

507 Final Project

Meghana Bhat (meghanab)

Title: Exploratory Data Analysis: COVID Misinformation Networks on Twitter

Project Code

GitHub Repository: <https://github.com/meghana-RB/TwitterNetworkAnalysis/tree/master>

README at: <https://github.com/meghana-RB/TwitterNetworkAnalysis/blob/master/README.md>

To start the program, `TwitterProject.py` is run. Interaction with the program is through a CLI. Users can choose one of the following options to view the Analysis and Visualization associated:

1. Word Frequencies from Users' Descriptions - a bar graph with the frequencies of terms in the 'Description' of Twitter Users. The Users include both authors of the tweets and ten of their followers
2. Word Frequencies from Tweet Content - a bar graph of the frequencies of terms in the text of the tweets fetched by querying for a list of keywords
3. Retweet Statistics - a stripplot of the number of retweets of the list of tweets collated from querying for a list of keywords, grouped by the keywords
4. Primary Node Statistics - a bargraph of the Tweet Count, and a side-by-side bar graph of the Followers and Following of the Primary Nodes
5. Network Graph - a visualization of the Primary and Secondary nodes with undirected edges connecting the nodes

Details of the Packages used and the initial setup to run the program are specified in the README file.

Data sources

1. Twitter API (v2RecentSearchPython) with OAuth2 - to access the recent tweets containing the keywords in the list ["scamdemic", "stop complying", "no masks", "#covidhoax", "do not comply", "#novaccineforme"] - (ten tweets from each category)

Sample of data collected:

```
{
  "author_id": "3031590441",
  "id": "1469821638782996480",
  "text": "#covidisflu #covidhoax #covidfarce \n\nI, for one, want to know what the government are really hiding \ud83e\udd14 https://t.co/ctOyBZBIPH"},
  {
    "author_id": "116534235",
    "id": "1469803658112081925",
    "text": "@geertwilderspvv @Gerard1945X Um... Wat dacht je van de #CovidHoax?\n\nEen ergere oorlogsverklaring dan die is er niet. Wereldwijd. En Fleur maar zeuren om de gifboosters.\n\nHoud toch op.\n\n#ControlledOpposition\n\n#AbolishTheState"},
    {
      "author_id": "1004817203609423873",
      "id": "1469783884590202889",
      "text": "@julianHjessop @spectator The #CovidHoax and the #ClimateChangeHoax are following exactly the same leftist playbook: Make outrageous predictions about a harm that will occur if the leftists aren't granted unprecedented powers, and when those predictions fail to come even close to being true, make more."},

```

Purpose:

- To get the Text of the tweets to find frequently used words
 - To get the IDs of the Authors to form a Network
 - Number of records collected 6 Keywords * 10 Tweets per keyword ~ 60 Tweets
 - Fields: author_id, tweet_id, text
2. Twitter API (v2UserLookupPython) with OAuth2 - to access the authors, follower counts, location, public metrics, descriptions, and IDs of the tweets retrieved from the first API call

```
"data": {
  "username": "KeithCullen9",
  "id": "888635633350004738",
  "description": "Student of politics, philosophy, psychology and stats. Tin foil hat wearing defender of civil liberties. PhD. from the school of hard knocks. Proudly Prole.",
  "created_at": "2017-07-22T05:43:39.000Z",
  "url": "",
  "public_metrics": {
    "followers_count": 479,
    "following_count": 213,
    "tweet_count": 2168,
    "listed_count": 0
  },

```

Purpose:

- To get the description of the author's Twitter Bios
 - To check the popularity of the author and the frequency of activity on Twitter
 - Number of records collected: ~60 records corresponding to each Tweet
 - Fields: username, id, description, follower_count, following_count, tweet_count
3. Twitter API (v2FollowersLookupPython) with OAuth2 - to access the followers with IDs, follower counts, location, public metrics, descriptions of the followers of the authors from the IDs retrieved in the preceding API call

```
{
  "data": [
    {
      "public_metrics": {
        "followers_count": 478,
        "following_count": 812,
        "tweet_count": 11087,
        "listed_count": 0
      },
      "created_at":

