**Kaggle: How ISIS uses Twitter**

https://www.kaggle.com/fifthtribe/how-isis-uses-twitter

This project was used purely for academic purposes

Beverly Rice

Springboard Capstone 2: Final Report

15 March 2018

**I. Introduction**

Background: The Islamic State of Iraq and the Levant (ISIS) also referred to by its Arabic acronym for al-Dawla al-Islamiya al-Iraq al-Sham (Daesh), is a Salafi jihadist extremist group involved in global terrorism, human rights abuse, and “ethnic cleansing.” They observe strict Salafist doctrine and seek to establish themselves as a caliphate, an Islamic state led by a group of religious authorities under a direct successor to the Prophet Muhammad. They demand the allegiance of all devout Muslims and actively promote their organization and message through social media. One platform commonly used by Daesh is Twitter, an online news and social networking site where users post and interact with one another through what may be private or public messages capped at 280 characters called “tweets.” In October of 2017 there were reportedly 330 million active users posting/discussing a wide variety of topics to include personal experiences, current events, and entertainment (actors, artists, movies, etc). It has become a revolutionary platform on which individuals and groups can promote themselves and voice their support and opposition to political decisions, world affairs, and various campaigns. Daesh uses Twitter to provide a constant stream of propaganda in an effort to acquire supporters for their cause world-wide.

Problem: Identify influencers in the pro-ISIS twitter network and conduct sentiment analysis on which clergy pro-ISIS individuals support versus denounce/dislike. For this particular project, an ‘influencer’ is considered an active twitter user connected to 50 or more pro-ISIS individuals.

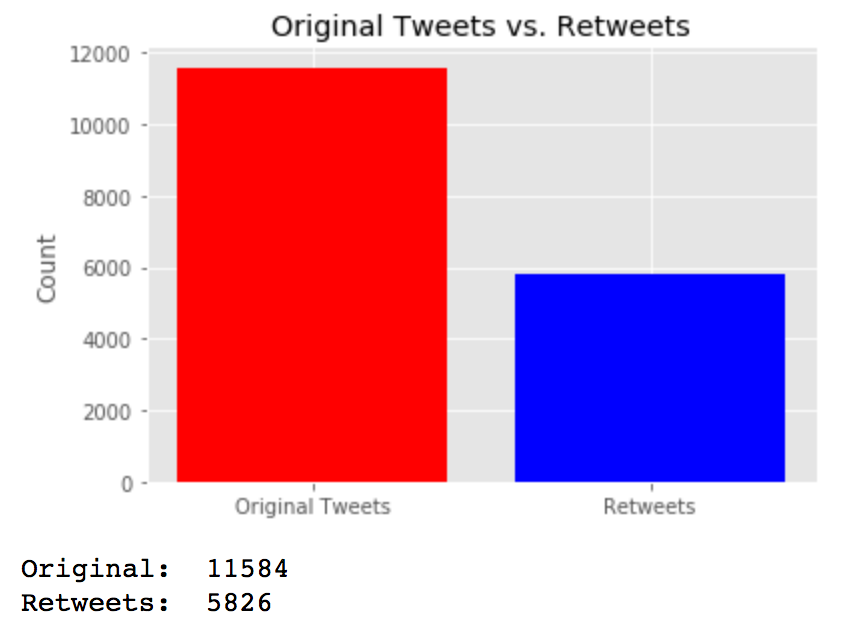
Client: Fifth Tribe is one of several companies that serve on the Private Sector Committee of the Board of the Global Community Engagement and Resilience Fund, which is the first global effort to support local, community-level initiatives aimed at strengthening resilience against violent extremism. In December of 2014, the company won the anti-ISIS “Hedaya Hack” organized by Affinis Labs and hosted at the “Global Countering Violent Extremism Expo” and since then, have been actively involved in working with the open-source community and community content producers in developing counter-messaging campaigns and tools.

