Unit 6 Assignment: Video Games Development Analysis

Samuel Loos

Jon McKeeby

IT527

Real World Scenario

The gaming industry is fiercely competitive, with game development studios constantly striving to launch successful titles that resonate with players worldwide. A game development studio is gearing up to release a new title into the market. The studio's executives are acutely aware of the importance of strategic decision-making in ensuring the success of the launch. They have tasked a data analyst with conducting market analysis and providing insights to inform their decision-making process throughout the development and release window.

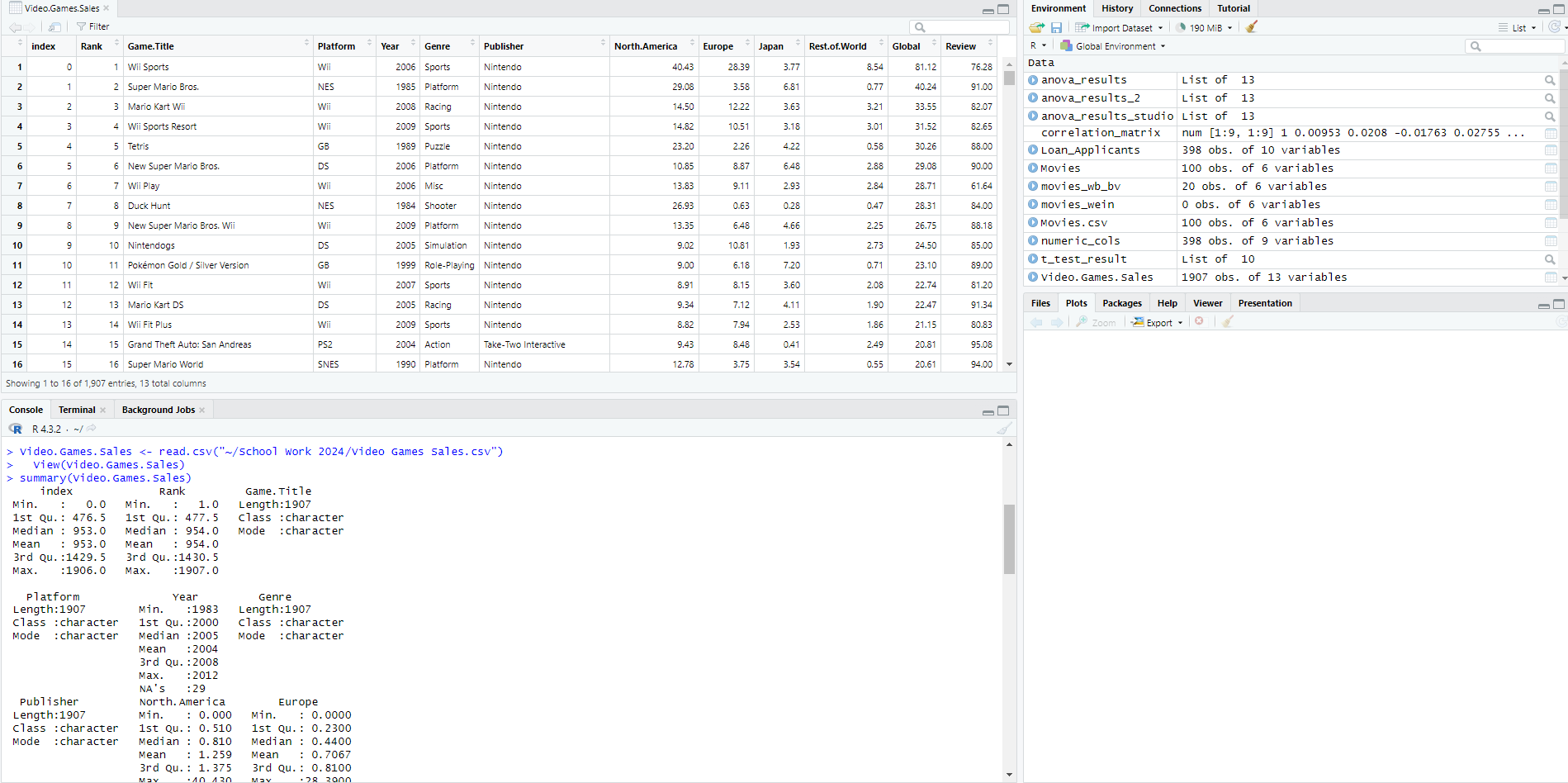
Problem Statement

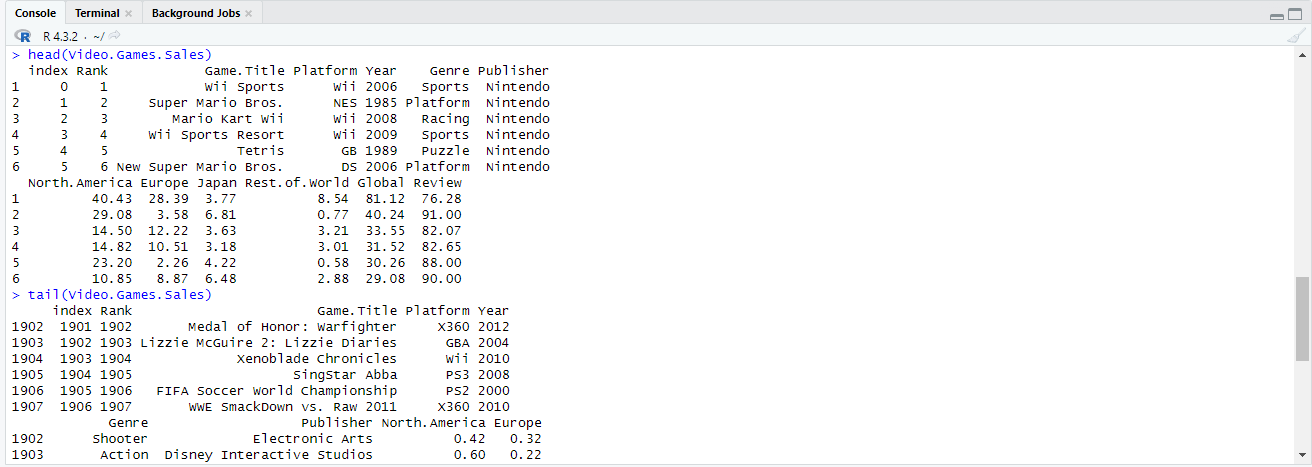
The game development studio is facing the challenge of launching a new title with the highest chance of success based on regional preferences, platform popularity, and genre trends. The studio's executives need to leverage data-driven insights to identify target regions with the greatest demand for the game genre, select the most profitable platforms for release, and devise effective marketing strategies to maximize sales and market penetration.

