INFO 4310

HW2 Writeup

February 24, 2025

Team Members:

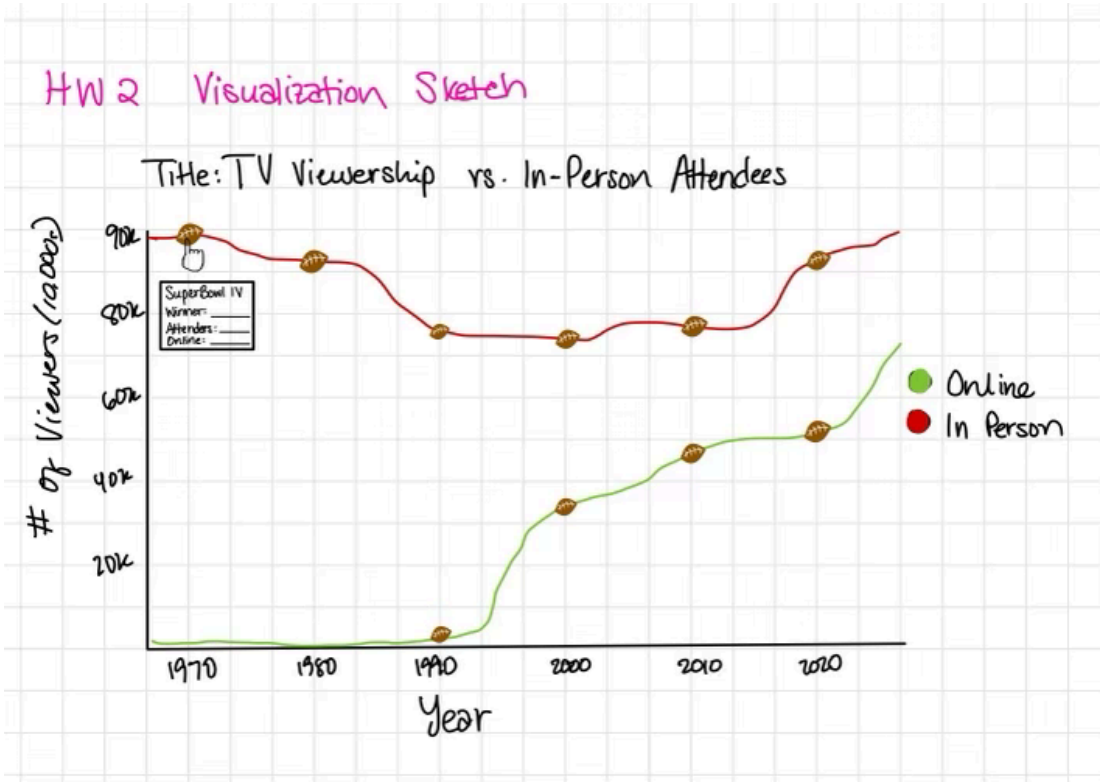
Celeste Naughton (can82), Ming Demers (mtd64), Goretti Muriithi (gm552), Sander Rodman (ahr68)

Dataset Exploration

Our project focuses on the history of the Super Bowl and how key features such as attendance and cost have changed over the years. Knowing the Superbowl is a massive event that brings worldwide attention, especially for the halftime show and extremely high ad prices, we wanted to examine the data surrounding it. We believe that the Superbowl is starting to cater more towards the upper class than the average American Football fan and wanted to see how the data reflects that.

The dataset we chose to explore was primarily the “Superbowl Game Records” Dataset from Kaggle. We combined this dataset with Average Median Household Income data by year from Statistica and manually scraped values for data regarding the Average Cost of Tickets adjusted for inflation from *GoBankingRates.com*. This assignment came during a time after the 52nd Superbowl had just ended. Each row represents a Super Bowl game, and each column is a feature of the game, the Winning team and Average U.S. Viewers, for example. The time series aspect of our data was critical not only in conveying the overall trends of the Super Bowl but also in the interactions by allowing the users to hover on the line graph to see the statistics of a single Super Bowl and switch plots to see how different attributes compare to each other.

