

Visualizing Video Game Features: Wii Recommend

Wii Design

Mario Aleman	mhbecer@emory.edu	Computer Science & Mathematics
Dusty Dang	dusty.dang@emory.edu	Computer Science & Mathematics
Karelin Marroquin	kmarroq@emory.edu	Computer Science & Spanish

Dataset:

The dataset that we used consists of three datasets obtained from Kaggle and Metacritic Games. The Metacritic dataset has over 14 thousand rows and it contains nominal attributes to identify the game, such as the title, platform, genre, and developer as well as some numerical attributes, such as the critic score, the user score, and the release date. Our video game sales dataset has over 16 thousand rows and it contains the same basic attributes, but it includes sales data from North America, Europe, Japan, and the total global sales. The last dataset we used had the same basic information, but it also included new attributes for gameplay, such as the average playtime for the overall game and the average playtime in story mode based on playstyle. Our final dataset was created with Pandas by merging the three datasets based on title, platform, and year as we deemed this a key to identify each game. We preserved the columns for title, user score, critic score, genre, platform, and developer as they were common attributes among all of the datasets. We also decided to preserve the North America and global sales data as well as the game play time for all playstyles and for leisure playstyles regarding the overall gameplay and the story mode gameplay. Although utilizing sales data and gameplay data would give us more numerical values to work with in our scatterplot, this generated various rows with null values for those attributes. We resolved this by omitting those values from the plot if the user chooses that filter, but since both the Metacritic and game sales datasets are the largest, we still had many rows to work with.

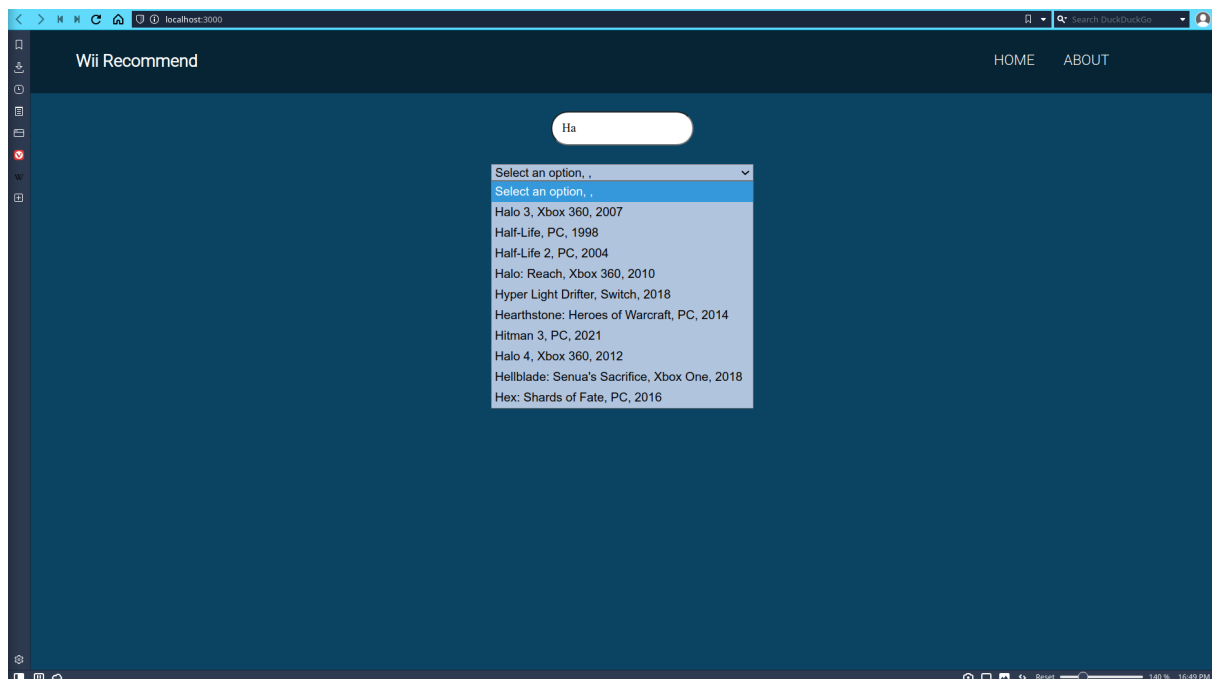
Contributions:

Mario	Dusty	Karelin
Implemented the search feature	Created the webpage	Cleaned and merged datasets
Created the similarity metric	Implemented the scatterplot with interactivity and color encodings	Implemented the table
Deployed React App to GitHub pages	Organized webpage UI and layout	Organized meetings and documentation

