Noelle Gray

Visualizing Information

Final Project

Description of Data: I am using a Kaggle csv file, named “All Playstation 4 Games” which contains about ten thousand games and about fourteen different and usable characteristics.

https://www.kaggle.com/shivamb/all-playstation-4-games

Outline

For the final project, I created an outline of topics that I would like to cover:

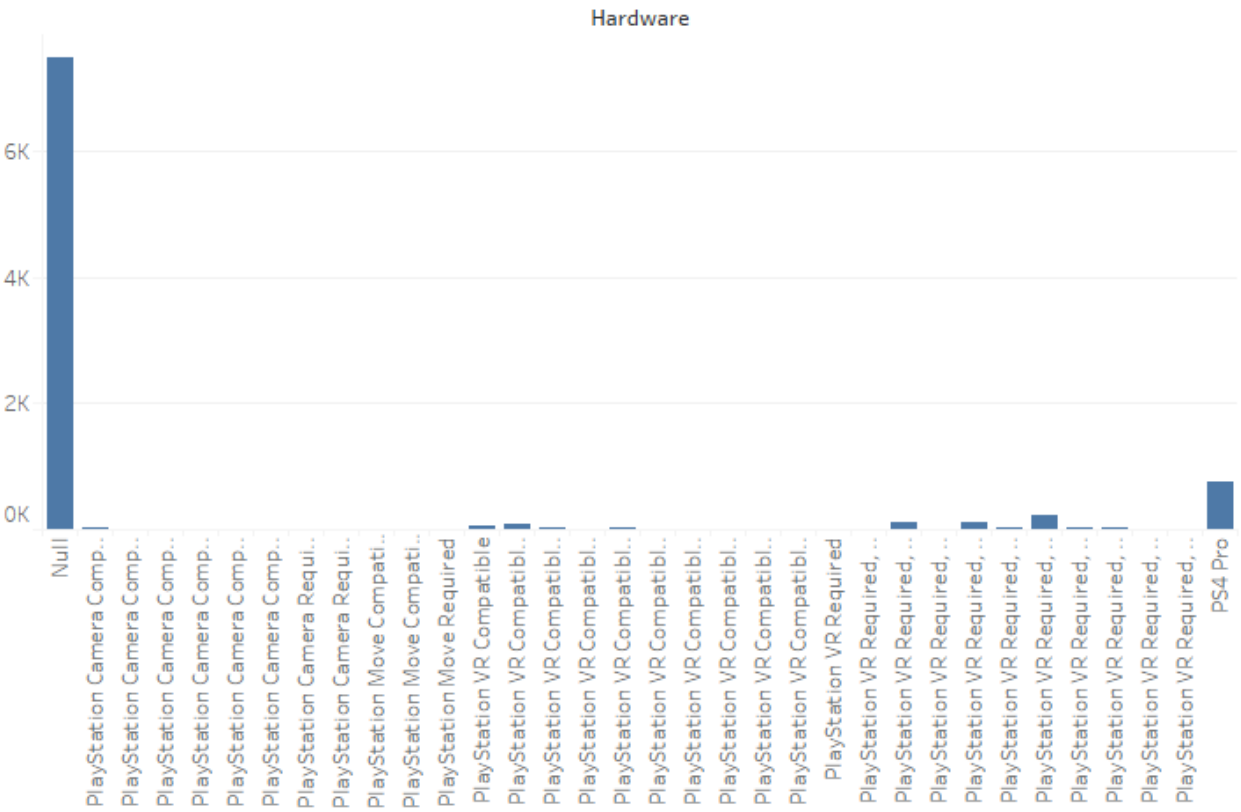
- Number of Releases Each Year From 2013-2021
- Physical vs. Digital Distribution of PS4 Games
- Number of Different Genres of PS4 Games

Design Process and Rationale

As I perused through the data, I noticed that many games had “null” values. This would be better illustrated when I took the data in Tableau to find out what visualizations would be feasible for me to create in D3, and which would not.

