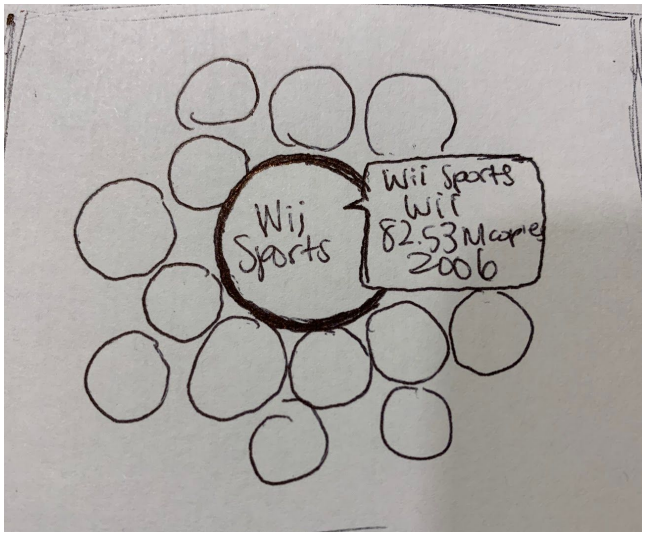
Yasmine Hejazi, Joseph Chou, Vanessa Hsu
5/14/19
Info 474 a3

Data domain: In this project, we will be working with video game sales (available on Kaggle here). This data contains information such as the rank, year, platform, genre, publisher, and sales (or copies sold) of each video game up to October 2016. The data was generated by a scrape of vgchartz.com. VGChartz, or Video Game Charts, has a methodology of which explains that data is collected through consulting with publishers and manufacturers, polling end users and retail partners, studying resell prices, and more. All data is checked against manufacturer firms and public data of tracking firms to ensure accuracy.

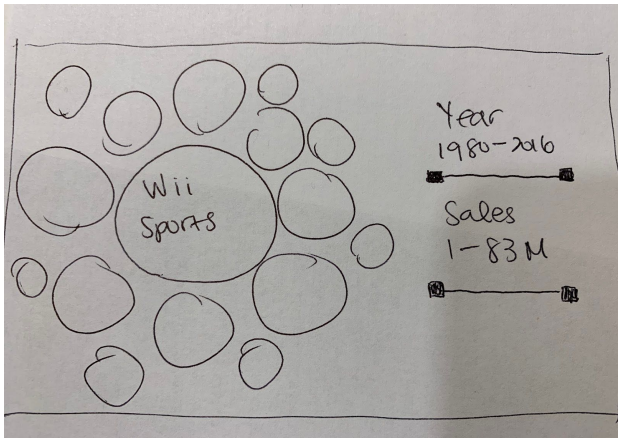
Interactive visualization techniques: For this domain, we think a bubble chart will be a good match in visualizing a high-level comparison of sales between games. Users will be able to quickly interpret which games have sold more copies based on the position and size of each bubble. Bubble size will depend on the number of video game copies sold globally and higher ranked games (based on global sales) will be positioned towards the center of the chart. We would like to have color encoded to Publisher (i.e. Nintendo) to give a quick insight into what publishers tend to sell more copies globally. A color legend will be paired with this encoding. To keep the bubble chart clean and not over-informative, we limit the chart to top 75 video games by rank.