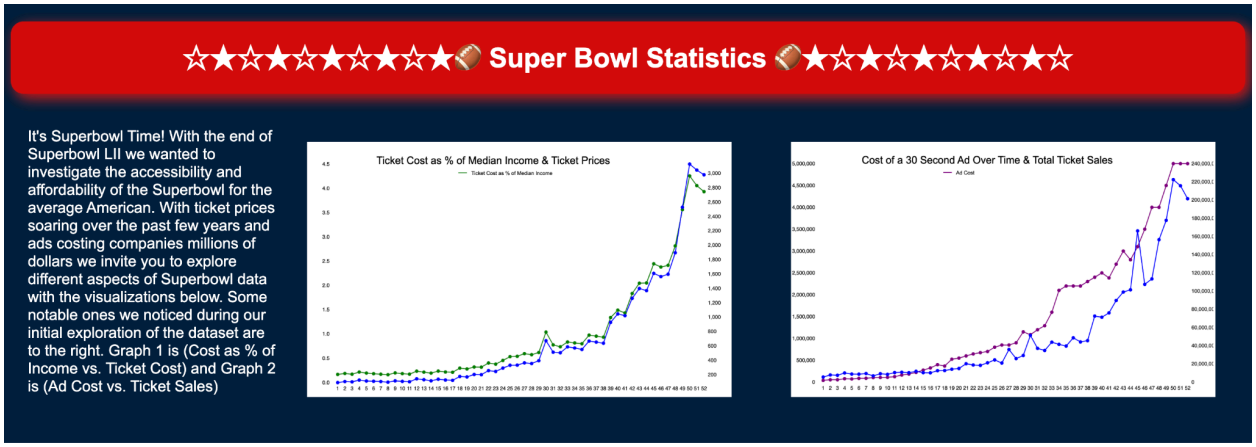
When planning our interactions, we felt that a line plot with the ability to do mouseovers and hover for different points would be the most ideal. This dataset is quite large and because it covers 52 years of history, it was beneficial to visualize it in this way. With the ability to hover and gain additional information about each piece of data, our visualization makes it easy to interpret a lot of information at once and quickly understand how they relate to each other.

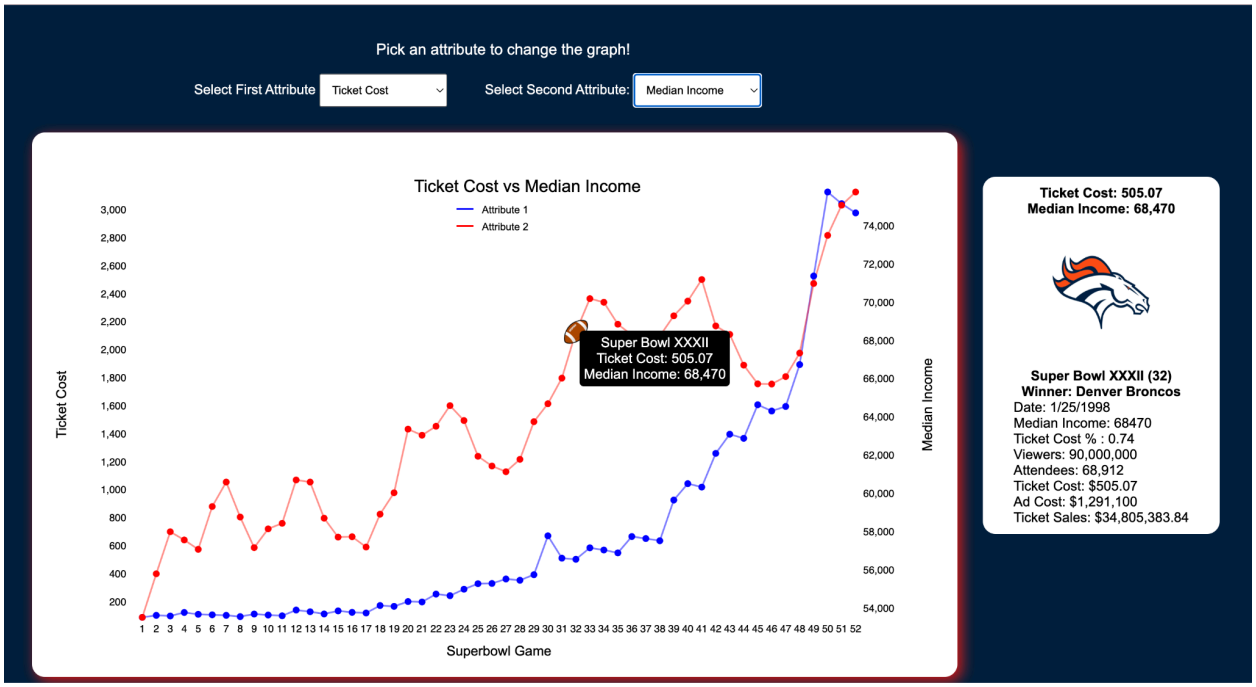


Final Visualization

Our final interactive visualization application allows users to select two different key attributes within the dataset and see how they are graphed for comparison to help further explore the dataset. To introduce the data, we display some static sketches that we discovered while initially exploring our dataset. This allows users to understand some of the unique trends and relationships within the dataset relating to Superbowl costs and affordability. The visualization also offers the ability to hover over data points with a single node being visualized as a football

