CSCI 3415 – Principles of Programming Languages
Dr. Doug Williams
Programming Assignment 2 – Python
Due Nov. 7, 2018 by midnight

The Marvel Universe

From The Marvel Universe Social Network:

> Marvel Comics, originally called Timely Comics Inc., has been publishing comic books for several decades. "The Golden Age of Comics" name that was given due to the popularity of the books during the first years, was later followed by a period of decline of interest in superhero stories due to World War ref. In 1961, Marvel relaunched its superhero comic books publishing line. This new era started what has been known as the *Marvel Age of Comics*. Characters created during this period such as Spider-Man, the Hulk, the Fantastic Four, and the X-Men, together with those created during the Golden Age such as Captain America, are known worldwide and have become cultural icons during the last decades. Later, Marvel's characters popularity has been revitalized even more due to the release of several recent movies which recreate the comic books using spectacular modern special effects. Nowadays, it is possible to access the content of the comic books via a digital platform created by Marvel, where it is possible to subscribe monthly or yearly to get access to the comics. More information about the Marvel Universe can be found here.
>
> The Marvel Comics character collaboration graph was originally constructed by Cesc Rosselló, Ricardo Alberich, and Joe Miro from the University of the Balearic Islands. They compare the characteristics of this universe to real-world collaboration networks, such as the Hollywood network, or the one created by scientists who work together in producing research papers. Their original sources can be found here. With this dataset, the authors published the paper titled: "Marvel Universe looks almost like a real social network".

Install Python

Python is an interpreted high-level programming language for general-purpose programming. Created by Guido van Rossum and first released in 1991, Python has a design philosophy that emphasizes code readability, notably using significant whitespace. It provides constructs that enable clear programming on both small and large scales. In July 2018, Van Rossum stepped down as the leader in the language community after 30 years.

Python features a dynamic type system and automatic memory management. It supports multiple programming paradigms, including object-oriented, imperative, functional and procedural, and has a large and comprehensive standard library.

Python can be downloaded and installed from the official Python web site at https://www.python.org/. Due to incompatibilities (at the source level) between versions 2 and 3 of the language, there are two current releases of the language – 3.7.1 (released 2018-10-20) for version 3 of the language and 2.7.15 (released 2018-05-01) for version 2 of the language. We will use version 3 of the language. As with most

modern languages, binary installers are available for Windows, Linux/UNIX, Mac OS, and others. Please see the download page(s) for information on installing Python on your platform.

Having two current releases of the Python language (as well as different versions of the myriad of packages available for the language) can be problematic. For example, many Linux distributions use python2 and python3 as the names of the executables for the two versions of the languages and python as a link to the 'preferred' version for that distribution – which is often still python2 for various reasons. Many users choose to use Anaconda (https://www.anaconda.com/) as an alternative means of installing Python. Anaconda manages multiple versions of the language and packages.

Part II – Get the Marvel Universe Data

I have posted the zip file containing the three comma separated value (CSV) files that define the social graph for the Marvel universe on the course's Canvas web site. These three files are:

- nodes.csv – Contains two columns, node and type. Each row defines a node in the social network. Node is the name of the node and type is either hero or comic.
- edges.csv – Contains two columns, hero and comic. Each row defines an edge from a hero node to a comic node in the social network.
- hero-network.csv – Contains two columns, hero1 and hero2. Each row defines an instance of two heroes appearing in a single comic.

Or, you can download these files directly from https://www.kaggle.com/csanhueza/the-marvel-universe-social-network [registration required].

Specifics

Write a Python program to read the Marvel universe social network from the csv files nodes.csv, edges.csv, and hero-network.csv.

Compute and print the following statistics for the social network:

- Number of heroes
- Number of comics
- Minimum, mean, and maximum books per character
- Minimum, mean, and maximum characters per book
- Minimum, mean, and maximum partners per hero

Report

You must include a short write up about your program. It must include a description of the program – essentially what is contained in this document, your design, the implementation details (including the code), sample runs, and conclusions. It should also cite any sources you used in developing your program.