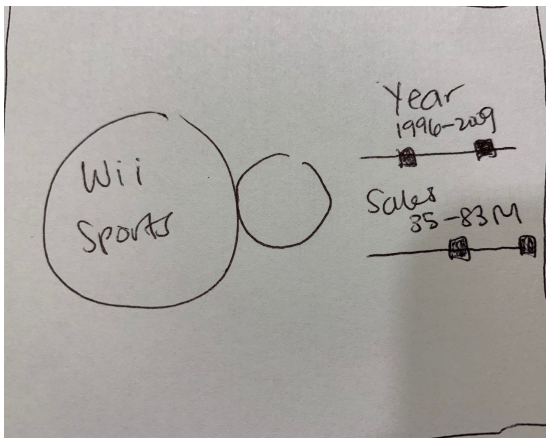
Bubble charts are great for visualizing situations and patterns quickly through frequency (Dymnikova 2015). It isn't great at explaining the *why* or *how* of a situation, but our goal for this project is to create a fast, high-level overview of the top video game sales. A bubble chart is essentially similar to a word cloud, but without the drawback of a word cloud where word length, letter style, and fonts can affect the perception of the word size. That saying, a bubble chart has the pros of a word cloud where it is fast, it reveals and highlights essential information, and it's engaging (Hein 2019).

We would like to include interactivity through tooltips and filtering widgets. We believe tooltips are essential to this chart because the bubbles can be very small and therefore cannot show essential information in the chart. It will be challenging to fit the text of video games with long names into the bubble, never mind other information such as year, genre, platform, and sales. We plan to include this information in a hover tooltip. As the user hovers over a bubble to reveal the tooltip, the bubble will also be outlined to highlight it from the rest.

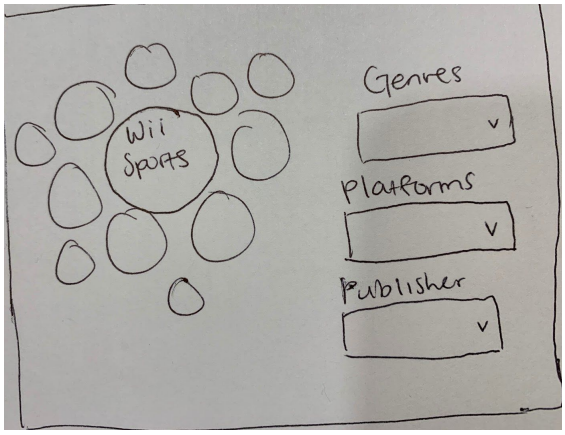


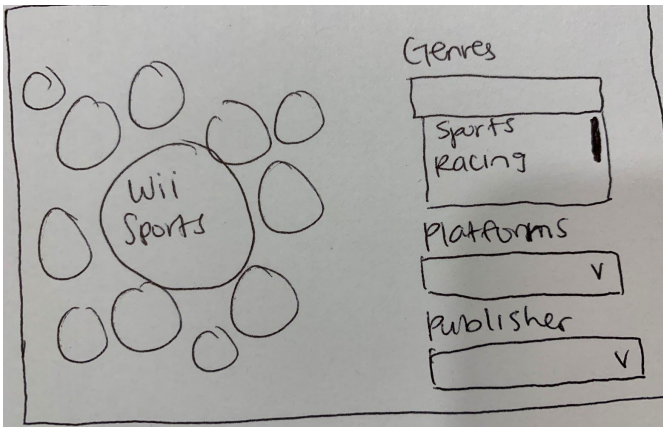
Next, we want to include filtering widgets to allow the user to run dynamic queries and do some more exploring with the data. One widget we would like to include is a two-handle slider for year and sales. This will allow the user to choose a range of these variables that they are most interested in and dive deeper into the data.

