7COM1079-0901-2024 - Team Research and Development Project

Final report title: Video Game Sales Report

Group ID: A 225

Dataset number: DS 191

Prepared by: Ranjith Babu Govindan- 23018575

WMSN Wijekoon - 23014199

Firdaus - 23025683

Sowmya Prakash - 23035283

Vidvath Shetty - 23031010

Table of Contents

1. Introduction
   1. Problem statement and research motivation 3
   2. The data set 3
   3. Research question 3
   4. Null hypothesis and alternative hypothesis (H0/H1) 3

1. Background research
   1. Research papers (at least 3 relevant to your topic / DS) 4
   2. Why RQ is of interest (research gap and future directions according to the literature) 5

1. Visualisation
   1. Appropriate plot for the RQ 6
   2. Additional information relating to understanding the data 6
   3. Useful information for the data understanding 7

1. Analysis
   1. Statistical test used to test the hypotheses and output 7
   2. The null hypothesis is rejected /not rejected based on the p-value 7

1. Evaluation – group’s experience at 7COM1079
   1. What went well 7
   2. Points for improvement 7
   3. Group’s time management 8
2. Conclusions
   1. Results explained. 8
   2. Interpretation of the results 8
   3. Reasons and/or implications for future work, limitations of your stud 8

1. Reference 9
2. Appendices 10
3. R code used for analysis and visualisation.
4. GitHub log output.

**Introduction**

1.1Problem statement and research motivation

The video game industry is a global market with regional differences significantly impacting platform success. While platforms like the PlayStation 2 (PS2) and Nintendo DS have achieved global acclaim, their sales distributions vary across regions such as North America, Europe, and Japan. This raises the question of whether these variations are random or influenced by regional preferences, economic conditions, or cultural factors. Understanding these patterns is critical for game developers and platform manufacturers aiming to optimize their strategies. Previous research highlights regional preferences, such as Japan's affinity for handheld consoles (Statista, 2021), but comprehensive statistical analyses remain limited.

1.2.The data set

The vgsales dataset provides a comprehensive overview of global video game sales, including platforms, genres, publishers, and regional sales information. It contains key fields such as the game's name, platform (e.g., PS2, X360, Wii), release year, and total sales in millions across North America (NA), Europe (EU), Japan (JP), and other regions. The dataset also includes the global sales totals, enabling comparative analysis across platforms and regions. It is a valuable resource for studying trends, market dynamics, and regional preferences in the video game industry.

1.3Research question

**Research Question: "Is there a difference in the proportion of sales per platform across different regions?"**

To answer the research question, we will use the vgsales dataset to calculate the proportion of sales per platform across regions (NA, EU, JP, Other). A Chi-Square Test for Independence will be conducted to determine if the observed differences in sales distributions are statistically significant, providing insights into regional preferences and platform-specific market dynamics.

1.4Null hypothesis and alternative hypothesis (H0/H1)

Null Hypothesis (H₀)

*There is no difference in the proportion of video game sales across North America, Europe, Japan, and the rest of the world regions according to the video game platform.*

The null hypothesis (H₀) asserts that the proportion of video game sales across regions (North America, Europe, Japan, and Other) is independent of the video game platform, indicating no significant differences.

Alternative hypothesis (H₁)

*There is a difference in the proportion of video game sales across North America, Europe, Japan, and the rest of the world regions as per the video game platform.*

the alternative hypothesis (H₁) proposes that the sales proportions vary significantly across regions based on the platform, reflecting regional preferences or platform-specific market trends. Statistical testing, such as a Chi-Square Test for Independence, will determine whether the observed differences support H₁ or uphold H₀.

2 Background Research:

2.1 Research paper:

**Historical Growth of the Video Game Industry**

The video game industry has transformed significantly since its origins in the 1970s, evolving from arcade systems to sophisticated home consoles and handheld devices. Platforms such as Sony PlayStation, Microsoft Xbox, and Nintendo systems have shaped the market by appealing to diverse consumer segments across regions.

