

**HOW CAN NINTENDO IMPROVE SALES?**

MSBA 320 C2: Tools for Business Analytics

Instructor: Dr. **Siamak Zadeh**

**By**

**Veena Priyadharshini Chintala**

Golden Gate University

536, Mission Street

San Francisco, CA- 94105

**TABLE OF CONTENT**

1. Introduction……………………………………………………………………………… 3

2. Data Collection ……………………………………………..…………………………… 3

3. Descriptive Analysis ………………………………………………………......................4

3. Explanatory Data Analysis ………………………………………………………............ 6

4. Multiple linear regression and Diagnostics ………………………………………...........12

5. Two sample t-test…………………………………………………………………….……17

6. ANOVA …………………………………………………………………..………………19

7. Time series Analysis…………………………………………………………………...… 20

6. Forecast Analysis…………..…………………………………………………………….. 22

8. Conclusions ……………………………………………………………………………… 29

9. Recommendations …………………………………………………………………………30

10. References ……………………………………………………………………………….. 31

**INTRODUCTION**

Nintendo is Japanese based organization which has developed multiple Videogame and Portable consoles which can be used with external devices. Nintendo has sold over 7 million video games as of 2015. Although the company had few game releases before Nintendo they were not a huge hit, but Nintendo Entertainment System (NES) was a huge hit in 1983 and this restarted the industry of video games after the video game crash in the year 1983. Nintendo Entertainment System was a huge success internationally.

Nintendo presented a first handheld device which had a touchscreen on it (DS) and a first console device with motion controlled inputs (Wii) in the year 2001. They were very successful worldwide and became best sellers of all the time. In 2010, Nintendo released a console game with stereoscopic 3D capabilities called 3DS, this saw very good sales in the world market. In 2012, they released Wii U, which was not as successful as 3DS and also their sales were significantly lower than expected.

Nintendo video games are very successful worldwide. Here I will be analyzing the sales of DS series and Wii series Nintendo games in Europe, Japan, and North America. I will also be providing suggestions to the company to improve their sales in North America, Europe and Japan based on the results.

**DATA COLLECTION**

I have acquired this dataset from Kaggle. The dataset consists of 203 rows of Nintendo Video game sales at various regions like North America, Europe, and Japan.

The different fields in the dataset include:

Platform - Platform of the game's release (i.e. DS, WiiU, Wii, 3DS)

Year - In which year was the game's release

Genre - Genre of the game

Publisher - Publisher of the game

NA\_Sales - North American sales (in millions)

EU\_Sales – Europe sales (in millions)

JP\_Sales - Sales in Japan (in millions)

Global\_Sales - Total worldwide sales.

Critic\_score - Aggregate score compiled by Meta critic staff

User\_score - Score by Meta critic's subscribers

Developer - Party responsible for creating the game

Rating - The ESRB ratings.

**DESCRIPTIVE ANALYSIS**

The Descriptive analysis helps us to understand the data and discover patterns, and make assumptions for further analysis. It also helps us to understand the relationship between explanatory and independent variables with the dependent variables.

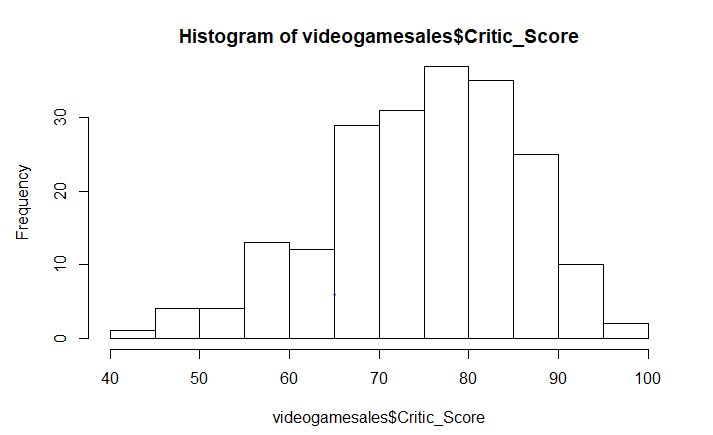
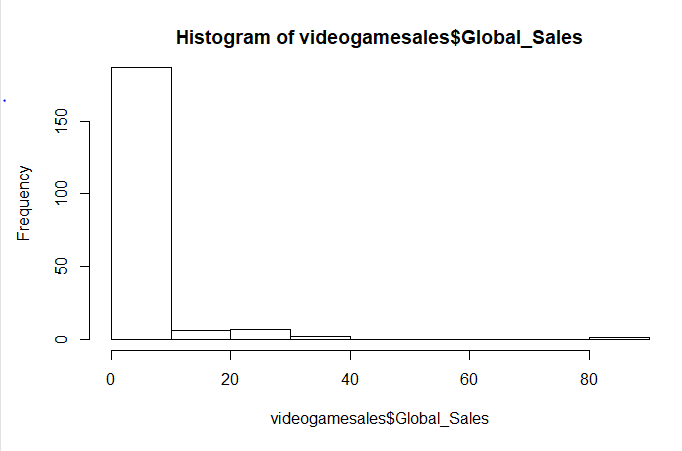
For this analysis, I am including Histogram, Boxplot and Scattered plot to understand the data set.

Summary results for North America Sales, Europe Sales, Global sales and Japan Sales.

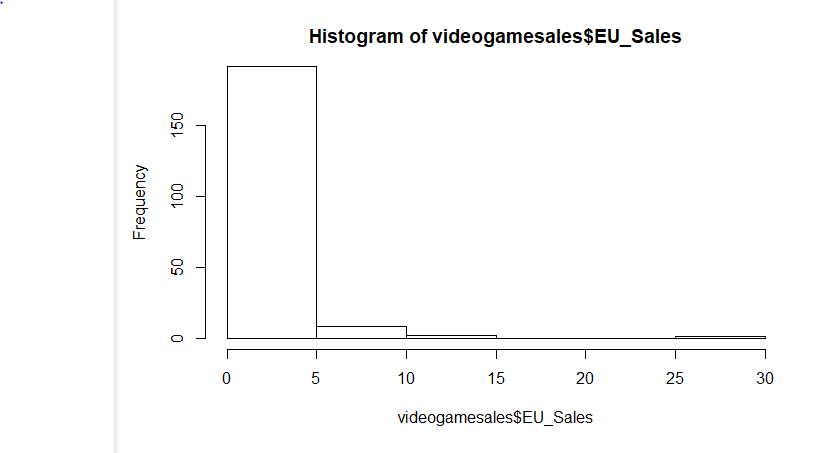
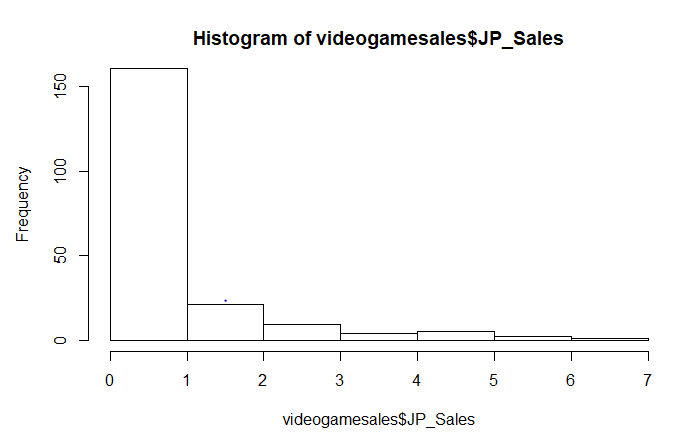
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **VARIABLE** | **NA SALES** | **EU SALES** | **JAPAN SALES** | **GLOBAL SALES** |
| **MIN** | **0** | **0** | **0** | **0** |
| **1ST QUADRANT** | **0.195** | **0.030** | **0.08** | **0.45** |
| **MEDIAN** | **0.47** | **0.2** | **0.29** | **1.06** |
| **MEAN** | **1.49** | **1.08** | **0.719** | **3.579** |
| **3RD QUADRANT** | **1.1** | **0.95** | **0.79** | **3.1** |
| **MAX** | **41.36** | **28.96** | **6.5** | **82.59** |

