CS/INFO 3300 - Project 2

Project Documentation on "Visualizing Game of Thrones"

Team members

Allie Meng (awm79@cornell.edu)

Giri Kuncoro (gk256@cornell.edu)

Josselyn Tsai (jt566@cornell.edu)

1. Data Description

Data Collection

Game of Thrones (GoT) is an American fantasy drama television series that is very popular nowadays. Using the [BeautifulSoup](#) package, a library to pull data out from HTML files, we wrote Python script to collect the major characters of GoT data from its [wikia](#) web pages into a csv file. In the end, we collected total of 37 major characters and 8 additional characters that are quite important in the story. The data includes the character name, allegiance, alive or dead status (death is one of the most important events in GoT), causes of death, number of appearance, origin, religion, culture and portrayal.

Data Cleaning

We then added the gender, link to their profile image and manually cleaned the raw data as below:
- Removed the comma separated number for season data
- Removed the "(see below)" hyperlink for appearance data
- Removed "house" words for allegiance data

We initially thought putting age, religion and culture of the characters as the main nodes in visualization will be great, but not all characters are provided with this data on wikia. We also thought putting the killer of the characters will be interesting, but we decided not to since it will make the edges line more complicated among characters.

2. Visualization Description

Static Features

We mapped the data from the wikia to visual elements in three main ways: color/color gradient, relationship paths, and scaled tick marks (bars). In our main circle graph, each house, cause of death, and character gets its own tick. Houses are green and the size and shade (color) of the tick mark denotes the number of major characters in that house. Causes of death are red and the size and shade of the tick mark denotes the number of major characters that died in that manner. Characters are blue and the size and shade of the tick mark denotes the number of episodes that character has appeared in. You can see the exact value of these tick marks by using the color gradient keys provided or by hovering over the tick mark with your mouse.

Every character has paths to the houses they have allegiance to and a cause of death (if they are deceased). The color of these paths run on a gradient from the color of one tick mark to the other. This way, you can visually see the most important relationships more prominently than less important ones. For example, paths from House Baratheon (a major house containing a lot of key characters in Game of Thrones) are more prominent than paths from House Tarth (a minor house that one character is originally from but is not important to the story) because of the darker color associated with houses that have a higher number of allegiances.

Dynamic Features

By hovering over a house or cause of death, you can see the exact number of characters associated with that house/cause of death. You will also see all paths that connect characters to that house/cause of death (all other paths will disappear). Similarly, if you hover over a character, you will see a small popup containing some information about that character: name, picture, status (dead or alive), cause of death, number of episodes they appear in, seasons they appear in, house allegiances, culture, religion, and actor who portrays them. You will also see all paths from that character to houses/cause of death (all other paths will disappear).

In addition to our circle graph, we also included two small widgets on the left of the page. The first one has a graph that show the number of deceased and the number of alive characters (percentages are shown). If you hover over these two bars you can see all the characters who have died or are still alive on the graph (equivalent of hovering over all the causes of death tick marks at once). You can also click the button under the graph to show all deceased. The second widget is the same as the first but shows male/female characters and highlights male/female characters with their house allegiances.

3. The Story

Our data is from the Game of Thrones wikia, which is based off of the television series on HBO (not the books), and is updated to where the series is in real time. (So a character who maybe eventually dies in the books but is still alive thus far in the television series will have a status of "alive.") This also means that some characters are different (Talisa Stark does not exist in the books).

That being said, in addition to organizing lots of character data, our visualization answers interesting questions about the television series such as: How many major characters have died thus far? Is there a large gender discrepancy in Game of Thrones? Which houses have the most major characters?

The results are pretty much what we expect from Game of Thrones. That is, nobody is safe, since the deceased includes a huge number of important characters, and that while there are twice as many male characters as female characters (a characteristic of the time period of Game of Thrones), there is still a huge number of strong female characters. Despite being outnumbered two to one, many of the top/most occurring characters are female.

Lastly, we included some text around the graph and in the two widgets to balance the visualization aesthetically and provide some light hearted fun facts!

References:

[1]
http://gameofthrones.wikia.com/wiki/Category:Major_Characters?display=exhibition&sort=mostvisited

[2]
http://www.theguardian.com/world/interactive/2013/apr/30/violence-guns-best-selling-video-games