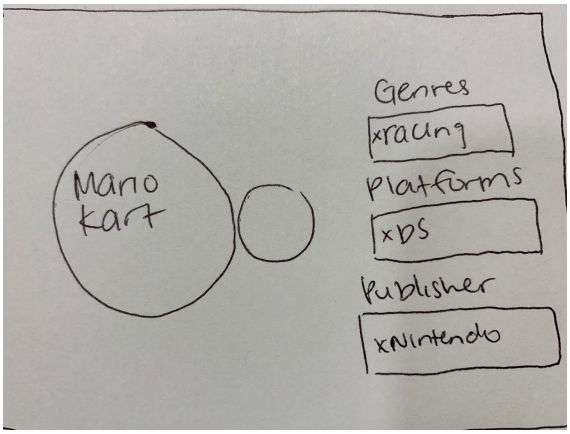




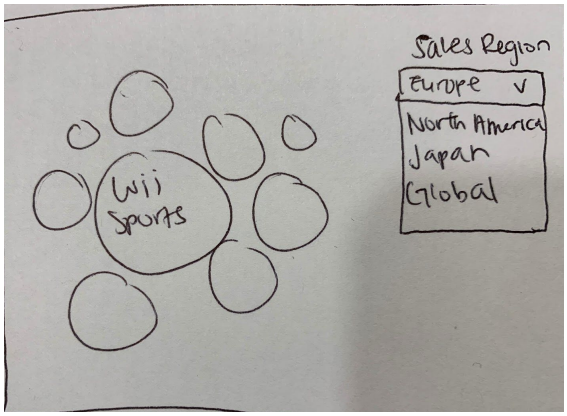
We would also like to include drop-down multi-selecting filters for the user to filter in or out values of genre, platform, and publisher.

