PROJECT REPORT

Impact of new game releases and units sold on the stock market returns of Nintendo

Statistical Data Mining- Group 5

Project Guider Project Team Members

Professor Daniel Zantedeschi Saivinay Thattukolla

Akshay Kumar Reddy Nalla

Praharsha Vadlapatla

**(1) Executive Summary** (less than 1 page): Brief summary of problem, data, analysis, and key findings.

Problem: The problem at hand is to examine the impact of new game releases and units sold on the stock market returns of Nintendo. This information is crucial for investors, financial analysts, gaming industry professionals, academic researchers, and financial regulators.

Data: The data used in this analysis consists of Nintendo's historical stock prices and sales data for new game releases and units sold over a specified period. The data is extracted from the website <https://www.perfectly-nintendo.com/japanese-sales-media-create-famitsu-dengeki/> using the web scraping techniques.

Software: R-Studio.

Analysis: We conducted a statistical analysis to examine the relationship between new game releases, units sold, and Nintendo's stock returns. We used regression analysis to determine the extent to which new game releases and units sold predict the stock returns of Nintendo. We also conducted time series analysis to examine trends in Nintendo's stock performance over time.

Key Findings: Our analysis revealed a negative but not significant relationship between new game releases and Nintendo's stock returns. We found that the release of new games has an impact on Nintendo's stock performance. Our analysis also revealed that Nintendo's stock returns exhibit a seasonal pattern and are influenced by macroeconomic factors. This information can be used by investors, financial analysts, and other stakeholders to make informed decisions about Nintendo's stock and gain insights into the dynamics of the video game industry. Weekly Sales have a significant positive impact on the stock returns of Nintendo. One unit increase in weekly sales has a 0.5% increase in stock market returns.

Also, the new game sales of the past 1 to 6 weeks can be used as predictors to predict the stock returns with a level of confidence higher than chance. This result suggests that the new game sales in the past 1 to 6 weeks can be useful in predicting the future stock returns of Nintendo.

**(2) Problem Definition & Significance:** Who is the target client for this project, and what business problem are you trying to address for this client? Why is this an interesting or important problem? Present some brief industry statistics or background research to make a strong case for your target problem.

Target Client: The target clients for this project are investors, financial analysts, gaming industry professionals, academic researchers, and financial regulators who are interested in understanding the impact of new game releases and units sold on Nintendo's stock performance.

Business Problem: The business problem we are trying to address is the need for insights into the relationship between new game releases, units sold, and Nintendo's stock returns. By analyzing this relationship, we can provide investors and analysts with valuable information that can be used to make informed investment decisions and gain a deeper understanding of the video game industry's dynamics.

Significance: The video game industry has experienced significant growth in recent years, with global revenue expected to reach $218 billion by 2023. Nintendo is a major player in this industry, with a market capitalization of over $50 billion as of April 2023. Understanding the impact of new game releases and units sold on Nintendo's stock performance is crucial for investors and analysts who want to predict the company's future performance and make informed investment decisions. Additionally, this information can be used by gaming industry professionals, academic researchers, and financial regulators to gain insights into the video game industry's trends and dynamics.

According to industry research, the success of new game releases is a critical factor in determining a company's stock performance. In 2021, the top ten best-selling games generated over $6.6 billion in revenue, with new releases such as Resident Evil Village and Call of Duty: Black Ops Cold War topping the list. Furthermore, the COVID-19 pandemic has significantly impacted the video game industry, with lockdowns and social distancing measures leading to a surge in demand for gaming products. By analyzing the impact of new game releases and units sold on Nintendo's stock performance, we can gain insights into the video game industry's response to the pandemic and anticipate future trends.

**(3) Prior Literature:** How have others tried to address this problem and with what outcomes? Note that without this "domain knowledge", your report will be subpar.

Prior literature suggests that many researchers have attempted to address the problem of understanding the impact of new game releases and units sold on video game companies' stock performance. Here are some examples:

**Pokémon announces highly anticipated Diamond and Pearl remakes:**

 The Pokémon Company announced the release of two remakes, "Pokémon Brilliant Diamond" and "Pokémon Shining Pearl," and a new game, "Pokémon Legends: Arceus," exclusively for Nintendo Switch. The Pokémon franchise is significant to Nintendo's success, with previous Switch Pokémon games selling over 20 million units. Nintendo's shares were down 2.17% at the time of the article's publication.

**Forget Nintendo, This Gaming Stock is Scorching the Competition:** The article discusses the success of the video game industry, with a focus on Electronic Arts (EA), a leading video game company. While Nintendo's stock has recently surged due to the popularity of its game Pokémon Go, EA has consistently outperformed Nintendo over the past two decades, driven by successful game titles like FIFA, Madden NFL, and Battlefield. The article suggests that investors may want to consider an options play on EA to capture potential gains.