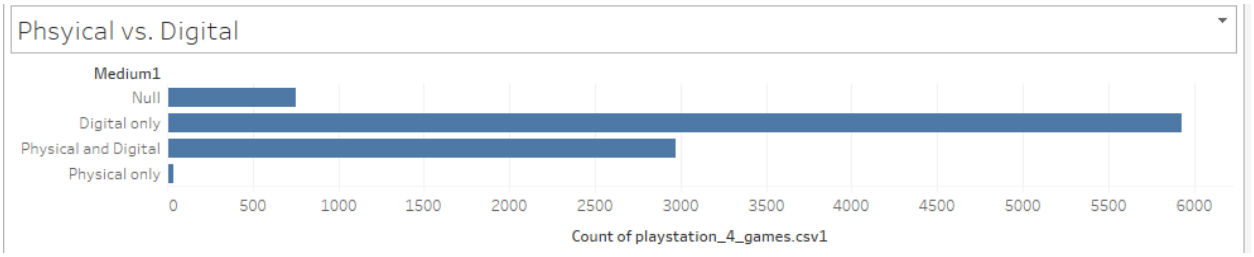
For example, I wanted to see what hardware each Playstation game was compatible with; however, when I made this graph, over 8,000 different games that had null compatibility with any of the hardware in the data table. Of course, this is not necessarily accurate, especially when the console of the “Playstation” is a category itself. I had to contend with this while I created the designs and for some of the characteristics of the games.

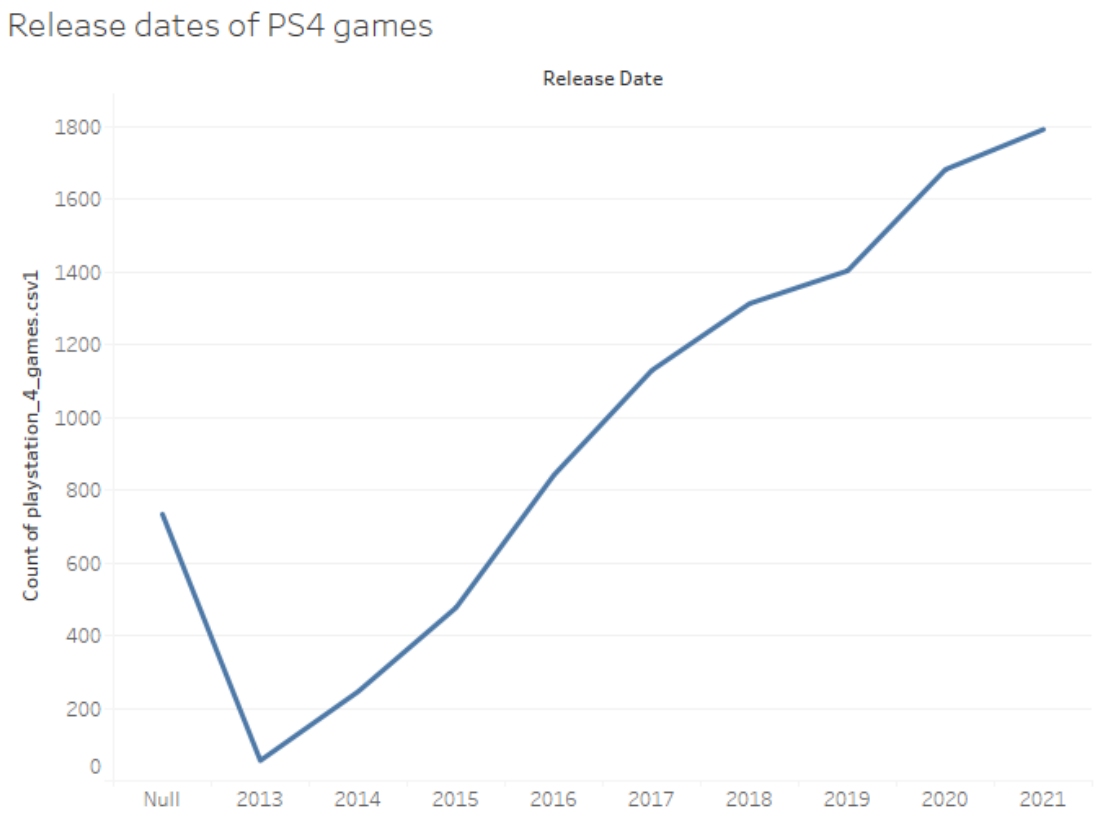
Example:



In the end, I came up with three different visualizations that help us explore some of the unique facets of the Playstation 4's lifespan, which have been listed above.

Example of Designs I wanted to Recreate:





As aforementioned, I began designing the visualizations that I wanted to create in Tableau since I figured this would be the most accurate way to test the data, see how missing data affected potential visualizations, and quickly find out which visualizations yielded the most interesting results. One facet of Tableau's resources that I took for granted was the ability to the easily sort through the data to obtain one column.