**II. Data**

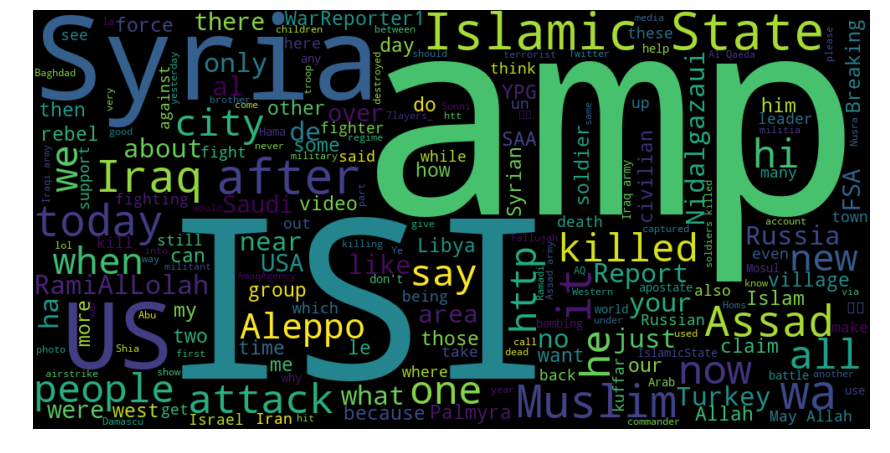
Data Source: Fifth Tribe has scraped over 17,000 tweets from 100+ pro-ISIS Twitter subscribers from all over the world since the November 2015 Paris Attacks.

Data Set: There are exactly 17,410 tweets in the data set. The name, twitter username, description, location, followers, number of statuses, and time, accompany each tweet. The name column is complete with no missing variables. Some names are very similar to the username while others differ significantly. Of note, some are written in Arabic. The username variable includes any combination of letters, numbers and underscores and is unique to every member. The description variable has missing values and when available, provides a brief synopsis of the individual tweet. The location is also missing values. The followers variable lists solely the number of total followers the user has. The numberstatuses variable lists the total number of statuses/tweets published by the user. The time variable is the time the tweet was completed and published. Minimal cleaning was needed. Missing values in the description and location variables were filled with ‘missing.’

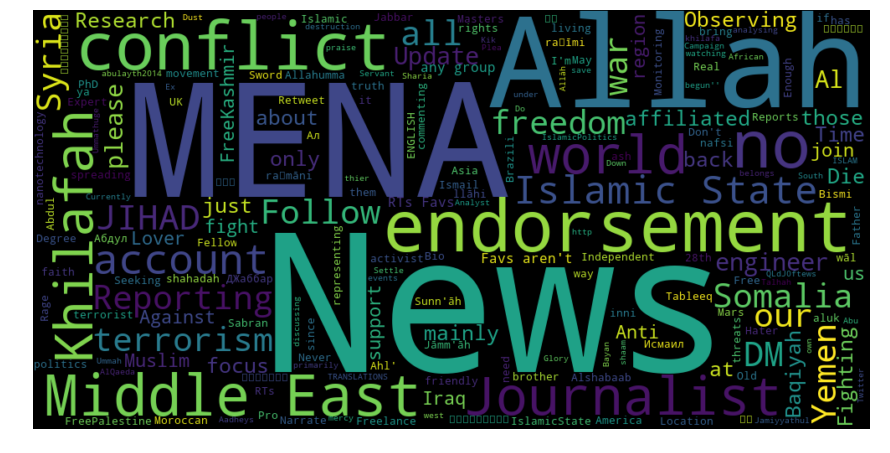
**III. Exploratory Data Analysis**



Word cloud of all tweets:



Word cloud of all unique user descriptions:



Top 10 tweeters who referenced/tagged another member

(username, #tweets w/tagged user):

Top 10 Tweeters: [('Uncle\_SamCoco', 1618), ('mobi\_ayubi', 1078), ('WarReporter1', 713), ('warrnews', 666), ('RamiAlLolah', 655), ('melvynlion', 642), ('MaghrabiArabi', 454), ('\_IshfaqAhmad', 432), ('AsimAbuMerjem', 289), ('NaseemAhmed50', 252)]

Top 10 tweeters who were tagged by another member

(username, #tweets they were tagged in):

Top 10 Receivers(tagged): [('RamiAlLolah', 573), ('Nidalgazaui', 339), ('WarReporter1', 256), ('7layers\_', 116), ('ScotsmanInfidel', 79), ('sparksofirhabi3', 76), ('Conflicts', 72), ('DidyouknowVS', 72), ('MaghrebiQM', 72), ('Uncle\_SamCoco', 70)]

