Maturity Through the Years: A Visualization of Netflix Maturity Ratings

- **Basic Info.** The project title, your names, e-mail addresses, UIDs, a link to the project repository.

Repo Link: https://github.com/Nathan-Joe/dataviscourse-pr-FlixHunters

Name: Gregorius Jonathan Sukiman
Email: u1346718@umail.utah.edu
uID = u1346718

Name: Tayler Tolman
Email:u0964081@umail.utah.edu
uID = u0964081

Name: Albert Frei
Email u1278125@umail.utah.edu
uID: u1278125

- **Background and Motivation.** Discuss your motivations and reasons for choosing this project, especially any background or research interests that may have influenced your decision.

As we are all movie and tv show enthusiasts the topic of collecting media data and sorting it was of interest to us. While rooting through different sources of data we stumbled upon a cache of Netflix history. Netflix has recently been in the news and a hot topic among the streaming community. To this day it is probably the most popular streaming site of all time. Along with that we want to see how maturity rating has changed over time in Netflix, and if certain directors/show creators are tied to certain genres/maturity ratings.

- **Project Objectives.** Provide the primary questions you are trying to answer with your visualization. What would you like to learn and accomplish? List the benefits.

With the current data we want to answer the following questions: How has the composition of maturity ratings changed throughout the years? For example, have rated R movies increased in volume recently? Do certain directors feel tied to a genre/maturity rating? Some benefits include:

Cultural knowledge of the history of maturity ratings. Historic trends of streaming media that may be able to predict future trends. Knowledge of the most prolific directors and their means of success.

- **Data.** From where and how are you collecting your data? If appropriate, provide a link to your data sources.

For data, we are going to collect our data from kaggle (https://www.kaggle.com/datasets/shivamb/netflix-shows). This website will give us the data in a csv file which we can download.

- **Data Processing.** Do you expect to do substantial data cleanup? What quantities do you plan to derive from your data? How will data processing be implemented?

Since we have a couple of data points which are null, we might do some data clean up so that these values don't cause us trouble. The format of the dates varies which we will have to clean up. Sometimes it is listed in the 'YYYY-MM-DD' format and other times the format is 'Month name, day, year'. We will need these to be consistent in order to work with them. We will need to transform the cvs values into Movie and Director objects. As for fields we are planning to use we will need movieTitle as the key for movies, directorName as the key for directors (not every movie or tvShow has a listed director). Year added is essential for our time scale. The rating field will be used to scale the maturity rating over the years and to see which show genres line up with which ratings.

- **Visualization Design.** How will you display your data? Provide some general ideas that you have for the visualization design.

Essential to our design is showing the relationship between time(in years) and the increase or decrease of certain maturity ratings. Additionally we need to display certain director information tied with their previous work in some way. The designs should work together and interact in some way with each other.

