# Table Of Contents

|  |  |  |
| --- | --- | --- |
| Serial Number | Title | Page number |
| 01 | Introduction to Video Games   * Type of Games * Major Regional Markets * Game Sales | 03 - 08 |
| 02 | INFO REGARDING ALL COLUMNS   * Some Important information about Gaming platforms | 08 - 10 |
| 03 | The reason why I chose this dataset   * AIM (Domain Understanding and Problem) | 10 - 11 |
| 04 | Questions for Analysis | 12 - 12 |
| 05 | Exploratory Data Analysis:   * Libraries used and approaches * Data Cleaning * Univariate Analysis * Bivariate Analysis * Multivariate Analysis * Distributions * Hypothesis Testing | 12 - 15 |
| 06 | Answers of Questions | 16 - 23 |
| 07 | Findings and Insights | 24 - 24 |
| 08 | Limitations | 25 - 25 |
| 09 | Recommendations   * General Recommendations * Company Specific Recommendations * Regional Specific Recommendations | 25 - 27 |
| 10 | Conclusion | 28 - 28 |
| 11 | Acknowledgments | 29 - 29 |

# Introduction to Video Games



A video game or computer game is an [electronic game](https://en.wikipedia.org/wiki/Electronic_game) that involves interaction with a [user](https://en.wikipedia.org/wiki/User_interface) [interface](https://en.wikipedia.org/wiki/User_interface) or [input device](https://en.wikipedia.org/wiki/Input_device) (such as a [joystick](https://en.wikipedia.org/wiki/Joystick), [controller](https://en.wikipedia.org/wiki/Game_controller), [keyboard,](https://en.wikipedia.org/wiki/Computer_keyboard) or [motion-sensing](https://en.wikipedia.org/wiki/Motion_sensing) device) to generate [visual](https://en.wikipedia.org/wiki/Visual) [feedback](https://en.wikipedia.org/wiki/Feedback) from a [display device,](https://en.wikipedia.org/wiki/Display_device) most commonly shown in a [video](https://en.wikipedia.org/wiki/Video) format

on a [television set,](https://en.wikipedia.org/wiki/Television_set) [computer monitor](https://en.wikipedia.org/wiki/Computer_monitor), [flat-panel display](https://en.wikipedia.org/wiki/Flat-panel_display) or [touchscreen](https://en.wikipedia.org/wiki/Touchscreen) on [handheld devices](https://en.wikipedia.org/wiki/Handheld_device), or a [virtual reality headset](https://en.wikipedia.org/wiki/Virtual_reality_headset). Most modern video games are [audiovisual,](https://en.wikipedia.org/wiki/Audiovisual)

with [audio](https://en.wikipedia.org/wiki/Sound) complement delivered through [speakers](https://en.wikipedia.org/wiki/Loudspeaker) or [headphones,](https://en.wikipedia.org/wiki/Headphones) and sometimes also with other types of sensory feedback (e.g., [haptic technology](https://en.wikipedia.org/wiki/Haptic_technology) that provides [tactile](https://en.wikipedia.org/wiki/Touch) sensations), and some video games allow [microphone](https://en.wikipedia.org/wiki/Microphone) and [webcam](https://en.wikipedia.org/wiki/Webcam) inputs for [in-game chatting](https://en.wikipedia.org/wiki/Voice_chat_in_online_gaming) and [live](https://en.wikipedia.org/wiki/Video_game_livestreaming) [streaming.](https://en.wikipedia.org/wiki/Video_game_livestreaming)

Video games are typically categorized according to their [hardware platform](https://en.wikipedia.org/wiki/Hardware_platform), which

traditionally includes [arcade video games,](https://en.wikipedia.org/wiki/Arcade_video_game) [console games](https://en.wikipedia.org/wiki/Console_game), and computer (PC) games; the latter also encompasses [LAN games,](https://en.wikipedia.org/wiki/Network_game) [online games](https://en.wikipedia.org/wiki/Online_game), and [browser games.](https://en.wikipedia.org/wiki/Browser_game) More recently,

the [video game industry](https://en.wikipedia.org/wiki/Video_game_industry) has expanded onto [mobile gaming](https://en.wikipedia.org/wiki/Mobile_gaming) through [mobile devices](https://en.wikipedia.org/wiki/Mobile_device) (such as [smartphones](https://en.wikipedia.org/wiki/Smartphones) and [tablet computers](https://en.wikipedia.org/wiki/Tablet_computers)), [virtual](https://en.wikipedia.org/wiki/Virtual_reality) and [augmented reality](https://en.wikipedia.org/wiki/Augmented_reality) systems, and

remote [cloud gaming.](https://en.wikipedia.org/wiki/Cloud_gaming) Video games are also classified into a wide range of [genres](https://en.wikipedia.org/wiki/Video_game_genre) based on their style of [gameplay](https://en.wikipedia.org/wiki/Gameplay) and [target audience.](https://en.wikipedia.org/wiki/Target_audience)

The first video game prototypes in the 1950s and 1960s were simple extensions of electronic games using video-like output from large, room-sized mainframe computers. The first consumer video game was the arcade video game Computer Space in 1971. In 1972 came the

iconic hit game Pong and the first home console, the Magnavox Odyssey. The industry grew quickly during the "golden age" of arcade video games from the late 1970s to the early 1980s but suffered from the crash of the North American video game market in 1983 due to the loss of publishing control and saturation of the market. Following the crash, the industry matured, was dominated by Japanese companies such as Nintendo, Sega, and Sony, and established practices and methods around the development and distribution of video games to prevent a similar crash in the future, many of which continue to be followed. In the 2000s, the core

industry centered on "AAA" games, leaving little room for riskier experimental games.