Needed Data

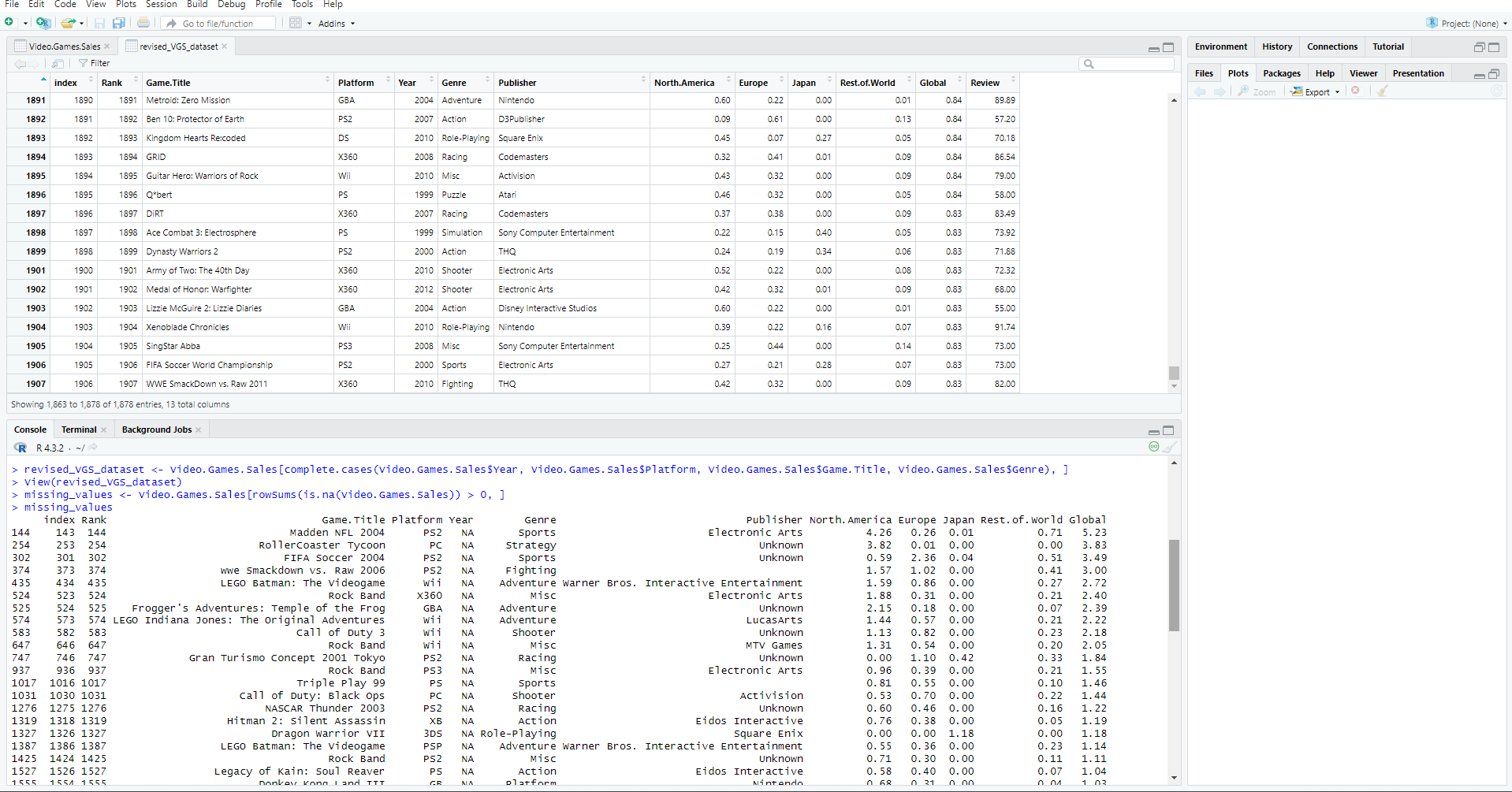
To address this challenge, the data analyst will utilize the "Discovering Hidden Trends in Global Video Games" dataset by Andy Bramwell. This dataset contains comprehensive sales data for video games from various regions, platforms, and genres (Bramwell, 2020). By analyzing key variables such as sales figures, game titles, regions, platforms, and genres, the analyst will gain valuable insights into emerging trends and consumer preferences in the gaming market (Bramwell, 2020). Additionally, the analyst will preprocess the data to ensure its suitability for statistical analysis and decision-making purposes (Bramwell, 2020).