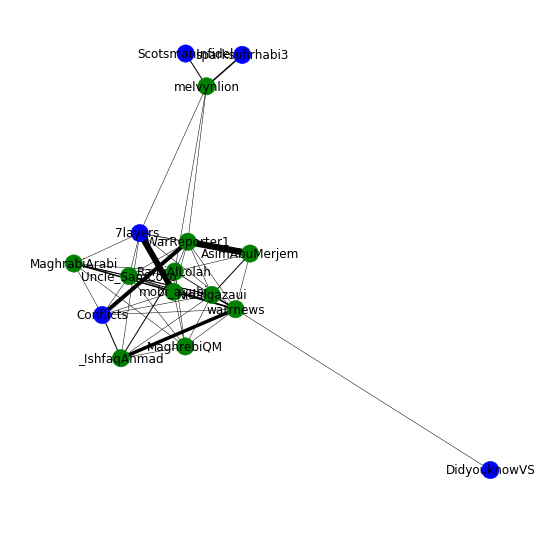
Network of Top 20 tweeters and tagged members in relation to each other:

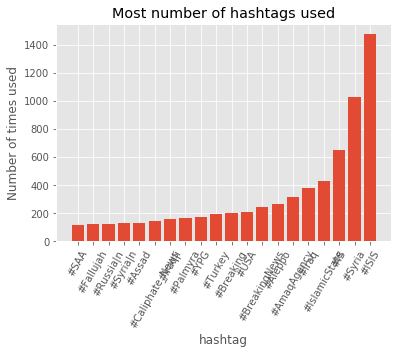
**Network of top 10 most active tweeters and top 10 individuals tagged the most**

Green = Tweeted and was Tagged by others

Red = Only tweeted

Blue = Only agged in someone else’s tweet

 Most common hashtags used:



#ISIS

#Syria

#IslamicState

#Iraq

#AmaqAgency

#Aleppo

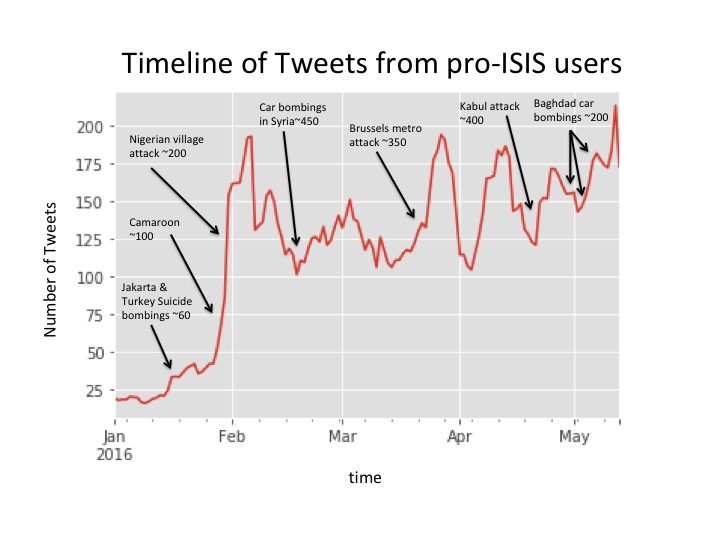
#Breaking News

#USA

#Breaking

#Turkey

#YPG

****

**IV. Data Wrangling**

First we start by defining “pro-ISIS. This is done using 4 Parameters ranked by weight

