DS210 Final Project Report
Belinda Ngo

For the final project, I explored Six Degrees of Separation using two different datasets. The first dataset represents the social network of Facebook users (https://snap.stanford.edu/data/egonets-Facebook.html - facebook_combined.txt.gz), and the second represents the social network of LastFM Asia users (https://snap.stanford.edu/data/feather-lastfm-social.html). While the Facebook dataset has ~4000 nodes, the LastFM Asia dataset has ~7000 nodes. Both are undirected graphs, where each line consists of a pair of connected nodes. Nodes are users, and each edge represents a friendship/connection.

In a separate module, I wrote a bfs function to execute Breadth-First Search algorithm on a graph. I initialized a dist map and a queue data structure. Initially, all distances are set to -1 to signify an unvisited node. The queue keeps track of the list of nodes to visit next. At the end of the function, the dist map contains the shortest distances.

In the main module, I wrote functions to read graphs from either a txt or csv file, to calculate the average distance between all edges within a graph, and to report the results.