Pick the Perfect Board Game for Today's Game Night!

INFO4310 HW2 Final Report

Group member: Xinyi Zhou (xz255), Curtis Xu (kx64), Flavia Jiang (yj472)

Purpose and Dataset Description

Dataset: Board Games (https://www.kaggle.com/datasets/andrewmvd/board-games)
This dataset contains detailed information on over 20,000 board games, including attributes such as game name, category, playtime, complexity score, number of players, overall rating, and year of publication. With its extensive features, the dataset serves as a valuable resource for game recommendation and discovery, enabling users to search for board games that best match their preferences.

Purpose of the Visualization: From the outset, we defined a clear objective: to help users find the most suitable board game for their group. Every design decision, from interaction features to filtering options, is centered around this goal.

Data Cleaning: We preprocessed the dataset in Python (*cleaning.ipynb*), selecting only relevant columns, removing missing values, and filtering out unreasonable entries. This ensures that users interact only with clean, reliable data, avoiding confusion from incomplete or inaccurate records. After the cleaning process, approximately 9,000 games remain.

Planned Interactions:

The primary challenge we aimed to address is the overwhelming number of board games—9,000 in total—which, as seen in the default scatterplot, cluster together, making exploration difficult. To resolve this, our visualization incorporates interactive filters for key attributes, including overall user rating, popularity, complexity score, number of players, and game type. These filters offer a streamlined, intuitive way for users to refine their search based on their preferences. By enabling dynamic filtering, the visualization improves usability, helping users efficiently identify board games that best suit their interests and group dynamics.