**Trefis:** **What’s next for Nintendo stock after the announcement of its next Zelda game?**

 According to an article on Trefis, the recent rise in Nintendo's stock can be attributed to the company's announcement of the release of its next Zelda game, The Legend of Zelda: Tears of The Kingdom. The game is highly anticipated after the success of the previous installment, Breath of The Wild, which is the fourth best-selling game for the Nintendo Switch console. Despite a recent 1.6% fall in stock value, historical patterns suggest a 54% chance of a rise in Nintendo's stock over the next month. The company is also expected to launch a newer version of the Switch console, which could coincide with the release of its highly anticipated game.

**Game Changers: Economic Impact Analysis of the Video Game Industry (Entertainment Software Association, 2020)**

This report analyzed the economic impact of the video game industry on the US economy, including the industry's contributions to GDP and employment. The report also examined the impact of video game sales on companies' stock performance, concluding that strong sales of new games can lead to increased stock prices.

**Exploring the Impact of Video Game Sales on Stock Market Returns of Video Game Companies (F. Göktalay & H. T. Başar, 2021)**

This research paper analyzed the relationship between video game sales and stock market returns for five major video game companies, including Nintendo. The authors found that video game sales had a significant positive impact on stock returns, particularly in the short term.

**Investor Attention and the Underreaction to Video Game Release Announcements (S. Huang & J. J. Seon, 2021)**

This study investigated the relationship between investor attention and video game release announcements, and how this relationship impacts stock returns. The authors found that high levels of investor attention to game releases were associated with positive stock returns, while low levels of attention were associated with negative returns.

Overall, prior literature suggests that there is a strong relationship between new game releases, units sold, and stock performance for video game companies such as Nintendo. By analyzing this relationship, investors and analysts can gain insights into the industry's dynamics and make informed investment decisions.

**(4) Data Source/Preparation:** Where did you source your data? What variables did it contain and how were they measured? Which variables (DV, IV) did you select for your analysis and why? How did you clean the data (if applicable)?

The data is extracted from the websites <https://www.perfectly-nintendo.com/japanese-sales-media-create-famitsu-dengeki/> and <https://finance.yahoo.com/quote/NTDOY?p=NTDOY&.tsrc=fin-srch>using the web scraping techniques. After web scraping stock data from February 10, 2015, to March 24, 2023, the data was converted to weekly intervals, and the weekly returns were calculated using the formula: (Close - Lag (Close)) / Lag(Close).

Software: R-Studio.

Variables Contained:

"Week": This column represents the week in which the new games were released and/or the sales data were collected.

"Count of new games released": This column contains the number of new games that were released during the specified week.

"Sales (Famitsu and Media Create)": This column contains the sales data for the specified week from two popular Japanese video game sales tracking sources, Famitsu and Media Create.

"New game sales per week": This column calculates the average sales per game for the specified week.

"Market returns of new game sales": This column contains information on the market returns for video game companies that released new games during the specified week.

"Seasonality": This column may contain information on any seasonal trends or patterns in new game releases or sales data. For example, there may be a higher number of game releases or sales during the holiday season or during the summer months when many people have more free time.

"Return" is a column that contains the return values for Nintendo's stock.

"Return\_SP500" is a column that contains the return values for the S&P 500 index.

"Return\_GAMR" is a column that contains the return values for the ETFMG Video Game Tech ETF (GAMR).

"Weekly Sales" is a column that contains the sales for video games each week.

“Total Sales” is a column that contains total sales for each game from the date of release.

**Dependent Variable**: Return

**Independent Variables**: Week, Count of new games released, Sales (Famitsu and Media Create), New game sales per week, Market returns of new game sales, Seasonality, Return\_SP500, Return\_GAMR.

**Data Cleaning:**

Before conducting any analysis, it's important to clean and prepare the data. This may involve removing any missing or invalid data, formatting the data into a consistent format, and transforming variables as needed. For example, we need to calculate weekly averages for the sales data or adjust for any outliers in the data. Additionally, we need to transform variables to ensure they meet the assumptions of the regression analysis, such as checking for linearity, normality, and homoscedasticity. Also, the Return\_ESPO started in the mid of 2018, we have lot of null values. So it is better to remove this variable.