The Quantitative variables distribution is shown in histograms below.

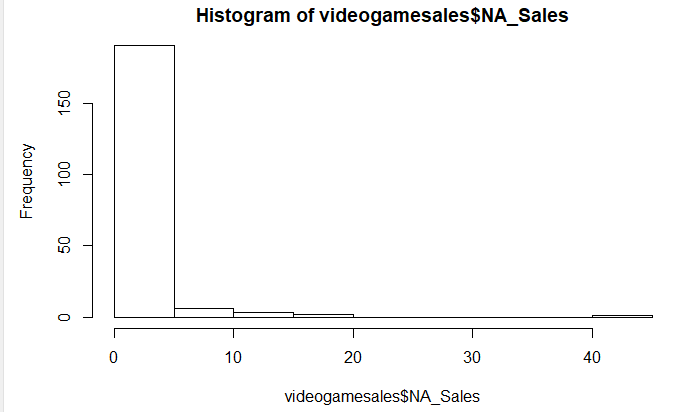
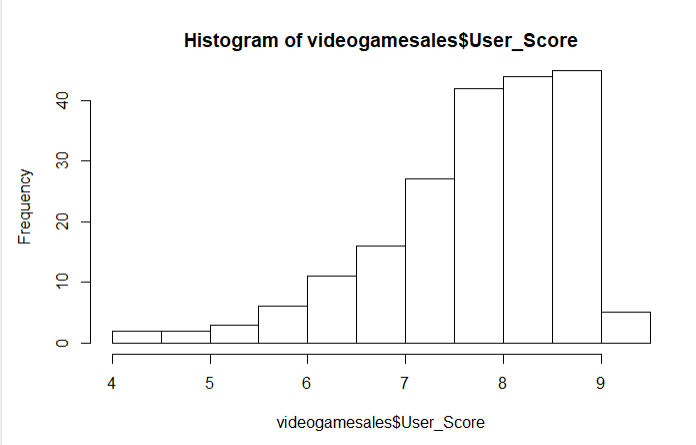
CRITIC SCORE GLOBAL SALES

JAPAN SALES EUROPE SALES



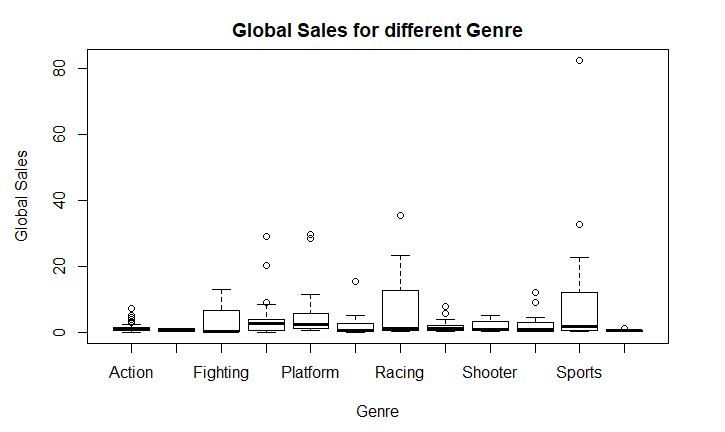
NORTH AMERICA SALES USER SCORE

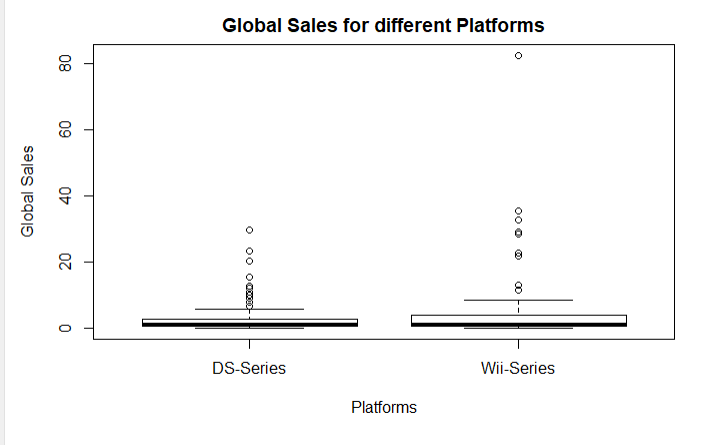
 

**EXPLANATORY DATA ANALYSIS**

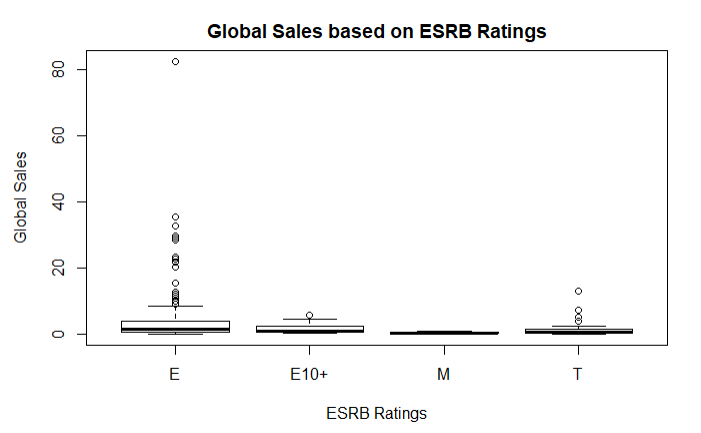
The following box plot represents the correlation between explanatory variables and the dependent variable. Boxplots provide a five number summary: minimum, first quartile, median, third quartile, and maximum.

