# Dreams in the time of COVID-19 - Apoorv Jha

## Introduction

On the 29th of March this year, I woke up- confused and shaken, from the weird dream I had just had. I brushed it off thinking it's maybe due to the tv show I was bingeing. Then the next day, it happened again, and then every day after that. This recurrence made me start wondering why I was having so many weird and vivid dreams lately. More importantly, was I the only one having weird dreams during this quarantine? I texted my friends and asked my brother about this, and they all affirmed that they had also been dreaming more often nowadays, and having really weird dreams. This corroboration intrigued me and I investigated further. A simple Google trend search for 'why am I having vivid dreams lately' showed that the volume of web searches for the phrase quadrupled just last week in the US and showed a +650% rise worldwide in the past 90 days. What really piqued my interest, as I venture down this rabbit hole, is the fact that everybody dreams![1] Considering how far science and technology have progressed, dreams still remain a mystery and we don't yet have a full understanding of them. I hope to find and visualize data and present my analysis about this topic.

## Sources of Data:

- Google search trends and relative search interest using Google Trends
- U.S. COVID-19 data:
  - https://data.humdata.org/dataset/nyt-covid-19-data
  - https://en.wikipedia.org/wiki/2020_coronavirus_pandemic_in_the_United_States
- Internet-using population data (https://www.internetworldstats.com/unitedstates.htm#WV)
- Link to code (https://github.com/microgel/DATA)

## Methodology

Data Collection:

- Collected the Google search trends for the past 90 days for "vivid dreams" in the US
- Collected the Google search trends for the past 90 days for "vivid dreams" in each US state where data was available
- Collected number of confirmed COVID-19 cases for each of those US states
- Collected data for the number of internet users in each state as total population for the state is not indicative of total internet users, making an accurate scaling for relative interest. (Entire population does not have access to internet and therefore would not typically use Google search)

Data Processing:

- Scaled the number of web searches with the number of internet users for that state and graphed the relative interest over time on a scale of 0 to 100
- Interest over time: Numbers represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means insufficient data for this term.
- Cleaned the data as there were 1721 null values on the number of confirmed cases axis for Number of confirmed COVID-19 cases
- Removed New York and New Jersey from the initial graph as the number of cases for these two states was much higher than other states and made the visualization difficult to read

[1] https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Understanding-Sleep

- Interest by subregion: See in which location your term was most popular during the specified time frame. Values are calculated on a scale from 0 to 100, where 100 is the location with the most popularity as a fraction of total searches in that location, a value of 50 indicates a location which is half as popular. A value of 0 indicates a location where there was not enough data for this term. In a relative scaling scenario, a higher value means a higher proportion of all queries, not a higher absolute query count. So a tiny state where 80% of the queries are for "vivid dreams" will get twice the score of a giant country where only 40% of the queries are for "vivid dreams"

Trend Analysis and Visualization:

- Created a separate graph with data from New York and New Jersey using the same procedure
- Cleaned the data for 87 null values on the number of confirmed cases axis for Number of confirmed COVID-19 cases
- Then I organized the data and scaled by internet-using population in each of those states to visualize relative interest by state on a scale of 0 to 100
- This scaling is especially helpful to understand and analyze this data as through a relative scale, we can understand how much the search trends have gone up in proportion to their internet-using population and number of queries in the given period
- In other words, we can understand how the search for “vivid dreams” has increased relative to the number of searches for it comparable to other states since the first recorded COVID-19 case in the U.S.
- Then, I graphed the relative interest for “vivid dreams” and “weird dreams” in the U.S. in the past 12 months

Sentiment Analysis on Tweets:

- I wrote a python program utilizing Twitter’s tweepy API to fetch tweets (in English) with keywords #vividdreams #weirddreams #quarandreams [2] #quarandream #coviddreams from the U.S.
- Unfortunately, student access for tweepy only allows you to fetch tweets from the past 7 days (13th April 2020 to 21st April 2020) - so the tweets collected are limited to the past 7 days
- Next, I trained a basic sentiment analysis model using TextBlob and NLTK to determine the polarity for tweets based on emoticons - happy or sad, and certain keywords (example: nightmare, bad dream, happy, good memories, laughing, sad, crying, scared, etc.)
- While the scope of the data was limited to 7 days worth of tweets with these keywords, the results are relevant to show how many of these tweets had positive, neutral, or negative association with them
- Next, I cleaned the data for retweets and duplicates and removed tweets that had links using tweet-preprocessor
- I visualized the number of tweets from each sentiment category as a pie chart
- Note: The reason I have used a relative interest for these visualizations is that it helps us understand better how the search interest has been affected in comparison to regular trends