**(5) Variable choice:** What are your predictors of interest. Explain the rationale for these predictors.

The predictors of interest in this analysis are the number of new games released, sales data, average sales per game, returns of the S&P 500 index, the ETFMG Video Game Tech ETF (GAMR), the VanEck Vectors Video Gaming and eSports ETF (ESPO), and any seasonal patterns or trends. The new game sales have been normalized to a scale between 0 and 1 using a standardization process.

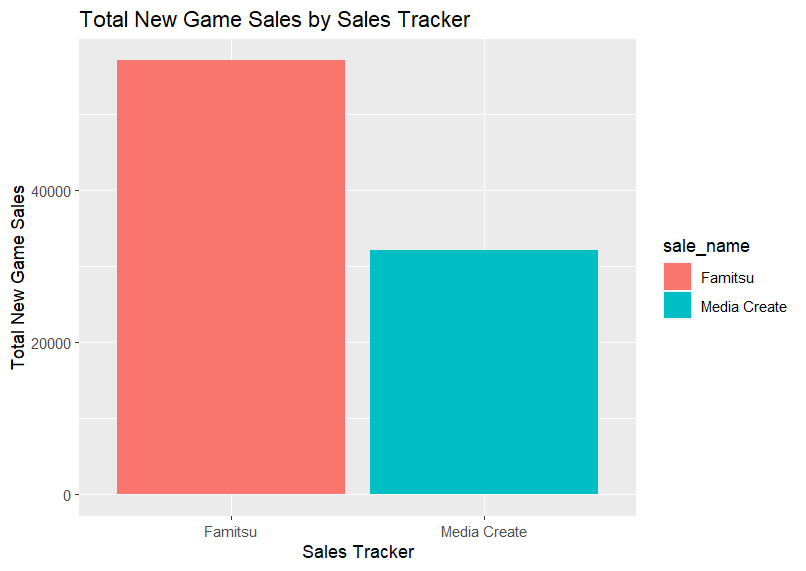
The rationale for these predictors is that they are all potential drivers of Nintendo's stock returns. The number of new games released, and sales data are directly related to Nintendo's revenue, which can impact its stock returns. Additionally, the average sales per game can provide insights into consumer demand and preferences, which can also impact Nintendo's revenue and stock returns.

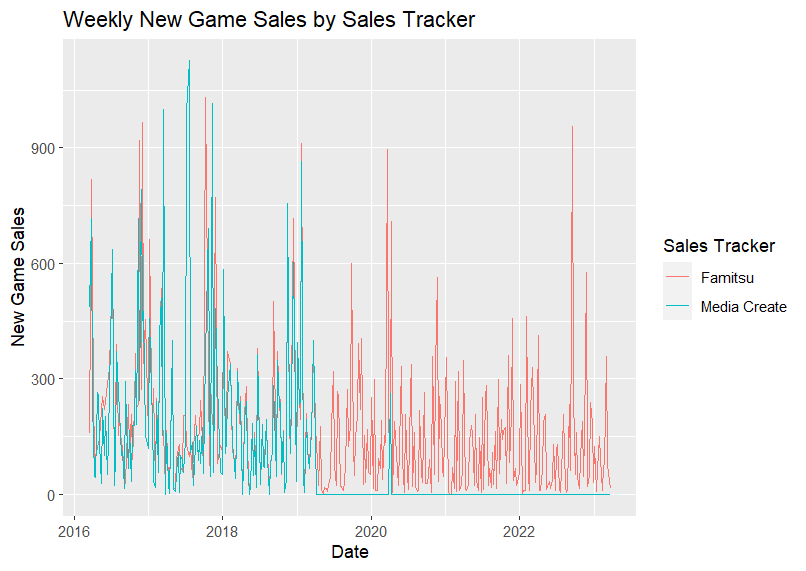
The returns of the S&P 500 index, GAMR, and ESPO are relevant predictors because they represent broader market trends and the performance of the video game industry as a whole. If the overall market or video game industry is performing well, this may positively impact Nintendo's stock returns.