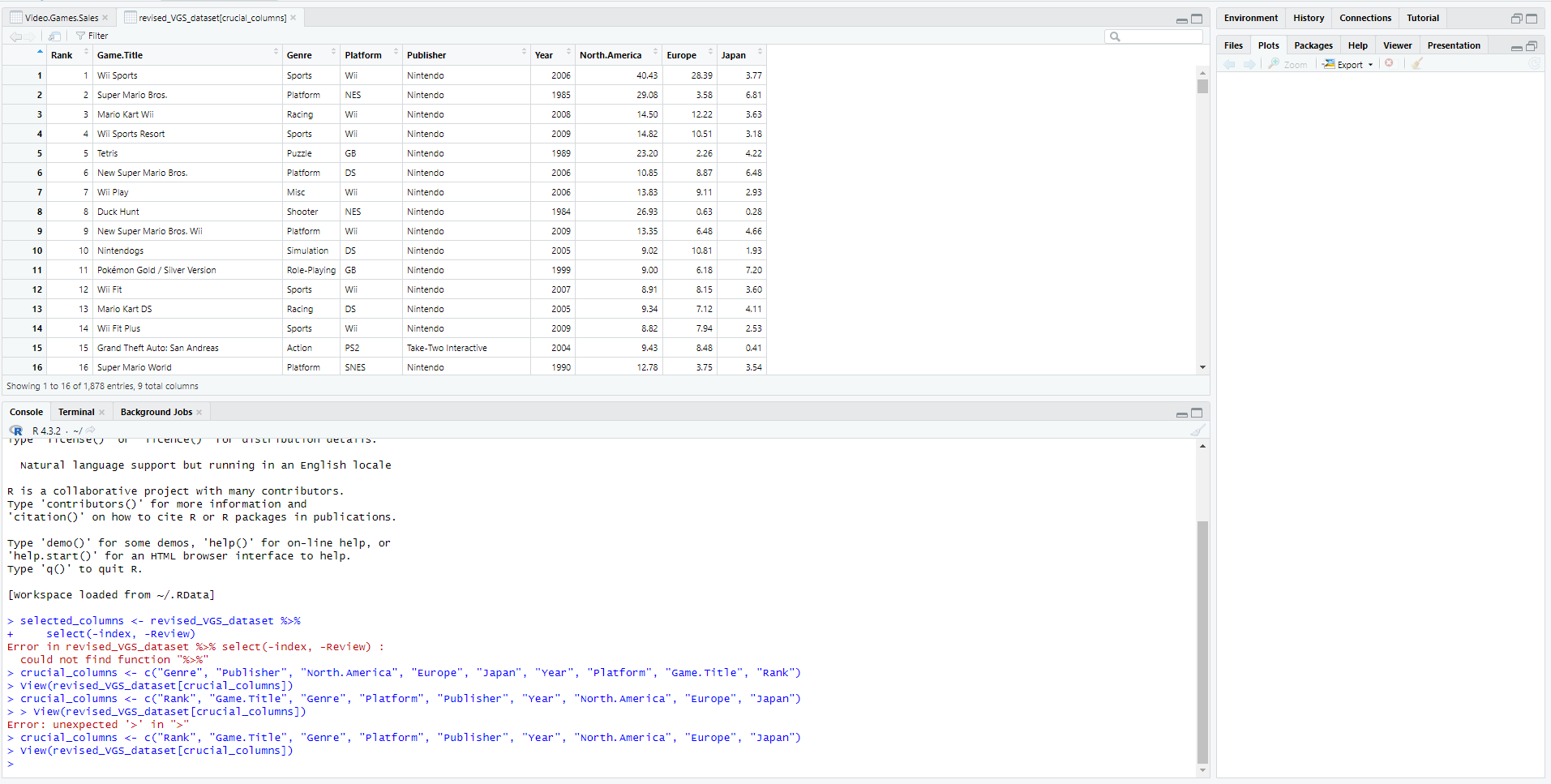
Dataset Visuals/Structure





This is the previously mentioned dataset and uses of the ‘head’ and ‘tail’ commands to familiarize with the data. Preparing the data by creating a subset was necessary to omit the missing values of the set.