```

```
'2019-10-26T19:08:59.000Z', 'name': 'David Ritter', 'id':
'1188170505574113280', 'url': '', 'description': 'Christian, father of 4,
conservative patriot, NRA, do-it your selfer, happily married, Harley
owner, retired.', 'username': 'DavidRi65339915'}},
```

Purpose:

- To get the description of the followers' Twitter Bios
 - To form a Network of the authors and their followers
 - Number of records collected: 60 authors * 10 followers ~600
 - Fields: username, id, description, follower_count, following_count, tweet_count
4. Twitter API (v2RetweetedByPython) with OAuth2 to access the retweet metrics like retweeted by usernames, descriptions, likes, IDs of the tweets retrieved from the first API call

```
{'data': [{ 'created_at': '2017-09-01T16:11:25.000Z', 'username':
'steph_judice', 'id': '903651515495313408', 'description': 'Proud
Louisiana Cajun . Love America, the flag, the military and law enforcement
and other conservatives. We need Pres. Trump. MAGA. Geaux LSU Tigers!',
'name': 'Steph Judice'}, { 'created_at': '2015-02-11T01:16:38.000Z',
'username': 'fanceyeagledance', 'id': '3029181299', 'description':
'LBGFJB', 'name': 'Fancy Eagle Dance'}, { 'created_at':
'2018-06-20T18:58:57.000Z', 'username': 'Media__Ferret', 'id':
'1009510935613132800', 'description': 'Writer; Social Media; No DMs,
unless work related. #BackTheBlue', 'name': 'Media Ferret'},
{ 'created_at': '2011-10-03T16:07:27.000Z', 'username': 'JohannaDenton',
'id': '384391011', 'description': 'Drive a big truck. Writer. Married to
soulmate. Love to read and spoil my chihuahuas. NRA member. MAGA', 'name':
'Trucker Jo'}, ..'meta': {'result_count': 19}}}
```

Purpose:

- To check the number of retweets of each of the tweets
- Number of records: ~60
- Fields: retweeted_by(username, id, descriptions), number_of_retweets

The Data for the API calls was cached in [twitter_cache.json](#)

Data Structures:

README with Description:

<https://github.com/meghana-RB/TwitterNetworkAnalysis/blob/master/README.txt>

- Python file that **constructs** a Graph of the network of Nodes from the data collected **and** **reads** the json of your graphs or trees.-
https://github.com/meghana-RB/TwitterNetworkAnalysis/blob/master/Graph_Json.py

Graph.py

```
class Node:

    connectedTo = []
    def __init__(self, key, name, url=None, description=None, public_metrics=None, isPrimary = False, associated_keyword = None):
        self.connectedTo = []
        self.id = key
        self.name = name
        self.url = url
        self.description = description
        self.public_metrics = public_metrics
        self.isPrimary = isPrimary
        self.associated_keyword = associated_keyword

    def setConnectedTo(self, connectionsList):
        self.connectedTo = connectionsList

    def addNeighbor(self, id_neighbor):
        self.connectedTo.append(id_neighbor)

    def getId(self):
        return self.id

    def getWeight(self, nbr):
        return self.connectedTo[nbr]

    def getConnections(self):
        return self.connectedTo

    def getId(self):
        return self.id
```