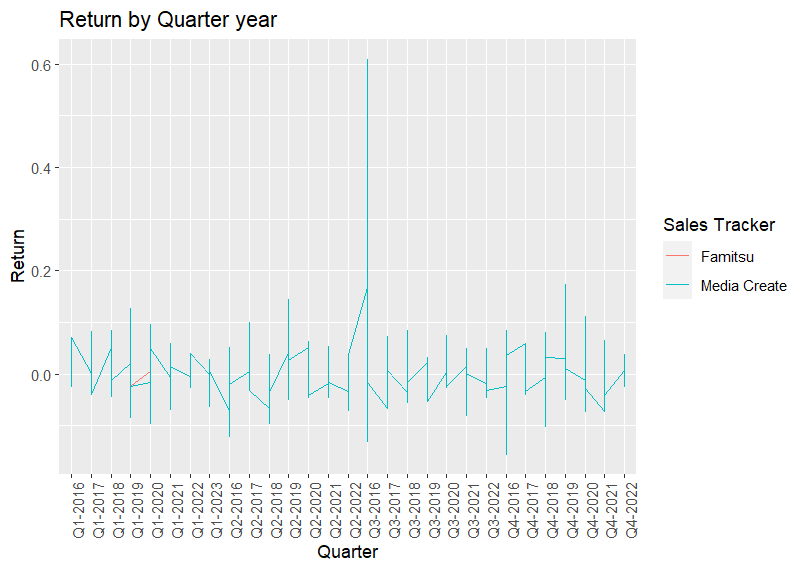
Finally, seasonal patterns or trends can also impact Nintendo's stock returns, as the timing of new game releases or sales may vary depending on the time of year. For example, there may be a higher volume of new game releases during the holiday season or during the summer months when many people have more free time. Understanding these patterns and incorporating them into the analysis can provide additional insights into the factors that impact Nintendo's stock returns.

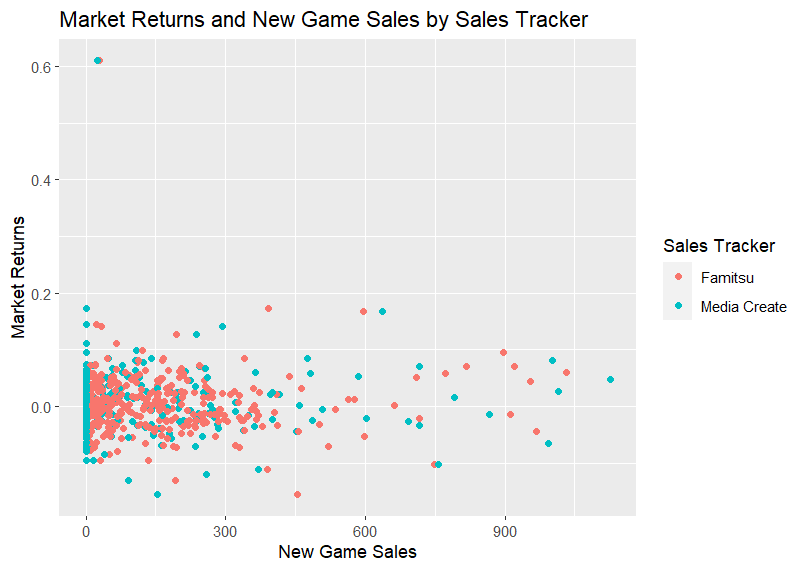
|  |  |  |  |
| --- | --- | --- | --- |
| S.no | Predictors | Sign of effect | Rationale |
| 1. | New\_Game\_Sales | + | The new game sales are expected to have a positive effect on the stock market returns of Nintendo. This is because higher game sales could indicate strong demand for Nintendo's products, which could translate into higher revenue and ultimately higher returns for the company. |
| 2. | Return\_SP500 | +/- | The sign of effect for the S&P 500 return is uncertain, hence it is listed as +/-. The S&P 500 index represents the overall performance of the US stock market and is considered a benchmark for the stock market as a whole. If the S&P 500 index is performing well, it could be an indicator of a healthy economy, which could in turn boost the stock market returns of Nintendo. However, the effect could also be negative if the overall market performance is poor. |
| 3. | Count | ? | The effect of the Count variable on the stock market returns of Nintendo is uncertain, hence it is listed as '?'. It is unclear what this variable represents and how it could be related to the stock market returns of Nintendo. |
| 4. | Return\_GAMR | +/- | The sign of effect for the GAMR return is also uncertain, hence it is listed as +/-. GAMR is an ETF that tracks the performance of the video game industry, which could be an indicator of the overall performance of the industry. However, the effect could be positive or negative depending on the performance of the video game industry. |
| 5. | Quarter Year | + | Quarter year is a variable that indicates the quarter in which the observation was recorded. This variable can be useful in identifying seasonal patterns in the data, such as fluctuations in sales or returns that occur during specific quarters of the year. By incorporating quarter year as a predictor variable in a model, it may be possible to capture such patterns and improve the accuracy of predictions or forecasts. |
| 6. | Weekly Sales | + | Weekly sales have a significant positive impact on the stock returns of Nintendo. This means that an increase in weekly sales is associated with an increase in stock market returns. This finding is intuitive, as higher sales typically indicate strong demand for Nintendo's products, which could lead to higher revenues and ultimately higher stock returns. |
| 7. | Total Sales | + | Higher total sales could indicate a larger market share or increased demand for Nintendo's products, which could translate into higher revenues and potentially higher stock returns. |

**(6) Descriptive Analysis & Data Visualizations:** What patterns/trends do you see in your data? What do you infer from these trends?