Interaction Storyboard and Design Intention

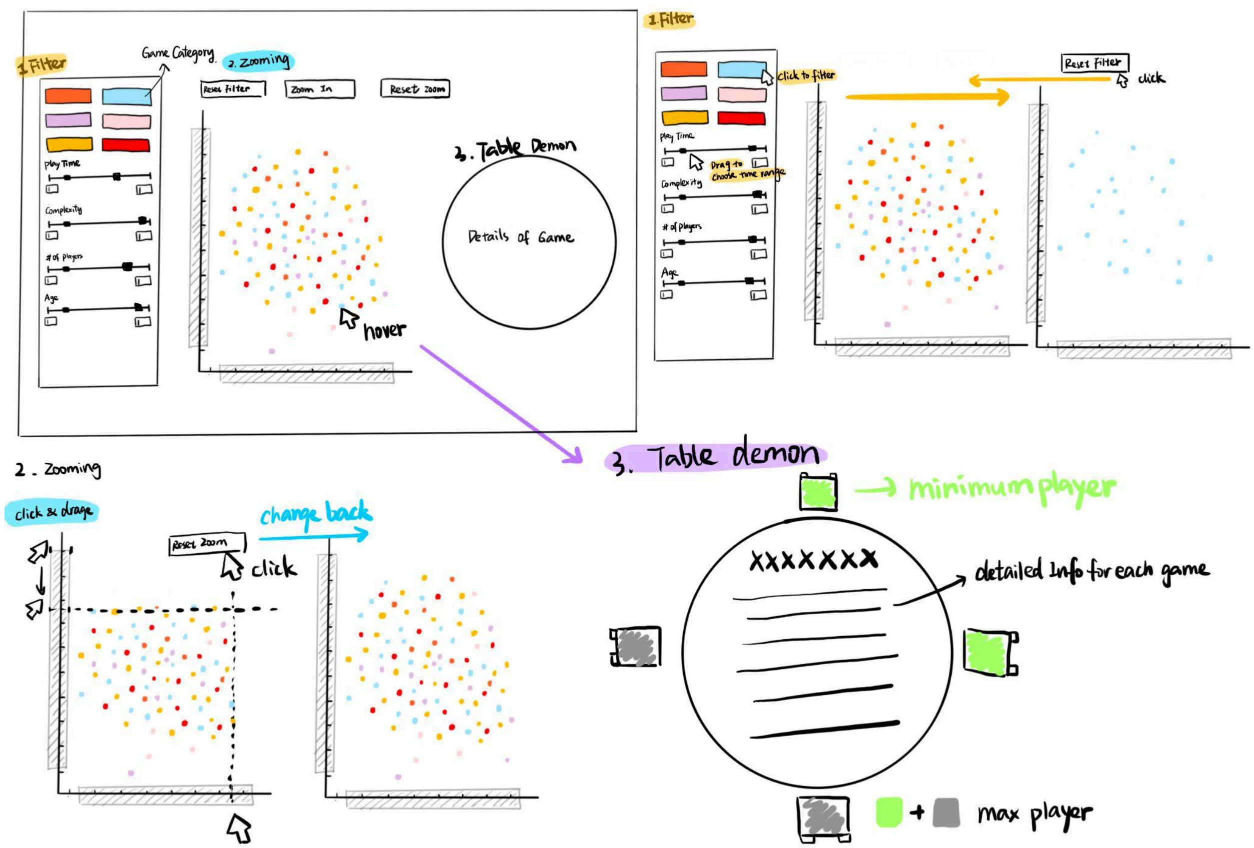

Figure 1 Interaction Storyboards. *The sketch shows all the interactions we made, including data filtering, zoom-in functions, and how to display details*

1. Data Filtering

The first interaction design enables users to filter games through **category tabs**, helping them quickly find games that match their preferences. We identified eight distinct categories—Thematic Games, Family Games, Customization Games, Party Games, Wargames, and Children's Games—allowing users to include their preferred types while filtering out others. To enhance flexibility, we incorporated **interactive scroll bars** for refining searches based on playtime, complexity, number of players, and recommended age. Overall, we made the data filtering interaction, the combination of categorical filtering and adjustable parameters, to create a more intuitive and personalized game discovery experience.

2. Axis Brushes, Box Brushes, and Zoom Interactions

In the second part of the interaction, we visualized the dataset using a scatter plot, where each dot represents an individual game and is color-coded based on its category type. However, with over 9,000 board games plotted, data points naturally cluster together, making it challenging to distinguish individual games. To address this, we incorporated **interactive axis brushes** on both the vertical and horizontal axes, allowing users to filter games by selecting specific areas of interest. Once an area is brushed, users can further refine their exploration by clicking the zoom button, which enlarges the selected dots and expands them to fill the entire scatterplot area. If users want a quick zoom into games without detailed filtering, they can also use **box brushes** to quickly draw a box on the scatterplot and then zoom into the selected area. This dynamic zooming functionality and axis brushes make it easier to identify and examine specific game points, improving clarity and user interaction within the visualization.

3. Table Demonstration for Details

In the third phase of the interaction, we designed a dynamic table visualization that presents detailed game information in an engaging way. When hovering over a point on the scatterplot, an interactive game table appears, surrounded by chairs representing the game's player capacity. The total number of chairs corresponds to the maximum number of players, while light green chairs indicate the minimum number of required players. Additionally, key details such as the game's publication year, rating, and complexity are displayed within this visual setup. This interactive table is designed to simulate the experience of playing a board game with friends, enhancing user immersion. By using chairs as a visual cue, we not only provide a clear representation of player limitations but also evoke the social nature of board games, making the data more tangible and intuitive.