In my analysis, I tried to study the distribution of NA, Europe, Japan and Global sales across Genre, Ratings, and Platform. From the below graphs, we can see the variations in the sales of video games across different countries based on genre, platform, and ESPR ratings.

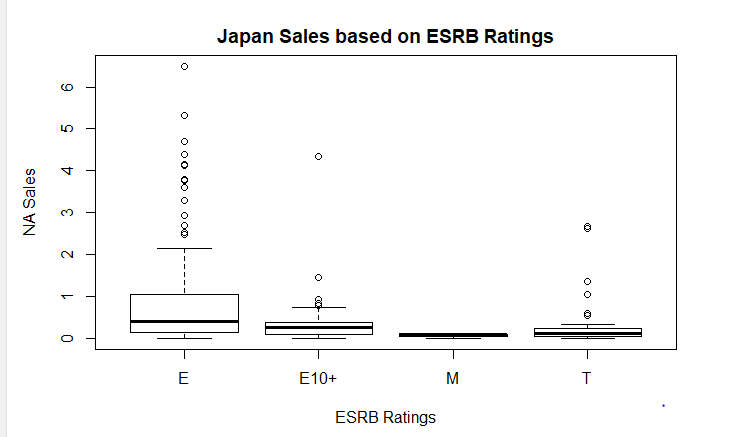
 The above graph shows that racing and sports games have the highest sales around the globe.



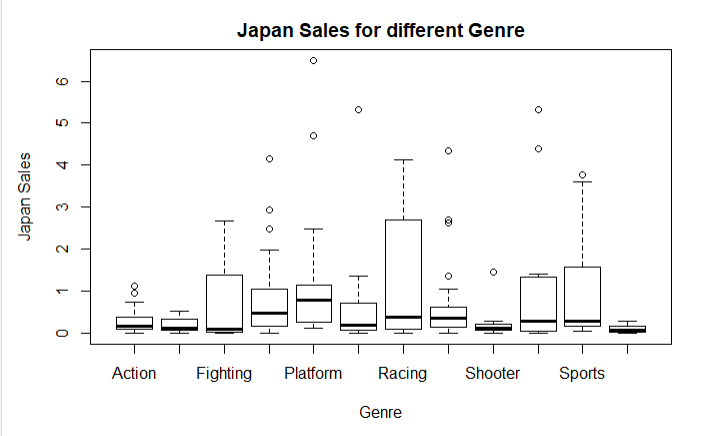
The above graph shows that the Wii series have more sales than DS-series globally.



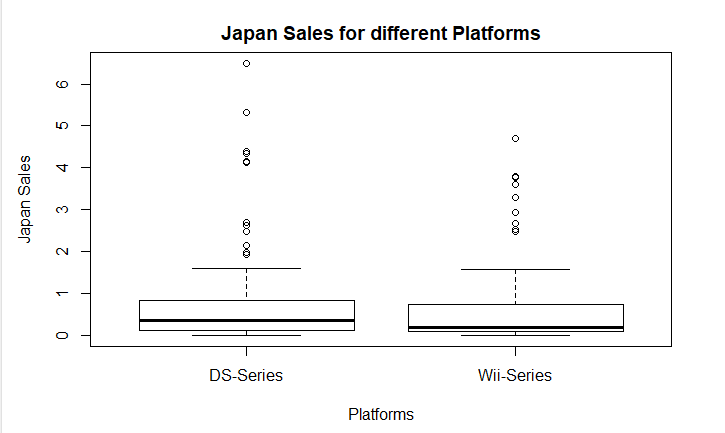
Above graph shows that the games which can be played by all age group have more sales globally.



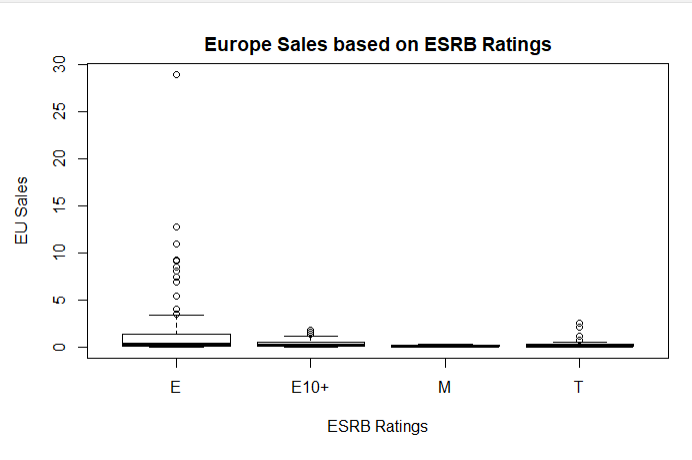
Above graph shows that the games which can be played by all age group have more sales in Japan than other categories.



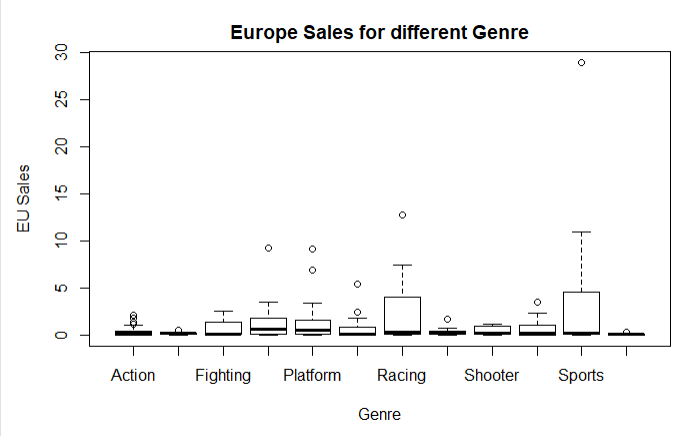
From the graph, we can see that racing games have the highest sales in Japan.



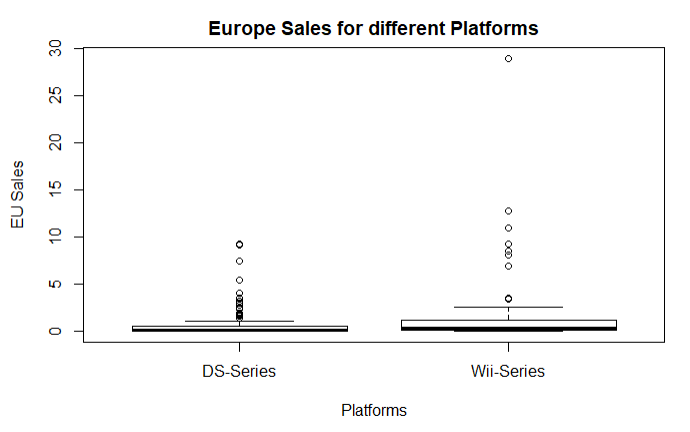
The above graph shows DS-series have highest sales in Japan compared to Wii Series.