Coupled with the availability of the Internet and digital distribution, this gave room for

independent video game development (or "indie games") to gain prominence into the 2010s. Since then, the commercial importance of the video game industry has been increasing. The emerging Asian markets and proliferation of smartphone games are altering player

demographics towards casual gaming and increasing monetization by incorporating games as a service.

Video games require a platform, a specific combination of electronic components or computer hardware and associated software, to operate. The term system is also commonly used. Games are typically designed to be played on one or a limited number of platforms, and exclusivity to a platform is used as a competitive edge in the video game market. However, games may be developed for alternative platforms than intended, which are described as ports or conversions. These also may be remastered - where most of the original game's source code is reused and art assets, models, and game levels are updated for modern systems – and remakes, where in addition to asset improvements, a significant reworking of the original game and possibly from scratch is performed.

* **PC games** – PC games involve a player interacting with a personal computer (PC) connected to a video monitor. Personal computers are not dedicated game platforms, so there may be differences running the same game on different hardware.



* **Home console –** The [PlayStation 2](https://en.wikipedia.org/wiki/PlayStation_2) is the [best-selling video game console](https://en.wikipedia.org/wiki/List_of_best-selling_game_consoles), with over 155 million units sold. A [console game](https://en.wikipedia.org/wiki/Console_game) is played on a [home console,](https://en.wikipedia.org/wiki/Home_video_game_console) a specialized

electronic device that connects to a common [television set](https://en.wikipedia.org/wiki/Television_set) or [composite video monitor](https://en.wikipedia.org/wiki/Composite_monitor).



* **Handheld Console** – A handheld game console is a small, self-contained electronic device that is portable and can be held in a user's hands.



* **Arcade video game** – An arcade video game generally refers to a game played on an even more specialized type of electronic device that is typically designed to play only one game and is encased in a special, large coin-operated cabinet that has one built-in console, controllers (joystick, buttons, etc.), a CRT screen, and audio amplifier and speakers.



* **Browser Game** – A browser game takes advantage of standardizations of technologies for the functionality of web browsers across multiple devices providing a cross-platform environment.
* **Mobile Game** – With the introduction of smartphones and tablet computers standardized on the iOS and Android operating systems, mobile gaming has become a significant platform.



* **Cloud Gaming** – Cloud gaming requires a minimal hardware device, such as a basic computer, console, laptop, mobile phone, or even a dedicated hardware device connected to a display with good Internet connectivity that connects to hardware systems by the cloud gaming provider.
* **Virtual Reality** – Virtual reality (VR) games generally require players to use a special head-mounted unit that provides stereoscopic screens and motion tracking to immerse a player within a virtual environment that responds to their head movements.



* **Emulation** – An emulator enables games from a console or otherwise different system to be run in a type of virtual machine on a modern system, simulating the hardware of the original and allowing old games to be played.



# Major Regional Markets

The industry itself grew out of both the United States and Japan in the 1970s and 1980s before having a larger worldwide contribution. Today, the video game industry is predominantly led by major companies in North America (primarily the United States and Canada), Europe, and Southeast Asia including Japan, South Korea, and China. Hardware production remains an area dominated by Asian companies either directly involved in hardware design or part of the production process, but digital distribution and indie game development of the late 2000s have allowed game developers to flourish nearly anywhere and diversify the field.

# Game sales

According to the market research firm Newzoo, the global video game industry drew

estimated revenues of over $159 billion in 2020. Mobile games accounted for the bulk of this, with a 48% share of the market, followed by console games at 28% and personal computer games at 23%.

Sales of different types of games vary widely between countries due to local preferences.

Japanese consumers tend to purchase much more handheld games than console games

especially PC games, with a strong preference for games catering to local tastes. Another key difference is that, though having declined in the West, arcade games remain an important sector of the Japanese gaming industry. In South Korea, computer games are generally preferred over console games, especially MMORPG games and real-time strategy games.

Computer games are also popular in China.

***INFO REGARDING ALL COLUMNS IN MY DATASET***

1. Rank - Depicts the video game which has the highest overall sales.
2. Name - Names of the Video Games.
3. Platform - In which the Video Game is supported.
4. Year - The year in which the video game is released.
5. Genre - The most widely used game classifying system categorizes games into eight genres: action, adventure, fighting, puzzle, role-playing, simulation, sports, and strategy.
6. Publisher - The company that released the game.
7. NA Sales - Sales of Video Games ever recorded in North America.
8. EU Sales - Sales of Video Games ever recorded in Europe.
9. JP Sales - Sales of Video Games ever recorded in Japan.
10. Other Sales - Sales of Video Games ever recorded in Every other country except North America, Europe, and JP.
11. Global Sales - Sales of all Video Games ever recorded in all Countries worldwide.

# Some Important Information about Gaming Platforms

* WS -- The WonderSwan is a handheld game console released in Japan by Bandai. It was developed by Gunpei Yokoi's company Koto Laboratory and Bandai.
* 3DO -- The 3DO Interactive Multiplayer, also referred to as simply 3DO, is a home video game console developed by the 3DO Company.
* DC -- Sega's Dreamcast is the first console of the generation and had several features to show an advantage from the competition, including Internet gaming as an optional feature through its built-in modem, and a web browser.
* SAT -- The Sega Saturn is a home video game console developed by Sega and released on November 22, 1994.
* SCD -- The Sega CD plays CD-based games and adds hardware functionality such as a faster central processing unit and graphic enhancements like sprite scaling.
* 2600 -- The Atari 2600 is a home video game console developed and produced by Atari, Inc. Released in September 1977.
* PS -- PlayStation Network's services are dedicated to an online marketplace

(PlayStation Store), a premium subscription service for enhanced gaming and social features (PlayStation Plus), music streaming (PlayStation Music, based on Spotify), TV streaming (PlayStation Vue), and formerly a cloud gaming service

