Syndia Iglesias
Prof. Kontothanassis
DS 210
December 15, 2022

Six Degrees of Separation

This report contains a summary of the the steps and results I acquired whilst attempting to understand six degrees of separation. More specifically, I wanted to find the average distance between pairs of vertices. The data I worked is called "Social Circle: Google +" and can be found under the Standford University website[1]. The dataset consists of social circles from Google +, where each node represents a user profile.

Data Preparation

The data contains a total of 107,614 nodes and 13,673,453 edges. For this project, I used Visual Studio Code as the code editor and downloaded the data directly from Standford's website. The text file originally contained two columns of large numbers, where the column on the left depicted a number identifying a node and the one on the right indicating another node that it was connected to. Since the numbers identifying each node were very large, and I would not need the specific number to find the degrees of separation, I converted them into smaller numbers starting from 0. Next, I created a directed graph to be able to map the nodes and find the different social circles there are.

Understanding the Data

To understand the data a bit further, I iterated through the vertices to find the different components of the data. In this case, components are different groups or islands of vertices that are connected, which can be depicted as different social circles of the graph. Using breadth-first-search, I found 335 different social circles, with an average distance between pairs of vertices of 2. This lets us know that usually, within social circles, the average amount of friends between one person and the other is 2 in this specific graph. This goes hand in hand with the six degrees of separation theory that states "any person on the planet can be connected to any other person on the planet through a chain of acquaintances that has no more than five intermediaries".[2]

---

[1] https://snap.stanford.edu/data/ego-Gplus.html
[2] https://www.techtarget.com/whatis/definition/six-degrees-of-separation#:~:text=Six%20degrees%20of%20separation%20is,no%20more%20than%20five%20intermediaries.