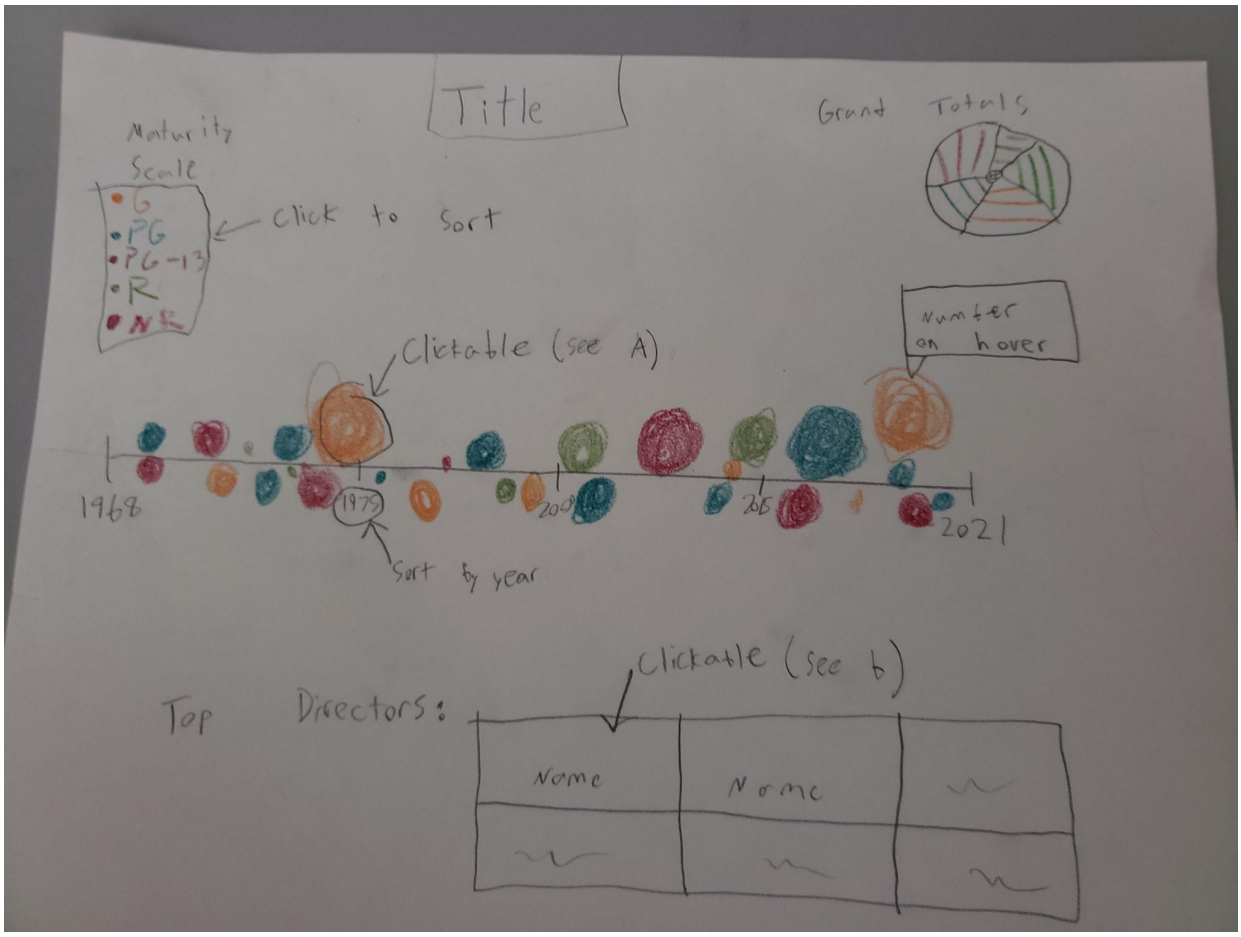
- Develop **three alternative prototype designs for your visualization**. Create **one final design that incorporates the best of your three designs**. Describe your designs and justify your choices of visual encodings. We recommend you use the Five Design Sheet Methodology.

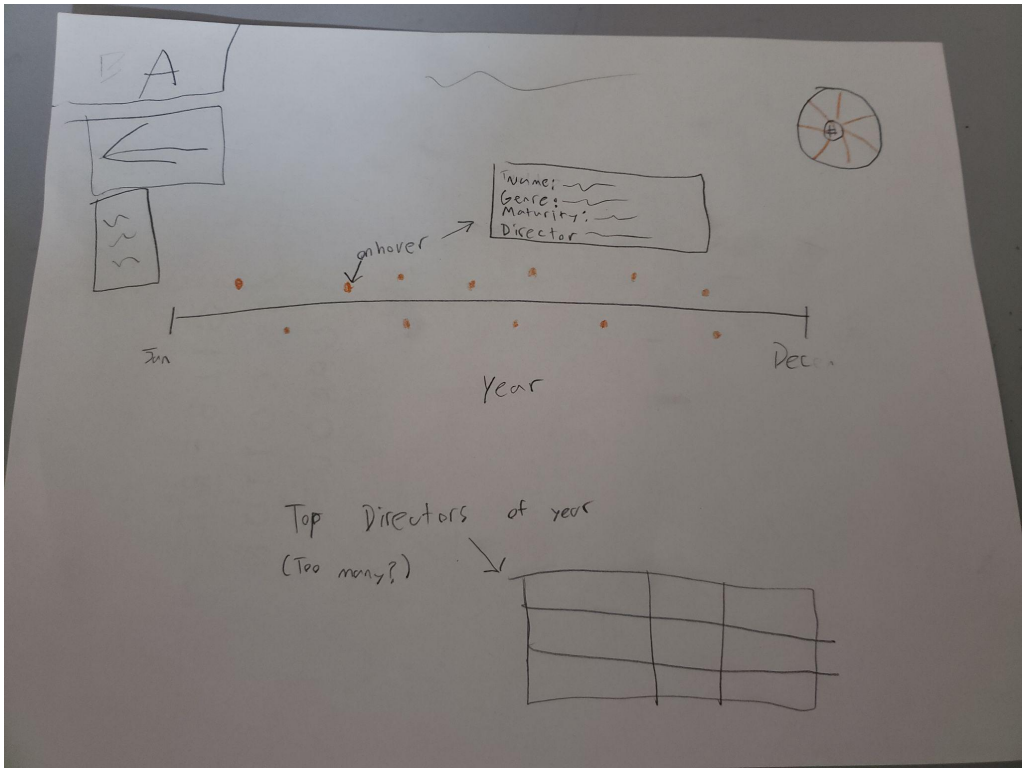
Design 1:

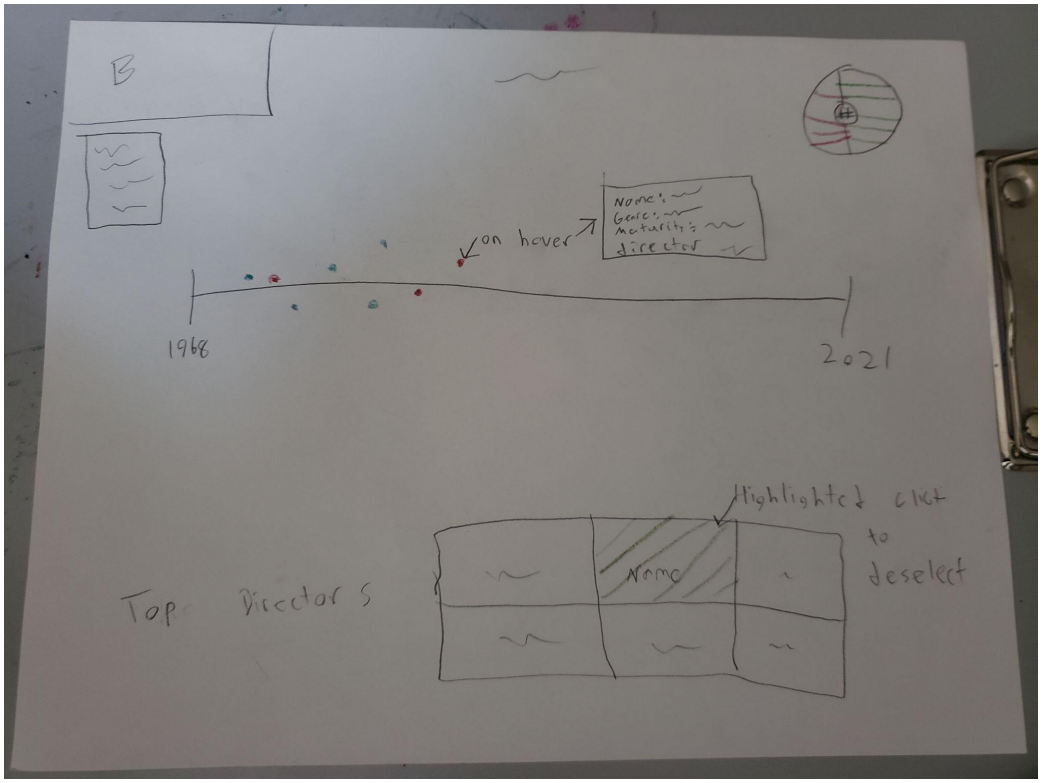
The first design features a world map on the left side. Here, we can use filters to pick what we want to show in the world map. If we did not pick anything, the world map will represent both movies and tv shows but it will be different by using color scale. If we pick one of the filters, we would most likely use dark colors to represent the highest count and light colors to represent the lowest count. As for the table on the right side which is the number of contents by country. If we do not pick a filter, that is what we are expecting to show, in which we could see which country has the highest contents or we could see how many TV shows or movies in each country. But if we pick one of the filters, it will just show the movies or the TV shows. As for the bottom left table, this will have a couple filters which will represent the genres, directors or even ratings, this could be in terms of both movie and TV shows or it could either be one of them. This also has some additional features, that is when we click

a certain country in the map, it will represent only the statistics that we have for that country. And another feature is that when we click the bar on the bar chart table, we will show some description of the movies. So, the users could see what the movies are about without having to google it up first.

Design 2:



