Kael Ecord

Intro to Informatics

Professor Leon Johnson

October 18th, 2023

Project Proposal

Option 1 – Friends TV Show Data:

The Friends tv show is one of the most popular tv shows of the early 2000s. The data that I would like to look at here contains information on every line said throughout the 10 seasons of the show. The goal of this project would be to create various data visualizations showcasing the most popular characters and who they are most often speaking to and about. The show features the same main characters in all the episodes, so I am curious to investigate what characters have had interactions with characters that the other characters haven’t. Another goal aside from building a network of characters in the show would be to use NLP to look at classifying the emotion of each line (utterance) in the show and use this to find the characters that display each emotion the most. The goal of the web app would be to make this investigative experience interactive for the user by providing various visualizations that can be filtered and clicked on to gain more specific information that they are looking for.

[character-mining/json/friends\_season\_01.json at master · emorynlp/character-mining · GitHub](https://github.com/emorynlp/character-mining/blob/master/json/friends_season_01.json)

Example of what the json data looks like:

A white screen with black text

Description automatically generated

Option 2 – NBA Shot Location Data:

This data set contains shot location (x and y coordinates) for the NBA from January 1998 through the 2019 season. The goal of this project would be to use this information to create shot chart visualizations for teams’ players and games. Doing this in a web app would allow the user to filter down to see information for specific matchups, specific players, or even league wide for certain seasons. This data is valuable for coaches who are scouting opponents and evaluating their own team’s performance. If you can find weaknesses in your own defense, then you have a better chance at being able to fix the issues before the next game.

Data Documentation can be found here with what an example shot chart might look like:

[sportsvizsunday/june-2020-nba-shots-1997-2019 | Workspace | data.world](https://data.world/sportsvizsunday/june-2020-nba-shots-1997-2019/workspace/project-summary?agentid=sportsvizsunday&datasetid=june-2020-nba-shots-1997-2019)

The csv is pretty large so this is a screen shot of what the data looks like:

A screen shot of a computer

Description automatically generated

A screenshot of a table

Description automatically generated