When I ported the data into D3, I had a really difficult time trying to recreate the tables I had created in Tableau through calculations only. For example, in the "Physical vs. Digital Distrubution of PS4 Games", I realized I needed to figure out which of the 10,000 games were physical, digital, both, etc. I assumed that meant I needed to run the data through a loop to categorize them, but from there, I couldn't figure out what to do. To circumvent this issue, I loaded the data into D3 manually rather than using a function to upload it. This way I would not have to sort through to find which categories I wanted to examine for each table; I could deal with the values I wanted.

Additionally, I manipulated some of the data in the genre table so that they were filtered by which genres had over 70 games categorized. This way, I could compare the larger numbers much more easily, as these were my focus. Many of the genres had subcategories like RPG, Turn Based, Strategy rather than just Strategy for example.

The designs of the visualizations did not change much once they were on the web, except for the line chart of the releases. I decided to make it a bar graph instead so the "null" value would be less out of place. I altered the color of the bars to be something that complemented the site's

design and would be dark, legible, and visible. I added a white stroke so each bar could be differentiated from each other. I also added in ticks and labels for each bar's categorization myself. Each visualization's ticks vary depending on the maximum value of the data.

I don't really have a specific order for the visualizations; I do not have them in a numbered list on the home page. A viewer is welcome to look at them in any way they would like. I think all of them display interesting information that does not necessarily overlap or coalesce with another visualization other than pertaining to PS4.

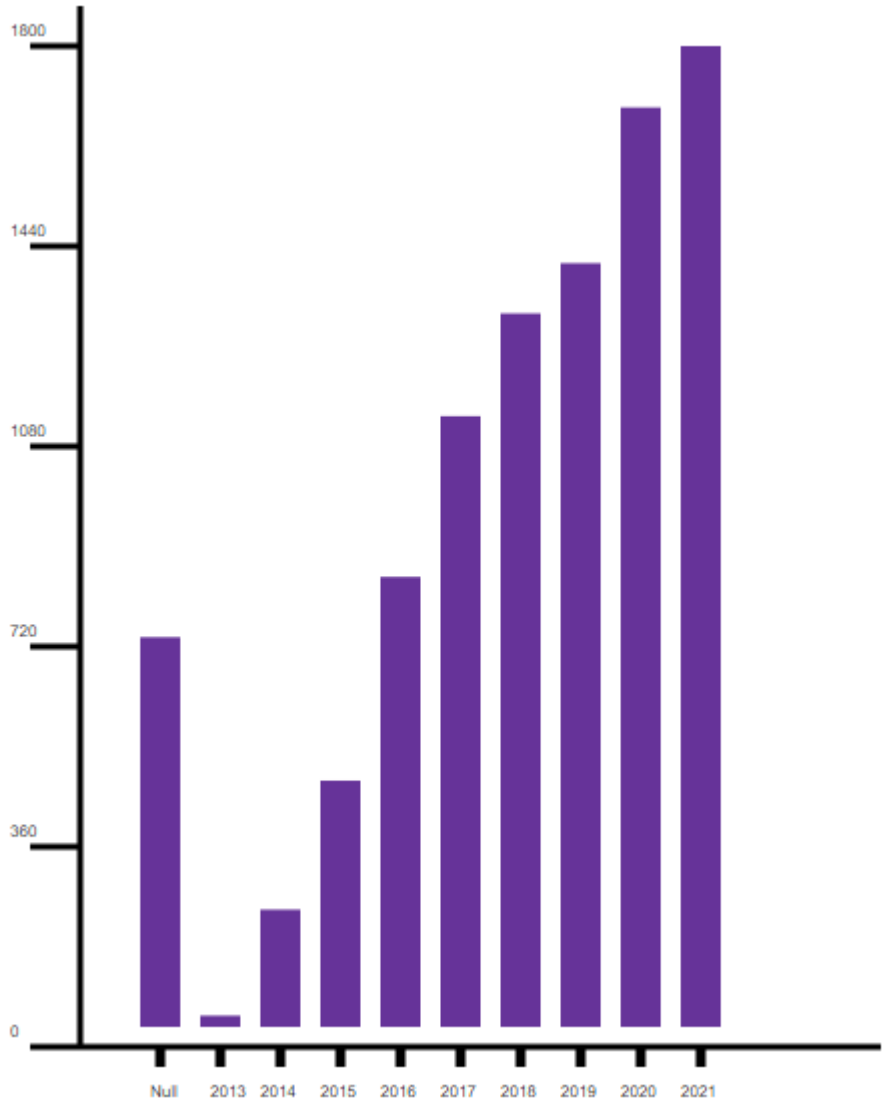
Findings

- Number of PS4 Game Releases Each Year From 2013-2021

The question that prompted this visualization was if the number of releases of PS4 games decreased due to the release of PS5 in 2020. However, that does not appear to be the case. In fact, we can see a gradual increase of releases year-to-year, even if it's not as dramatic. It only slags for one year in 2019. We can exclude the null, but I left it in the graph for a reason.