1. Tagged #AmaqAgency – an ISIS affiliated news network (4x multiplier )

2. Tagged #ISIS, #IS, #IslamicState (3x multiplier)

3. Retweeted someone identified as pro-ISIS (2x multiplier)

4. Tweeted around the time of an ISIS terrorist attack (1x multiplier)

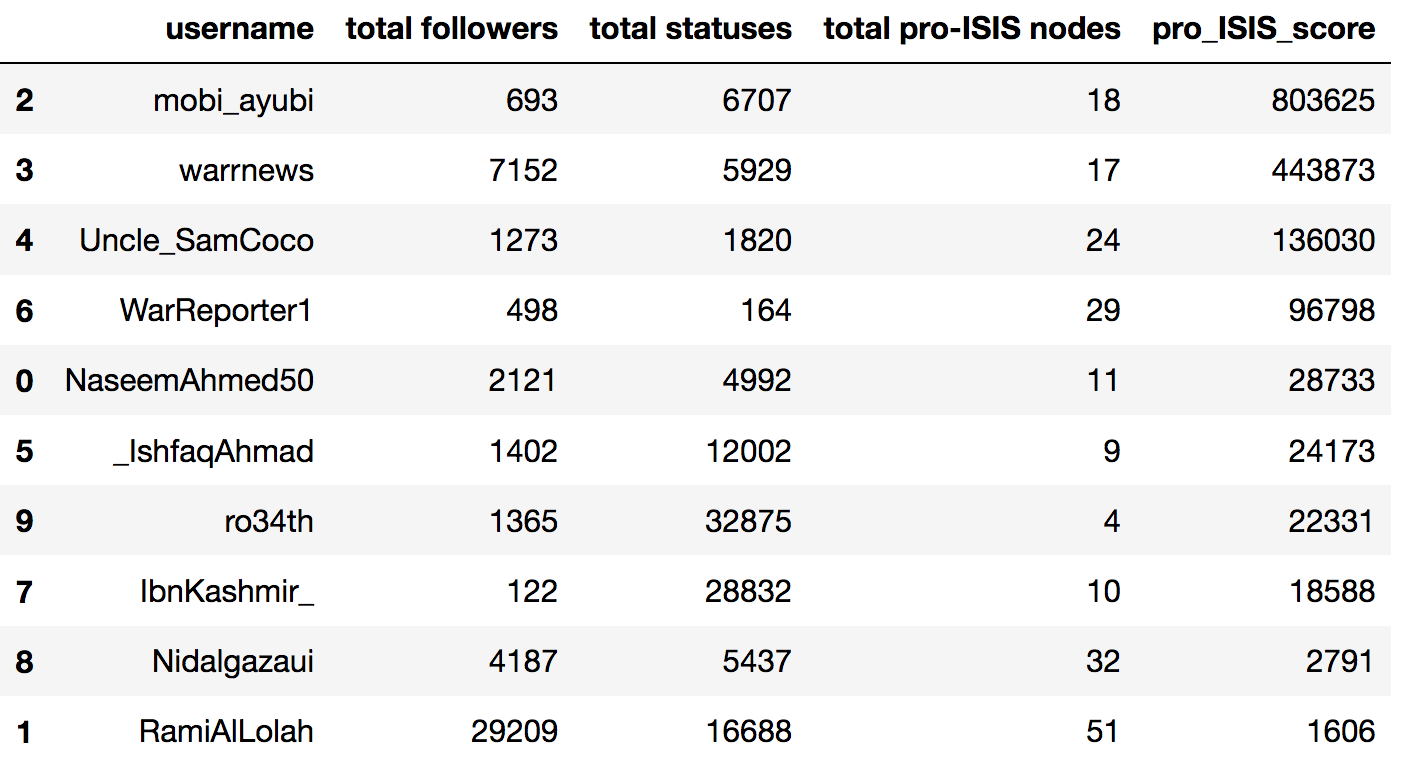
Each parameter has an associated multiplier in determining likelihood that the member is pro-ISIS. If a member exhibits/falls into a category, that single instance will be multiplied by the multiplier. All instances will be scored and added to determine whether an individual is pro-ISIS on a high, med or low scale. For instance:

twitterUSER3 tagged AmaqAgency once, ISIS twice and retweeted a pro-ISIS user seven times. twitterUSER3's score would be: 4(1) + 3(2) + 2(7) + 1(0) = 24

