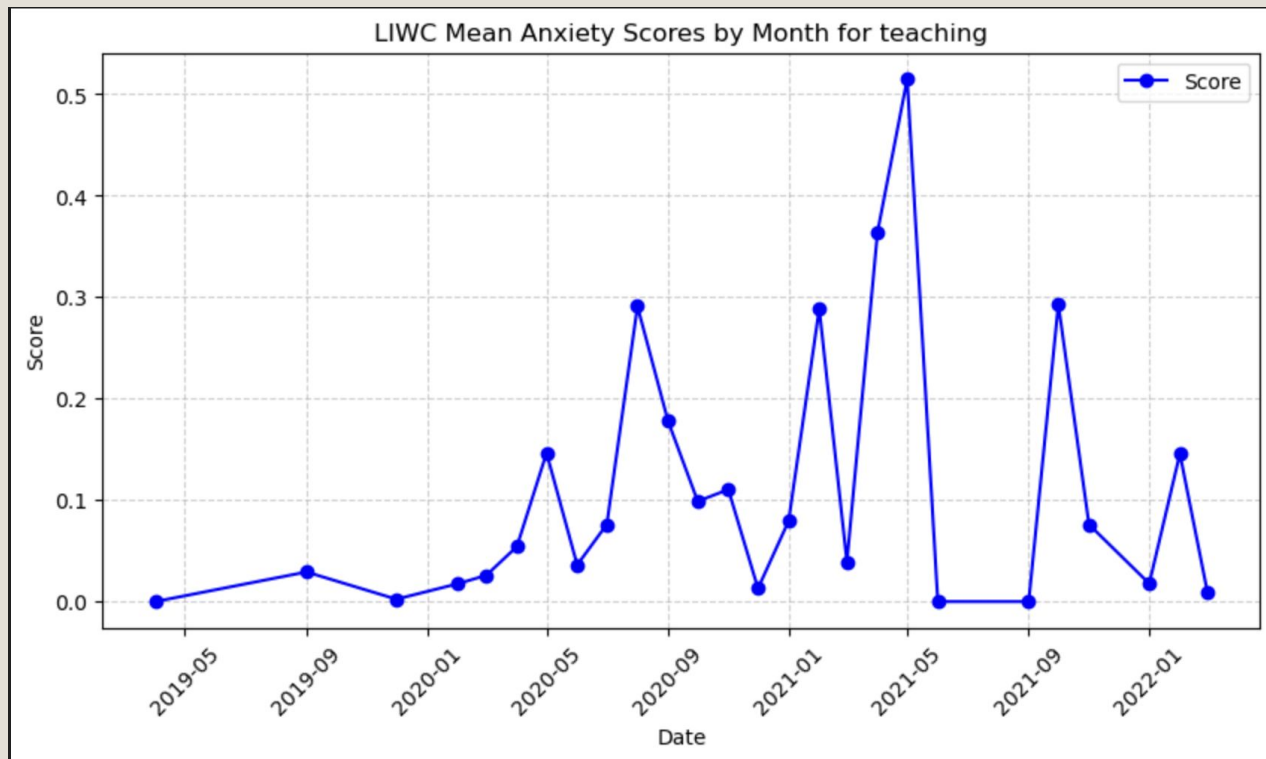
- A large increase in anxiety in early 2020
- COVID might lead to growing concerns about health

# Data Analysis: LIWC



- Anxiety scores were high for 2020
- Anxious sentiment peaked in July 2021
- Generally consistent with the health anxiety subreddit

# Data Analysis: LIWC



- Anxiety started to increase in 2020
- A surge in anxiety scores around May 2021

# Conclusions

- Observable increase in anxiety in the first half of 2020
- The level of anxiety on the nursing and teaching subreddits do not seem to be necessarily higher than the health anxiety subreddit
- The nursing subreddit can display a higher level of anxiety than teaching
- Anxiety reached a high point in the middle of 2021 across all 3 subreddits
- During the first half of 2021, while anxiety rates declined in the general American population, they increased for teachers and nurses



# Take Home Points

## Workflow

Could have started with data collection earlier to address potential issues

The team worked well and each of us made good contribution to different aspects of the project

## Room for Improvement

Data Collection: Reddit scraping was not able to gather all intended dates.

Visualizations/Analysis: The scraped data limitations meant it would be less beneficial to combine comments with their posts - something we would aim for in future research.

# Thank You

---