

Statement of Work

Team:

- Samuel Carpenter
- Jason Marks
- Evan Wieland
- Nick Albrecht

Content:

For our project we would like to construct an idealized version of what it would take to create a simple social media platform. This would require tables that contain user information, relations between users, posts, post comments, post tags, among other things. If the table count requirement is not met, additional features can be added to the social media platform that would increase the table count.

Sources:

There are plenty of readily available data sets containing social media data. The intention is to use as much as we can from public data sets and generate what isn't as public, such as who a person is following. The data sets that will be used will be coming from this list: <https://github.com/shaypal5/awesome-twitter-data>.

Scope:

The span of time the data will represent will depend on what specific data set we choose to go with, but it will likely be limited at about 1-2 years.

Queries:

- What percentage of users post more than 10 times a day?
- What top 10 users have the greatest number of users they follow and have the greatest number of followers?
- Many, many more...