* PSP -- The PlayStation Portable (PSP) is a handheld game console developed and marketed by Sony Interactive Entertainment.
* PCFX -- The PC-FX is a 32-bit home video game console developed by NEC and Hudson Soft. It was released in 1994 and discontinued in February 1998
* PS2 -- The PlayStation 2 (PS2) is a home video game console developed and marketed by Sony Interactive Entertainment.
* PSV -- The PlayStation Vita (PS Vita, or Vita) is a handheld game console developed and marketed by Sony Interactive Entertainment.
* PS3 -- The PlayStation 3 (PS3) is a home video game console developed and marketed by Sony Interactive Entertainment. The successor to the PlayStation 2, it is part of the PlayStation brand of consoles.
* TG16 -- The TurboGrafx-16, known as the PC Engine outside North America, is a home video game console designed by Hudson Soft and sold by NEC Home

Electronics.

* PS4 -- The PlayStation 4 (PS4) is a home video game console developed by Sony Interactive Entertainment. Announced as the successor to the PlayStation 3, it is part of the PlayStation brand of consoles.
* PC -- A personal computer game, also known as a computer game or abbreviated PC game, is an electronic game ·
* N64 -- The Nintendo 64 (N64) is a home video game console developed by Nintendo.
* GB -- The Game Boy is an 8-bit fourth-generation handheld game console developed and manufactured by Nintendo.
* Wii -- The Wii is a home video game console developed and marketed by Nintendo.
* WiiU -- Wii is a series of simulation games published by Nintendo for the game console of the same name, as well as its successor, the Wii U.
* DS -- The Nintendo DS is a handheld game console manufactured by Nintendo and released worldwide from 2004 to 2005. DS is an acronym for "developer system" or "dual screen" and introduced new features characteristic of handheld games.
* 3DS -- The Nintendo 3DS is a handheld game console produced by Nintendo. The console was announced in March 2010 and unveiled at E3 2010 as the successor to the Nintendo DS.
* NES -- The NES features several groundbreaking games, such as the 1985 platform game Super Mario Bros. and the 1986 action-adventure game The Legend of Zelda.
* SNES -- The Super Nintendo Entertainment System, commonly shortened to Super Nintendo, Super NES or SNES is a 16-bit home video game console developed by Nintendo.
* GC -- The Nintendo GameCube is a home video game console developed and released by Nintendo.
* GBA -- The Game Boy Advance (GBA) is a 32-bit handheld game console developed, manufactured, and marketed by Nintendo as the successor to the Game Boy Colour.
* XB -- The Xbox is a home video game console manufactured by Microsoft that is the first installment in the Xbox series of video game consoles.
* XONE -- The Xbox One is a home video game console developed by Microsoft. Announced in May 2013.
* X360 -- The Xbox 360 is a home video game console developed by Microsoft. As the successor to the original Xbox, it is the second console in the Xbox series.
* GG -- Guilty Gear, subtitled The Missing Link in Japan, is a 2D fighting video game developed by Team Neo Blood.

# The reason why I chose this dataset

Video game culture is a worldwide new media subculture formed around video games and game playing. As computer and video games have increased in popularity over time, they have had a significant influence on popular culture. Video game culture has also evolved hand in hand with internet culture as well as the increasing popularity of mobile games. Many people who play video games identify as gamers, which can mean anything from someone who enjoys games to someone passionate about them. As video games become more social with multiplayer and online capabilities, gamers find themselves in growing social networks. Gaming can both be entertainment as well as competition, as a new trend known as electronic sports is becoming more widely accepted. In the 2010s, video games and discussions of video game trends and topics can be seen in social media, politics, television, film, and music. The COVID-19 pandemic during 2020–2021 gave further visibility to video games as a pastime

to enjoy with friends and family online as a means of social distancing.

Moving beyond its origins as a niche hobby, video game culture has surged into a global phenomenon that resonates across various aspects of society. As technology advanced and gaming platforms diversified, the influence of video games on popular culture has grown exponentially. This evolution has gone hand in hand with the rise of internet culture and the

the soaring popularity of mobile games. The term "gamer" encapsulates a wide spectrum, ranging from casual enthusiasts to dedicated aficionados deeply immersed in the gaming world.

The impact of video games extends far beyond entertainment, with the industry shaping markets and trends over the years. The proliferation of multiplayer and online gaming has fostered expansive social networks, transforming solitary gaming into a communal

experience. Moreover, the emergence of electronic sports (eSports) has forged a new frontier, where players compete professionally in tournaments and championships, captivating audiences globally.

Throughout the 2010s, the ubiquity of video games permeated various facets of contemporary life, with discussions on gaming trends becoming prevalent in social media, politics,

television, film, and music. As the world grappled with the challenges posed by the COVID-19 pandemic in 2020–2021, video games garnered even greater attention as a means of connecting with friends and family during a time of physical distancing.

**Aim:**

*To comprehensively understand the trajectory of video gaming trends and their profound market impact, one must delve into the vast array of games released over the years. These releases, spanning various genres and platforms, have been orchestrated by numerous*

*publishers, each contributing to the evolving landscape of the gaming industry. By examining the chronology of these releases and the strategies adopted by publishers, we can gain*

*valuable insights into the industry's growth and transformation. From seminal titles that revolutionized gameplay to the strategic decisions that defined specific years of release, the narrative of video game culture intertwines with the dynamic interplay of creativity,*

*technology, and market forces.*

Understanding the trajectory of video gaming trends and their profound market impact requires delving into the extensive history of game releases. These releases, encompassing a wide array of genres and platforms, have been orchestrated by numerous publishers,

collectively shaping the dynamic landscape of the gaming industry. By examining the chronology of these releases and the strategies employed by publishers, we can uncover invaluable insights into the industry's growth and transformative journey.