Final Interactive Visualization Application

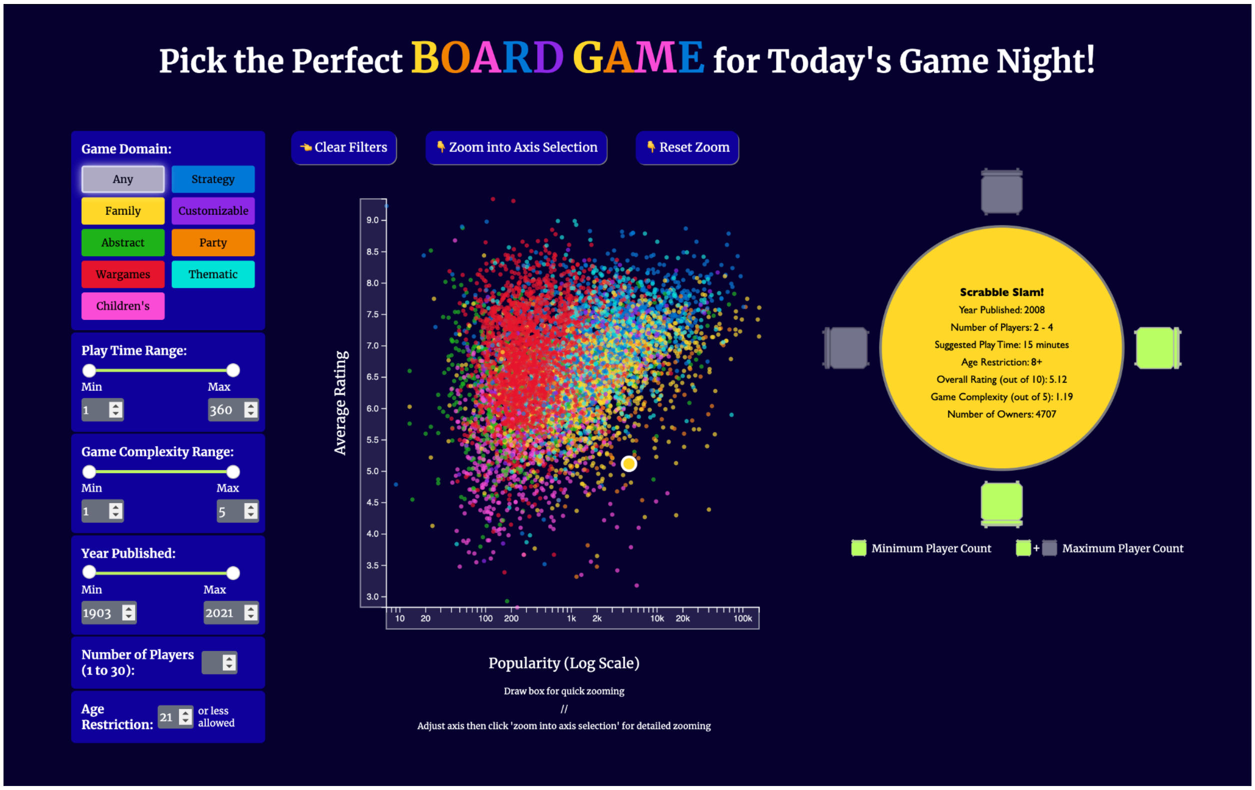

Figure 2 Interactive Visualization Application. *The image demonstrates our final interactive visualization.*

This tool enables users to explore board game options through an interactive filtering system. The central scatterplot helps users quickly identify games within their preferred rating and popularity range using multiple zoom methods. For more specific preferences, such as game genres or playtime, the toolbar on the left provides detailed filtering options. After applying filters, hovering over a data point reveals a table displaying key game details, including the title, minimum and maximum player count, and other relevant information.

Issues and Trade-offs

- **Zoom-in methods:** We initially implemented axis brushes as our primary filtering method. During class critiques, we received positive feedback on this approach, along with suggestions to incorporate a box-brush-based zoom method for a more direct and intuitive user experience. After careful consideration, we decided to retain both zoom methods. While box-zoom offers users the flexibility to select specific areas, axis-zoom provides precision by allowing users to adjust axis values for detailed data filtering.
- **Table Demonstration for Game Details:** While the Game Details table presentation looks fancy, creative, and appealing, it has some limitations. For example, the circular nature limits flexible text layout. The long text must be wrapped to fit into a circular form. However, since most of the information in our dataset is short, this isn't a huge problem. For some of the longer text, we wrote code to logically twist it into different rows to fit the circular table.

- **Filter toolbar:**
    - **Game Domain:** Since this is the only categorical variable, the most intuitive way to represent it is through distinct colors. We carefully selected nine colors with the highest possible contrast to maximize clarity. However, we acknowledge potential accessibility issues, which are difficult to avoid. To address this, we allow users to filter by specific categories, enabling them to focus on smaller, more manageable subsets of data.
    - **Playtime Range:** Users can specify both minimum and maximum playtime, offering flexibility in filtering. We considered using a single input field, but this approach was too rigid—playtimes of 210 and 211 minutes are nearly identical, yet the number of games available for each may differ significantly. Moreover, a dual-range selection allows users to customize their search, whether they want to set a maximum, a minimum, or both. Additionally, we provide both slider and input field options to accommodate different interaction preferences.

- **Game Complexity Range & Year Published Range:** These follow the same reasoning as the playtime range, providing flexibility with dual-range selection and multiple interaction methods.
- **Number of Players:** For this filter, we use a single input field instead of a range. Typically, when selecting a game, a group already knows the number of participants, making a single input more intuitive. By default, the field is left empty to ensure all data points remain visible, giving users an overview of the game distribution before they apply filters. However, we realized that this approach removes hints about the valid input range. To address this, we added "(1 to 30)" to the label, providing users with clear guidance on the acceptable values.
- **Age Restriction:** Similarly, we use a single input field here. When choosing a game, groups are generally aware of the youngest participant's age. For example, if the youngest player is 12, they can simply enter 12 to filter out games with a higher minimum age requirement.