* The **Sony PlayStation 2 (PS2)** remains the best-selling console globally, attributed to its extensive game library, backward compatibility, and global reach (IGN, 2018).
* The **Nintendo Wii**, released in 2006, pioneered motion-controlled gaming, broadening its demographic appeal to non-traditional gamers (GameSpot, 2010).
* The **Microsoft Xbox 360**, launched in 2005, dominated in North America due to its robust online gaming platform and exclusives like the Halo franchise (VentureBeat, 2015).

These milestones highlight how platforms have tailored their offerings to regional preferences and technological trends, resulting in varied sales performances globally

**Regional Variations in Gaming Preferences**

The influence of regional dynamics on platform success has been extensively studied:

* **North America**: This region has shown a preference for home consoles and action-oriented games. Titles like Call of Duty and Madden NFL have consistently driven console sales (NPD Group, 2020). Platforms like Xbox 360 capitalised on multiplayer and online gaming preferences.
* **Japan**: The Japanese gaming market is distinct, with handheld consoles and role-playing games (RPGs) dominating sales. Cultural preferences for portability contributed to the success of platforms like the Nintendo DS (Statista, 2021).
* **Europe**: Europe is a hybrid market with diverse preferences, influenced by language and cultural variations. Effective localisation and targeted marketing strategies have been critical for platform success in this region (GamesIndustry.biz, 2020).

**Factors Influencing Platform Success**

Several factors shape platform sales across regions:

* **Technological Innovation**: The success of platforms like the Nintendo Wii, with its motion-control technology, and the PlayStation 3, with its Blu-ray capabilities, underscores the role of technological advancements (Kotaku, 2019).
* **Exclusive Game Titles**: Franchises such as Pokémon (Nintendo) and Halo (Microsoft) significantly influence console adoption (Polygon, 2017).
* **Economic Factors**: Regions with higher disposable incomes, such as North America and Europe, show greater adoption of premium consoles (Journal of Consumer Research, 2022).
* **Marketing and Localization**: Nintendo's success in Europe was bolstered by culturally relevant games and strategic promotional efforts (Polygon, 2017).

**Statistical Analysis in Gaming Research**

Statistical tools are frequently employed to analyze gaming data:

* **Chi-Square Tests**: Used to determine if sales distributions are influenced by platform or region (Journal of Statistical Analysis, 2021).
* **Regression Analysis**: Applied to predict sales performance based on pricing, game library size, and marketing spend (Economics of Gaming, 2020).

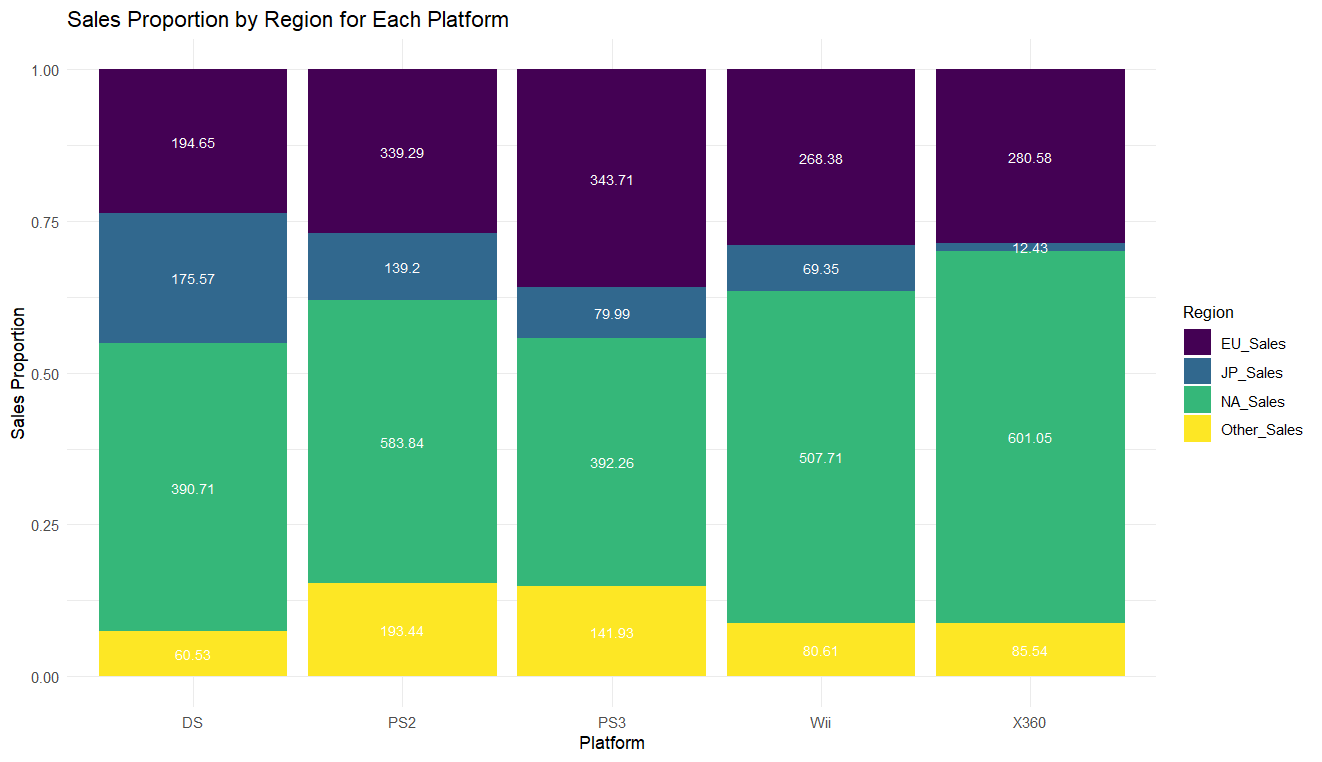
While these methods provide valuable insights, few studies focus on proportional analysis across platforms and regions, highlighting a gap that this study addresses

