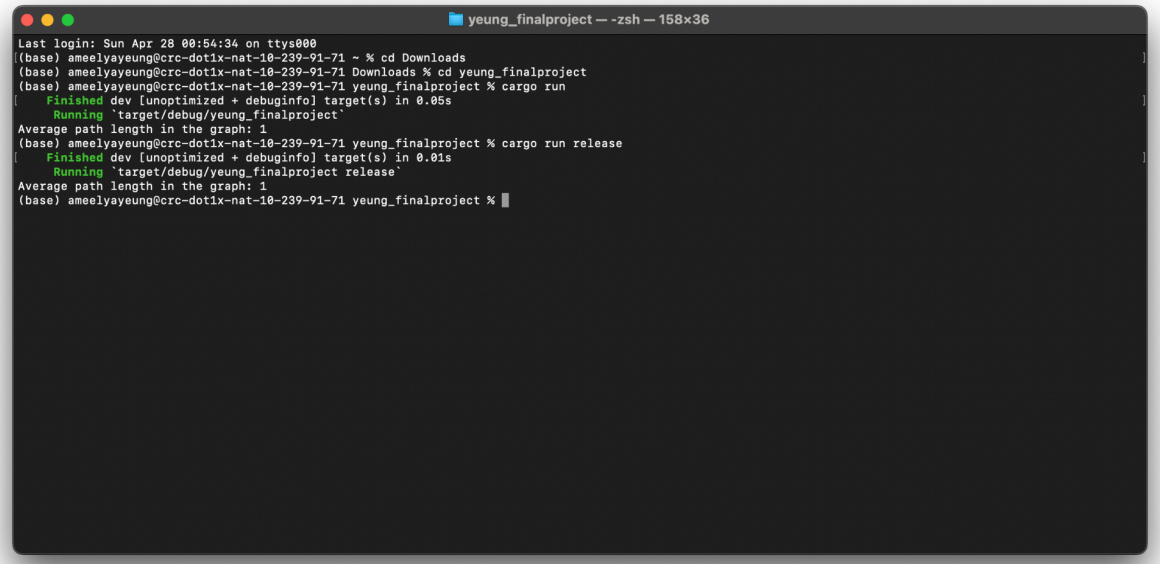
Ameelya Yeung
DS210 - Final Project

Result:



My project is a take on the six degrees of separation theory. I used a dataset that identifies different Facebook users and their connections or better known as “friends” on the social media site. The code reads each line in the text file and splits it to get node identifiers. Each node represents individual people and the edges would represent the connections or friendships between the people. So each vertex that was added to the graph corresponds to a person and its edges would be the connections between two people. I used a breadth-first search to calculate the distances from one start node (so it starts at one person) to all the other nodes in the graph. To achieve this, I used HashMaps to track distances through the graph and so that it would update the distance for each vertex. I also used the “neighbors” function to access any adjacent vertices. I then computed the average path length across the graph to find the average “distance” it would take to be connected to another node or person.

Some areas that could use some improvement were the optimization of the code. It was a little slow to run in both cargo run and cargo run release as shown above. The function to calculate the average path length was not ideal for larger datasets. But as a whole, the program is a measure of the overall connectivity of the network and a short path length suggests that people are very close together and emphasizes the “small-world” that we live in. The result of the program is consistent with the six degrees of separation theory because there are less than 6 separations to each “node” in the graph which means everyone was, on an average, one person away from another. This also speaks to the efficiency of social media in terms of spreading information or misinformation because a small “path distance” implies that information is easily spread. This aspect would influence dynamics such as viral marketing and public news.