

Homework 2: Video Game Sales Data and Ratings Exploration

Link to dataset: https://www.kaggle.com/datasets/thedevastator/video-game-sales-and-ratings?select=Video_Games.csv

Link to code: <https://github.com/sid-valecha/stat436-hw2-shiny/blob/main/app.R>

Introduction

In this project, I created a Shiny application to visualize and explore video game sales data. The dataset includes information about game titles, platforms, release years, genres, and critic/user scores. The goal of this app is to allow users to analyze the sales distribution and trends in video games across regions, as well as compare critic and user scores across different genres.

The interface was designed to give users maximum flexibility in interacting with the data, providing a wide variety of filtering options and multiple visualizations. This allows users to explore the data from multiple perspectives and discover trends in video game performance, reviews, and sales over time.

Interesting Findings

- **Global Sales Trends:** The industry saw slow growth in the early 1980s and 1990s, but rapidly accelerated near 2010 after which sales started to decline. This trend indicates a peak period for the gaming industry during this time.
- **Sales by Region:** North America consistently had the highest sales, followed by Europe, Japan, and the “Other” regions. While sales in North America and Europe dominate the market, Japan remains a key player with specific genres like role-playing games.
- **Critic vs User Scores:** The relationship between critic and user scores shows a positive correlation, but is not perfectly linear. The regression line suggests that higher critic scores tend to align with higher user scores, however, there is a wide spread in user ratings, suggesting that users may be more divided in their opinions than critics. For instance, some games with lower critic scores still received relatively high user scores, especially in niche genres. This suggests that user perception doesn’t always align with professional critic reviews.
- **Critic Scores:** The box plots for critic scores across different genres revealed that critic ratings tend to be relatively consistent across genres, with most games falling within a similar score range. There are fewer outliers and variability in critic scores, indicating that critics generally agree on the quality of games across different genres.

Interface Creation

The interface was built to allow users to:

- **Select multiple genres** of video games using a multi-select dropdown, giving users the ability to filter by one or more genres at a time.
- **Specify the year range** for games released via a slider, allowing users to analyze trends over a specific period.
- **Explore different visualizations**, including sales trends, critic vs. user scores, and regional sales distributions.

Data Preparation:

- **Removed rows** where genre was missing
- **Missing values** for Critic_Score and User_Score were removed.
- **Non-numeric values** such as “tbd” in the User_Score and Critic_Score fields were handled by converting them to NA and filtering them out.

Interface and Design Choices:

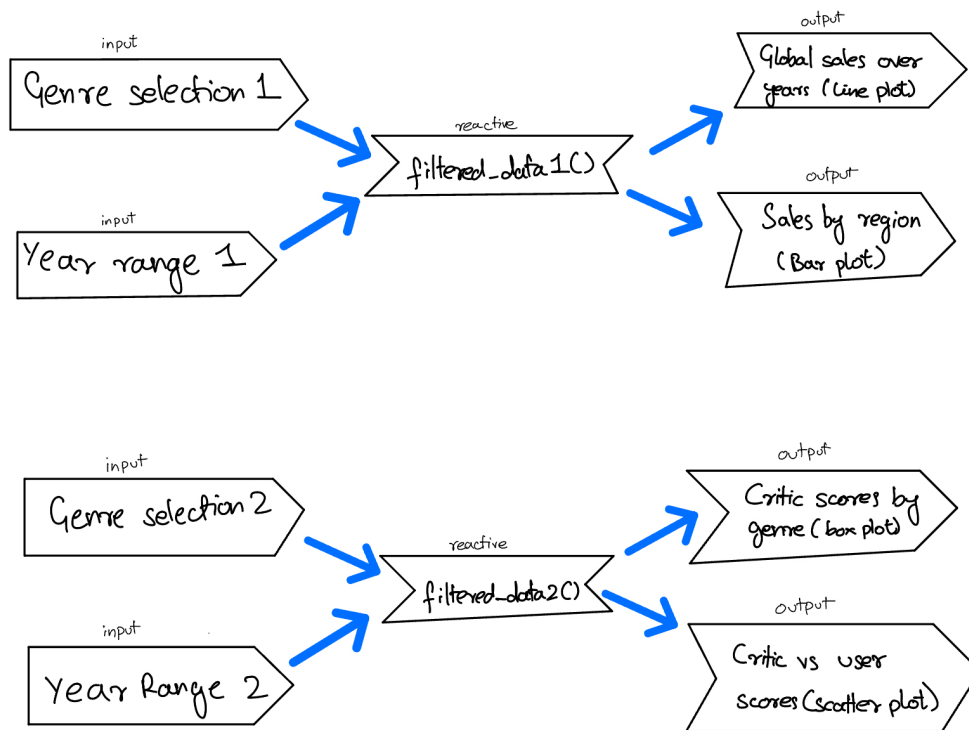
The app provides four key visualizations:

1. **Global Sales Over Years:** A line plot that tracks global sales trends for selected genres over time. This allows users to see how game sales evolved across different time periods and identify peak periods for the gaming industry.
2. **Sales by Region:** A bar chart that breaks down total sales by region—North America, Europe, Japan, and Others. This chart shows that North America and Europe dominate video game sales, with Japan following, especially in role-playing genres.
3. **Critic Scores by Genre:** A box plot that shows the distribution of critic scores across genres. This highlights the consistency of critic scores, revealing that most genres receive similar ratings with few major deviations.
4. **Critic vs User Score:** A scatter plot comparing critic and user scores. This plot reveals a positive but non-linear relationship between critic and user ratings, where critics tend to be more consistent, while user ratings show more variability.

Reactive Graph Structure:

The app uses **reactive expressions** to dynamically update the visualizations based on user input. When users adjust the genre selection(s) or change the year range, the underlying data is re-filtered, and the visualizations are regenerated in real time. This reactive structure ensures the app remains responsive and interactive.

The **sales trends**, **critic score distributions**, and **critic vs. user score comparisons** are all dependent on the user's input, making the app super customizable for different analysis needs.



Conclusion

This project provided me with a great opportunity to explore Shiny's capabilities for data visualization. The app's flexible interface allows users to easily interact with the data and extract meaningful insights about video game sales trends, regional preferences, and the relationship between critic and user scores. The results of the visualizations revealed important trends in the video game industry, in terms of sales as well as ratings.