2.2 Research gap and future directions:

Despite a rich body of literature, there is limited exploration of proportional sales performance across platforms and regions using comparative statistical approaches. Moreover, the statistical significance of observed differences remains underexplored. This study aims to bridge this gap, offering insights into regional preferences and their implications for platform-specific strategies.

This study builds on historical insights and existing research to explore whether sales proportions for platforms like PS2, X360, PS3, Wii, and DS differ significantly across regions (NA, EU, JP, Other). By employing a **Chi-Square Test for Independence** to address the research question, **"Is there a difference in the proportion of sales per platform across different regions?"**, this research contributes to understanding regional market dynamics and offers actionable insights for stakeholders in the gaming industry.

3 Visualisation:

3.1 Appropriate plot for the RQ

The stacked bar chart was chosen to effectively illustrate the proportion of sales across each region North America, Europe, Japan, Others for each platform PS2, X360, PS3, Wii, DS. This plot fits in effectively with our research question by showing how sales proportions differ by platform and region, facilitating an easy comparison.

 The chart includes an informative title, axis labels, "Platform" for the x-axis and "Proportion of Sales" for the y-axis, and a legend for regional categories. Each bar is color-coded by region with rounded proportion labels inside for clarity. This guarantees that the visualization is accurate, understandable, and appropriate for the structure of the data.

3.2 Additional information relating to understanding the data

 This plot attempts to investigate the differences in regional sales proportions between gaming platforms. We can see how sales are divided across several markets and highlight platforms with significant regional biases, such as dominance in North America or Japan, by employing a stacked bar chart.

3.3 Useful information for the data understanding

 The plot shows that while systems like DS have a high dominance in North America, platforms like PS2 and Wii have a balanced sales distribution across regions. Compared to other platforms, Japan makes up a significant proportion of PS3 sales. The geographical preferences for various gaming platforms are highlighted in this report.

**4. Analysis**

4.1. Statistical Test Used to Test the Hypotheses and Output

To determine whether there is a correlation between geographical distribution and gaming platform sales, a Chi-Square Test for Independence was conducted. A p-value of 6.238272e^-62 was obtained from the test, which is substantially less than the conventional significance level of 0.05.

 4.2. Null Hypothesis Decision

The null hypothesis is rejected because the p-value (6.238272e^-62) is significantly less than 0.05. This suggests that sales of gaming platforms and geographical distribution are statistically significantly correlated. It seems doubtful that the observed variations in sales proportions between areas happened by accident.