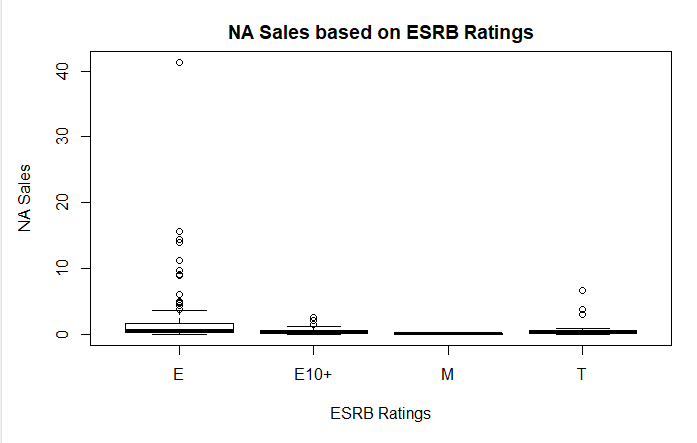
Above graph shows that the games played by all age group have more sales in Europe.



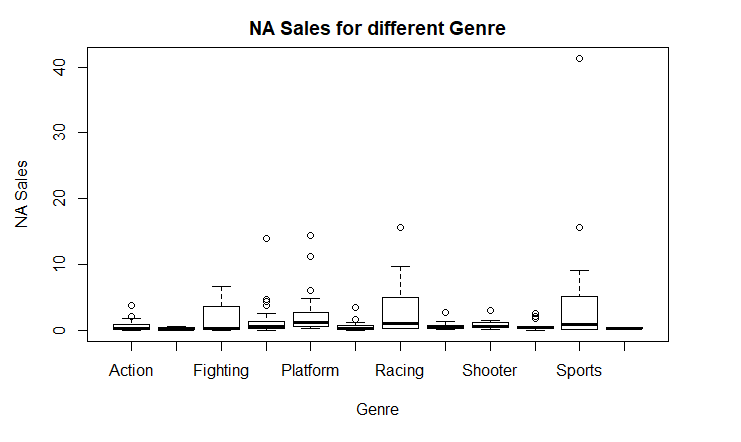
Sports and Racing games have higher sales in Europe compared to other Genre.



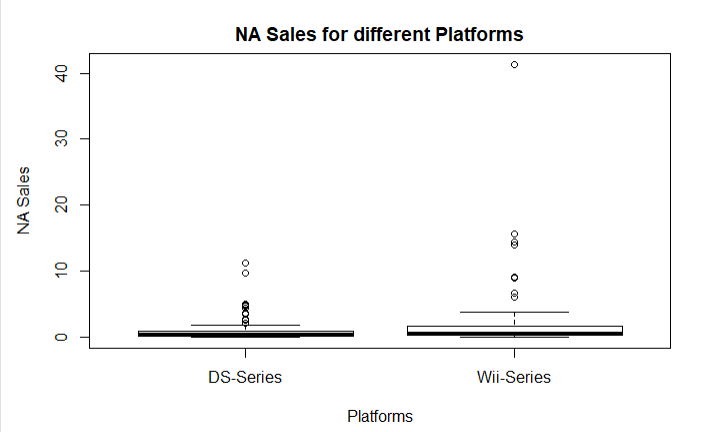
Above graph shows that Wii series have more sales in Europe than the DS series.



Games that can be played by all age group has more sales in North America than others.



Above graph shows that racing and sports-related games have more sales in North America



We can see from the graph that, Wii series has more sales in North America than the DS series.

**MULTIPLE LINEAR REGRESSION AND DIAGNOSTICS**

According to Crawley (2015), multiple linear regression is a statistical method used to predict the value of a response variable based on the values of multiple explanatory variables. The multiple linear regression is used to identify the independent variables (explanatory variable) that have a significant effect on the dependent variable (response variable) and hence provide an equation to calculate the response variable from the independent variable.

In this dataset, North America sales, Europe sales, Japan sales, and Global sales are dependent variables and rating, Platform, Critic score, Genre, User score are independent variables. Multiple linear regression analysis will help us to explain how significant the relation between dependent and independent variables is. The result shows that only Genre, Platform and Critic Score has a significant effect on North America sales, Europe sales, and Global Sales. But in the case of Japan, Genre, Critic Score, and Rating has a significant effect on the sales of Video games.

Based on the multiple linear regression results only 3 variables are significant for the variations in the sales. Hence, I created the first model with dependent variable NA sales, Europe sales, Japan sales and Global sales with the independent variables individually. In the second model, I only included the dependent variables with their corresponding independent variable which were shown significant from the multiple linear regression analysis. I created AIC (Akaike’s Information Criterion) for all the four models to choose the best model. The model with the lowest AIC score is the best fit for further analysis. The result showed that the second model is the best regression model in all the four cases. Thus, I used the second model to conduct my research.

**Two regression models for NA sales**

|  |  |  |
| --- | --- | --- |
|  | Dependent Variable | Independent Variable |
| Model1 | North America sales | Platform, Genre, Critic score, User score, rating |
| Model2 | North America sales | Platform, Genre, Critic score |

**Two regression models for Europe sales**

|  |  |  |
| --- | --- | --- |
|  | Dependent Variable | Independent Variable |
| Model1 | Europe sales | Platform, Genre, Critic score, User score, rating |
| Model2 | Europe sales | Platform, Genre, Critic score |

**Two regression models for Japan sales**

|  |  |  |
| --- | --- | --- |
|  | Dependent Variable | Independent Variable |
| Model1 | Japan sales | Platform, Genre, Critic score, User score, rating |
| Model2 | Japan sales | Rating, Genre, Critic score |

**Two regression models for Global sales**

|  |  |  |
| --- | --- | --- |
|  | Dependent Variable | Independent Variable |
| Model1 | Global sales | Platform, Genre, Critic score, User score, rating |
| Model2 | Global sales | Platform, Genre, Critic score |

Lastly, I conducted regression diagnostic to check if the model works well with the study. Regression diagnostics are used to evaluate the model suppositions and inspect if there are perceptions with an immense, undue effect on the examination (Kabacoff, 2011). This is important because irregularities in the data or misspecification of the associations between the indicators and the response variable can lead to an incorrect model. Linear regression model has 4 assumptions and 4 diagnostic plots. I created the diagnostic plots for all the four sales groups

The four assumptions are:

1. Linearity: The relationship between the dependent variable and the mean of the independent variable is linear.

2. Homoscedasticity: The variance of the residual is the same for any value of the dependent variable.

3. Independence: Observations are independent of each other.

4. Normality: For any fixed value of the dependent variable, the independent variable is normally distributed.

There are 4 diagnostic plots showing residuals:

1. Normal Q-Q: This plot shows if residuals are normally distributed.

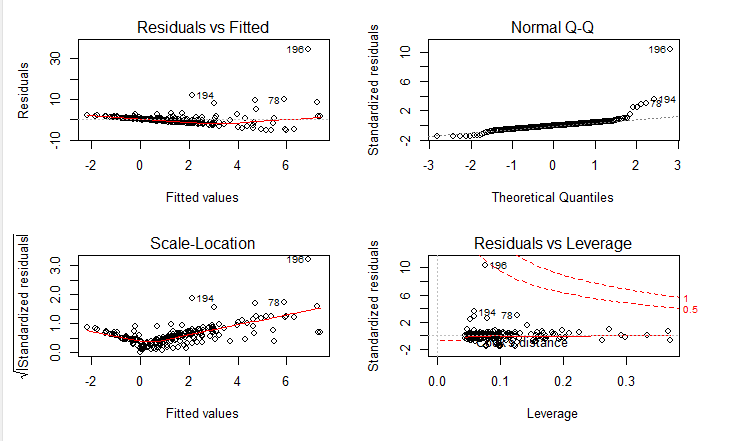
2. Residuals vs Fitted: This plot shows if residuals have non-linear patterns.

3. Scale-Location: This is to check the assumption of equal variance.

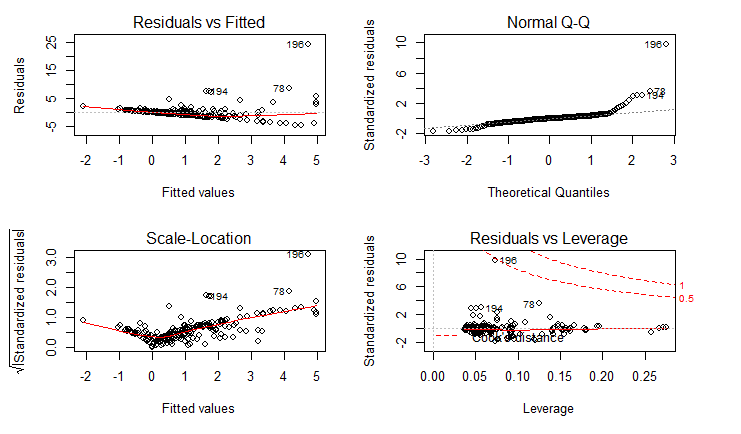
4. Residuals vs Leverage: This plot helps us to find influential cases.

The result for NA sales, Europe sales, Japan sales, and Global sales shows that the points almost lay on the line, so we can say that the independent variables are normally distributed.

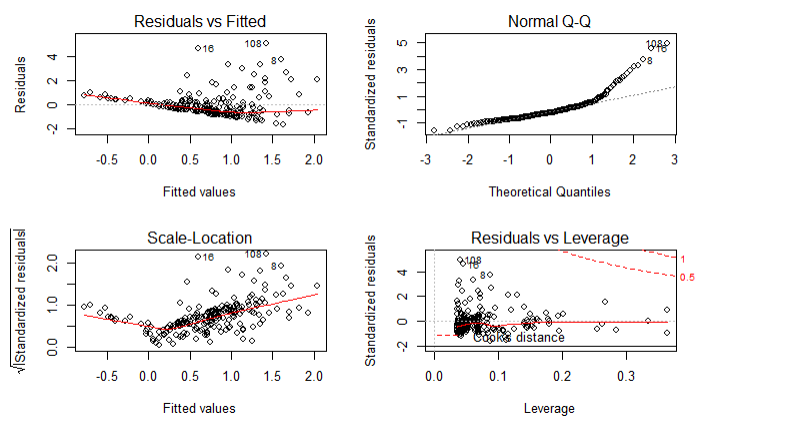
**North America Sales Regression Diagnostic**



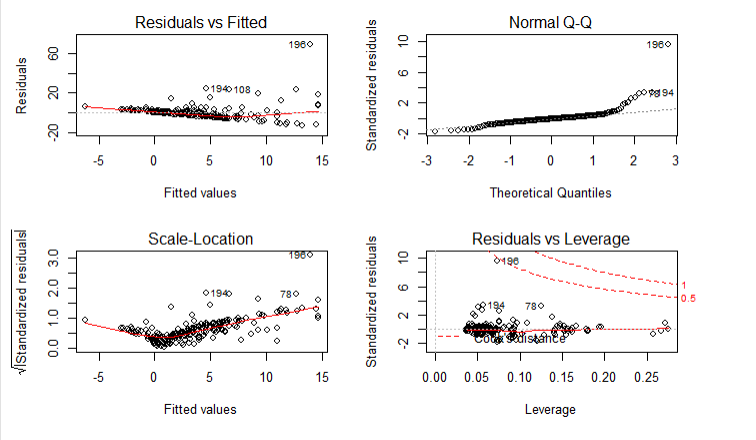
**Europe Sales Regression Diagnostic**

****

**Japan Sales Regression Diagnostic**



**Global Sales Regression Diagnostic**

****

**TWO SAMPLE T-TEST**

Independent 2 sample t-test and confidence interval are statistical methods for comparing the difference in means for 2 populations. This means that using two-sample t-test we can examine the relationship between a numeric outcome variable (Y) and a categorical explanatory variable (x, with 2 levels). The two sample t-test is conducted with a 95% confidence interval on NA sales, Europe sales, Japan sales and Global sales on the explanatory variable platform.

**Null Hypothesis: There is no relationship between sales and platforms.**

The goal of this test is to check if platforms have any effect on the sales of video games across different regions. Here the dependent variable is NA sales, Europe sales, Japan sales, and Global sales and the independent variable is a platform which has 2 levels (DS series and Wii series).

**The result of the two sample t-test for all the 4 groups is:**

**Null Hypothesis: There is no relationship between NA sales and platforms.**

|  |  |  |
| --- | --- | --- |
| **Mean DS series** | **Mean Wii series** | **P – value** |
| 0.9213913 | 2.2353409 | 0.02939 |

**Null Hypothesis: There is no relationship between Europe sales and platforms.**

|  |  |  |
| --- | --- | --- |
| **Mean DS series** | **Mean Wii series** | **P – value** |
| 0.7482609 | 1.5128409 | 0.08015 |

**Null Hypothesis: There is no relationship between Japan sales and platforms.**

|  |  |  |
| --- | --- | --- |
| **Mean DS series** | **Mean Wii series** | **P – value** |
| 0.7802609 | 0.6394318 | 0.3695 |

**Null Hypothesis: There is no relationship between Global sales and platforms.**

|  |  |  |
| --- | --- | --- |
| **Mean DS series** | **Mean Wii series** | **P – value** |
| 2.641391 | 4.803636 | 0.08773 |

From the above chart, we can see that only NA sales have a p-value less than 0.05 hence we reject the null hypothesis and accept that there is a significant difference in the means. And all others including Europe, Japan, and Global sales have p-value greater than 0.05, hence we accept the null hypothesis for these regions that there is no significant difference between the means. This implies that the platform has a significant effect only on North American sales, whereas the platform has no effect on Europe, Japan, and Global Sales.

