CSE 522 Visualization Final Project Proposal

Visualization of Marvel Comic Universe

Chinmaya Dehury-112077747

cdehury@cs.stonybrook.edu

1.Introduction

Choosing a good topic for doing visualization project was an interesting exercise. Since there were so many interesting topics available containing a large amount of intriguing data. The first idea that came to mind was to present data about COVID 19. But then I didn't really want to present data about a topic which is over saturated with data visualization in the current time. Also, I wanted to present on a topic which was more entertaining and upbeat and had an universal appeal. Being an ardent fan of comic books and movies I am always curious to know facts and trivia regarding them. So, I thought it was interesting idea to present visualization on such a comic book universe. So, after some research I came to know about the Marvel Universe dataset. It contained a lot of data about arguably most famous comic book universe. And it contained just the right mix of data that would be fit perfectly on a dashboard.

2.Dataset

A quick search on Kaggle returned over 300 results on Marvel comic book universe. Of those data I found some of which were interesting and would look good when visualized. Here are some of the datasets:

1. Marvel Universe API: Marvel studios provide API support to developers. And each developer can make 3000 calls to the API per day. I plan to use this freedom to gather data from the original and most authentic source.
2. The Marvel Universe Social Network: The dataset contains heroes and comics, and the relationship between them.
3. Marvel Superheroes: This data set contained stats and info on each hero in the marvel universe.
4. The Superhero Ability Scoreboard: This website contained fan-given scores on superheroes abilities (intelligence, strength, speed, durability, combat). I plant to collect raw data from them.
5. I also plan to collect some images for the comic book characters. For this I plan to crawl marvel fandom wikia page and gather them.

3. Project Description

3.1 Technology stack

I plan to use D3 for visualization, Python for data analysis, Flask will be used for setting up server and passing data to the front end. The front end will be made using HTML, css and vanilla javascript.

3.2 Deliverables

1.Dashboard

I plan to create a dashboard for all comic book characters. And choosing a character would bring a picture of the character a short description and graphs and charts regarding its abilities. For example on of the chart can be a bar chart showing the character power rating.

2.Marvel Social Network

I plan to visualize the marvel social network using parallel coordinates visualization technique. The idea is to show the social network in graph format. But I have not finalized on the technique.

3.Aggregated Data:

 I also plan to categorize the heroes based on their superpowers and abilities and plan to show charts like pie charts and scatter matrix based on them. Using these charts we will be able to draw conclusions about the whole of the datasets and it would be possible to see the trends in the dataset as whole.

4. What Result I plan to get

From this visualization I plan to find trends in comic book characters. I will try to answer questions like what is the majority of the comic book characters strengths. How many of them are evil vs how many are good? Based on gender does the alignment of the character is biased to one particular spectrum of the morality.
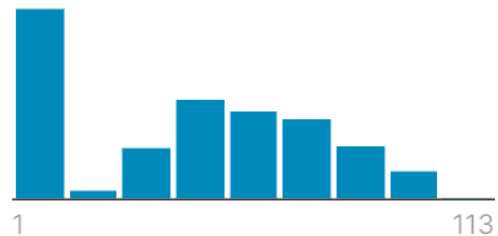
From the social network graph I plan to find which characters in the marvel universe are most central to it. What are the cliques or social groups among the superheroes?

5. Current Progress:

I have downloaded all the datasets that I plan to use and cleaned the dataset of invalid entries and entries having empty entries. I have also started the medium blog containing periodic updates. Preliminary data analysis. I have generated some of the preliminary charts to see how the data looks like.
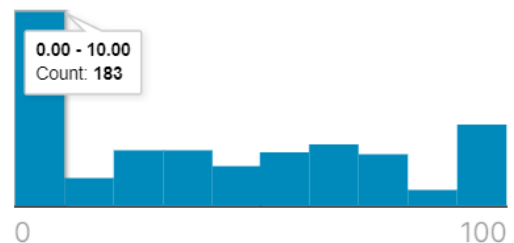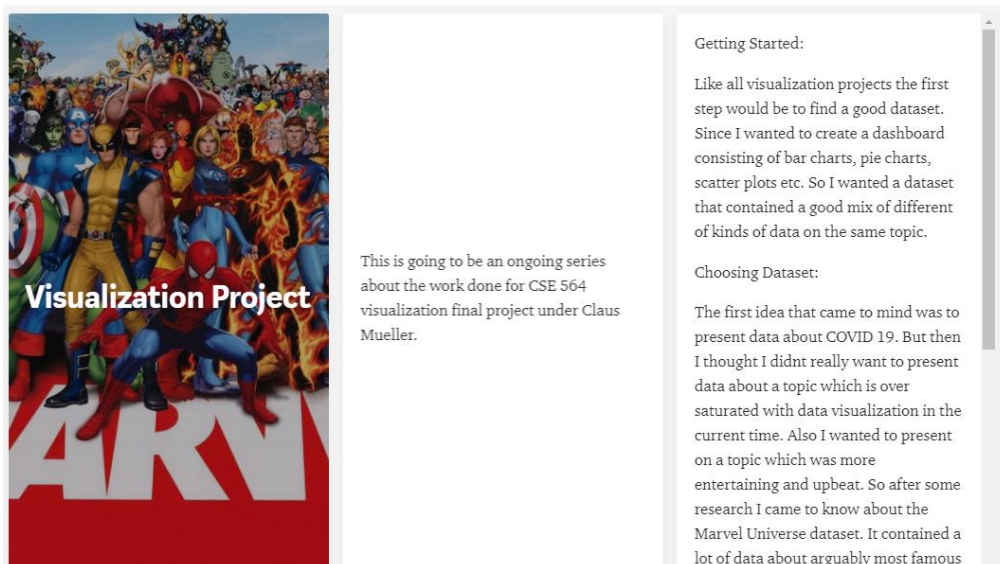

This is the intelligence rating of different comic book characters.

This is the power rating of different comic book characters.

Screenshot of the medium blog

5. References:

1. https://www.kaggle.com/csanhueza/the-marvel-universe-social-network

2. https://www.kaggle.com/dannielr/marvel-superheroes#marvel_dc_characters.csv

3. https://www.superherodb.com/characters/

4. https://developer.marvel.com/account