4x #AmaqAgency 3x #ISIS 2x RT a pro-ISIS user 1x Tweeted around time of attack

**Most influential pro-ISIS users**

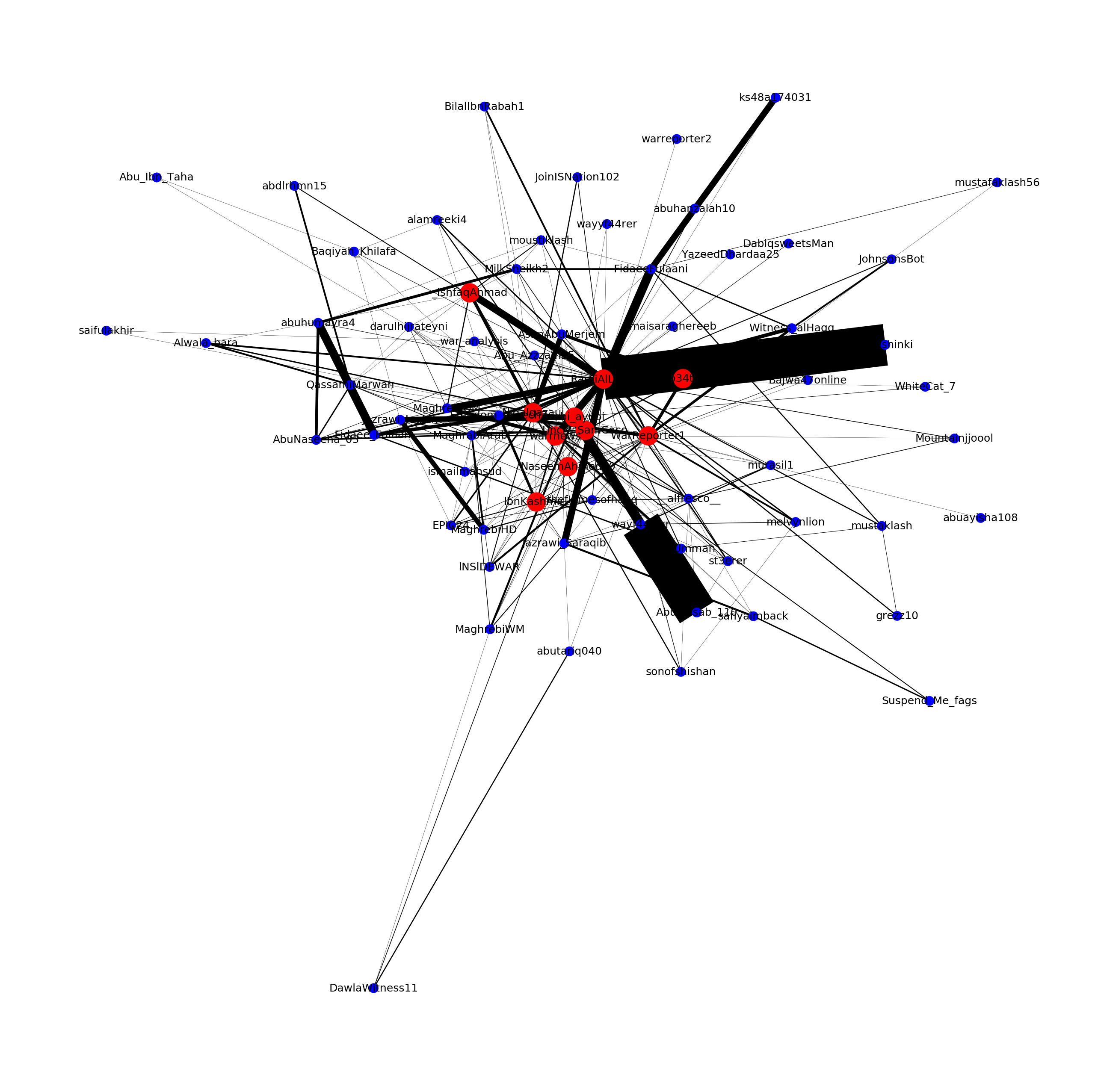
Based on total followers, total statuses, total pro-ISIS nodes, and pro\_ISIS\_score