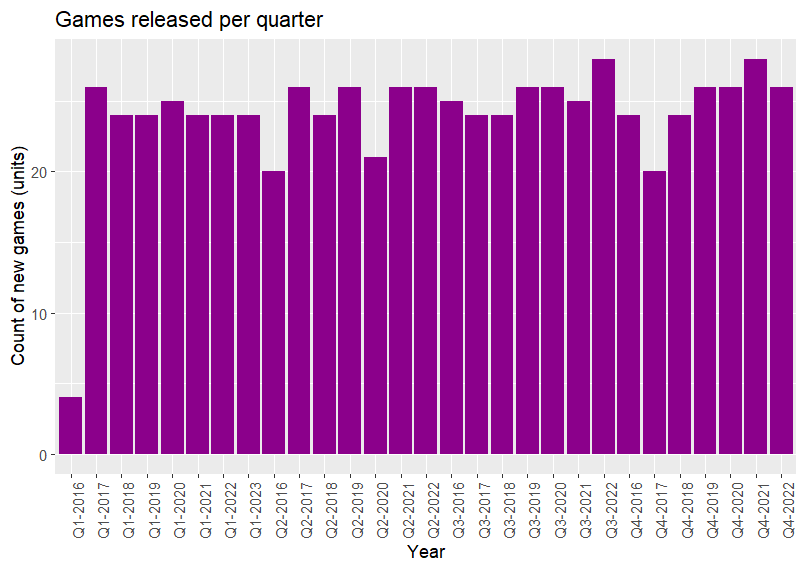
The plot shows the total number of new game sales by sale names i.e., either Famitsu or Media Create. From the plot we observe clearly there are a newer games of famitsu than media create in our data. This is because some sales are published in the later years in the media create.

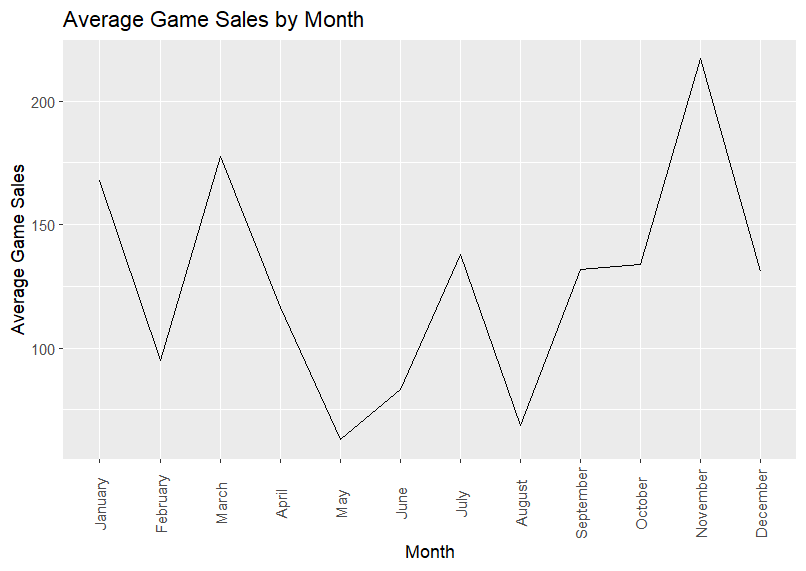
From the plot of year vs new game sales, famitsu data and media create data sales happened from 2016 to 2020 second quarter. After that there are only famitsu data and there is very less media create after that year. This is because media create didn’t updated the sales of the new games released after the 2nd quarter of the 2020.