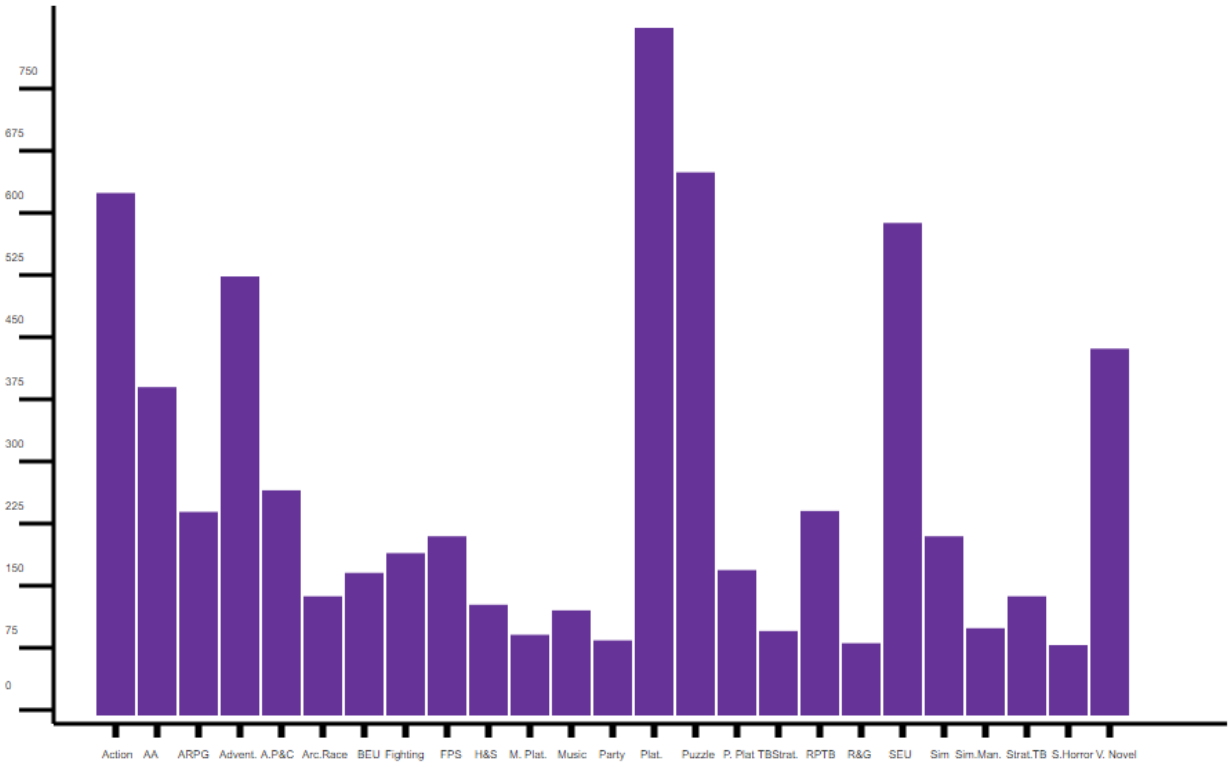
One of the other interesting findings from this graph—that I didn't really expect—was how many null values there were in the data set. There are about 700 of them. This could of course be interpreted as the author or compiler of the data not completing this section completely; however, I think there might be a more compelling solution.

On Playstation 4, especially in its collection of games that are available for digital sale and download, there are many games that were released many years before the PS4 ever came out. For example, the original Final Fantasy VII is available for sale on the Playstation Store; this game came out in 1997, and there are many other video games like this. I'd like to assume that this would be the case for many, if not all of the null value games, but this would take a thorough investigation of each game in the list that has a null value for the release year.