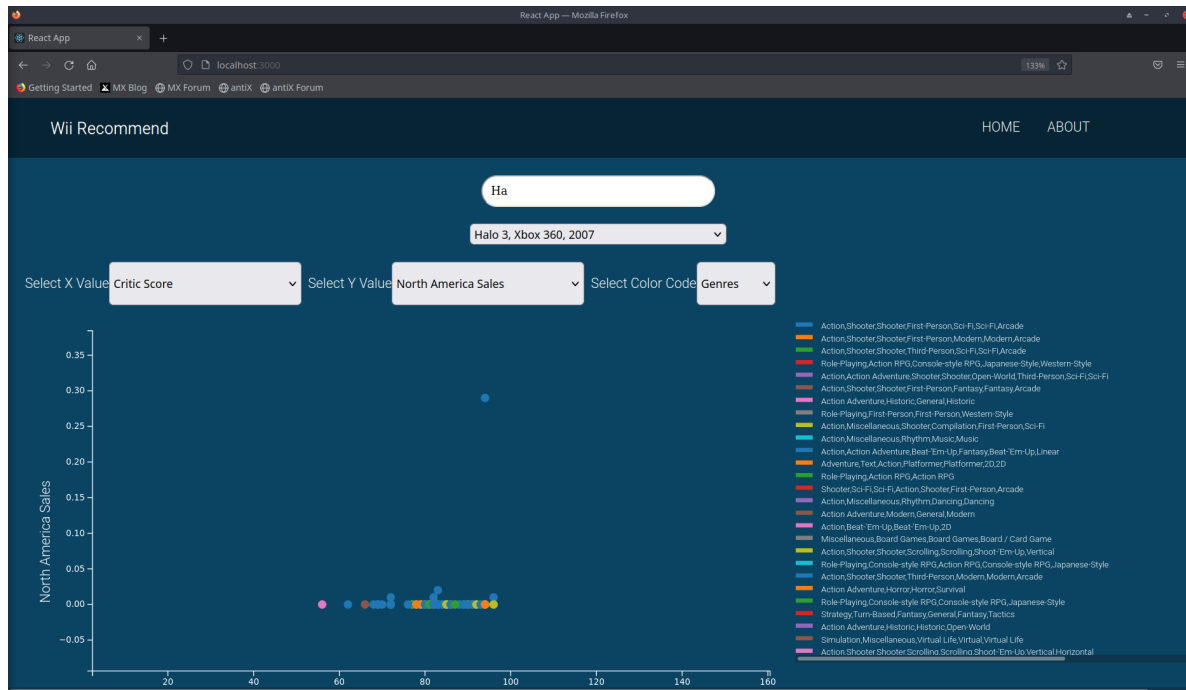
Reflections:

Although our initial progress with the visualization implementations in D3 was slow, we managed to create all of the features that we had initially envisioned. We had proposed a tool that would allow users to search our database for their favorite video games and that would allow them to obtain a top number of recommendations with which they could interact in the scatterplot. Within the scatterplot, they would be able to filter by features that they consider important to customize their results, and upon selecting a point on the scatterplot, they would be able to populate a table that shows all of the attributes for that particular game. We were able to implement all of the visualization components and the general user flow, so we did not have to diverge from our original plan. Upon testing our implementation, we did consider implementing a more sophisticated similarity metric by utilizing machine learning similarity algorithms that pair the features and assign a weight and a score. However, our dataset was not labeled and this implementation was out of the scope of the class. One difficulty that we encountered during the initial stages was regarding our dataset. Our visualization relies on a scatterplot to show patterns and trends, but we did not have many numerical attributes to work with. We resolved this by merging multiple datasets, but we still had many null values that we had to omit from the scatterplot. Our team had some downfalls with communication during the initial stages, but we were able to resolve that and work together to merge the visualization components and test the overall design.

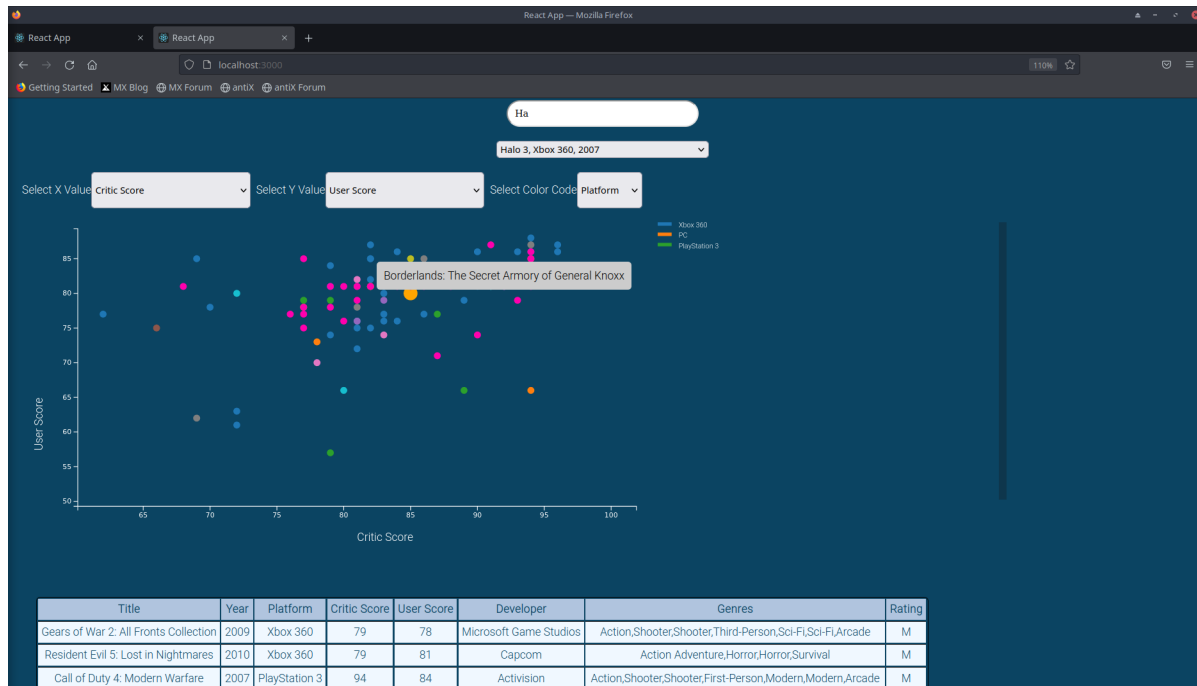
Interface:



Search bar with top ten results for "Ha".



Top search recommendations for Halo 3 on Xbox 360 with default filters of critic score and sales for North America and the color encoding for genre.



Search recommendations have been filtered for critic score and user score and color encoding has been changed to platform. Tooltip functionality is shown and magenta dots have been previously viewed. User has clicked on games of interest to populate the table.