From pioneering titles that revolutionized gameplay mechanics to strategic decisions that defined specific years of release, the narrative of video game culture is intricately intertwined with the interplay of creativity, technological advancements, and market dynamics. Analyzing how games have evolved provides a window into consumer preferences, technological advancements, and market trends. This historical exploration illuminates pivotal moments

that have propelled the industry forward and sheds light on the symbiotic relationship

between innovation and consumer demand. Ultimately, the intricate tapestry of the gaming industry's evolution showcases its resilience in adapting to change while consistently captivating a global audience.

# Questions for Analysis

1. How is sales distribution across the global market?
2. The platforms that support video games?
3. Number of games that support these Platforms?
4. Trends in Video Game releases each year.
5. How many Games were released in each year?
6. What is the distribution of genre in Video Games?
7. What is the trend of Global Sales each year?
8. How is the Global Market distributed?
9. What is the top-sold Genre in the North American market?
10. What is the top-sold Genre in the European market?
11. What is the top-sold Genre in the Japanese market?
12. What is the top-sold Genre in Other Country Markets?
13. What is the top-sold Genre in the Global Market?
14. Top Publishers who have released more Video Games?
15. Correlation Between Global Sales, NA Sales, EU Sales, JP Sales, and Other Sales?
16. To understand which variables are directly influencing global sales.
17. What are the most common game genres in the dataset, and how many games belong to each genre?
18. What are the top-selling game genres across different market areas?
19. How do the sales performances vary across North America, Europe, Japan, and other regions?
20. Which genres of the highest publisher have the highest sales in specific market areas?

***Exploratory Data Analysis:***

# Libraries used and approaches

* **Pandas:** This library provides data structures and tools for working with tabular data, such as data frames, series, and indexes. It also supports various operations such as filtering, grouping, merging, reshaping, and aggregating data. pandas is based on numpy, which is another library for numerical computing.
* **sklearn.decomposition (PCA):** A library used for Principal Component Analysis (PCA), a technique for reducing the dimensionality of the data used in multivariant analysis PCA is one of the many techniques to perform multivariant analysis.
* **numpy (np):** A library used for numerical computing in Python. It is not used in the provided code snippet.
* **warnings:** A library used to handle warnings in Python. In this document, it is used to ignore warnings.
* **matplotlib. pyplot (plt):** A library used for creating static, animated, and interactive visualizations in Python. It is not used in the provided code snippet.
* **seaborn (sns):** A library based on matplotlib used for statistical data visualization. It is not used in the provided code snippet.

***Data Cleaning***

The data cleaning steps taken to handle missing values, outliers, and other data quality issues are as follows:

1. **Finding Null Values:** The first step in data cleaning is to identify the presence of any null values in the dataset. The code

*df.isnull().sum()*

is used to calculate the total number of null values in each column.

1. **Calculating the Percentage of Missing Values:** To understand the impact of missing values on the dataset, the percentage of missing values is calculated for each column using the formula

*percent\_missing = df. isnull().sum() \* 100/ len(df)*

. This helps in deciding whether to impute or remove the missing values based on their proportion in the dataset.

1. **Displaying Null Valued Rows:** To further analyze the missing values, the rows containing null values are displayed using a null mask, and a new DataFrame ‘*ds*’ is created to store these rows. This helps in understanding the pattern of missing values and deciding the appropriate method to handle them.
2. **Handling Outliers:** The document also deals with the presence of outliers in the dataset. Outliers can significantly affect the analysis and modeling process. In this case, the document uses log transformation to handle the outliers in the sales columns (NA\_Sales, EU\_Sales, JP\_Sales, Other\_Sales, and Global\_Sales). Log transformation is a common technique to reduce the impact of outliers and make the data more normally distributed.
3. **Visualizing the Distributions:** After applying the log transformation, the document visualizes the distribution of the sales columns using histograms and Seaborn's distplot function. This helps in understanding the distribution of the data and confirming if the log transformation has effectively handled the outliers.

In summary, the data cleaning process in the document involves identifying and analyzing missing values, calculating their percentage, handling outliers using log transformation, and visualizing the distributions to ensure data quality.

# Univariate Analysis

In this analysis, univariate analysis was conducted by examining the distribution of video game sales variables such as NA Sales, EU Sales, JP Sales, Other Sales, and Global Sales. This was done using techniques like histograms and density plots to visualize the distribution of the data.

By analyzing these individual variables, we can understand their skewness, central tendency, and spread. For example, we can determine if the data is normally distributed or if it is skewed to one side. This information can be useful in identifying outliers, understanding the variability in sales, and making informed decisions based on the characteristics of each variable.

***NOTE:*** *Univariate analysis provides insights into individual variables and their characteristics. It helps us understand the data at a granular level. However, it is often necessary to combine univariate analysis with other techniques like bivariate or multivariate analysis to gain a comprehensive understanding of the data and draw meaningful conclusions.*

# Bivariate Analysis

In this analysis, bivariate analysis was conducted by creating scatter plots to visualize the relationships between different sales variables. The scatter plots show the relationship between global sales and other sales, NA sales, and JP sales.

This information can be useful in understanding the market dynamics and identifying potential opportunities or areas of improvement for companies.

***NOTE****: Bivariate analysis provides insights into the relationship between two variables, but it does not establish causation.*

# Multivariate Analysis

In this analysis, multivariate analysis techniques were utilized to examine the interactions and correlations among multiple variables. One of the techniques mentioned in the report is dimensionality reduction using Principal Component Analysis (PCA).

PCA is a statistical technique that reduces the dimensionality of a dataset while retaining most of the important information. It does this by transforming the original variables into a new set of variables called principal components. These principal components are linear combinations of the original variables and are chosen in such a way that they capture the maximum amount of variation in the data.