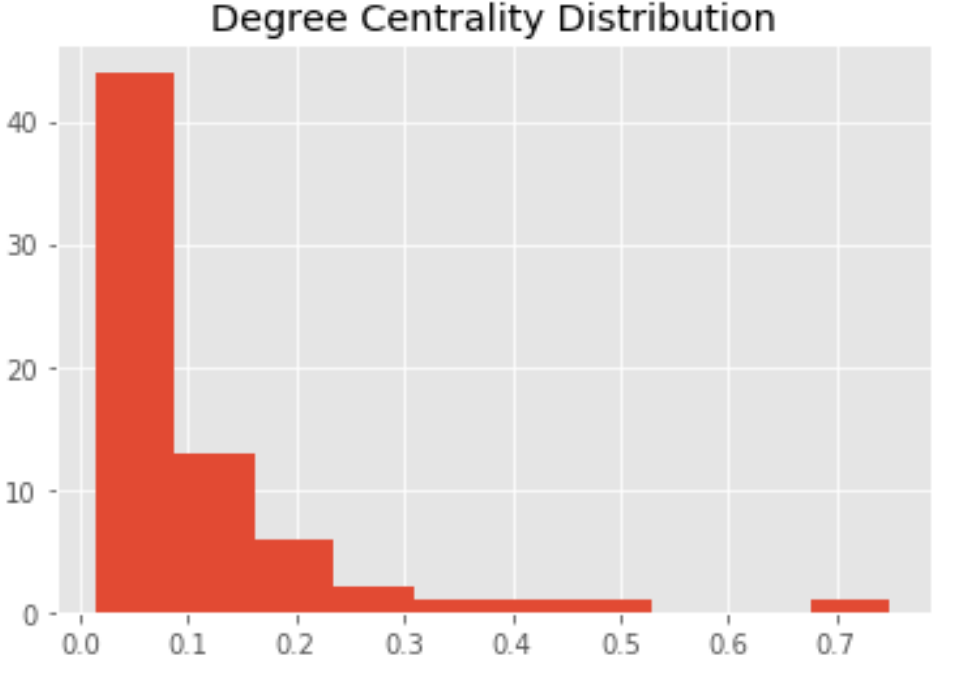
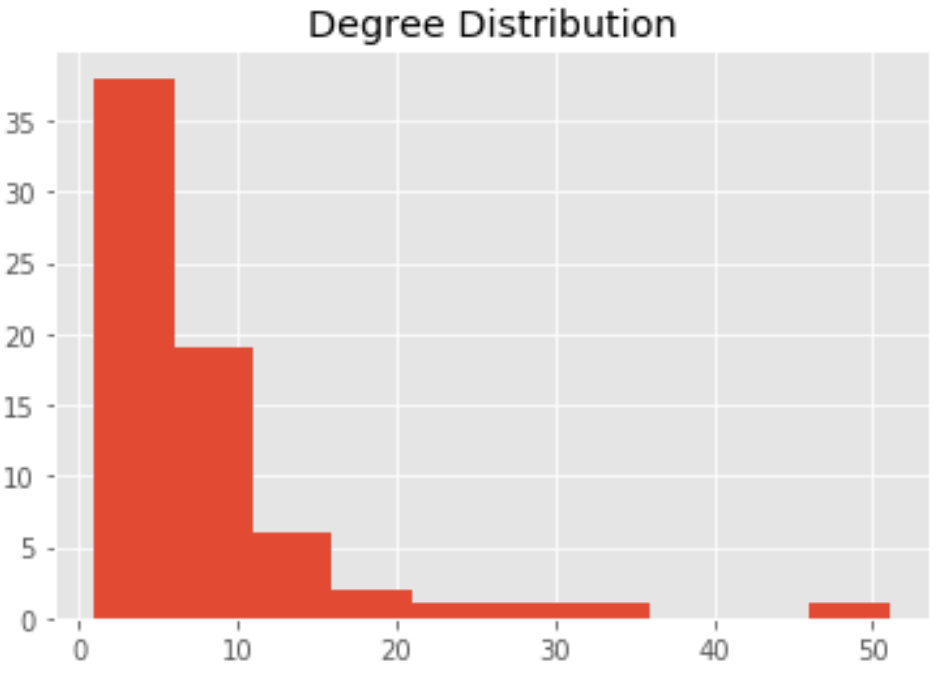
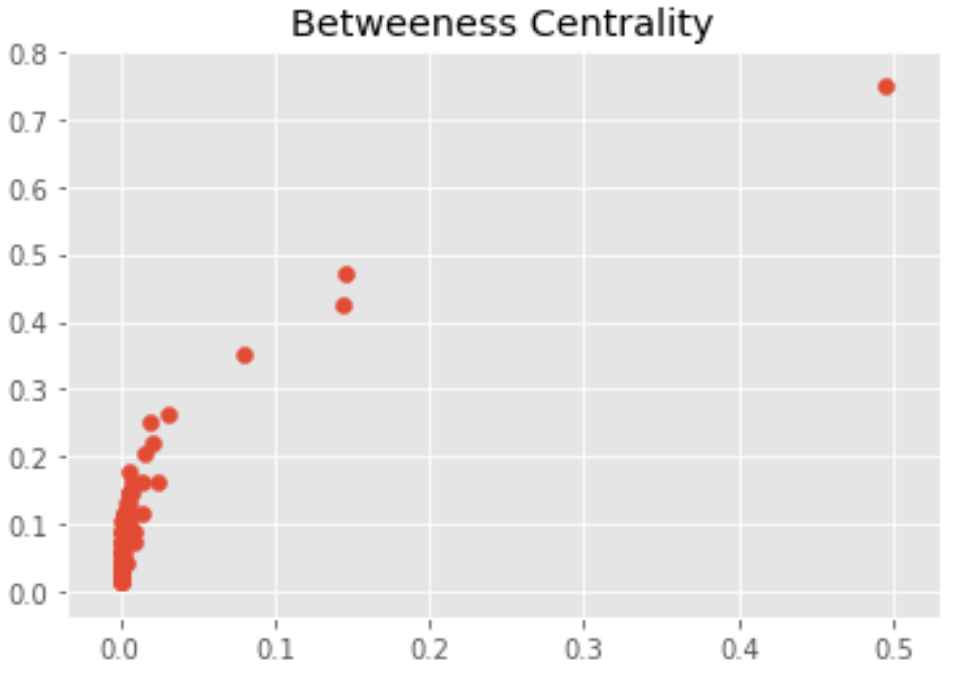
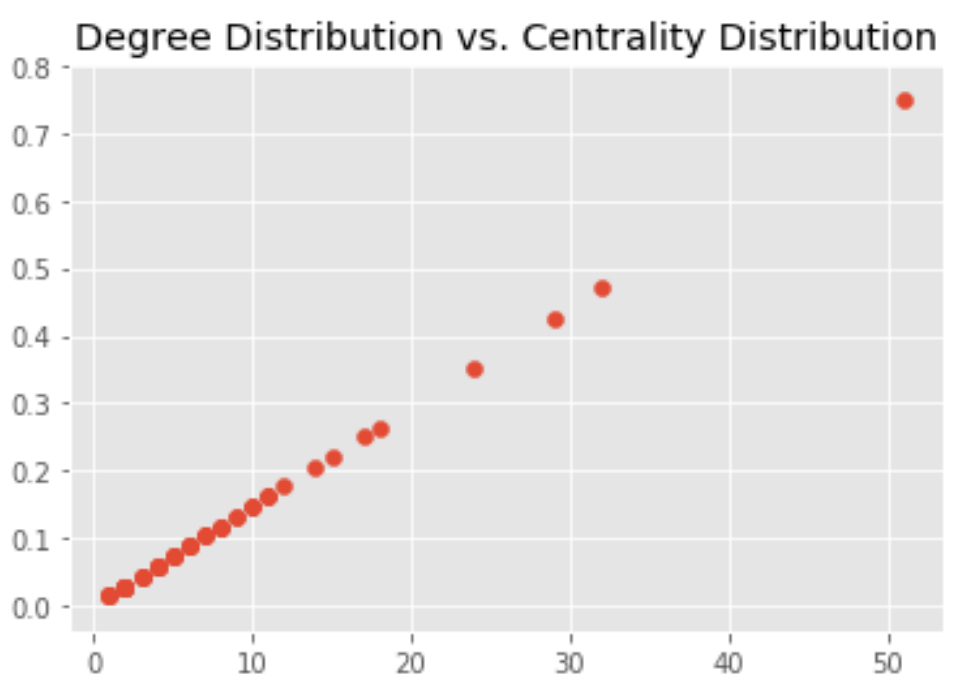