**ANOVA TEST**

Anova is a statistical method used for comparing the means of two or more independent population. In this project, one-way ANOVA is conducted to check if there is any significant difference in the NA, Europe, Japan and Global sales due to the independent variables including Developer or Genre.

**Null Hypothesis: Developer has no effect on NA, Europe, Japan, and Global Sales.**

The results show that the p-value is greater than 0.05 and has a very low F- value, hence we accept the null hypothesis and have evidence to believe that Developer has no effect on sales of video games.

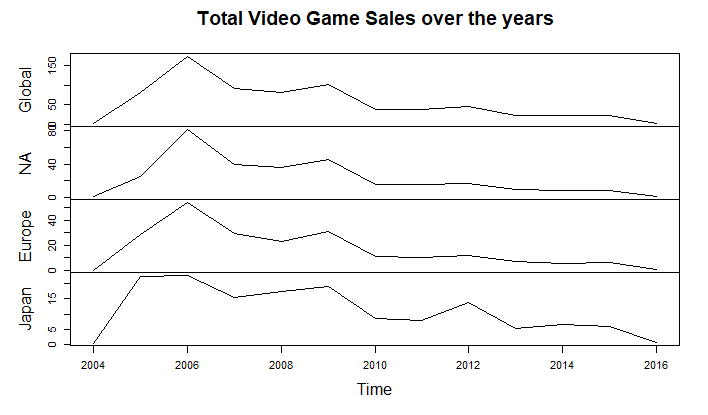
**Null Hypothesis: Genre has no effect on NA, Europe, Japan, and Global Sales.**

The results show that the p-value is less than 0.05 and has high F- value for NA, Europe, and Global sales, hence we reject the null hypothesis for these regions and have evidence to believe that Genre has a significant effect on NA, Europe and Global sales of video games. But, in case of Japan, the p-value is greater than 0.05 and has a very low F- value, hence we accept the null hypothesis and have evidence to believe that Genre has no effect on Japan sales.

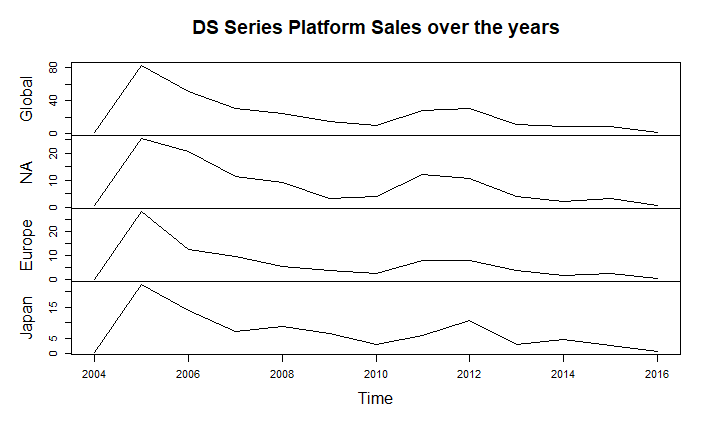
**TIME SERIES AND FORECAST ANALYSIS**

Any data analyses over regular intervals of time are described as time series analysis. Time series analysis is very important for the business to know the trend and also to perform the forecast analysis.

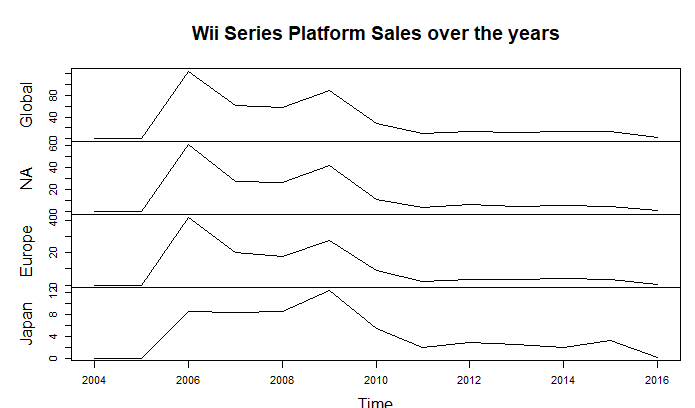
In this dataset, I have analyzed the total sales across all regions from the year 2004 to 2016. Time series analysis is also done on DS series sales and Wii series sales across the regions from the year 2004 to 2016.



From the above graph, we can see that North America and Europe follows the trend of global sales but Japan has a little different trend. But, at the end in the year 2016, it is seen that video game sales are declining.



From the above graph, we can see that NA and Europe are following the trend of global sales but Japan has a little different trend. But, at the end in the year 2016, it is seen that the video game sales are declining for the DS series platform.

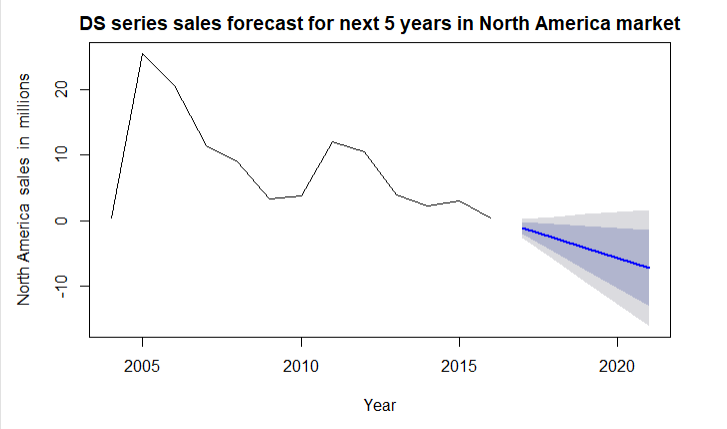


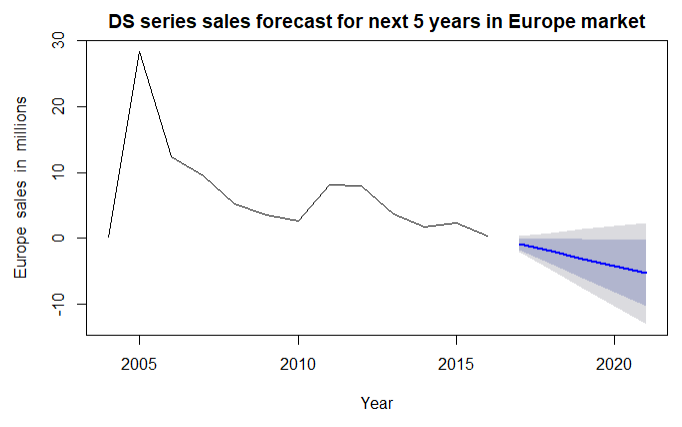
From the above graph, we can see that NA and Europe are following the trend of global sales but Japan has a little different trend. But, at the end in the year 2016, it is seen that the video game sales are declining for the Wii series platform.