Visualizations:

1. Relative search Interest for "vivid dreams" and "weird dreams" in the U.S. in the past 12 months
2. Relative search interest for "vivid dreams" with number of confirmed COVID-19 cases in each U.S. state (except New York, New Jersey) Since first recorded case in the U.S.
3. Relative search interest for "vivid dreams" with number of confirmed COVID-19 cases in New York and New Jersey since first recorded case
4. Relative interest for "vivid dreams" by U.S. state since first recorded case [Map]
5. Sentiment Analysis of Tweets collected (Positive, Neutral, or Negative)

[2] https://www.independent.co.uk/life-style/quarandreams-coronavirus-dreams-lockdown-isolation-nightmares

Observations:

- RSI for “vivid dreams” on Google in the U.S. in (1) has shot up from 44 on 19th Jan. 2020 to a RSI of 100 on 16th Apr. 2020
- RSI for “weird dreams” on Google in the U.S. in (1) has shot up from 26 on 19th Jan. 2020 to a RSI of 50 on 16th Apr. 2020
- 34 out of 42 states visualized in (2) had RSI for “vivid dreams” of 100 when they had confirmed COVID-19 cases within that state
- All 33 of these states hit RSI of 100 within 44 days since the first case confirmed case in that state; 50% of those 34 states had RSI of 100 when they had more than 1000 cases in that state
- Both, New York and New Jersey had RSI of 100 when they had more than 50000 and within 42 days since first confirmed case in that state
- Rhode Island, West Virginia, and South Dakota had highest RSI by sub-region; They have an an internet-using population percentage of 81.0%, 70.5%, and 72.9% respectively
- 306 out of 490 tweets collected with relevant keywords were unique; 360% more tweets had negative sentiments than tweets with positive sentiments

Analysis:

I find these observations very interesting. They indicate how search trends for “vivid dreams” and “weird dreams” have shot up since the COVID-19 pandemic. While this data is just limited to Google search trends and tweets from the U.S., 88.3% of Americans use Google as their primary search engine[3]. This suggests the likelihood of these observations being truly indicative of the internet-using population in the U.S. to be dreaming more often/vividly/weirdly.

The tweets are really interesting to analyze as a theme emerges that these dreams that people are having have more negative sentiments than positive. Two possible theories emerge here. First being that alcohol use is tied to dreaming more often/vividly and having nightmares[4]. Alcohol sales in the U.S. have gone up by at least 55% since the first case has been recorded in the U.S.[5] The other theory is suggested by a research paper from Sleep Research Society titled ‘A Systematic Change in Dreams after 9/11/01[6],’ reported “a highly significant increase in central image intensity, as well as central image proportion.” This is interesting as this suggests how events like 9/11 changed the way people dreamed for a time, making their dreams more intense and memorable in the days after the attacks. It seems feasible that the coronavirus pandemic, which has personally impacted almost everyone on Earth, could have a similar impact. A recent dream survey[7] conducted by Deirdre Leigh Barrett, an assistant professor of psychology at Harvard Medical School, seems to confirm[8] that the incidence of vivid dreams has increased as the virus has spread around the world, corroborating the data and observations I have made. Furthermore, dreamjournal.net - world’s largest archive of dream journals[9] - reports a spike in journal entries and most dreams in the past month to be revolving around common themes of the pandemic, quarantine, bugs, resisting authority, and missing friends and family.

Both these theories seem interesting to me. This analysis opens up a window to how underlying emotional distress can have adverse effects. Considering the scale of this pandemic, it is crucial to think about how an entire population of people having nightmares can have unforeseen consequences in the future[10]. On the other hand, how quarantining has increased alcohol consumption and can impact the mental and physical health of these people. More data with variables such as alcohol users reporting dream frequency can help us understand more about this issue.

What seems to be really fascinating to me is that the COVID-19 pandemic presents an unprecedented and extraordinary opportunity to conduct dream research and delve deeper into understanding this succession of images, ideas, emotions, and sensations that usually occur involuntarily in the mind during our sleep!

---

[3] https://gs.statcounter.com/search-engine-market-share/all/united-states-of-america
[4] https://health.clevelandclinic.org/why-you-should-limit-alcohol-before-bed-for-better-sleep/
[5] https://news.usc.edu/168549/covid-19-alcohol-sales-abuse-stress-relapse-usc-experts/
[6] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2225570/
[7] https://www.surveymonkey.com/r/B8S75CN
[8] https://www.latimes.com/lifestyle/story/2020-04-07/coronavirus-quarantine-dreams
[9] http://www.dreamjournal.net/
[10] https://time.com/5821896/coronavirus-nightmares-dreams/