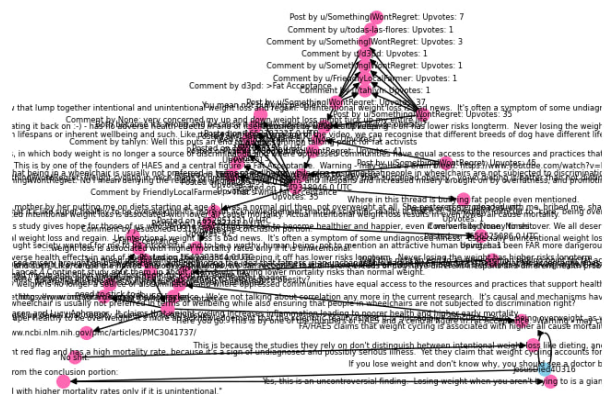


Final Project Proposal: Post and Comment Search for r/Obesity

Data Source(s):

1. The [r/obesity subreddit](#) via the Reddit API
 - a. The r/obesity subreddit is “a subreddit in which the medical and social impact of obesity is discussed. Peer reviewed articles preferred”.
 - b. Get at least 500 posts or comments with:
 - i. Author
 - ii. Date posted
 - iii. Content/body
 - iv. Number of upvotes
 - v. Times shared
 - vi. Url to peer-reviewed article (posts only)

Reddit Comment Tree - r/obesity

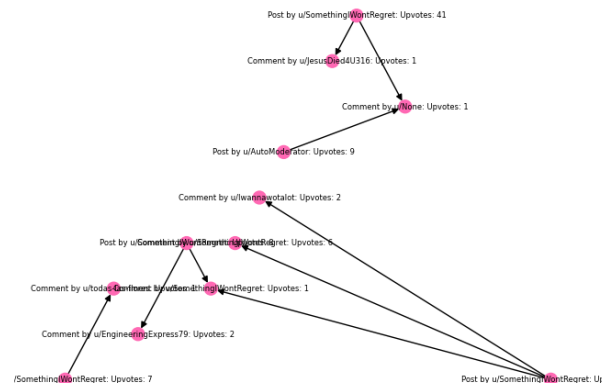


(examples of trees with data from the top 5 r/Obesity posts)

2. [Nutrition, Physical Activity, and Obesity - Behavioral Risk Factor Surveillance System](#) Dataset

- a. 88629 rows of data containing percentage of obesity per area/region by year and other conditions

Reddit Comment Tree - r/obesity



Processing and Visualization

The trees shown above are examples of how I would like to present/visualize the data. I would like to perform a sentiment analysis on the comments to gauge public sentiment, and use a tree structure to visualize trends in sentiment change over time. This may identify correlations between obesity in the U.S. and the attitudes of those who left comments under r/Obesity posts.