Radiya Rafat
DS210 Final Project Write Up

For my final project, I chose to do six degrees of separation. I wanted to construct a graph and Breadth-First Search (BFS) model to analyze the relationship between products based on shared reviewers. Thus, using the BFS algorithm, I could output the percentage of nodes I could reach within six degrees of separation from a starting ASIN (product ID).

To begin, the project reviews data from a JSON file in the 'data.rs' file, with different attributes, such as ASIN, reviewer ID, reviewText, and more. I used data.rs to define the structure for reviews and for reading the data file. When writing data.rs, I encountered issues regarding how I wrote my code. This was because Rust had a difficult time reading my JSON file. To combat this issue, I had to make changes, such as utilizing 'serde' which is a serialization and deserialization library that allows you to convert Rust data structures into various formats, including JSON. In doing so, I was able to avoid issues related to the struct portion of my code to effectively call my data file.

Following 'data.rs', I was able to construct a graph in my 'graph.rs' file, with nodes representing products and edges representing the shared reviewers between the products. My graph function iterated over every review and created a graph where each node is an ASIN and each edge is a hashmap of reviewers who have reviewed the product, with the count of reviews as values.

Within my 'graph.rs' file, I also implemented my BFS to explore the graph from a starting point. I initiated my graph from a specific ASIN, counting the percentage of how many products are reachable within six degrees of separation from the starting point.

In my 'main.rs' file, I initiated the file to import my 'data.rs' and 'graph.rs' files to read my data, construct my graph, and execute the BFS algorithm through user commands. Below is an implementation of my code and the percentages each starting ASIN was able to reach with six degrees of separation. Thus, when I input different ASINs, I get varied results depending on the BFS and the relationship from my starting point. For example, for an ASIN of B0010ZBORW, The percentage of nodes reached with six degrees of separation: 1.20%. Another example showcases that for an ASIN of B00006L9LC, The percentage of nodes reached with six degrees of separation: 11.73%

```
(base) radiyarafat@crc-dot1x-nat-10-239-190-27 project % cargo run B0010ZBORW
    Finished dev [unoptimized + debuginfo] target(s) in 0.09s
     Running `target/debug/project B0010ZBORW`
The percentage of nodes reached with six degrees of separation: 1.20%
(base) radiyarafat@crc-dot1x-nat-10-239-190-27 project % cargo run B000X7ST9Y
    Finished dev [unoptimized + debuginfo] target(s) in 0.09s
     Running `target/debug/project B000X7ST9Y`
The percentage of nodes reached with six degrees of separation: 0.05%
(base) radiyarafat@crc-dot1x-nat-10-239-190-27 project % cargo run B0010ZBORW
    Finished dev [unoptimized + debuginfo] target(s) in 0.10s
     Running `target/debug/project B0010ZBORW`
The percentage of nodes reached with six degrees of separation: 1.20%
(base) radiyarafat@crc-dot1x-nat-10-239-190-27 project % cargo run B0000530HU
    Finished dev [unoptimized + debuginfo] target(s) in 0.08s
     Running `target/debug/project B0000530HU`
The percentage of nodes reached with six degrees of separation: 0.12%
(base) radiyarafat@crc-dot1x-nat-10-239-190-27 project % cargo run B00006L9LC
    Finished dev [unoptimized + debuginfo] target(s) in 0.08s
     Running `target/debug/project B00006L9LC`
The percentage of nodes reached with six degrees of separation: 11.73%
(base) radiyarafat@crc-dot1x-nat-10-239-190-27 project % cargo run B00021DJ32
    Finished dev [unoptimized + debuginfo] target(s) in 0.09s
     Running `target/debug/project B00021DJ32`
The percentage of nodes reached with six degrees of separation: 0.10%
(base) radiyarafat@crc-dot1x-nat-10-239-190-27 project % cargo run B0006O10P4
    Finished dev [unoptimized + debuginfo] target(s) in 0.08s
     Running `target/debug/project B0006O10P4`
The percentage of nodes reached with six degrees of separation: 0.07%
(base) radiyarafat@crc-dot1x-nat-10-239-190-27 project % cargo run B0009RF9DW
    Finished dev [unoptimized + debuginfo] target(s) in 0.09s
     Running `target/debug/project B0009RF9DW`
The percentage of nodes reached with six degrees of separation: 9.27%
(base) radiyarafat@crc-dot1x-nat-10-239-190-27 project %
```