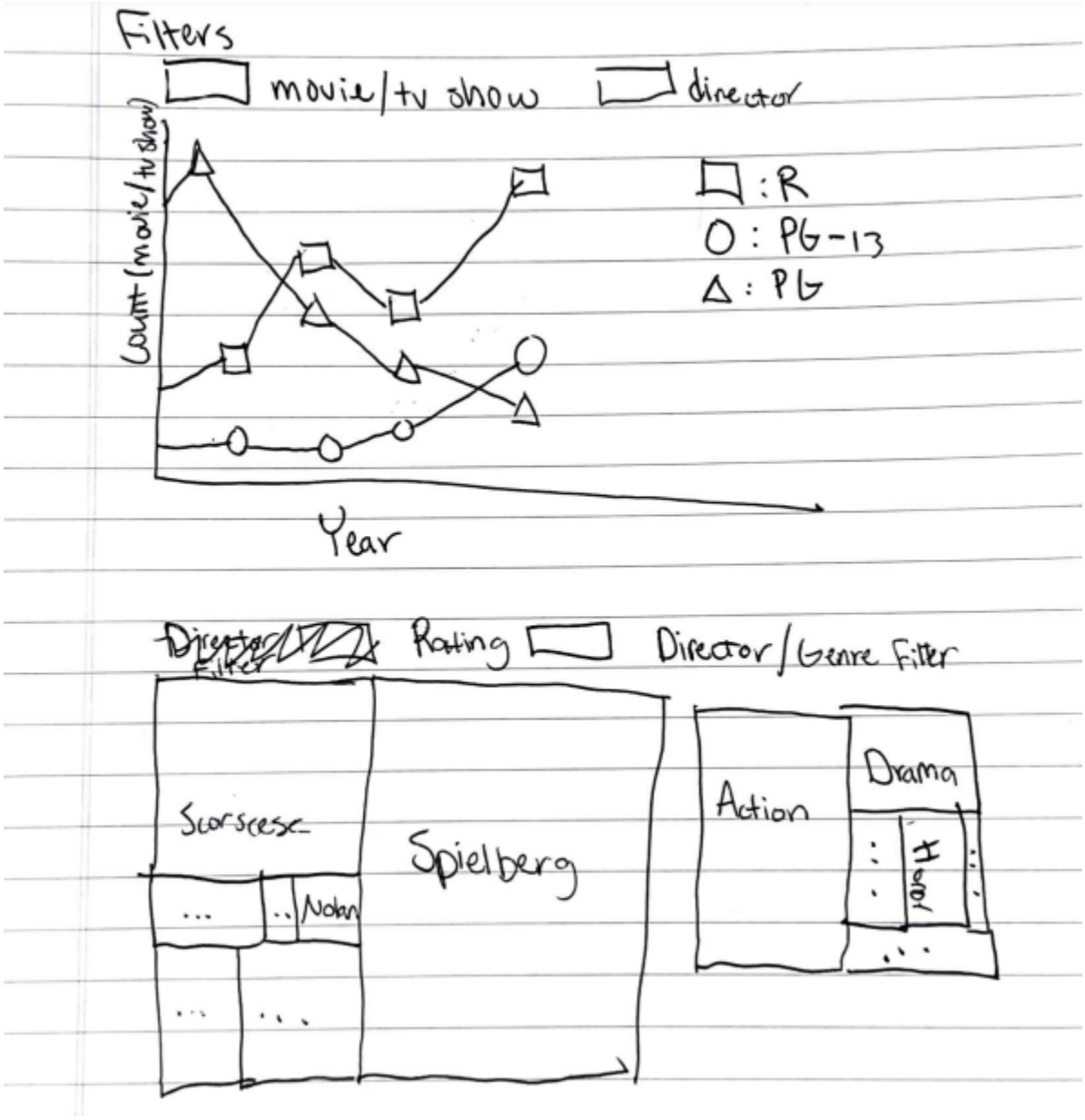




Design two features bubbles that represent each maturity rating throughout the years. The bigger the bubble the more of a single genre was created that year. Clicking on the legend will sort by maturity. Top directors of all time are listed at the bottom. Selecting a certain name will highlight that

directors films across the years. At the top right is a context sensitive donut chart highlighting what is shown in the bubble graph. Tooltips provide further insight on highlighting indicated movies or tv shows.

Design 3:

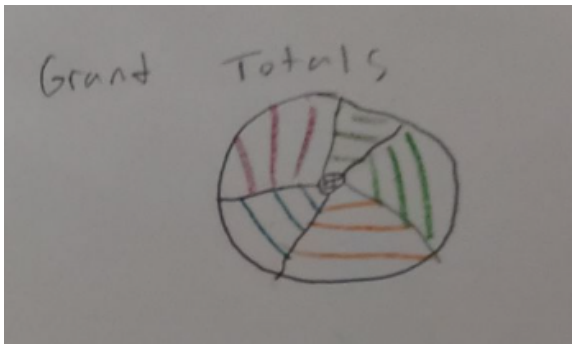


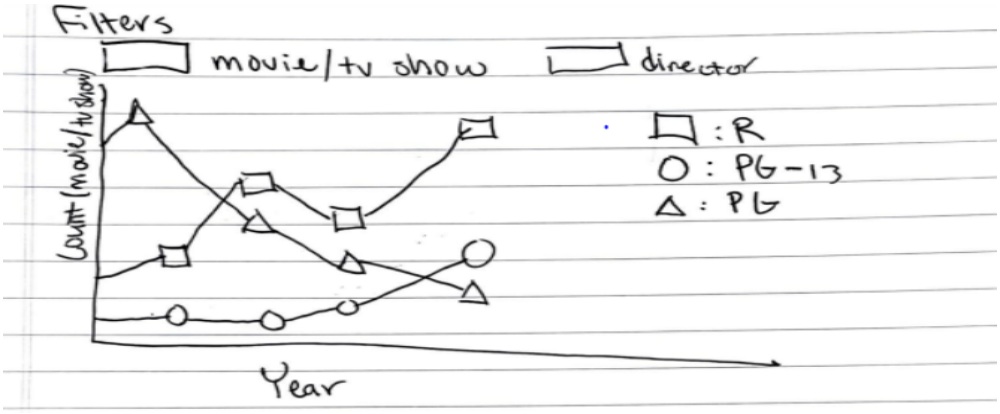
The third design uses a line graph to show the trends in different ratings over the years on Netflix. There are filters that allow the user to switch between seeing trends in movie ratings vs TV ratings and also a filter that would allow the user to filter on a specific director to see how a particular directors ratings have trended over the years. This is assuming there are enough data points for single directors. Below that there is a treemap that will show the part per whole of a single rating

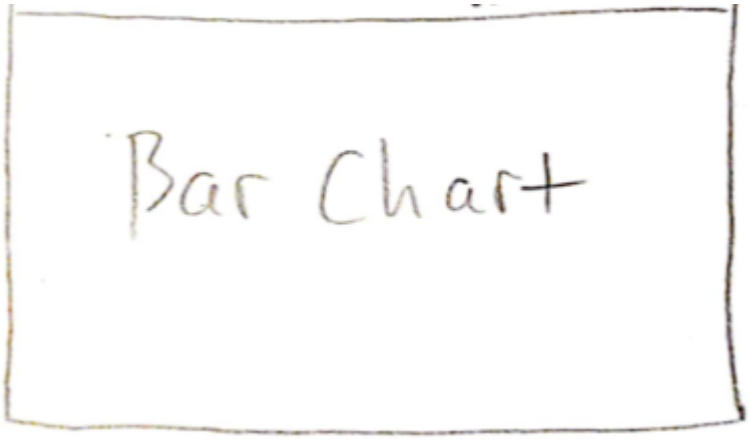
made up of directors and also the option to switch to a single rating made up of genres. This way we can ask questions like which director has the most R rated movies on netflix or which genre has the most PG-13 ratings. The filters would allow the user to select a rating and whether or not they are viewing that rating by director or by genre.

Final Design

Title: Netflix Maturity Through the Years







: X Director Y Number of Movies, each single bar color separated by number of movies with certain maturity rating

This concept takes some of the best features from our three prototypes. First the line graph. This will show the count of each maturity rating in each year shown. The graph will be double encoded, with line color and shape indicating the maturity ranking for each year. Sorting here includes deselecting or selecting parts of the legend and further year sorting. Secondly the directory bar chart. This will be a simple bar chart showing the most prolific directors. Each bar will be further separated by color indicating the amount of each maturity rating consistent with the director. Sorting here would likely involve showing the authors movies on the line graph. Lastly, the context sensitive donut chart in the top right. This will adapt to filtering on the director bar chart as well as the maturity rating line graph. This will show totals with the color encoding in mind and the grand total in the middle of the donut chart. Possibly it will also allow filtering by clicking on it. We will also include context sensitive tooltips displaying movie information on hover if time allows.

- **Must-Have Features.** List the features without which you would consider your project to be a failure.

Maturity distribution line graph, showing the increase or decrease of maturity ratings throughout the years. Director connections with maturity ratings, including highest number of movies made by director and maturity ratings common among directors.

- **Optional Features.** List the features which you consider to be nice to have, but not critical.

Addition of tv-show maturity ratings and data. Showing maturity trends by director. Website links to movie reviews and search field added. Tooltips for individual movie information. Further interaction to show specific movies/directors along the line graph.

Schedule:
Week 1: Data Wrangling, initial set up of web page with basic html js css

Week 2: Addition of line graph, including legend

Week 3: Addition of director bar graph including legend

Week 4: Addition of context sensitive donut chart, filtering added

Week 5: Add interactions between all graph elements, including tooltips for sorting on movie elements.

Week 6: Padding to fix bugs and polish features. Gap time to include optional features if time allows.