Twitter_Users_Description.csv

Tweet_Description.csv

ID	Description
8.89E+17	@tatiana
4.51E+09	@overit_miss Nothing. If not for liberals and fake news I would not know there was a ScamDemic? https://t.co/Vcr6KcY7Q1
1.12E+18	@dtite67 @clairlemon @szabosolicitors Uh-huh, the scamdemic? You sound just like a Herman Cain who hasn't paid the consequences yet of being stupid.
1.54E+09	@Corleone_Roller Thatâ€™s exactly how the commie pricks want you. Donâ€™t be a #Doomer or a #FluFlake and get caught up in their #FearPorn driven #SCAMdemic #LiveYourLife itâ€™s better than their BS.
1.42E+18	@LegendaryEnergy No one has been more wrong about the scamdemic than @scottadamssays
3.31E+08	If health of people were important to government, theyâ€™d be handing out vitamin D, C, zinc & ivermectin like candy. Fortunately the rabble appears to becoming wise to the scamdemic. https://t.co/9y7psB0zdd
1.47E+18	@JohnRuddick2 SCAMDEMIC
2.94E+09	@KendrickPerkins ðŸ””,ðŸ””-ðŸ”” NOT U TOO!!ðŸ”” #DoYourHomework #Scamdemic shows who donâ€™t #Go2College ðŸ””-â€”
1.34E+18	ðŸ””#Stop Complying https://t.co/RtVhavNWof
7.38E+17	@Corleo
1.24E+18	@EssexPR @Charlot62189447 Stop complying.
9.46E+17	Grow a brain, stop complying with this bullshit https://t.co/plv0nMsmye
1.37E+09	Stop complying, and watch how quickly this crap disappears https://t.co/ywwCfu575t
3.06E+08	@bilohash @ColinDMello its a bit late but welcome to our group. Im vaxxed by the way too. and Ive been warning about this for over a year. the govt knows its losing the narrative. we have to keep fighting, stop complying
1.24E+18	@CanteringClark Stop fing complying
1.37E+18	@KevinKileyCA It will end when people stop complying.
5.45E+08	@rtroth2946 @MarcPDumont Wrong. Let's stop blaming the 12% not vaxed. And who's not complying? Masks & other restrictions r reality we all live with. Its a virus, it's going to spread. And it spreads equally in vaxed
7.65E+08	@KHWallar @patrickdoucette I think its gonna get way worst before it gets better,but we still need to hold the line and to Stop Complying !!

Retweet_Data.json

```
[{"meta": {"result_count": 0}}, {"data": [{"created_at": "2017-09-01T16:11:25.000Z", "username": "steph_judice", "id": "903651515495313408", "description": "Proud Louisiana Cajun . Love America, the flag, the military and law enforcement and other conservatives. We need Pres. Trump. MAGA. Geaux LSU Tigers!", "name": "Steph Judice"}, {"created_at": "2015-02-11T01:16:38.000Z", "username": "fanceyeagledance", "id": "3029181299", "description": "LGBFJB", "name": "Fancy Eagle Dance"}, {"created_at": "2018-06-20T18:58:57.000Z", "username": "Media_Ferret", "id": "1009510935613132800", "description": "Writer; Social Media; No DMs, unless work related. #BackTheBlue", "name": "Media Ferret"}, {"created_at": "2011-10-03T16:07:27.000Z", "username": "JohannaDenton", "id": "384391011", "description": "Drive a big truck. Writer. Married to soulmate. Love to read and spoil my chihuahuas. NRA member. MAGA", "name": "Trucker Jo"}, {"created_at": "2011-03-05T18:42:25.000Z", "username": "jwogadyjr", "id": "261327445", "description": "In life, it is important to know when to stop arguing with people and simply let them be wrong!! - #Patriot #MAGA #KAG #FATHER #AllLivesMatter #ProFace #IFBAP", "name": "Joe O'Grady #PatriotParty"}, {"created_at": "2021-02-03T04:36:49.000Z", "username": "DefenderofMank1", "id": "1356823788172763137", "description": "Christian. Man. American. Latin0. Warrior. Truth dealer. \nJustice seeker. Pure Blood. Proud member of the LGBFJB community.", "name": "Defender of Mankind"}, {"created_at": "2020-12-20T20:20:34.000Z", "username": "Rlohr5", "id": "1340753148399542275", "description": "Never argue with stupid people, they will drag you down to their level, and then beat you with experience! #NotMyPresident\ud83d\ude0e #TermLimitsForCongress\ud83d\udc00 #VoterID\ud83d\uddef3", "name": "A-Clear-View\ud83d\udc41"}, {"created_at": "2019-01-31T19:03:54.000Z", "username": "Lisa50267387", "id": "1091049449890164736", "description": "Compacency is not an option\n\n#SaveAmerica #ConstitutionalConservative", "name": "Lisa"}, {"created_at": "2017-01-11T16:20:27.000Z", "username": "RogerYostJr1", "id": "819217419075981312", "description": "Happily married love trout fishing, hunting, family. grandfather on 1/15/19, BAMA fan 50yr.sROLL TIDE follow #@codeofvets, Back the Blue, profile pic is my wife", "name": "Roger Yost Jr"}, {"created_at": "2018-01-20T06:49:52.000Z", "username": "KathyOC63783924", "id": "954606889236496384", "description": "LET'S GO BRANDON!!!", "name": "Kathy - Just a Patriotic American"}, {"created_at": "2011-12-16T02:41:29.000Z", "username": "DSteampipe", "id": "438004703", "description": "LGBFJB,I follow Back / Fare elections a must ,low energy cost is prosperity . Global equalization & aggressive climate control is regressive., LGBFJB", "name": "DavGoo"}, {"created_at": "2020-03-03T15:27:00.000Z", "username": "PatFish62304572", "id": "1234862609758900224", "description": "\ud83c\uddfa\ud83c\uddef8 \ud83c\uddee\ud83c\uddef1 Married
```