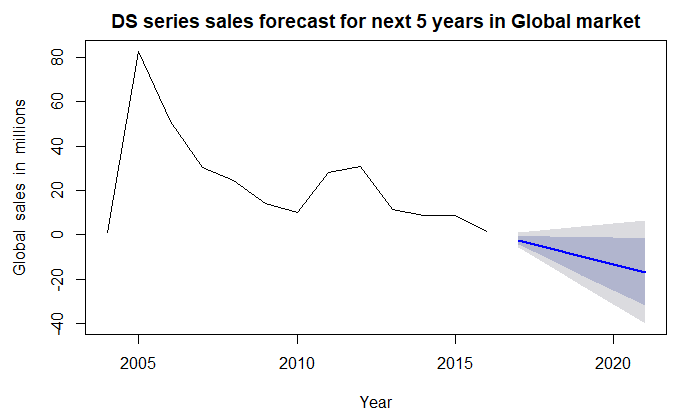
**FORECAST ANALYSIS**

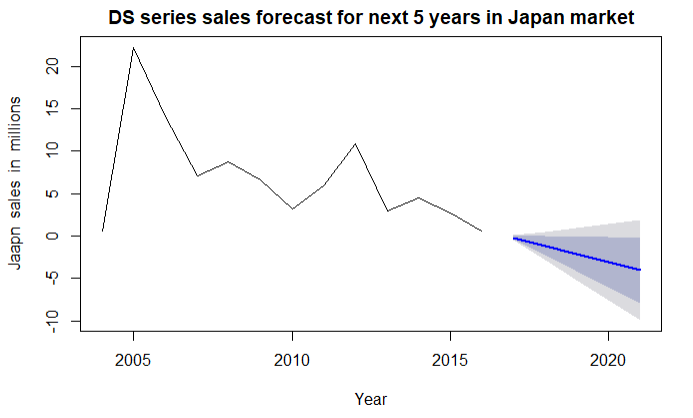
Forecasting is the process of using statistical methods to predict the future values based on the historical trend in the time series. In this project forecast analysis helps us to predict the sales of video games for the next 5 years across the two different platforms (DS and Wii series). The forecast is also done on Total sales of video games for the next 5 years across different regions (NA, Europe, Japan, and Global). It is required to install forecast package to do forecast analysis.

**DS Platform series Forecast Analysis**



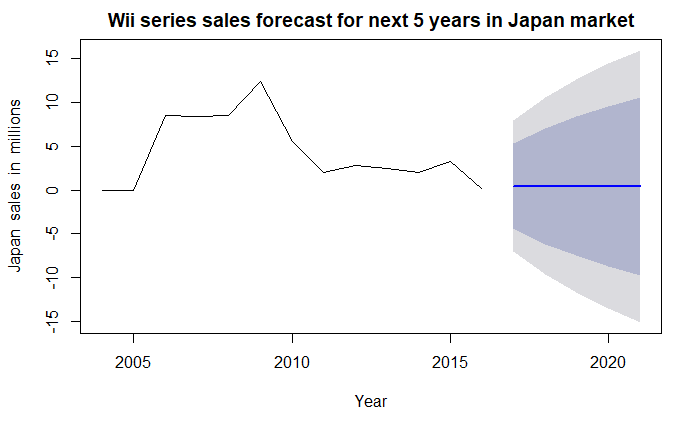
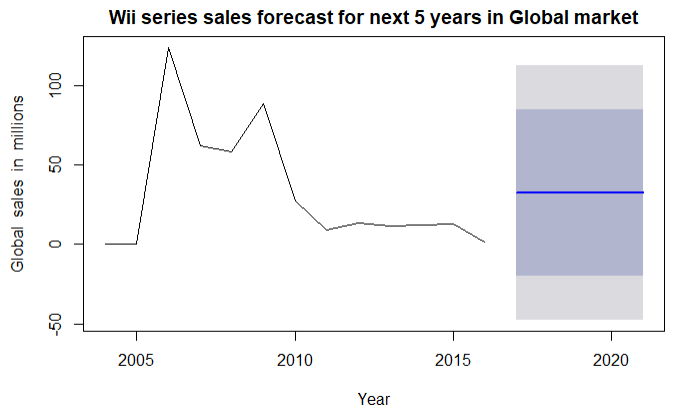
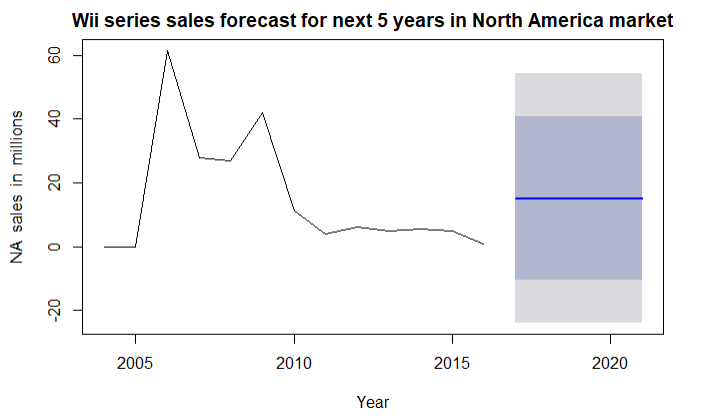
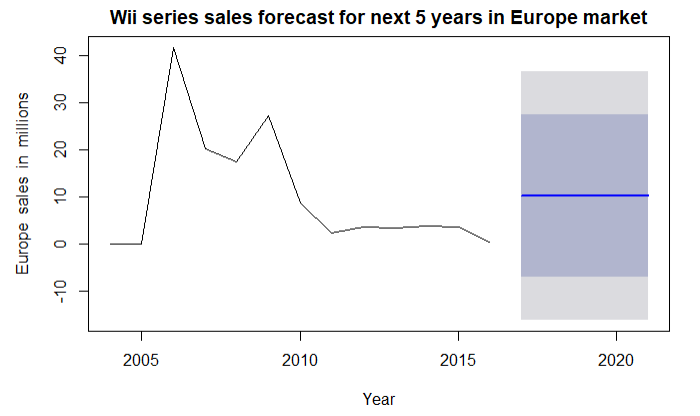