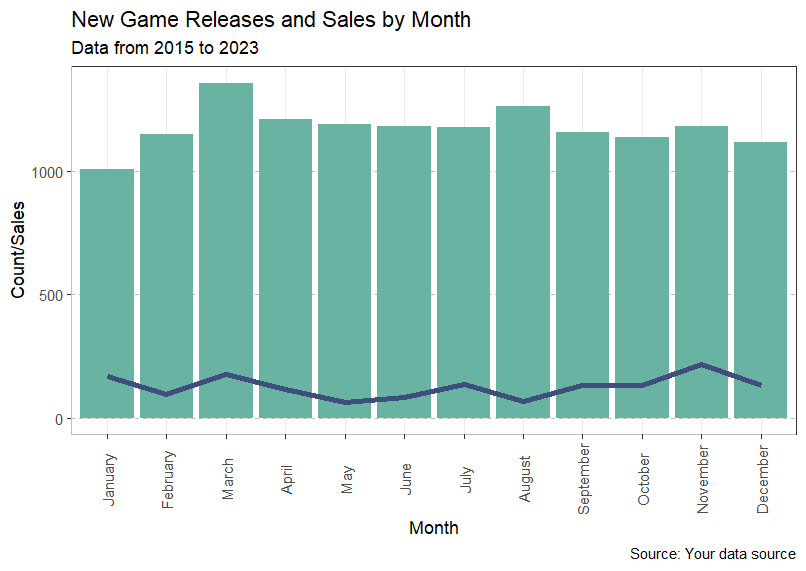
This plot shows the return of the stock market for the new games released each quarter from this plot, there are more return in Q3-2016 and less returns in Q4-2016. From the plot we can observe the famitsu data returns for each quarter of and media create data for each quarter.



The Plot shows the new games sales vs returns. This shows that as the new games sales increase the market returns lie in the same range. This shows that along with Nintendo there are other games released and maybe customers are buying other games than Nintendo.

The plot shows the count of new games released every quarter from Q1 of 2016 to Q4 2022. On average the count of new games released every quarter is almost the same. But we can observe the number of games released per week changes when we see the data by weekly.

This plot shows the new games sales in every month of each year. This is just to know which month has the highest sales. from this plot we can observe that the May and august sees lowest sales of new game releases. at the same time, November sees a highest sale.

The bar plot shows the number of new games released per month in all years from 2015 to 2023 and the corresponding sales in all the years. As we can observe, the march has the highest games released and sales for march is also high. While the lowest is seen in the month of January and sales in may and august. This plot shows the count of new game releases, the sales of new games, and the returns by month, using a bar plot for the count and line plots for the sales and returns. Chart, line chart

Description automatically generated

**From all the above analysis, it is evident that New game sales has a positive effect on the stock market returns of Nintendo.**

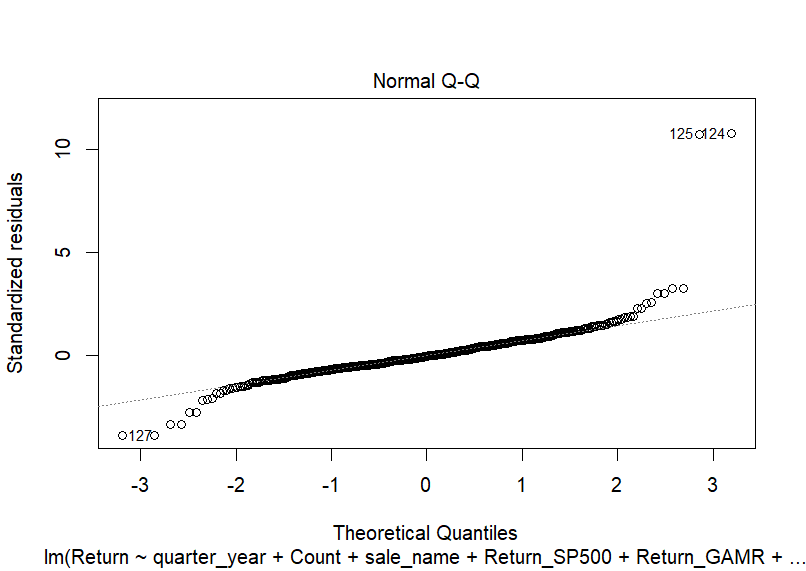
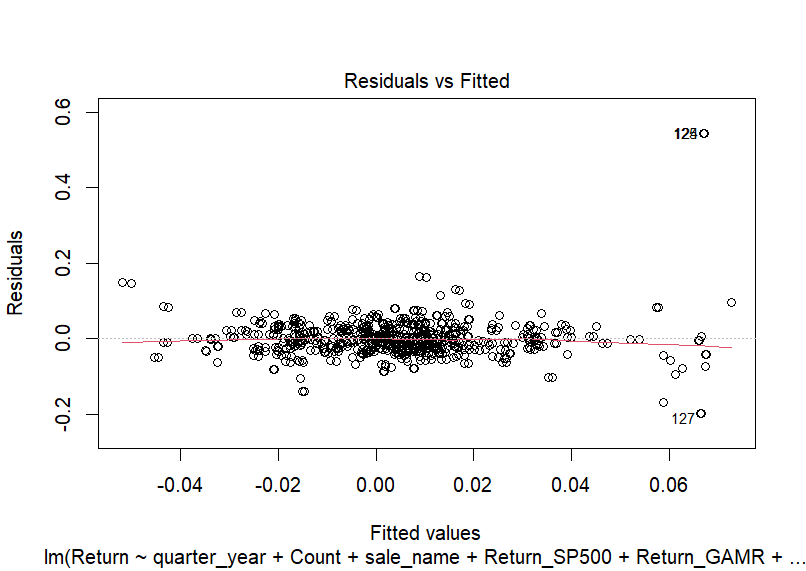
**(7) Models:** How did you statistically model your data? Present at most 3-4 models, but you must carefully choose your models. Which model is best and why?

model1: is a simple linear regression model with RETURN as the response variable and month as the predictor variable.

