

The dataset that our group has selected details all the connections between a large selection of Twitter users and the social circles that are generated from these connections. This dataset was taken from the Stanford Large Data Network Collection. Specifically, the data was compiled by J. McAuley and J. Leskovec for the purposes of their publication *Learning to Discover Social Circles in Ego Networks*. Each node in the dataset represents one user with each edge representing a follow from one user to another. In traversing this dataset, we plan to use a Breadth First Search. Not only will it be helpful to find all the vertices/nodes connected to each other vertex/node, but it may also assist in our Betweenness Centrality algorithm, one that measures the “centrality” of each node based on criteria such as number and length of edges. We hope to display that output of said algorithm graphically, first by running all the nodes and outputted data into a Graph Coloring algorithm to easily identify the centrality of nodes visually and then through a Force Directed Graph algorithm that should graphically display the centrality and density of different node clusters (i.e. social circles). Our end goals that we hope to achieve for this project are to learn how to graphically display complex data, learn to implement a betweenness centrality algorithm, and to implement graphical analysis to real-world applications.