Primary_Nodes.json


```
[{"data": {"username": "KeithCullen9", "id": "8886356335004738", "description": "Student of politics, philosophy, psychology and stats. Tin foil hat wearing defender of civil liberties. PhD. from the school of hard knocks. Proudly Prole.", "created_at": "2017-07-22T05:43:39.000Z", "url": "", "public_metrics": {"followers_count": 479, "following_count": 213, "tweet_count": 2168, "listed_count": 0}, "name": "\u2b50\ufe0f My freedom is not your choice \ud83c\udef38", "data": {"name": "POWERLIFT555", "url": "", "id": "4509862400", "description": "MAGA..PRESIDENT TRUMP IS THE MAN!!! I WILL BLOCK ALL LIBS\ud83d\ude0e\ud83d\udecd", "public_metrics": {"followers_count": 8867, "following_count": 8938, "tweet_count": 362879, "listed_count": 7}, "username": "Kegan545", "created_at": "2015-12-09T18:22:25.000Z"}}, {"data": {"url": "", "description": "", "id": "1124164582065336321", "public_metrics": {"followers_count": 4, "following_count": 96, "tweet_count": 941, "listed_count": 0}, "username": "Dan_From_Reddit", "created_at": "2019-05-03T04:11:37.000Z", "name": "Dan"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"id": "1422183968057794568", "name": "HeroCallPoker", "created_at": "2021-08-02T13:14:44.000Z", "url": "", "username": "HeroCallPoker", "public_metrics": {"followers_count": 8, "following_count": 82, "tweet_count": 1748, "listed_count": 0}, "description": ""}, {"data": {"created_at": "2011-07-07T22:48:22.000Z", "url": "", "name": "ComradeZim #BLM #BDS #M4A #UBI #XSD #FreePalestine", "id": "331278040", "public_metrics": {"followers_count": 157, "following_count": 341, "tweet_count": 3460, "listed_count": 1}, "description": "Eco-Socialist, sustainable geek, pinhead, substacker #PTF - Pass The Fentanyl. Don\u0219t take fentanyl. Expressing frustration of our f\u0219d up society.", "username": "ZimInSeattle"}}, {"data": {"public_metrics": {"followers_count": 0, "following_count": 10, "tweet_count": 47, "listed_count": 0}, "created_at": "2021-12-08T10:28:31.000Z", "name": "BMF", "id": "1468527831416127491", "url": "", "description": "", "username": "BMF071"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "id": "1543799401", "public_metrics": {"followers_count": 272, "following_count": 415, "tweet_count": 21819, "listed_count": 3}, "username": "dtite67", "created_at": "2013-06-24T17:52:07.000Z", "name": "Dave"}, {"data": {"url": "", "description": "", "
```

Interaction and Presentation Options

The project aims to study the network composition and summarize metrics related to the queried tweets. The utility for the users is to visualize the metrics as a form of preliminary/exploratory data analysis around the popularity of tweets that could potentially spread Covid Misinformation. The program uses a CLI to present the options to users. Users can select from one of the following options:

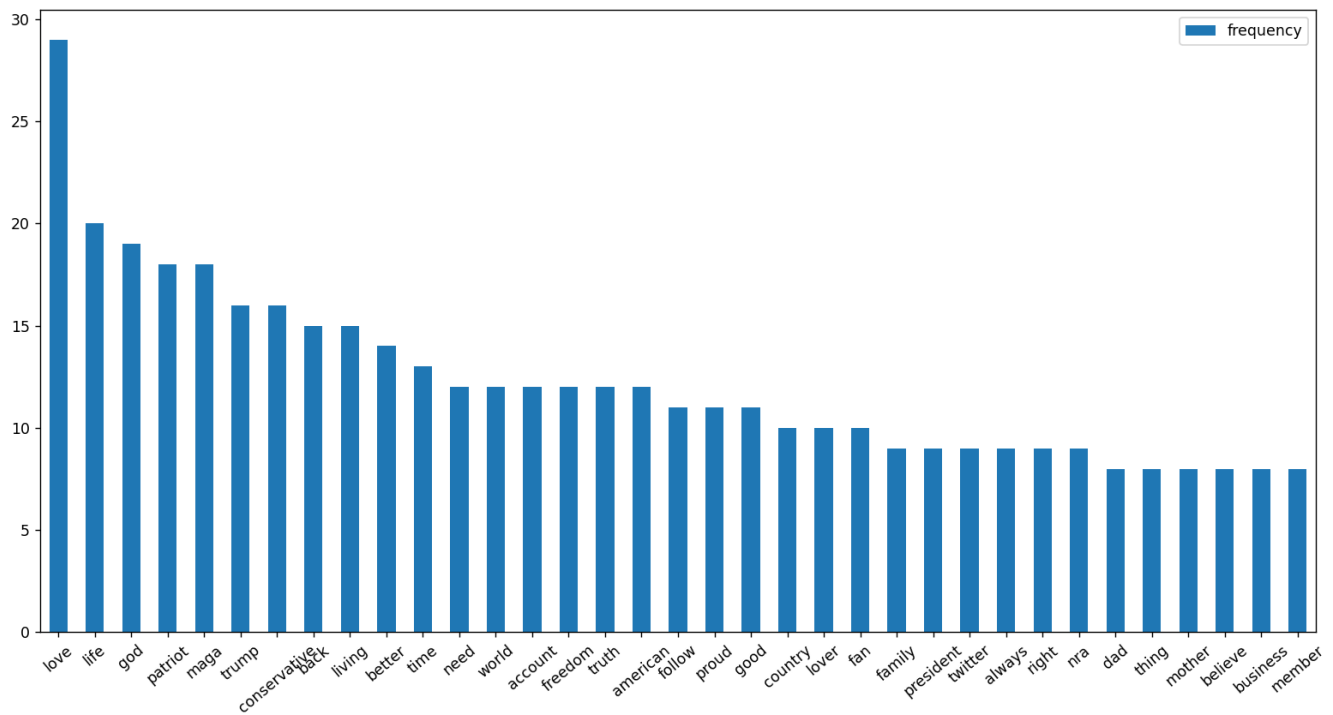
Enter the option for the type of the Analysis component you want to view.

- ```

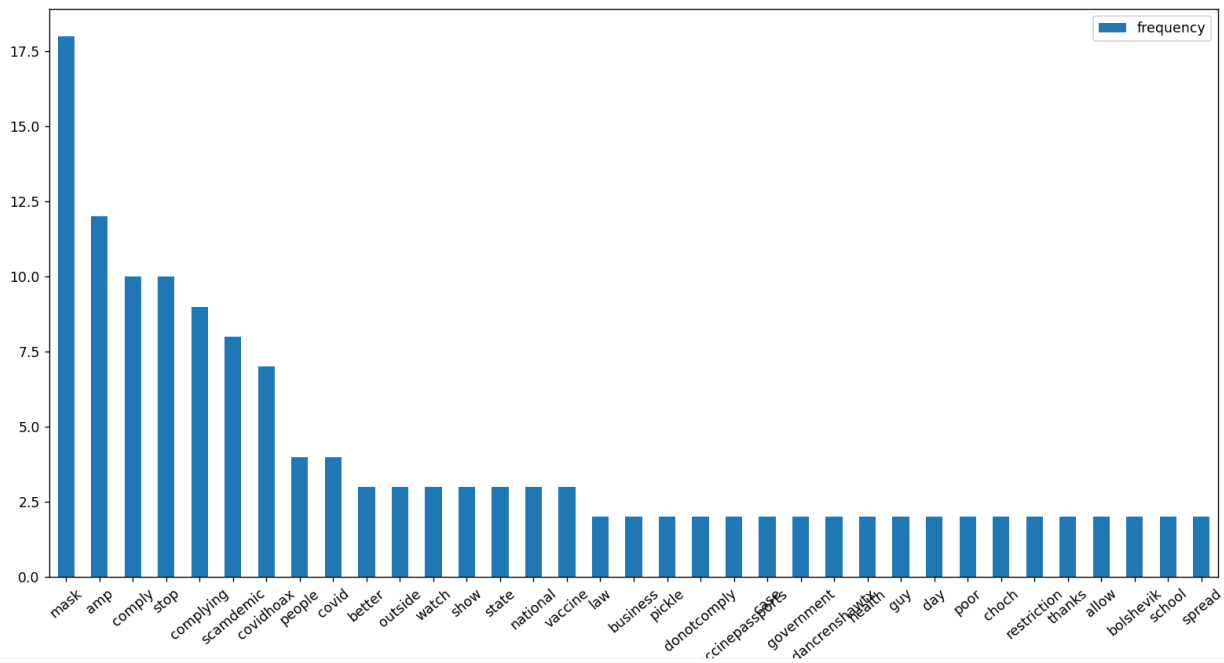
1. Word Frequencies from Users' Descriptions
2. Word Frequencies from Tweet Content
3. Retweet Statistics
4. Primary Node Statistics
5. Network Graph
6. EXIT