By applying PCA, I aimed to identify the underlying patterns and structure in the dataset. It helps in reducing the complexity of the data and visualizing the relationships among variables in a lower-dimensional space.

# Distributions

To examine distributions in the report, techniques such as histograms, density plots, or probability density functions have been utilized. Histograms are graphical representations that showcase the frequency or count of observations within different intervals or bins, allowing for a visual understanding of the data's shape and spread.

By analyzing distributions, I aimed to gain a deeper understanding of the characteristics and behavior of the variable under investigation. This information can be valuable in identifying outliers, comprehending the skewness or symmetry of the data, and making informed decisions based on the distributional properties of the variable.

***Hypothesis Testing***

**Hypothesis 1:** The average video game sales in North America are higher than the average video game sales in Europe.

* Null Hypothesis (H0): mu\_NA <= mu\_EU
* Alternative Hypothesis (Ha): mu\_NA > mu\_EU

A two-sample t-test was conducted to compare the means of the two groups. The results were:

* t-statistic: 15.831371100049914
* p-value: 3.0384925458015177e-56

The t-statistic of 15.83 and p-value of 3.0384925458015177e-56 provide strong evidence that the mean video game sales in North America are significantly higher than the mean video game sales in Europe. Thus, the results support the first hypothesis.

**Hypothesis 2:** The variance of video game sales in Japan is higher than the variance of video game sales in North America.

* Null Hypothesis (H0): var\_JP <= var\_NA
* Alternative Hypothesis (Ha): var\_JP > var\_NA

An F-test was conducted to compare the variances of the two groups. The results were:

* F-statistic: 760.1394504077623
* p-value: 1.8417918080111082e-165

The F-statistic of 760.14 and p-value of 1.8417918080111082e-165 provide even stronger evidence that the variance of video game sales in Japan is significantly higher than the variance of video game sales in North America. Thus, the results support the second hypothesis.

# Answers of questions

1. **How is sales distribution across the global market?**

Global Market is distributed into 4 markets majorly with shares

North America Market - 49.3%

European Market - 27.3%

Japanese Market - 14.5%

Other Market (all other Countries) - 8.9%

And North America has the highest shares

1. **The platforms that support video games?**

**The platforms are:**

DS, PS2, PS3, Wii, X360, PSP, PS, PC, XB, GBA, GC, 3DS, PSV, PS4, N64, SNES, XOne, SAT, WiiU, 2600, NES, GB, DC, GEN, NG, SCD, WS, 3DO, TG16, GG, PCFX

The description of these platforms is on page number 8

1. **Number of games that support these Platforms?**

|  |  |
| --- | --- |
| **Platform** | **Count of Games Support** |
| DS | 2163 |
| PS2 | 2161 |
| PS3 | 1329 |
| Wii | 1325 |
| X360 | 1265 |
| PSP | 1213 |
| PS | 1196 |
| PC | 960 |
| XB | 824 |
| GBA | 822 |
| GC | 556 |
| 3DS | 509 |
| PSV | 413 |
| PS4 | 336 |
| N64 | 319 |
| SNES | 239 |
| XOne | 213 |
| SAT | 173 |
| WiiU | 143 |
| 2600 | 133 |
| NEWS | 98 |
| GB | 98 |
| DC | 52 |
| GEN | 27 |
| NG | 12 |
| SCD | 6 |
| WS | 6 |
| 3DO | 3 |
| TG16 | 2 |
| GG | 1 |
| PCFX | 1 |

1. **Trends in Video Game releases each year.**

* Increasing Trend (2005-2009): The number of video game releases steadily increased from 941 in 2005 to 1473 in 2007, reaching a peak of 1473 in 2007. This period saw a significant rise in the gaming industry.
* Decreasing Trend (2009-2016): After 2009, there was a gradual decline in the number of game releases. The numbers decreased from 1473 in 2007 to 344 in 2016. This decline could be attributed to various factors, such as market saturation or shifts in gaming platforms.
* Stabilization (2016-2020): After 2016, the number of game releases stabilized at around 300-400 per year. There might be a new balance in the industry, with a consistent number of games being released annually.
* Significant Drops (2017-2020): There was a sharp drop in 2017 with only 3 releases, and in 2020, there was only 1 release. These years might represent a shift in the industry, possibly due to emerging trends like mobile gaming or changes in consumer preferences.

1. **How many Games were released each year?**

The number of games released each year shows that the video game industry has grown significantly over time, reaching its peak in 2007-2011. The number of games released has declined since then, possibly due to market saturation, higher production costs, or changing consumer preferences.

The video game industry boomed from 2007-2011 with over 1000 games per year. Several games dropped after 2011, maybe due to oversupply, expense, or demand shift. See the table for more details.

|  |  |
| --- | --- |
| **Year** | **Number of Games** |
| 2007 | 1473 games |
| 2009 | 1431 games |
| 2008 | 1428 games |
| 2010 | 1259 games |
| 2011 | 1139 games |
| 2006 | 1008 games |
| 2005 | 941 games |
| 2002 | 829 games |
| 2003 | 775 games |
| 2004 | 763 games |
| 2012 | 657 games |
| 2015 | 614 games |
| 2014 | 582 games |
| 2013 | 546 games |
| 2001 | 482 games |
| 1998 | 379 games |
| 2000 | 349 games |
| 2016 | 344 games |
| 1999 | 338 games |
| 1997 | 289 games |
| 1996 | 263 games |
| 1995 | 219 games |
| 1994 | 121 games |
| 1993 | 60 games |
| 1981 | 46 games |
| 1992 | 43 games |
| 1991 | 41 games |
| 1982 | 36 games |
| 1986 | 21 games |
| 1989 | 17 games |
| 1983 | 17 games |
| 1990 | 16 games |
| 1987 | 16 games |
| 1988 | 15 games |
| 1985 | 14 games |
| 1984 | 14 games |
| 1980 | 9 games |
| 2017 | 3 games |
| 2020 | 1 game |

1. **What is the distribution of genre in Video Games?**

Action, sports, and misc games dominate the video game market. Role-playing, shooter, adventure, racing, platform, simulation, and fighting games have moderate popularity. Strategy and puzzle games have the lowest demand. Each genre appeals to different types of players and preferences.

