DS210 Final Project: Marvel Universe Network Analysis

Purpose

For my DS210 final project, I decided to try and create a graph that represents the intricate connections between Marvel characters and their appearances. In the end, I settled on creating a degree distributions graph. My final goal was to use network analysis to discover interesting patterns in relationships among the many characters in the Marvel Universe, including comics, films and TV shows alike.

Dataset

I started this project by sourcing data from Kaggle, specifically choosing to use the data from the 'edges.csv' file. To make it easier to use the data on my end, I converted it from its original CSV format into a TXT file.
Dataset Link: https://www.kaggle.com/datasets/csanhueza/the-marvel-universe-social-network

Implementation

I chose to use HashMap and Hashset through Rust. First, I used HashMap to effectively group characters together based on their comic appearances. I chose to use HashSet in order to ensure that each and every character would be represented as a unique entry in order to maintain the accuracy of my analysis. In order to run the project, you must first run “cargo build” to compile the code, and then follow it with “cargo run” in order to see the output.
For my graph, I focused on two main computations. The first one I chose was determining the degree distribution of the vertices and the second one was calculating the number of neighbors each character has at a distance of 2. The first part of the output provides an insightful look into the social network, by calculating the number of different marvel characters connected to each other within two degrees of separation. For instance, if the output read "Hero: DEATHWATCH/STEPHAN L, Neighbors at distance 2: 1737," it means that the character DEATHWATCH/STEPHAN L had connections at a distance 2 with 1737 other characters. Secondly, if the output says “Characters with 83 connections: 13,” it means that 13 characters in the social network have exactly 83 connections to other characters. My code also outputs the minimum number of connections for a hero, which was 1, and maximum number of connections for a hero, which was 1345. This means that every character is connected to another character, since the minimum number is 1. It also shows that one of the characters is connected to 1345 other characters, showing how deep the Marvel Universe storyline runs.

Results

The results from this project provide a look into the structure of the Marvel Universe and the interactions with the many characters in the world. For example, it provides insight to the central characters in the networld based on the connections between other characters. It also indicates that some supporting characters may have more of an influence than initially presumed in the storyline. It provides new information about the influence of the characters. It is interesting to see the storyline turned into numbers.

```
Hero: SPITFIRE/LADY JACQUE, Neighbors at distance 2: 2849
Hero: PRISM, Neighbors at distance 2: 1852
Hero: ARACHNE, Neighbors at distance 2: 2262
Hero: SHOT/, Neighbors at distance 2: 247
Hero: DR. STRANGE/STEPHEN, Neighbors at distance 2: 3268
Hero: CITIZEN V II/HELMUT, Neighbors at distance 2: 3359
Hero: GHOST, Neighbors at distance 2: 2181
Hero: ELEJEA, Neighbors at distance 2: 671
Hero: UNION JACK III/JOEY, Neighbors at distance 2: 3499
Hero: T'CHAKA, Neighbors at distance 2: 2250
Hero: FIREFIST II/, Neighbors at distance 2: 1162
Hero: SUNFIRE/SHIRO YASHID, Neighbors at distance 2: 684
Hero: REQUIEM/, Neighbors at distance 2: 1290
Hero: MACHETE/FERDINAND LO, Neighbors at distance 2: 1794
Hero: MYSTIQUE/RAVEN DARKH, Neighbors at distance 2: 3446
Hero: SUNDRAGON/PAMELA DOU, Neighbors at distance 2: 1343
Hero: HOOPSNAKE/, Neighbors at distance 2: 1104
Hero: VOSTOK, Neighbors at distance 2: 3220
Hero: TYRANT, Neighbors at distance 2: 847
Hero: SPYMASTER II, Neighbors at distance 2: 1990
```

```
Characters with 56 connections: 16
Characters with 196 connections: 2
Characters with 841 connections: 1
Characters with 125 connections: 8
Characters with 46 connections: 8
Characters with 32 connections: 28
Characters with 37 connections: 30
Characters with 185 connections: 1
Characters with 802 connections: 1
Characters with 275 connections: 1
Characters with 286 connections: 1
Characters with 268 connections: 1
Characters with 288 connections: 2
Minimum number of connections for a hero: 1
Maximum number of connections for a hero: 1345
```


Sources:
https://www.programiz.com/rust/hashmap
https://web.mit.edu/rust-lang_v1.25/arch/amd64_ubuntu1404/share/doc/rust/html/book/second-edition/ch08-03-hash-maps.html
https://www.programiz.com/rust/hashset
https://www.programiz.com/rust/hashset