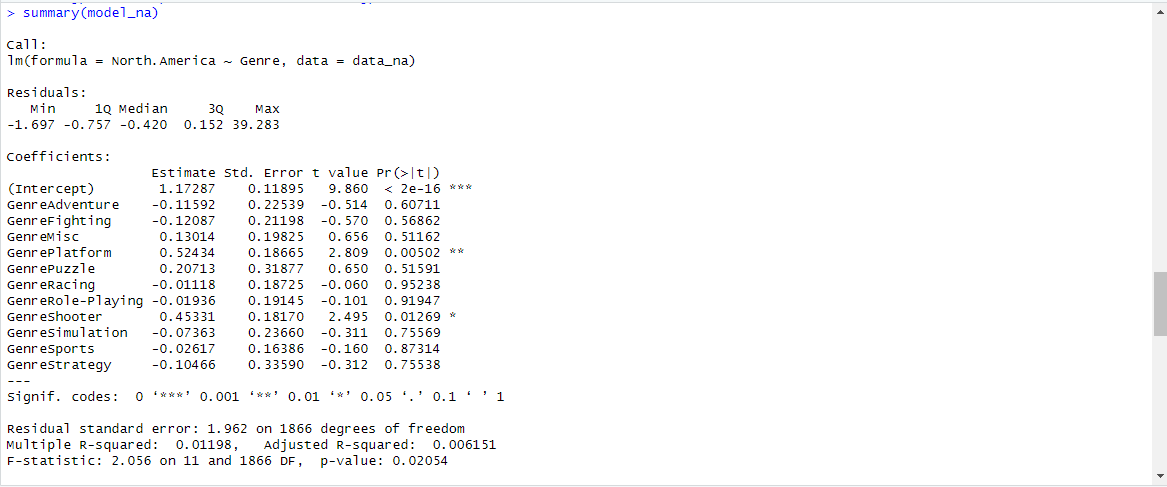


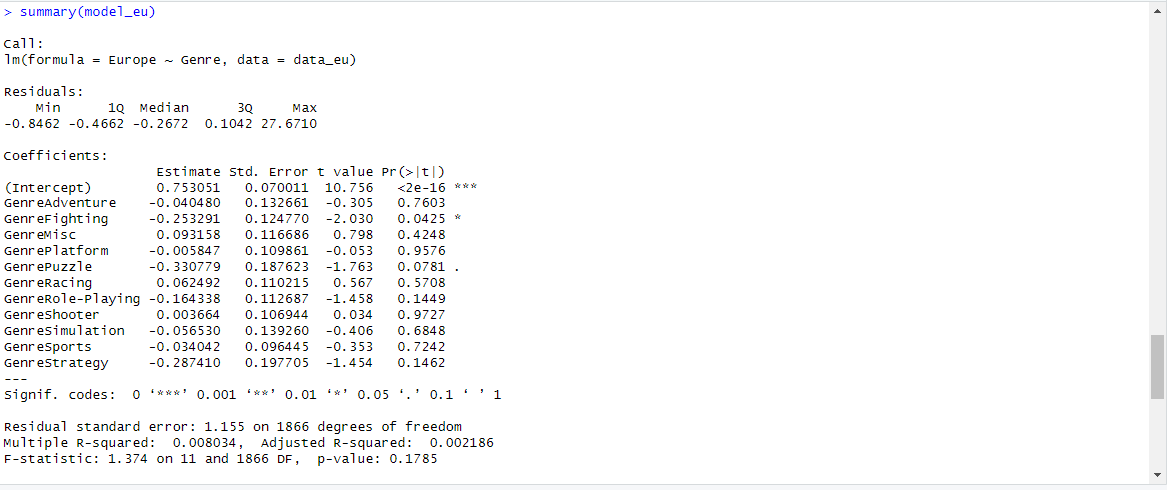
This is a screenshot of the revised\_VGS\_dataset that omits any missing values. Confirming the number of complete entries, which went from 1907 to 1878, was crucial to have an accurate analysis of the video games sales data. Lastly, the dataset omits columns not crucial to the analysis. The process provides a clearer presentation of the data that is needed for further analysis.

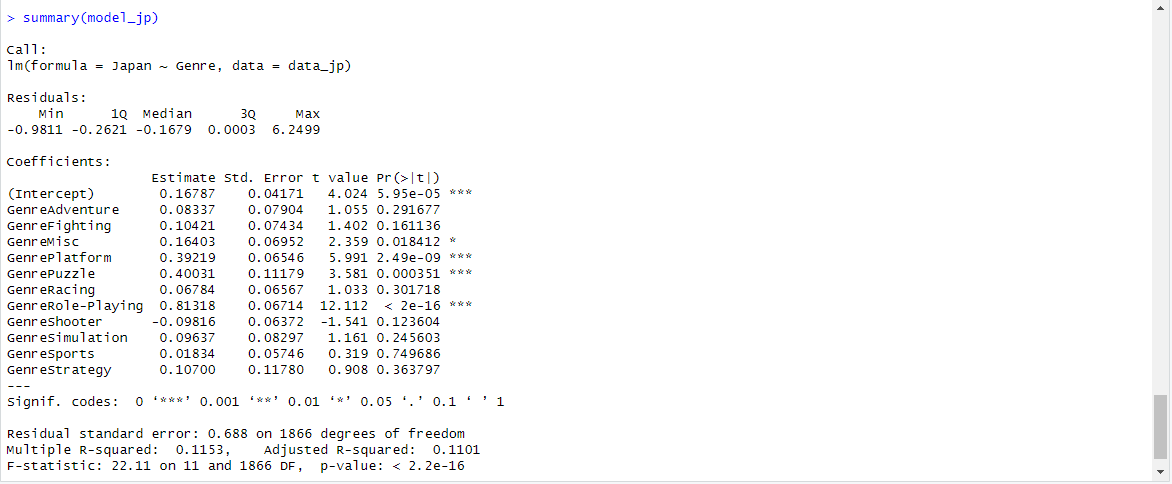
That data that is needed from the current dataset are Genre, Regions that include North America, Japan, and Europe, Publishers, and their Rankings. Genres will be important for early development to guide the team to creating a game that is the most receptive to players and for sales goals. The Regions columns will inform the team how well previous titles have done in the past. Rankings and Publishers are important because this gives the team how competitive the game needs to be and how to create their game based on previous titles from competitors.