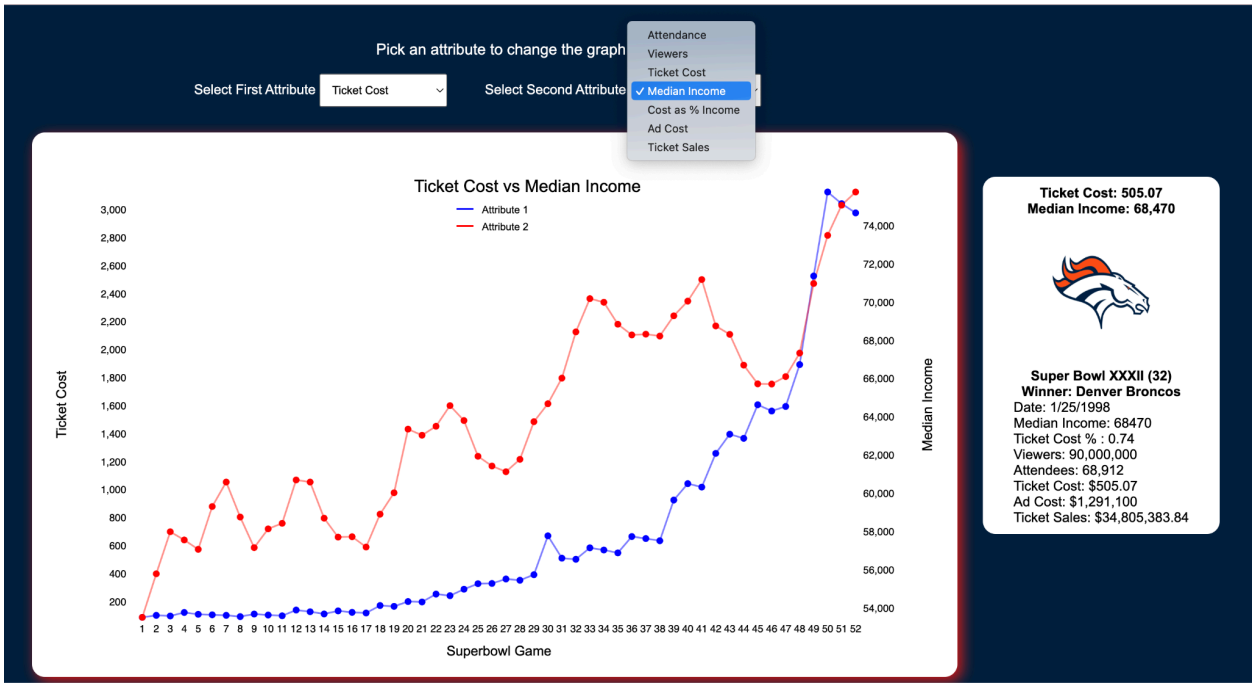
icon to learn more information about the two selected attributes. The hover then shows all of the related information of that node on the sidebar (Superbowl Number, Date, Median Income of that Year, Ticket Cost as a Percentage of Median Income, Total Number of Viewers, Total Number of In-Person Attendees, Ticket Cost, Ad Costs, and Total Ticket Sales) Users can switch between attributes through the drop-down box near the top of the application.





Tradeoffs

Though our interactions aid users in understanding the data and Super Bowl trends, there are some potential drawbacks. First, the interaction of choosing plots does not give the user much flexibility in that only two line graphs can be displayed at a time. Some users might want to display more than two line graphs at a time, and our current implementation does not allow that. If we continue this project, we could alleviate this problem by allowing users to choose the



number of line plots on the graph. In addition, our application does not allow users to change the colors of the line graph; this can be significant to those who are color-deficient and struggle with the default colors. We could alleviate this problem by allowing users to change the color of each line plot.

Other trade offs include some duplicate information within the sidebar preview of information. It was difficult to put the specific attributes chosen at the bottom in bold, so it is duplicated to highlight which attributes are being viewed for that particular data point. There could also be a tradeoff with having two different y-scales on the same graph. This is helpful for

quickly identifying trends in the data but doesn't aid users in diving into a specific attribute. Having two y-scales could be alleviated with gridlines, but that also risks cluttering the graph.

Developmental Process

When starting our visualization process, we initially did a visualization on two attributes, looking at the relationship between in-person attendance and US home television viewership. As we began to explore other aspects, such as Ticket Prices vs. Median Income and Ad Revenue the idea began that there might be a better way to visualize the different attributes. This led to the creation of an additional aspect of interactivity which allowed users to select attributes of choice to further explore the dataset and see different trends and relationships within the Superbowl Dataset. At times it was challenging to correctly align and style the sidebar that appeared for each datapoint but it proved to be a tremendously useful tool for the interaction. Having the logo of each winning Super Bowl team helps users to also make connections and more easily interpret the data, going farther than our original tooltip design.

Our final design does not deviate far from our original sketches. We kept a lot of the design choices, such as two-line graphs and the hover interaction. However, some of the design choices were changed in the final product. For example, in our sketch, we had the x-axis be the year of the Super Bowl, but in the final, we decided to go with the Super Bowl number instead. We also decided to focus on how each point is a single Super Bowl and not the year it took place. There was a significant tradeoff with the suggestion of gridlines. We encountered some difficulties with correctly aligning them in the remaining timeline we had to complete the assignment.

Project Breakdown

Celeste - Drop down selection interaction, Report write up, Ad Revenue visualization, Data collection, Sidebar Preview creation, Football icon hover interaction, addition of static images and introduction of data. (~8 hours)

Sander - Report write-up, sourced original datasets (~2 hours)

Goretti - Initial sketches, GitHub setup, webpage styling, first visualization (U.S. viewership vs. In Person-Attendees) & initial interactivity, Data collection. Spent the most time on visualization (~3 hours)

Ming - Dual Axes on U.S. viewership vs. In-Person Attendees visualization