1. **What is the trend of Global Sales each year?**

* 1980s to early 1990s: Global sales started relatively low in the early 1980s, gradually increased, and then spiked in 1989 and 1994.
* The mid-1990s to early 2000s: After the spike in 1994, sales remained consistently high in the mid to late 1990s and early 2000s, peaking in 1998.
* Early 2000s: There was a slight decline in the early 2000s, followed by a significant increase in 2001 and 2002.
* The mid-2000s to 2010s: Sales remained high in the mid-2000s, reaching its peak in 2008, after which there was a notable decline. The sales gradually stabilized but remained relatively high until 2010.
* 2010s to 2020: From 2010 onwards, there was a general downward trend, with sales decreasing each year. There was a steep decline in 2013, followed by a slight increase in 2014, but the overall trend remained negative.
* 2016 to 2020: Sales continued to decrease significantly, reaching almost negligible levels in 2017 and remaining extremely low in 2020.

1. **How is the Global Market distributed?**

The global sales of video games have fluctuated over the years, with some peaks and valleys. The highest global sales were recorded in 2008 with 678.90 million units sold, followed by 2007 with 711.21 million units sold. The lowest global sales were recorded in 2017 with only 0.05 million units sold, followed by 2020 with 0.29 million units sold. Also, global sales have declined sharply since 2010, indicating a possible saturation or decline in the video game market.

1. **What is the top-sold Genre in the North American market?**

Top sold Genre in the North American Market is the Action genre with 20%

|  |  |
| --- | --- |
| **Genre** | **Sales Percentage** |
| Action | 20.0% |
| Strategy | 15.6% |
| Simulation | 13.3% |
| Shooter | 10.2% |
| Role-Playing | 9.3% |
| Racing | 8.2% |
| Platform | 7.5% |
| Misc | 5.1% |
| Fighting | 4.2% |
| Sports | 2.4% |
| Puzzle | 1.6% |
| Adventure | 0.026% |

1. **What is the top-sold Genre in the European market?**

Top sold Genre in the European Market is the Action genre with 21.6%

|  |  |
| --- | --- |
| **Genre** | **Sales Percentage** |
| Action | 21.6% |
| Strategy | 15.5% |
| Simulation | 12.9% |
| Shooter | 10.2% |
| Racing | 9.8% |
| Role-Playing | 8.9% |
| Platform | 7.7% |
| Fighting | 4.7% |
| Misc | 4.2% |
| Sports | 2.6% |
| Adventure | 2.1% |
| Puzzle | 1.9% |

1. **What is the top-sold Genre in the Japanese market?**

Top top-sold Genre in the Japanese market is the Platform genre with 27.3%

|  |  |
| --- | --- |
| **Genre** | **Sales Percentage** |
| Platform | 27.3% |
| Action | 12.4% |
| Strategy | 10.5% |
| Shooter | 10.1% |
| Role-Playing | 8.3% |
| Misc | 6.8% |
| Fighting | 4.9% |
| Adventure | 4.4% |
| Racing | 4.4% |
| Sports | 4.0% |
| Puzzle | 3.8% |
| Simulation | 3.0% |

1. **What is the top-sold Genre in Other Country Markets?**

Top sold Genre in Other Countries Markets is the Action genre with 23.5%

|  |  |
| --- | --- |
| **Genre** | **Sales Percentage** |
| Action | 23.5% |
| Strategy | 16.9% |
| Simulation | 12.9% |
| Racing | 9.7% |
| Role-Playing | 9.4% |
| Platform | 7.5% |
| Shooter | 6.5% |
| Misc | 4.6% |
| Fighting | 4.0% |
| Sports | 2.1% |
| Adventure | 1.6% |
| Puzzle | 1.4% |

1. **What is the top-sold Genre in the Global Market?**

Top sold Genre Globally is the markets is Action genre with 19.6%

|  |  |
| --- | --- |
| **Genre** | **Sales Percentage** |
| Action | 19.6% |
| Strategy | 14.9% |
| Simulation | 11.6% |
| Platform | 10.4% |
| Shooter | 9.3% |
| Role-Playing | 9.1% |
| Racing | 8.2% |
| Misc | 5.0% |
| Fighting | 4.4% |
| Adventure | 2.7% |
| Sports | 2.7% |
| Puzzle | 2.0% |

1. **Top Publishers Who Have Released More Video Games?**

These are the top 10 Publishers on the Number of games released in total

|  |  |  |
| --- | --- | --- |
| **Publisher** | **Count of games** | **Percentage** |
| Electronic Arts | 1409 | 17.1% |
| Activision | 975 | 11.9% |
| Namco Bandai Games | 932 | 11.3% |
| Ubisoft | 921 | 11.2% |
| Konami Digital Entertainment | 832 | 10.1% |
| THQ | 715 | 8.7% |
| Nintendo | 703 | 8.6% |
| Sony Computer Entertainment | 683 | 8.3% |
| Sega | 639 | 7.8% |
| Take-Two Interactive | 413 | 5.0% |

1. **Correlation Between Global Sales, NA Sales, EU Sales, JP Sales, and Other Sales?**