Statistical Modeling Technique

The statistical modeling technique used will be the linear regression technique. Linear regression is implemented to estimate the relationship between two quantitative variables (Bevans, 2020). The variables being used will the response variable, which are the sales figures for the regions, and the predictor variables genre and platform (Nayak, 2021).



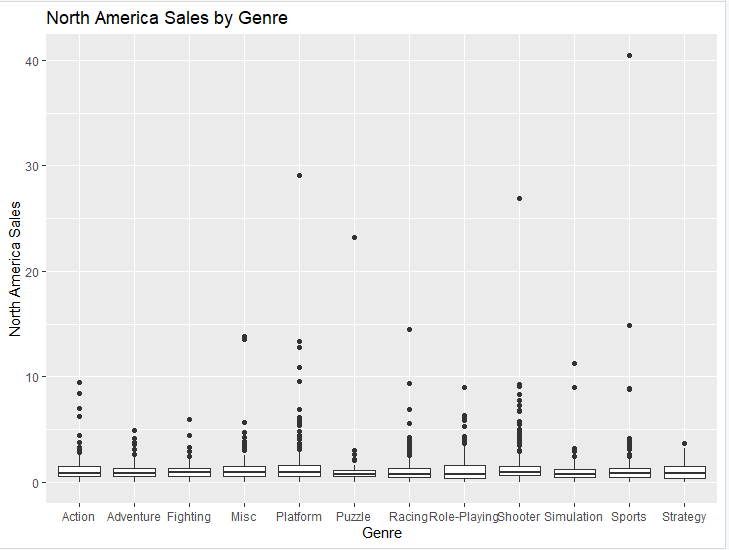


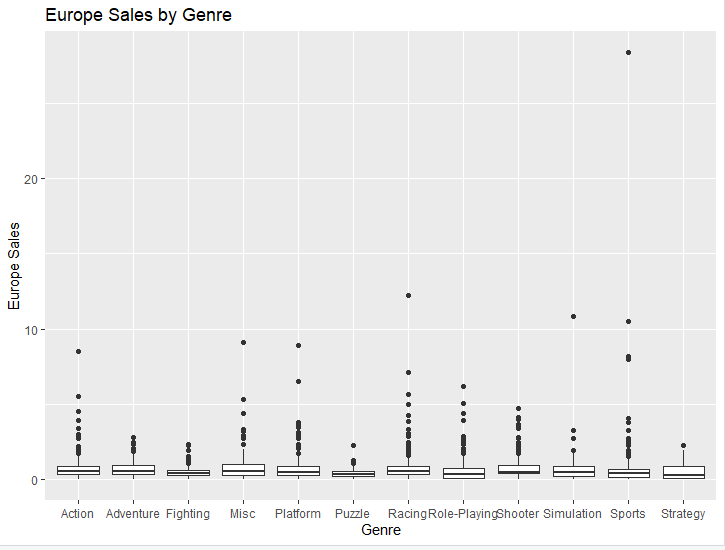


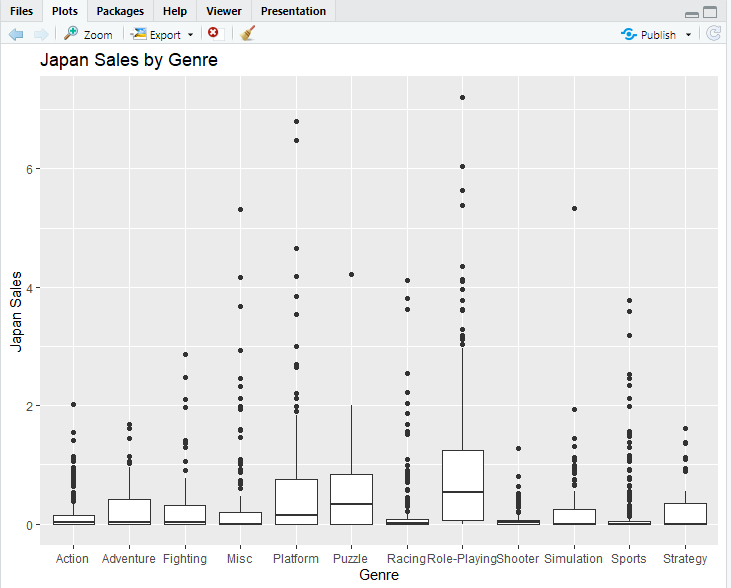
The pictures above explain the intercepts, coefficients, P-values, and R-squared for each region and their respective influences. The North America model suggests that the genre of a game has some influence on North American sales, with Platform and Shooter genres showing significant positive effects. The model's overall explanatory power is relatively low.