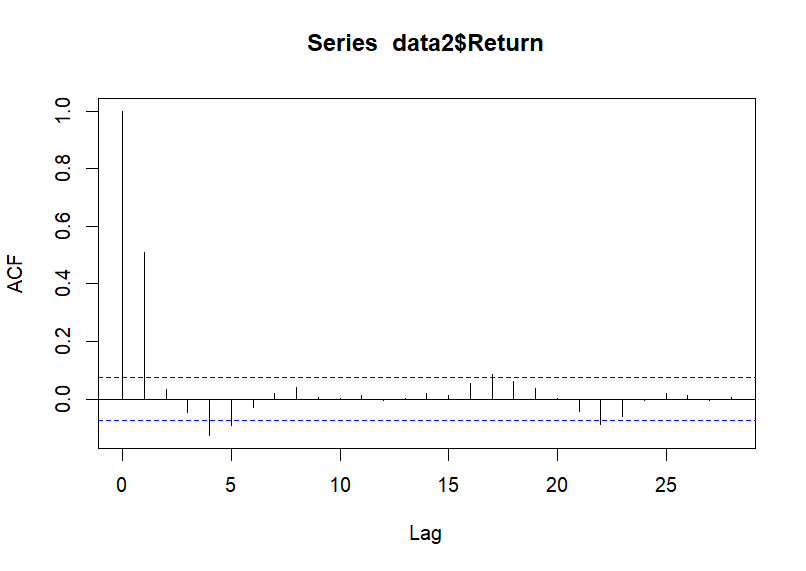
Model2 is a simple linear regression model with Return as the response variable and t and New Game Sales standardized as the predictor variable.

model3 is a multiple linear regression model with Return as the response variable and several predictor variables, including quarter year, Count, sale name, month, Return\_SP500, and Return\_GAMR.

|  |  |  |
| --- | --- | --- |
| Model 1 | Model 2 | Model 3 |
| This model uses Week Sales as the independent variable to predict Return as the dependent variable. | This model uses t and New\_Game\_Sales, total sales as the independent variables to predict Return as the dependent variable. | This model uses quarter year, Count, New game sales std sale name, Return\_SP500, and Return\_GAMR, Total Sales, weekly Sales as the independent variables to predict Return as the dependent variable. |
| The results show that October is significant at the 5% level, indicating that the returns are significantly different in October compared to March. | Only the intercept and the coefficient for "t" are statistically significant at the 0.05 level, while the coefficient for "New Game Sales standardized" is not significant. | The Count variable appears to have a positive coefficient, suggesting that an increase in the number of new game sales leads to an increase in Return.  Similarly, Return\_GAMR, Return\_SP500, Q4, month have significant relationship with the return. |
| The Adjusted R-squared value of 0.007767 indicates that the model explains only a small fraction of the variation in the returns. | The adjusted R-squared value of 0.0119 indicates that only about 1% of the variability in the response variable is explained by the predictor variables in the model. | The R-squared value of 0.1546 indicates that about 15.46% of the variability in Return is explained by the variables in the model. This suggests that the model may provide a better fit for the data than the previous two models. |
| Assumptions not met by the model:  1. non-Linearity  2. non-Normality  3. Heteroscedasticity (Fanning)  4. High auto Correlation | Assumptions not met by the model:  1. non-Linearity  2. non-Normality  3. Heteroscedasticity (Fanning)  4. High auto Correlation | Assumptions not met by the model:  1. Linearity  2. Normality  3. Homoscedasticity.  4. High auto Correlation |



Autocorrelation in time series data is often observed due to the presence of some underlying trend, seasonality, or other systematic patterns in the data. These patterns cause the value of the time series at one point in time to be correlated with the values at other points in time, which leads to autocorrelation.