* All the Sales NA Sales, EU Sales, JP Sales, and Other Sales are positively linearly co-related, whereas the JP Sales is a bit fluctuated from other sales due to differences in the preferred genre by consumers and differences in the direction of the graph.
* Co-relation does not show the relation between the variables it shows only linear co-relation between the variables

1. **To understand which variables are directly influencing global sales.**

* Variables that directly influence the global market are
* Consumer preference of Video Games
* Availability of technology to play video games
* Economic constraints of a Country
* Local beliefs and culture
* Availability of genre of video games
* Top Publishers

1. **What are the most common game genres in the data set, and how many games belong to each genre?**

The most common game genres in the data set are action, sports, and misc, which together account for more than half of the total number of games. Here is a table that shows the count of games for each genre:

|  |  |
| --- | --- |
| **Genre** | **Count** |
| Action | 3316 |
| Sports | 2346 |
| Misc | 1739 |
| Role-Playing | 1488 |
| Shooter | 1310 |
| Adventure | 1286 |
| Racing | 1249 |
| Platform | 886 |
| Simulation | 867 |
| Fighting | 848 |
| Strategy | 681 |
| Puzzle | 582 |

1. **What are the top-selling game genres across different market areas?**
2. The top-selling genre in North America is 'Action' with total sales of 877.83 million copies.
3. The top-selling genre in Europe is 'Action' with total sales of 525.0 million copies.
4. The top-selling genre in Japan is 'Role-Playing' with total sales of 352.31 million copies.
5. The top-selling genre in Other Countries is 'Action' with total sales of 187.38 million copies.
6. **How do the sales performances of different game genres vary across North America, Europe, Japan, and other regions?**

The sales of different game genres vary by region due to consumer preferences and tastes. North America and Europe prefer action, sports, and shooter games. Japan prefers role-playing and puzzle games. Other regions have lower sales and similar preferences as North America and Europe.

1. **Which genres of the highest publisher have the highest sales in specific market areas?**

Electronic Arts is the publisher with the greatest number of games released (1409) in count contributing 17.1% in all the games released

1. North America: The most popular genres in this market are shooter, sports, and racing games, which have high sales for games like Battlefield 3, Madden NFL 2004, and Need for Speed Underground. These genres reflect the preference for action, competition, and realism among North American consumers.
2. Europe: The most popular genres in this market are sports, simulation, and action games, which have high sales for games like FIFA 16, The Sims 3, and FIFA Soccer 13. These genres reflect the influence of European culture and lifestyle, such as football, social interaction, and diversity.
3. Japan: The most popular genre in this market is role-playing games, which have high sales for games like Final Fantasy X-2. This genre reflects the cultural and historical influences of Japanese mythology, folklore, and anime. Other genres have low sales in this market, possibly due to the dominance of local publishers and developers.
4. Other regions: The most popular genres in this market are similar to North America and Europe, such as action, sports, and shooter games. However, the sales performances of these genres are generally lower than the other markets, due to various factors such as economic development, internet access, gaming culture, etc.

***Findings and Insights***

* The most successful publisher in terms of global sales is Nintendo, having published almost all the top 10 games in the table, with a total of 369.37 million units sold globally.
* Sales Distribution: The data suggests that North America (NA\_Sales) has consistently higher sales compared to Europe (EU\_Sales), Japan (JP\_Sales), and other regions (Other\_Sales). This indicates that the North American market is a significant contributor to the overall sales of video games.
* Sports and Racing genres are popular in Europe while in North America the most popular genres are shooter, sports, and racing games, In Japan the most popular genre in this market is role-playing games and in Other Countries, the most popular genres are action, sports, and shooter games.
* The most popular genres are Sports and Platform, with Wii Sports and Super Mario Bros. being the best-selling games in each genre.
* The most popular platforms are Wii and NES, with four games each in the top 10.
* The distribution of release years shows that the number of games released has increased over time, with a peak around 2005-2010.
* The distribution of genres shows that Action, Sports, and Misc are the most common genres, accounting for approximately 45% of all games.
* The trend of global sales shows a significant increase in sales from the 1980s to the mid-2000s, followed by a decline in sales after 2010.
* The correlation heatmap indicates a strong positive correlation between sales in different regions (NA\_Sales, EU\_Sales, JP\_Sales, and Other\_Sales), suggesting that games that sell well in one region tend to sell well in other regions as well.
* There is a negative correlation between Rank and sales in different regions, indicating that games with higher sales have a lower rank (i.e., they are more popular).
* There is a weak negative correlation between Year and sales in different regions, suggesting that more recent games may have slightly lower sales compared to older games.

***Limitations***

* Missing data: The dataset has some missing values, particularly in the 'Year' and 'Publisher' columns. This could potentially impact the accuracy of the analysis and lead to biased results.
* Data quality: The dataset only provides information up to Jan 2020, which means that the analysis may not be representative of the current state of the video game industry.
* Confounding variables: The analysis does not consider potential confounding variables that could influence the sales of video games, such as marketing strategies, game quality, or cultural preferences. These factors could have a significant impact on the observed differences in sales across regions.
* Statistical tests: The analysis uses t-tests and F-tests to compare the means and variances of sales data between regions. While these tests provide some evidence for the hypotheses, they do not account for potential interactions between variables or the influence of other factors on the sales data.

***Recommendations***

***General Recommendations: -***

1. Develop games for cross-platform play: Cross-platform play allows players to play games on different devices and platforms. Companies can develop games that allow cross-platform play to tap into this growing market.
2. The report highlights that Europe is the second-highest region in terms of video game sales. Companies could explore opportunities to expand their presence in the European market by localizing their games, partnering with European publishers, or conducting targeted marketing campaigns to appeal to European gamers.
3. While it is important to cater to regional preferences, companies should also consider diversifying their game offerings across different genres. This allows them to appeal to a wider range of gamers and mitigate risks associated with relying too heavily on a single genre.
4. Conduct comprehensive market research to understand the preferences, interests, and demographics of gamers in different regions. This will help companies identify target markets, tailor their game offerings, and develop effective marketing strategies.