**Network of most influential pro-ISIS members**

Red = Influential member

Blue = Network node





**IV. Important Findings**

* The most influential pro-ISIS twitter members were mobi\_ayubi, warrnews, UncleSamCoCo, WarReporter1, NaseemAhmed50, \_IshfaqAhmad, ro34th, IbnKashmir\_, Nidalgazaui, RamiAlLolah
* RamiAlLolah had the highest degree centrality and betweeness centrality of all the pro-ISIS users defined by the data set
* The top 3 pairs of users were RamiAlLolah & Jazrawi\_Joulan, Nidalgazaui & Jazrawi\_Joulan, and Maghrabi & mobi\_ayubi
* There were a total of 140 cliques, the largest involving RamiAlLolah, Nidalgazaui, UncleSamCoco, WarReporter1, mobi\_ayubi, NaseemAhmed50, and warrnews

**IV. Further Research**

An additional data set that can be used in conjunction with this project is that provided by the ‘Tweets Targeting ISIS’ kaggle competition. The new set contains roughly 122,000 tweets collected on two separate days, which contained any of the following terms: ISIS, ISIL, DAESH, islamicstate, raqqa, mosul, and “Islamic state.” This data set would be a good backdrop against which to measure the pro-ISIS tweets.

Another resource that may be of good use to this project is tweepy. Theoretically tweepy could pull the friends of pro-ISIS members and provide more insight on any one individual’s network.