CS416 Narrative Visualization Project Essay
Lydia Brothers
07/28/2024
URL: https://lydiabrothers.github.io/index.html
https://github.com/LydiaBrothers/LydiaBrothers.github.io/

Message:

The purpose of this visualization is to show how the quality and quantity of Netflix Originals have varied over time and across different genres. I found it particularly interesting that documentaries have been prevalent and highly rated throughout 2014 to 2021. Additionally, although the quantity of Netflix Originals has increased and the content more varied, between 2019 and 2021 the overall IMDB score quality appeared to drop. This narrative is explored through three scenes. Through three scenes, initially the overall distribution of IMDB score is explored by genre, then the change in quantity of highly rated produced Netflix Originals is explored by year, then each data point is listed on the last graph for a deeper dive into the overall quality of the Netflix Originals generated each year. Through the annotations and descriptions, the audience is encouraged to draw their own conclusions from the data while some conclusions are provided by the author as well.

My dataset used for this project is based on:
https://www.kaggle.com/datasets/luiscorter/netflix-original-films-imdb-scores/data

The visualization can be viewed with the following URL:
https://lydiabrothers.github.io/index.html

Notes: It was tested in Google Chrome and Firefox. If your browser is resized mid visualization, you will need to reload the page to see the page update to the new dimensions.

Narrative Structure:

The narrative structure I selected was an interactive slide show. My visualization has navigation forward throughout the story, dissecting different dimensions of the data for insights. Each of these pieces has interactivity but is walked through one scene at a time. Next buttons allow the user to choose when they want to proceed to the next step of the narrative. Through the use of hovering with tooltips, annotation, and parameter filtering, additional insights can be gathered by the user at will along the guided narrative. Tangents can be investigated and details can be drilled down along the way, but moving from slide to slide will provide you with some key takeaways from the author needed to understand the next visualization.The opportunity to drill down and explore is before navigating to the next scene, as expected in an interactive slide show.

Visual Structure:

Scenes are crafted to be consistent for easy navigation and follow an order to encourage exploration. Each scene begins with some insights from the previous page, continuing the narrative thread throughout the story. The background color is black for each scene, and text and annotations are written in white. In the last scene, when color represents the range of IMDB score from good to bad, it is logically displayed with colors green, orange, red. A description is also given on each page about the type of graph, axes, and other quantities of interest to ensure understanding. A title gives additional high level context about the purpose of the scene. All buttons to move forward or back in the narration have consistent styling for quick understanding. Annotations also provide further insights into the creator's intended narrative and are used to guide the user. The visual structure ensures the viewer can understand the scene by being clear and providing instruction where needed to keep things simple while allowing the flexibility to explore relationships. The visual structure encourages the viewer to focus on important parts through the use of color and interactive triggers. This keeps the user engaged and in-tune with key attributes that influence the narrative. It helps the viewer transition to other scenes by having clear labeling on how to move to the next scene with an intuitive button.

Scenes:

For scene 1, the visual structure is a histogram which can be filtered on genre by the user. It is a clear overview of the kinds of IMDB scores Netflix Originals over the whole time period have received. The axes are kept static to better understand the scale of the subgenres of the data as a whole, leading the user to explore potential cases of Simpson's paradox. By understanding how subcategories of the data behave, we can make better conclusions about the data as a whole. This is scene one so the user has a baseline understanding for future scenes.

For scene 2, the structure is a bubble chart of the count of good (IMDB score >= 7.0) films premiered by that year, where the count is related to the size of the bubble. A slider is used to see how many good Netflix Originals had been created by that year, accumulating over time. Annotations provide additional information on this scene into the relative increases of particular genres year over year. This scene is provided second to give the user a better look into the growth of high quality Netflix Originals to grow appreciation for how the scaling up may have impacted the quality and to observe an increase in variation.

For scene 3, the structure is a violin plot mixed with individual observations represented as a scatter. This provides deeper insights into which specific productions are being considered in this analysis of Netflix Originals, as well as showing how the distribution of quality based on IMDB Score changed over time. This scene was left for last as the context gained from the last two scenes allows deeper exploration of the data. For instance, viewing the genre histograms irrespective of year, seeing the increase in count of certain genres, and overall observing the

scaling up of the program provide additional insights into the quality changes. This allows the user to investigate outliers in more detail.

Annotations:
- For Scene 1, annotations are used to highlight the median IMDB score (not always obvious using the histogram). When the user makes a filtering selection, the annotation will shift to providing the median for the genre selected. A tooltip is also used for each bin in the histogram to provide more clear information. All tooltip information follows a consistent template. This supports the message by helping the user understand that the median shifts based on the genre of production, either pulling up or pulling down the overall IMDB score. The annotations will change within this scene whenever the state changes.
- For Scene 2, annotations are used to describe the difference in count of Netflix Originals by genre between the previous selection and the current selection on the slider. This can be used to compare relative increases in genre for highly rated films. This is done for three genres: Documentary, Comedy, and Drama. A tooltip is also available to view the information provided by the bubble using a consistent tooltip template. This supports the message by showing there is great variation year over year and also influential growth which can impact the quality of the productions. The annotations will change within this scene whenever the state changes.
- For Scene 3, a tooltip is available which shows ample information about each individual observation, including the genre, runtime, and language. This supports the message by allowing deeper investigation into outliers and individual films and how that might impact the overall fairness to assess how 'high quality' Netflix Original productions have been.

Parameters:
- For Scene 1, the parameter is the genre. When the user selects a different genre, the data will be filtered down to the appropriate information of interest. The current state of the visualization is controlled by this parameter. By default the state begins including all genres, and then the user can drill down into 17 specific genres which will generate the corresponding state.
- For Scene 2, the parameter is the year. When the user uses the slider, the data for that year will accumulate and be added into the view for perusal. The current state of the visualization is controlled by this parameter, and by default it begins at 2015 as this is the earliest time period with multiple bubbles to display. The states then shift from 2015 to 2021 based on user selection.

Triggers:

One of the methods of interactivity with this narrative visualization are the 'Next' and 'Previous' buttons the user is able to select when they are ready to continue with the narrative. This allows them time to explore while separating the narrative into manageable chunks. This is an

affordance because they are clearly and intuitively provided to the user so they can navigate to the next scene.

Furthermore, in each scene the user is able to use tooltips by hovering to review information given in the visualization as well as additional context. This is an affordance that helps the user understand data elements more clearly.

In Scene 1, the state can be changed through a drop down list that the user can manipulate. In Scene 2, the state can be changed through a slider which adds additional yearly information into the view based on the year the Netflix Original premiered. These are both affordances that clearly indicate that there is variation to be observed through this trigger.