- Number of Different Genres of PS4 Games

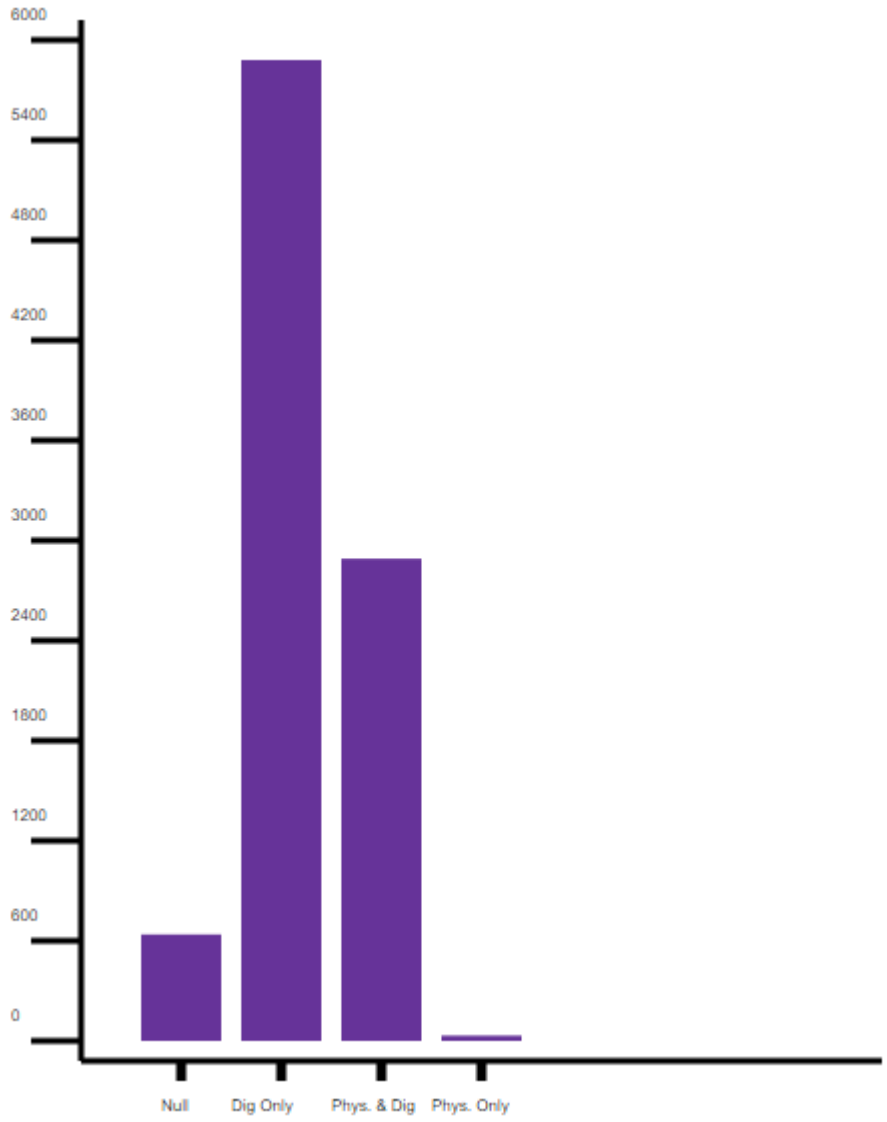
This visualization was my starting point. As aforementioned, this data is filtered by which genres had over 70 games; this way, I could ignore some of the genres that had some overlap or were extremely niche. I just had the simple question of which genre was the most popular for developers to create for the PS4. Looking at the graph, it's apparent that platforming games are extremely popular for the lifespan of the console. I had expected that Action and Adventure—or even FPS games would have been created most, but I was surprised by the results. We might assume that more funding toward advertising and marketing goes to big games like Call of Duty or Last of Us. There could be more indie games, that might not have as much fame, that rely on the platformer genre as well.



- Physical vs. Digital Distribution of PS4 Games

As the console and gaming industries progress, the necessity of owning of physical copy of a game has decreased—and now, it can even be considered undesirable. Physical copies can get scratched, damaged, or lost. If consoles and tech update, you will still be able to download a digital game on another device. There are many different methods to owning games digitally now, such as Steam and Epic Games, which will install them on your computer; however, is the trend reflective pertaining to home consoles, which were born off the premise of being able to play video games at home.

The question that prompted this visualization was if the PS4 reflected these trends toward a more digital industry. Here, I think this bar graph answers that question and emphasizes a more digital future for at-home gaming—at least for this console. Almost 6000 of the games that were in this data set were released digitally only. If you compare this number to the only 36 miniscule physical only releases—or even the 3000 digital and physical releases—it's apparent to see this shift.


6000
5400
4800
4200
3600
3000
2400
1800
1200
600
0
Null
Dig Only
Phys. & Dig
Phys. Only