Antonio Fernandez Jr,

NewM-N328 Visualizing Information

Final

Netflix TV vs Movies and MLB Stats

So initially I wanted to do my final project on a dataset from kraggle.com that was over Netflix, <https://www.kaggle.com/shivamb/netflix-shows> , here is the link for download. One major issue I found when working on the data set to get it ready for presentation one was that my entire data set was mainly qualitative data. This made it difficult for me to really graph anything without adding descriptive statistics or some other model of analyzation to the csv file itself before even coding it in d3. Below is a screenshot of the Netflix TV vs Movies.

A picture containing calendar

Description automatically generated

My next frame of thought was to completely start from scratch and get a new data set. So, I went back to kraggle.com I looked from some subject I knew would have quantitative statistics. Sports was the first topic that came to mind. Once I had selected sports I focused my attention to baseball. I looked up baseball dataset and found the following link; <https://www.kaggle.com/open-source-sports/baseball-databank> , The dataset has over 20 csv files so I decided to go with one specific stat/csv file, which was Home-Runs. A screenshot is available below

Table, Excel

Description automatically generated

As you can see I have a lot to work with. So my initial plan was to make a scatter plot and see where I can add interactions. But when I started coded I had some rough patches, I included designs called MLB V1 and V2 which are listed below and on my github Repository (<https://antferna.github.io/N220Fall2020/> ).

Chart, bar chart

Description automatically generated



As you can see based on my first version (MLB V1) I could not get the table to show properly but the fact that it had data populated was a good sign to so keep working and referring to d3 Book for help. Once I got to version 2 (V2) I was too close to my final version, but I had some kinks like the axis’s being on the wrong side. But as you can see, I limited the domain and range for the x and y axis and needed to include the title as well.

**Findings**

Based on my final version located at (<https://antferna.github.io/N220Fall2020/Assignments/Project%206%20-%20MLB%20Stats%20V3/index.html> ) I included an animation as my interaction as well when you hover over each circle it tells the player id, how many home runs he hit and the year he hit it in (all in the developer tab) Once hovered over the player’s circle goes back to normal.

Some interesting findings I had noticed is that there is a large number of people hitting home runs in between 1980 and the early 2000s this coincides with the steroid era in baseball so more players at the timer were hitting the ball further than ever before.

Another interesting finding was the low number of homeruns during the wars in the 1900s, this was due to the low number of men playing because of the draft.

TChart

Description automatically generated

The link to my viziualtion is below:

<https://antferna.github.io/N220Fall2020/Assignments/Project%206%20-%20MLB%20Stats%20V3/index.html>

The link to my YouTube Video is listed below:

<https://www.youtube.com/watch?v=F5tn8Ih75yM>