The Europe model suggests that the genre of a game has some influence on European sales, with Fighting genre showing a significant negative effect. However, the model's overall explanatory power is relatively low.

The Japan model suggests that the genre of a game has a notable influence on Japanese sales, with certain genres like Platform, Misc, Puzzle, and Role-Playing showing significant positive effects. The model's overall explanatory power is moderate.







The linear regression models provide insights into the relationship between genre and sales in each region. Analyzing the boxplots can discern trends in sales by genre, identifying which genres tend to have higher or lower sales within each region. The variability in sales, as indicated by the spread of the boxplots, offers insights into the consistency of sales across different titles within a genre (GamingScan, 2021). Outliers in the boxplots highlight exceptional cases of games with unusually high or low sales, allowing for a deeper exploration of factors contributing to their success or failure. Comparing the boxplots between regions facilitates a comparative analysis, revealing potential differences in genre performance across regions (Oza, 2022).

The Findings

The analysis conducted using linear regression models and boxplots has shown crucial insights regarding the relationship between genre and sales across different regions. Through the linear regression models, its found varying degrees of association between genre and sales, indicating that certain genres exhibit stronger correlations with sales performance in specific regions compared to others (GamingScan, 2021). These findings were further analyzed by the boxplots, which revealed distinct sales distributions across genres within each region.

Outliers in the data were identified, representing exceptional cases where games achieved unusually high or low sales. Delving into these outliers has provided valuable insights into the factors influencing their success or failure. If experts of the subject could provide surrounding factors, it could possibly include innovative gameplay mechanics or targeted marketing strategies (Oza, 2022)(VG Insights, 2022).

The comparative analysis of sales distributions between regions unveiled disparities in genre popularity and sales trends, highlighting the importance of tailoring marketing strategies and game development efforts to suit the preferences of target regions.

These findings underscore the significance of leveraging data-driven insights to inform strategic decision-making in game development and marketing, ensuring optimal market penetration and revenue generation in diverse global markets.

Recommendations

The recommended strategy would be to utilize the identified genre preferences and sales trends in each region to tailor marketing campaigns effectively, and to focus resources on regions where specific genres have demonstrated higher sales potential, optimizing advertising networks and ad messaging to reach the target audience (GamingScan, 2021)(VG Insights, 2022).

Consider the findings of the genre associations with sales information to streamline the game development process. Create ways to combine different aspects of popular genres to reach more audiences.

The analysis conducted on regional preferences and genre trends provides valuable insights for the game development studio seeking to launch a successful new title. Leveraging data-driven approaches, the studio can strategically align their efforts with market demands, optimizing their chances of success in diverse global markets.

References

Bevans, R. (2020, February 19). *Simple linear regression: An easy introduction & e*

*xamples*. Scribbr. https://www.scribbr.com/statistics/simple-linear-regression/

Bramwell, Andy (2020) *Discovering Hidden Trends in Global Video Games* [Dataset]. Kaggle.

<https://www.kaggle.com/datasets/thedevastator/discovering-hidden-trends-in-global-vide>

o-games/data

GamingScan. (2021, June). *2024 Gaming Industry Statistics, Trends & Data (Biggest Study)*.

https://www.gamingscan.com/gaming-statistics/

Nayak, B. (2021). *Linear Regression on Video Game Sales*. GitHub.

https://github.com/Ebulo/Linear\_Regression\_on\_Video\_Game\_Sales/blob/main/README.md

Oza, H. (2022, April 4). *7 roles of data analytics in video games development: HData Systems*.

Data Science & Business Intelligence Company. <https://www.hdatasystems.com/blog/7-roles-of-data-analytics-in-video-games-development>

VG Insights. (2022, January 23). *Games industry data and analysis*. Video Game

Insights. https://vginsights.com/insights/article/video-game-insights-2021-market-report