```

### Word Frequencies from Users' Descriptions:

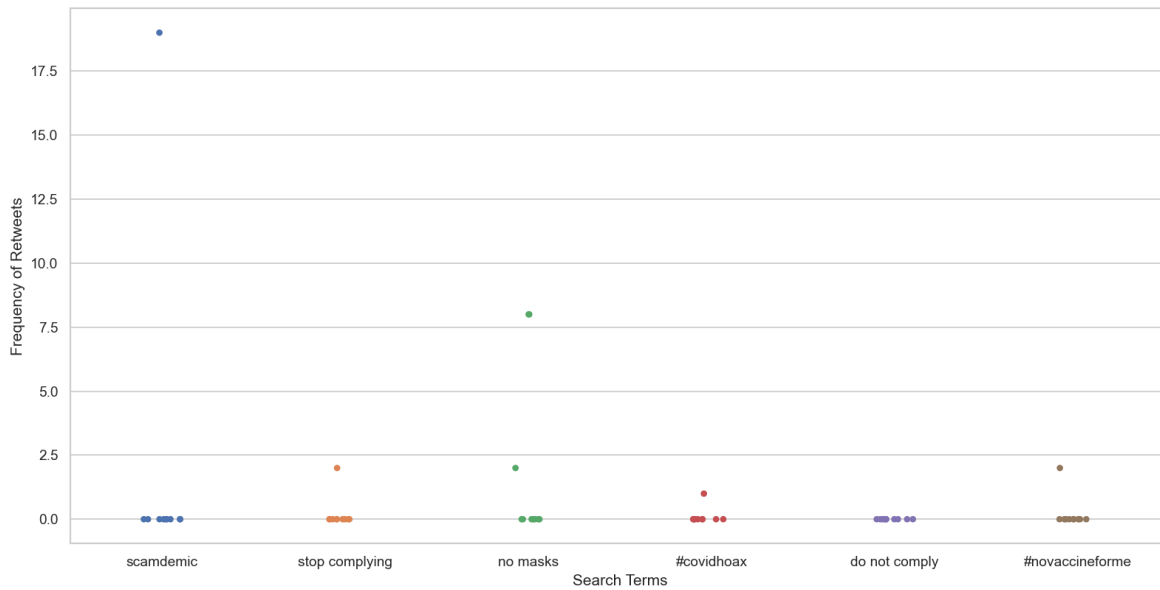


## Word Frequencies from Tweet Content