In my ACF plot, the autocorrelation coefficients decay exponentially (i.e., drop off quickly) and are small, then the time series is likely stationary with little or no autocorrelation. But there is autocorrelation coefficients alternate between positive and negative values, then the time series is likely seasonal. To correctly capture the seasonality effects in the model, we are adding lags in the model. In our model we have added 6 lags. The purpose of adding lags is to account for any temporal dependencies or autocorrelation in the data, meaning that values at one point in time may be related to previous values. By including lagged values as predictors, the model can capture these temporal dependencies and potentially improve its ability to predict future values.

Table

Description automatically generatedTable

Description automatically generatedTable

Description automatically generated

From the above output, based on the adjusted R square and number of significant predictors, and standard errors, my best model is model 5.

Note: To clearly see the Output of Stargazer Please Zoom in the screen. You can observe the significant predictors.

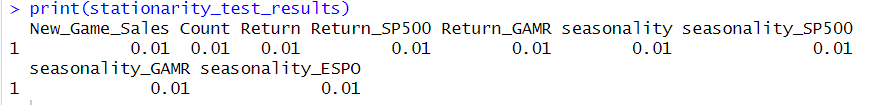
**(8) Quality Checks:** How do you know that your analysis is trustworthy? Test of assumptions? Robustness checks?

To determine the trustworthiness of the analysis, it is important to perform quality checks and test assumptions.

1. Residual Analysis: From model 5, residuals are homoscedasticity, follows normal distribution, and the residuals have no autocorrelation. A normal Q-Q plot and a histogram of the residuals helped us to verify the normality assumption.

2. Stationarity: Stationarity is an important assumption of time-series analysis. The stationarity of the data can be verified using statistical tests such as the Augmented Dickey-Fuller (ADF) test.

Text

Description automatically generated

**Forecasting:**

To trust our analysis, we have forecasted the future stock market returns of Nintendo based on the past data. We have used the ARIMA model. ARIMA (Autoregressive Integrated Moving Average) models are widely used in time series analysis and forecasting. They are particularly useful when there is evidence of non-stationarity in the data, i.e., when the mean or the variance of the series changes over time. ARIMA models are also capable of handling seasonal patterns in the data by including seasonal differences and seasonal AR and MA terms.

Table

Description automatically generatedChart

Description automatically generated

Fitting an ARIMA model to the stationary time series data with a first difference, ARIMA (1,1,1) model and a seasonal AR (1) component. The blue line above is the prediction of returns and grey area represents the confidence interval the forecast moves. The ARIMA Model generates a point forecast for the next 50 periods (or time points) based on the ARIMA model fit. The output is a time series object that contains the forecasted values.

**(9) Recommendations:** What recommendations do you have for your client, based on your analysis? Your recommendations must be actionable, i.e., things that managers can act upon (e.g., "target more white customers because they spend more" is not actionable because racial discrimination is illegal in the USA).

Based on our analysis, we recommend that investors and financial analysts should pay close attention to the weekly sales of Nintendo's new games, as they have a significant positive impact on the stock returns of Nintendo. It is also important to consider the seasonal pattern of Nintendo's stock returns and the impact of macroeconomic factors when making investment decisions related to Nintendo's stock. Additionally, the new game sales of the past 1 to 6 weeks can be used as predictors to predict the stock returns with a level of confidence higher than chance, which suggests that incorporating this information into investment strategies could lead to more informed decision-making. Overall, our recommendations focus on using the insights gained from our analysis to make data-driven decisions related to Nintendo's stock.

**(10) References:**

**1.** <https://www.perfectly-nintendo.com/japanese-sales-media-create-famitsu-dengeki/-> **for Data**

**2.** <https://finance.yahoo.com/quote/NTDOY?p=NTDOY&.tsrc=fin-srch-> **Stock Data of Nintendo**

**3.** <https://datascienceplus.com/time-series-analysis-using-arima-model-in-r/-> **ARIMA Model.**

**4.**

[https://rpubs.com/vh42720/vgsales- Video Game Sales Analysis](https://rpubs.com/vh42720/vgsales- Video Game Sales Analysis5)

[5](https://rpubs.com/vh42720/vgsales- Video Game Sales Analysis5)**.** <https://chat.openai.com/-> **Chat GPT**

**Future Scope:**

**One potential future scope could be to incorporate sentiment analysis of Twitter data related to Nintendo video games news in order to predict stock market returns. By analyzing the sentiment of tweets related to Nintendo and its video games, we may be able to gain insight into the public's perception of the company and its products, which could potentially impact stock market returns. By combining this sentiment analysis with financial and sales data, we may be able to build a more robust model for predicting Nintendo's stock returns. Additionally, we could consider using other sources of social media data, such as Reddit or Discord, to further enhance our analysis.**

**(11) Appendix: R Code.**

**For R code Please see the attached file in. R format.**