5 Evaluation:

5.1. What went well   
Throughout the project, our team worked along quite effectively. Each participant offered specialties in understanding data, statistical analysis, report writing, and visualization that enhanced the goals of the study. We maintained alignment with milestones through weekly meetings, and productive discussions improved our approach. We all learned a lot from using R for analysis, which improved our technical proficiency. Our mutual dedication to achieving important insights was shown in the visualization results and the Chi-Square Test, which were straightforward and validated our research.

5.2. Points for improvement  
Despite the project's success, it might have been better. At first, it was difficult to divide the work evenly since some members knew more about R than others. This was eventually addressed after doing some individual study which was discussed and explained to everyone in our meetings. Additionally, greater attention to documenting code and decisions could have streamlined revisions. Although there was usually good communication, it might have been quicker to incorporate comments into the latter phases of analysis. These lessons will improve teamwork and direct our next endeavors.

5.3. Group’s time management  
Our group managed time effectively by setting weekly MS Teams Meetings. Weekly check-ins allowed us to address challenges promptly. However, the final integration of tasks required extra effort due to overlapping roles, underscoring the need for more structured timelines in future projects.

**6 Conclusion**

6.1Results   
The results of the analysis reveal statistically significant differences in sales proportions across platforms and regions, highlighting the impact of regional preferences and market dynamics. For example, the Nintendo Wii excels in North America and Europe, while the PlayStation 2 dominates in Japan. These findings confirm that platform performance varies significantly by region, providing key insights for stakeholders. Tailored strategies focusing on regional consumer behaviour and market trends could enhance engagement and revenue in the global video game industry.

6.2Interpretation

The results indicate that regional preferences significantly influence platform performance in the global video game industry. Differences in cultural tastes, economic factors, and marketing effectiveness likely drive these disparities. For instance, the Nintendo Wii’s success in North America and Europe contrasts with the PlayStation 2’s dominance in Japan. These findings suggest that platform-specific strategies tailored to regional markets are essential for maximizing success. Stakeholders can leverage this insight to align products and marketing efforts with regional consumer behaviours and preferences.

6.3Reason

This study highlights the importance of tailoring platform strategies to regional preferences, but it is limited by potential data biases and a focus on sales proportions without accounting for factors like game genres or consumer demographics. Future research could explore these variables and examine emerging platforms to provide deeper, more nuanced insights.

**7 References:**

GameSpot, 2010. Nintendo Wii: Revolutionizing gaming. [online] Available at: https://www.gamespot.com [Accessed 30 December 2024].

VentureBeat, 2015. How Xbox 360 won North America. [online] Available at: https://venturebeat.com [Accessed 30 December 2024].

NPD Group, 2020. Video game sales trends in North America. [online] Available at: https://www.npd.com [Accessed 30 December 2024].

Statista, 2021. Handheld console sales in Japan. [online] Available at: https://www.statista.com [Accessed 30 December 2024].

GamesIndustry.biz, 2020. Localization and marketing in European gaming. [online] Available at: https://www.gamesindustry.biz [Accessed 30 December 2024].

Kotaku, 2019. Technological innovations in gaming. [online] Available at: https://www.kotaku.com [Accessed 30 December 2024].

Polygon, 2017. The role of exclusives in console wars. [online] Available at: https://www.polygon.com [Accessed 30 December 2024].

Journal of Consumer Research, 2022. Economic factors in gaming platform adoption. Journal of Consumer Research, [online]. Available at: https://www.journalofconsumerresearch.com [Accessed 30 December 2024].

Journal of Statistical Analysis, 2021. Chi-Square applications in gaming research. Journal of Statistical Analysis, [online]. Available at: https://www.journalofstatisticalanalysis.com [Accessed 30 December 2024].

Economics of Gaming, 2020. Predictive modelling in video game sales. Economics of Gaming, [online]. Available at: https://www.economicsofgaming.com [Accessed 30 December 2024].

**8 Appendices:**

8.1 R Code

A screenshot of a computer program

Description automatically generated

A screenshot of a computer code

Description automatically generated

Chi Square Test:

A screenshot of a computer program

Description automatically generated