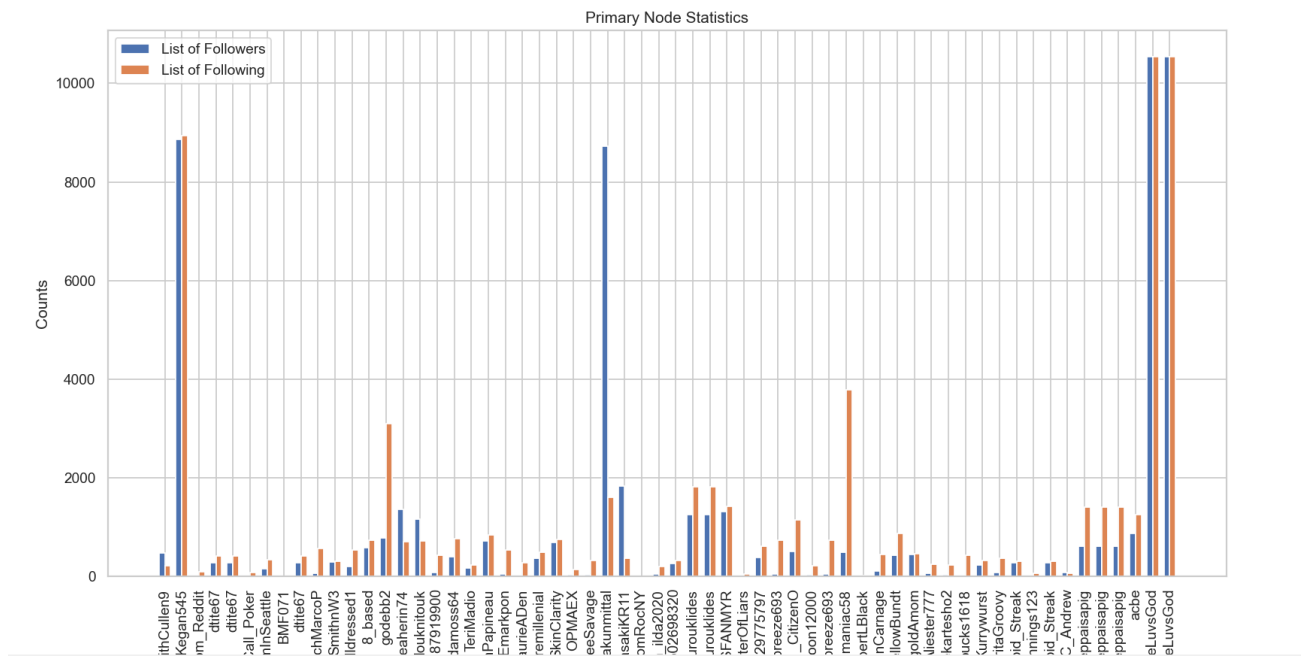


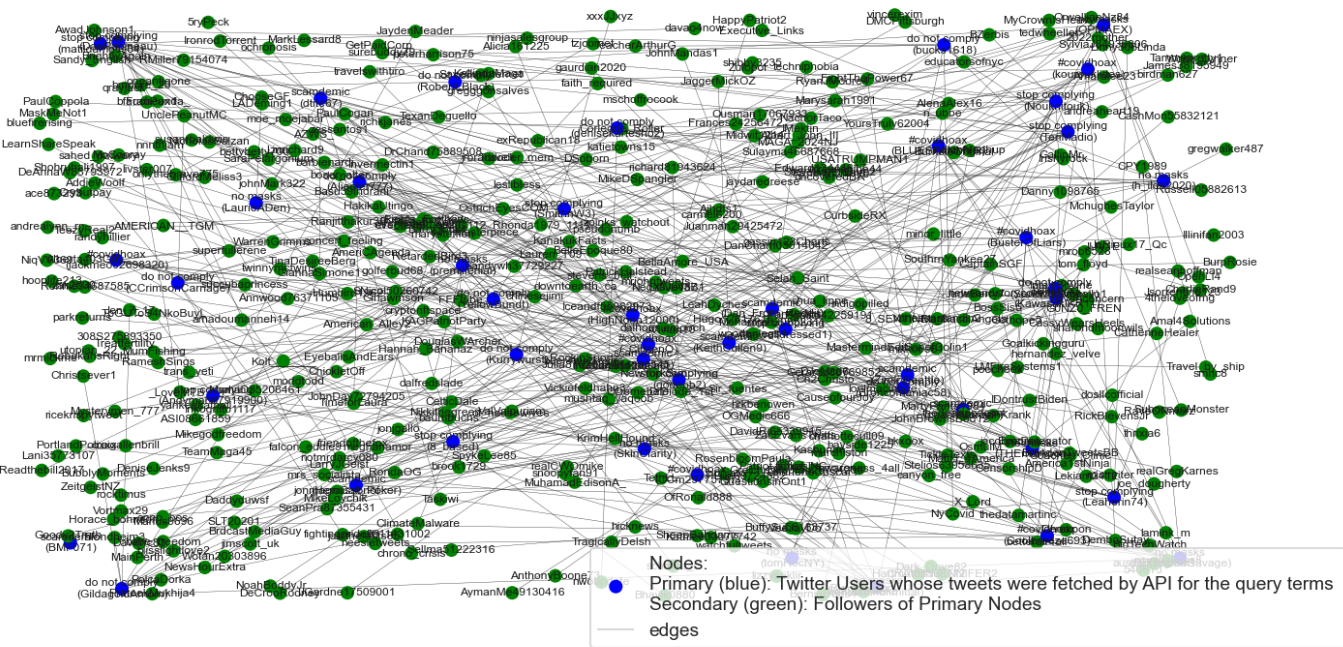
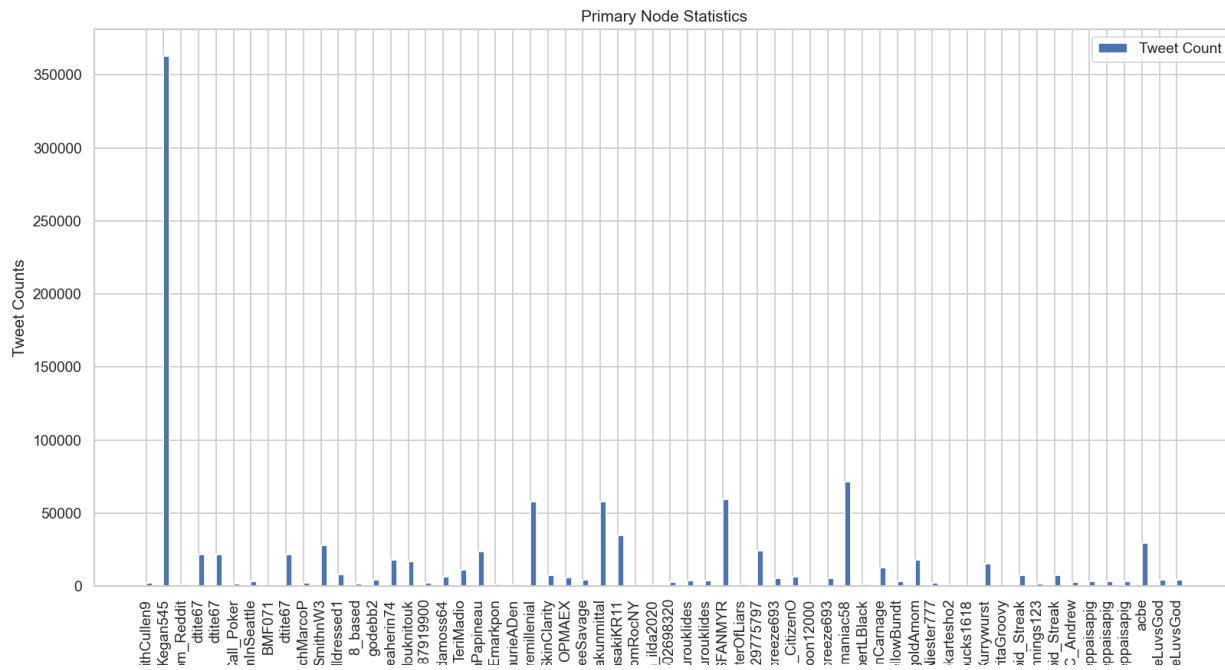
## Retweet Statistics





## Primary Node Statistics





**Demo Link** - <https://drive.google.com/drive/folders/1zXD8hNBwEflUBh5RfbmVR2SbhsATBpMn>