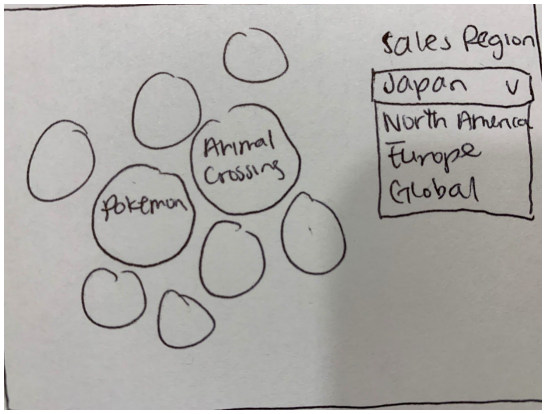




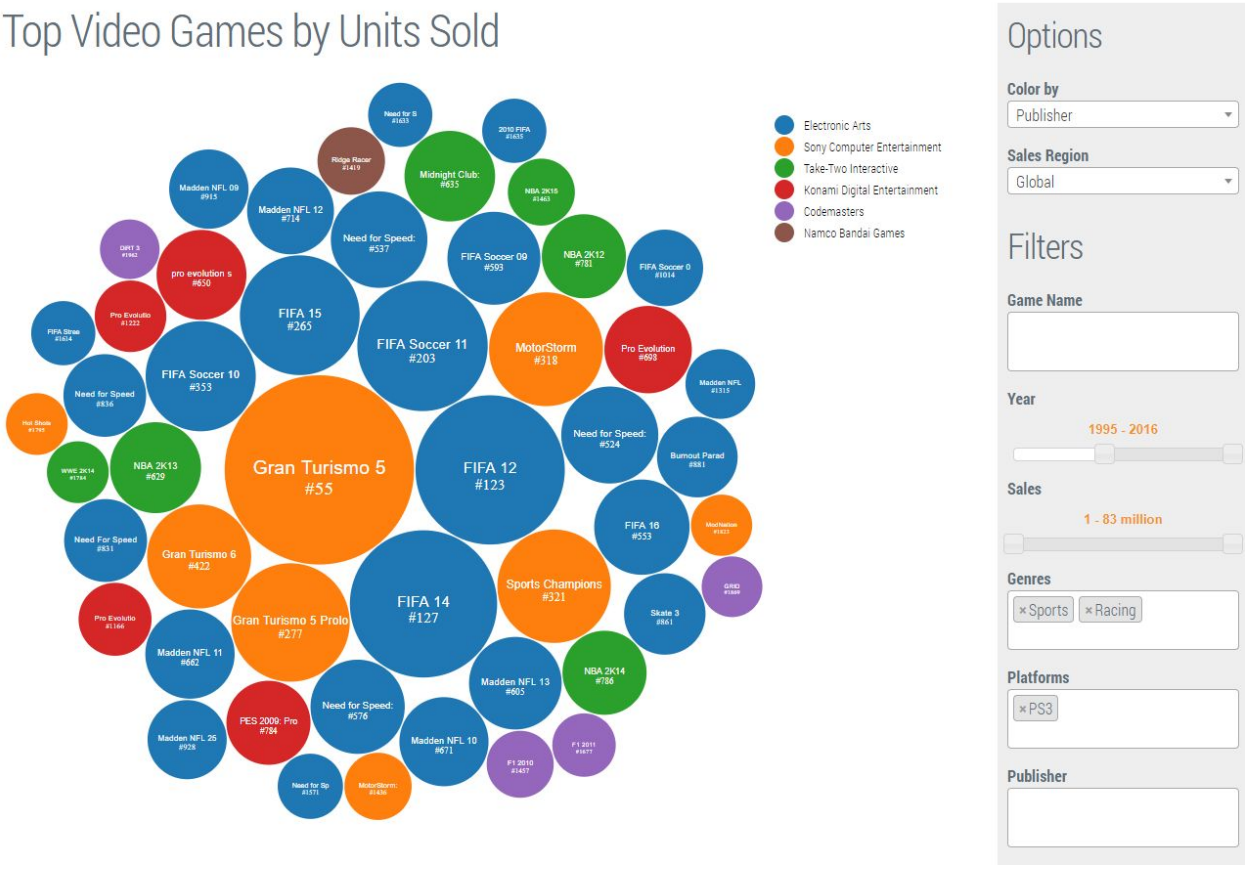


To allow users to explore high-level overviews of variables other than publishers, we would also like to implement coloring by genre, platform, and critic scores. Furthermore, we would like to include the option to change the sales region, allowing users to see patterns for each region and find differences between regions.





Final interactive visualization: Our completed visualization can be viewed here.



The visualization consists of a bubble chart, where the size of each bubble corresponds to the game's sales. We included interactivity through tooltips as well as various filters such as sliders and drop-downs. Hovering over a bubble displays a tooltip that shows the game's name, platform, number of copies, and year of release. We also included two-handle sliders for year and sales, so users can choose a range of values they want to see. Our drop-down

multi-selecting filters allow users to filter different genres, platforms, and publishers. Users are also able to color the bubbles by publisher, genre, platform, and critic scores. Hovering over part of the legend highlights bubbles that fall into that category, and clicking on the legend applies a filter for that category. If users choose a combination of filters that do not apply to any games, we inform them of this with a message in place of the bubbles.

Our final visualization includes all the features we included in our storyboard. One difference from the storyboards is the addition of an interactive legend, which we did not originally consider in our brainstorming, but found to be a helpful feature. We also decided to reduce the number of bubbles displayed from 75 to 50, as we found 75 nodes to look a bit too cluttered. Another addition not included in the storyboard is the text that appears when no games match the filters, as we found that a blank screen was not a clear way to communicate this.

Development process: To split up the work, Vanessa began by creating the storyboards for our visualization. Yasmine then set up the page and created the bubble chart, legend, and hover tooltip, Joseph implemented the filters, and Vanessa polished the visualization style and fixed bugs. Each of us spent about the same amount of time working on the project (~6 hours each).

The aspects that took the most time were figuring out how to implement the bubble chart, then determining how to transition the bubbles in response to filters. Implementing transitions was particularly time-consuming as it involved a lot of trial-and-error before figuring out that each node's children had to be transitioned separately from the node itself. Another challenge was determining how to set up the legends - to start out, we used a consistent legend only showing the top publishers, but found that this did not hold up after applying filters. Because of this, we decided to implement a dynamic legend that changes depending on the currently visible data, which was challenging to implement.

References:

Dymnikova, Mila (2015). Reframer strikes a chord diagram. Optimal Workshop. Blog. Retrieved from: https://blog.optimalworkshop.com/reframer-strikes-a-chord-diagram

Hein, Robert (2019). The Pros and Cons of Word Clouds as Visualizations. VisionCritical. Blog. Retrieved from: https://www.visioncritical.com/blog/pros-and-cons-word-clouds-visualizations