1. Abstract
2. Introduction
3. Background

* Twitter Introduction
  + What is it?
    - Twitter is a social media platform based on the sharing of short posts. It was originally founded as a microblogging organization, but has expanded into the world of social media. It serves as a way to facilitate discussion and share information and opinions to millions of users.
  + How does it work?
    - Users can share short posts known as “tweets” which other users can react to by “liking” the post or by further sharing the information by “retweeting.” Users can also “follow” other users to keep up-to-date with any information they share.
      * If we use mentions/hashtags in our analysis we can include more about these notions
  + Who uses it?
    - Politicians, celebrities, and everyday people

1. Dataset
   1. Hateful users (hateful)
   2. Baseline dataset (normal + others)
2. Analysis
   1. Statistical analysis
      1. General characterization
         1. **Favorites that a user get**
         2. **Tweet**
         3. **Retweeting others**
         4. Status length
         5. ~~Time between tweets~~
      2. Content analysis
         1. Hashtags
            1. Hashtags that two groups used most, top 10
         2. Mentions
         3. Number of URLs used per tweet
         4. Sentiment analysis
      3. Account evolution:
         1. **Followers count compare**
         2. ***Followees count***
      4. Correlation:
         1. Followers count vs Statuses/favorite count
         2. Follower count vs Retweeted count (this is different than retweet, calculate this from the edge list)
         3. Follower count vs Followee count (This reveals a users choice)
         4. Mention count vs Retweet count (not sure how to interpret this)
            1. Could help to identify influencers - those with a lot of mentions and retweets
            2. See if there is any correlation between how many times a user is mentioned in a tweet and the amount of retweets that person gets

For example, celebrities would have high counts in both instances

* 1. Network analysisLook at a bigger scale:
     1. Look at users with more than 1K, 10K, and 50K likes(favorites)
        1. Do they form weakly connected components?
           1. What is the percentage of hateful users in such a network?
        2. What about strongly connected components?
           1. We want to restrict the diameter

Reason: the opinions would not spread very far, maybe diameter <3?

* + - * 1. What is the percentage of hateful users in such a network?
        2. Calculate the PageRank for these strongly connected components
        3. Compare the users with top PageRank number

Are there any hateful users?

What rank do they have?

* + - 1. Use Hyperlink-Induced Topic Search (HITS) to find hubs
         1. Are there any hateful users?
  1. Influence estimation
     1. Look at a smaller scale:
        1. Find the distribution of strongly connected components
        2. Plot the covariance between the size of strongly connected components vs ratio of hateful users
           1. guess: Negative correlation
     2. How important are these hateful users vs others in this retweet network?
        1. Closeness centrality
        2. Betweenness centrality
        3. If we remove them, will there be a huge impact?
  2. Why are they labeled as hateful?

1. Related work(Traditionally, related work/studies will be placed at the very front of a paper. Maybe we can mention it in the introduction)
2. Conclusion