***Company Specific Recommendations: -***

**Nintendo:**

* As the dominant publisher in both North America and other countries, Nintendo can leverage its strong market presence to further capitalize on its success.
* Could consider expanding their game offerings in Europe and Japan, where they already have a significant presence, by developing localized content and marketing strategies to the preferences of gamers in those regions.

**Take-Two Interactive:**

* With Grand Theft Auto: San Andreas being one of the highest-selling games in the action genre, Take-Two Interactive can focus on further strengthening its position in the action game market.
* They could invest in developing sequels or new titles within the action genre to cater to the demand and capitalize on the success of their existing games.

**Sony Computer Entertainment:**

* Gran Turismo 4 has shown strong sales in Europe. Sony Computer Entertainment can leverage this success by continuing to invest in the racing genre and developing new titles or sequels within the Gran Turismo franchise.
* They could also explore partnerships with European racing teams or events to enhance brand visibility and appeal to the European market.

**Activision:**

* Call of Duty: Black Ops II has performed well in terms of other sales. Activision can build upon this success by continuing to invest in the first-person shooter genre and developing new installments or innovative gameplay features to keep the franchise fresh and engaging for players.
* They could also consider expanding their marketing efforts in regions outside of North America and Europe to tap into new markets.

***Regional Specific Recommendations: -***

**North America (Action genre):**

* Companies can capitalize on the popularity of the action genre in North America by developing and promoting action-packed games that offer thrilling gameplay experiences.
* Collaborating with popular action movie franchises or incorporating cinematic elements can attract North American gamers who enjoy high-intensity gameplay.
* Engaging with gaming influencers and leveraging social media platforms can help generate buzz and reach the target audience effectively.

**Europe (Sports and Racing genres):**

* Companies can focus on developing and marketing sports and racing games specifically for the European market. This could involve partnering with European sports leagues or teams to create authentic and localized gaming experiences.
* Investing in realistic graphics, immersive gameplay, and multiplayer features can enhance the appeal of sports and racing games in Europe.
* Conducting market research to understand specific sports preferences in different European countries can help tailor game content and marketing strategies accordingly.

**Japan (Role-Playing and Puzzle genres):**

* Companies can focus on developing role-playing and puzzle games that cater to the preferences of Japanese gamers.
* Incorporating rich storytelling, character development, and immersive gameplay mechanics can enhance the appeal of role-playing games in Japan.
* Collaborating with popular anime or manga franchises can help attract Japanese gamers who are fans of these genres.

***Conclusion***

This in-depth exploration of the video game dataset has illuminated several pivotal aspects of the gaming industry, offering valuable insights for stakeholders. Here is a concise summary of our findings:

1. Publisher Influence:

Electronic Arts, Activision, Namco Bandai Games, and Ubisoft are major players, collectively constituting a significant portion of the market share. Electronic Arts alone accounts for 17.1% of the games in the dataset, underlining its substantial impact on the industry.

1. Yearly Trends:

The gaming industry witnessed a surge in releases around 2007, with 1473 games. While the number of releases dipped in subsequent years, 2009, 2008, and 2010 still saw substantial game launches, indicating the industry's resilience and adaptability.

1. Platform Preferences:

DS and PS2 emerged as the leading platforms, each supporting over 2000 games. This suggests developers focus on platforms with high user bases and robust technological capabilities.

1. Global Market Dynamics:

North America stands as the powerhouse in the global gaming market, commanding a significant 49.3% share. European and Japanese markets follow, with 27.3% and 14.5% respectively. The rest of the world accounts for 8.9% collectively. This insight underscores the necessity for tailored marketing and localization strategies, given the varying market sizes and player preferences.

In essence, this analysis not only offers a snapshot of the gaming industry's past but also provides valuable guidance for future endeavors. Understanding the interplay between publishers, platforms, market regions, and historical trends is essential for making informed decisions. As the gaming landscape continues to evolve, these insights serve as a valuable compass, guiding businesses toward strategies that resonate with players and capitalize on emerging opportunities. The future of gaming lies not just in the games themselves but in the nuanced understanding of the dynamic factors shaping this vibrant industry.

***Acknowledgments***

I would like to express my sincere gratitude to the following individuals, organizations, and tools that significantly contributed to the success of this project:

**Shivangini Gupta Madam:** Your guidance and insights have been invaluable throughout the project. Your expertise and support have played a crucial role in shaping the direction of the analysis.

**Subhodeep Sir:** Thank you for your mentorship and continuous encouragement. Your teaching has greatly enhanced the quality of this project.

**Aman Sir:** I am thankful for your assistance and advice, which have been instrumental in overcoming challenges and refining the analytical approach.

**UPGRADE Organization:** I appreciate the resources and opportunities provided by UPGRADE. The learning environment and exposure to real-world applications have been instrumental in the development of this project.

**Kaggle Community:** The collaborative spirit and wealth of knowledge within the Kaggle community have been invaluable. Engaging in discussions and shared insights has enriched the analytical process.

**Medium Community:** The diverse perspectives and in-depth articles from the Medium community have been a valuable source of inspiration and learning.

**Tools:**

* Instabase-Converse: The functionality and ease of use provided by Instabase-Converse significantly facilitated with greater understanding of my report
* Jupyter Notebook: The flexibility and interactivity of Jupyter Notebook were essential in iteratively developing and presenting the code and analysis.
* YouTube: Various tutorials and educational content on YouTube have been instrumental in expanding my knowledge and resolving technical challenges.

I extend my heartfelt thanks to everyone mentioned above for their support and contributions to this project.