Development Process and Change

Development Schedule:

1. **Basic Framework and Scatterplot:**
   We began by setting up the basic framework and implementing the scatterplot, followed by adding functionalities such as zooming.
2. **Table Diagram:**
   Our first thought was to utilize the game table for detailed information descriptions, and then we thought about how to make the table more interesting by adding chairs that not only clearly display information about the number of players, but also give users a more interactive sense of play.
3. **Filter Toolbar:**
   We implemented a function that dynamically responds to changes in filter values. Each time a user modifies a filter, the function is triggered, ensuring real-time updates. The filtered dataset is then passed to the main update function, which redraws the chart to reflect the selected criteria. This approach maintains smooth interactivity, allowing users to refine their search and immediately see how their choices impact the visualization.

4. **Overall Integration and Adjustment:**
   After completing the three main development phases, we began testing the system's overall integration and making layout adjustments. These included adding the text tutorial for the zoom function, incorporating a legend for the table diagram, and refining button functionalities for a smoother user experience.

Development changes:

- **Scatterplot's Axis:** Our initial design allowed users to freely change the x- and y-axis domains. However, given that the toolbar already includes multiple filtering options, we decided to use a fixed set of axes to maintain clarity and streamline the data filtering process.
- **Table Diagram:** Initially, we added background decorations to the table, but the patterns were too visually complex and distracted from the information presented. We also initially displayed game names in a single line, but longer names extended into the round table, disrupting readability. To resolve this, we implemented line breaks to keep the text neatly contained within the designated space.
- **Toolbar (including discussion of tradeoffs):**
  - Unlike the static representation in our storyboard, coding dual-range filters proved significantly more complex due to their interactive nature. A key challenge was preventing users from entering values outside the valid range. Allowing an invalid input, such as setting a complexity score maximum to 6 when the scale only goes up to 5, could mislead users into thinking higher scores exist. To prevent this, any invalid input is automatically reset to the default value.
  - Another major issue involved handling multi-digit inputs. If a user intended to enter "90" for playtime, the system initially recognized "9," triggering an unintended chart update. This caused unnecessary confusion. To address this, we implemented a 0.25-second buffer time, allowing users to complete their input before the chart responds. While this solution may not be perfect for all typing speeds, it significantly reduces interruptions and ensures a smoother user experience.

Trade-offs/Design choices:

- **Color theme:** we use a dark background and colorful color theme to simulate the atmosphere of a game night.
- **Scatter Plot Coloring:** Although some games belong to multiple domains, we chose to assign colors based solely on their primary domain. This decision was made because games with multiple domains represent only a small fraction of the overall dataset.
- **Color Choices for Chairs:** Initially, we considered using red to represent the minimum number of players. However, since a similar red was already used for game categories, we decided to choose a different color to avoid confusion. We then explored dark blue as the default color, as it aligned with the filter and button colors. However, we found that dark blue lacked sufficient contrast with the background. Ultimately, we selected light green. While it may not perfectly align with our theme color, it serves as an accent color that enhances visual clarity and helps the primary theme color stand out.

Team Member Contribution

Xinyi Zhou:

Responsible for finding potential datasets, some data cleaning, interactive storyboard sketches, and designing and coding the interactive table section. Completed the dataset and interactive storyboard portion of the report. Spent 5 hours on development. I spent the most time figuring out how to wrap long text into different lines and controlling the position and rotation of the chairs so that the chairs were placed around the table

Curtis Xu:

I was responsible for identifying and evaluating potential datasets, establishing the initial project framework, developing the scatterplot visualization, and implementing box-zoom and axis-zoom functions. Additionally, I assisted with layout integration, adjustments, and report writing. I dedicated a total of 5 hours to the development and 1 hour to the report. with the majority of my time spent implementing and debugging the Zoom functions.

Flavia Jiang:

I was responsible for data cleaning, implementing the toolbar and filters, and writing my sections of the report. Additionally, I contributed to improving the overall code efficiency of our project. I spent approximately 10 hours coding and 1 hour writing the report. The most time-consuming aspect was breaking down complex interaction tasks into smaller, more manageable functions to improve modularity and maintainability.