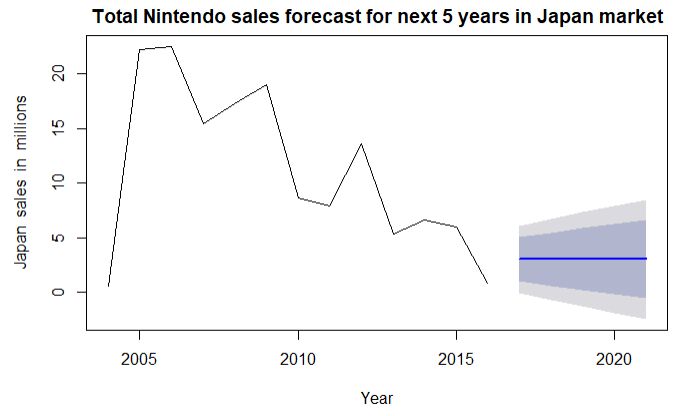
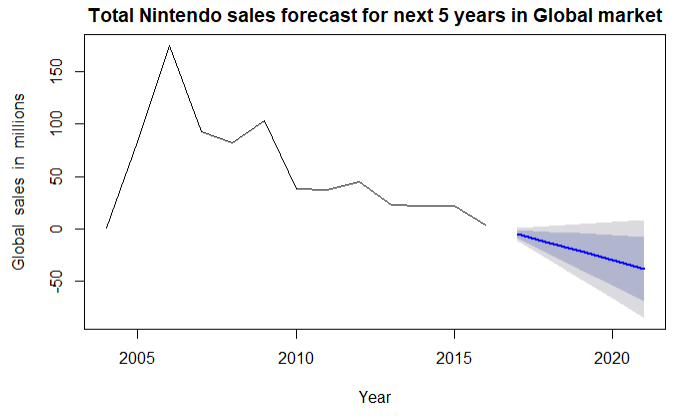
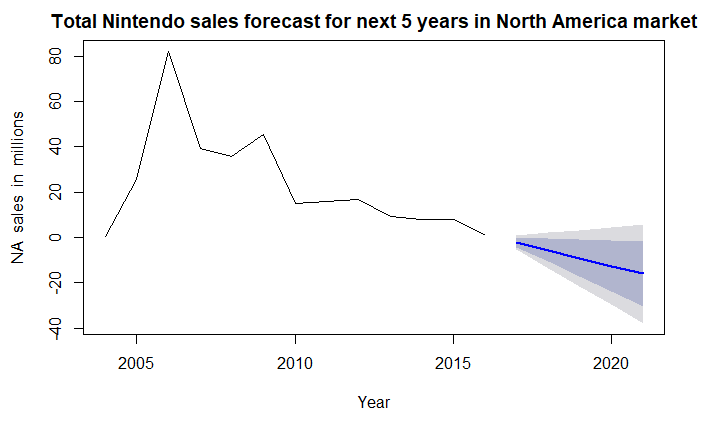
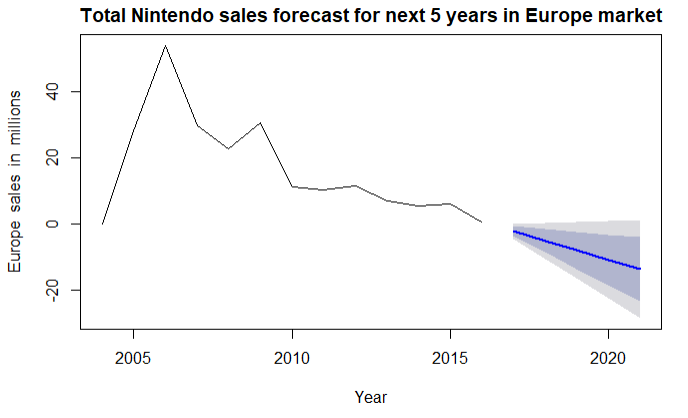
From the above graphs, we can see that the DS series sales are declining across all regions in the next 5 years.

**Wii Platform series Forecast Analysis**



From the above graph, we can see that Wii series sales are rising over the next 5 years in North America, Europe and in Global market. But in Japan, we can see that it is neither raising nor coming down. The area around the darkly shaded region is the place where the maximum sales of the video games reside for the next five years for the respective regions.

**Total Nintendo Video game sales Forecast Analysis**



From the above graph, we can see that total sales are declining across North America, Europe and in the Global market over the next 5 years. But in Japan, we can see that the sales are raising. The area around the darkly shaded region is the place where the maximum sales of the video games reside.

**CONCLUSION**

Video games have a lot of craze among people around the world. Nintendo is one of the bestselling games in the world. From my analysis, I found that Console games are more popular in North America, Europe and globally, but handheld games are more popular in Japan.

The tests show that racing and Sports games dominate the North American and Global market. Whereas Europe is dominated by Sports and Japan is dominated by racing games.

The tests show that the Developer has no influence on the sales of video games. The platform has a significant influence on North American sales but not in Japan, Europe or Global sales. It is found that the Critic score and Genre is more influential than User score to drive sales.

Finally, the time series plots show that the sales of Nintendo video games are declining from the year 2010 to 2016. The article, The Nintendo Switch (2017) also shows that Video game sales of Nintendo are declining from the year 2009 to 2016. The author also describes that PS4 and XBOX have more sales in the year 2012 and 2013 which might have taken up the Nintendo market (The Nintendo Switch, 2017). The Global video game sales forecast for next 5 years shows that the total Nintendo sales are improving in Japan but it is declining in Europe, North America and in the global market. The Forecast for the next 5 years shows that the sales of handheld games (DS series) is declining across all the regions and the sales of Console games (Wii series) are improving across all the regions. The Author Kain (2018) mentions that Nintendo was the best seller in July 2018 and it completely dominated the video game market over other companies. Russell (2018) points out that Nintendo console games have created a huge profit in 2018.

**RECOMMENDATIONS**

* Focus on building sports and racing Console games for North American and global market.
* Build Console games for North America, Europe, and Globally.
* Build racing games in both the platforms (Handheld and console) for Japan.
* Focus on sports games for Europe market.

**REFERENCE**

Kabacoff, R. (August, 2011). R in Action (2nd ed.). Manning Publications Co. : Publisher.

Crawley, M. R. (August, 2015). Statistics an introduction using R (2nd ed.). Wiley Publications

Co. : Publisher.

Russell, J. (June, 2018). Nintendo’s annual profit rockets by 500% after selling 15M Switch

consoles. Retrieved from <https://techcrunch.com/2018/04/26/nintendo-profit-switch-15-millon-sales/>

Unit sales of Nintendo's home consoles from 1997 to 2018 (in millions). (n.a). Retrieved from

<https://www.statista.com/statistics/227012/lifetime-unit-sales-of-nintendos-home-consoles/>

Kain, E (August, 2018). Nintendo Completely Dominated Video Game Sales In July. Retrieved

from <https://www.forbes.com/sites/erikkain/2018/08/23/nintendo-completely-dominated-video-game-sales-in-july-octopath-traveler/#244de9a62435>

The Nintendo Switch’s sell-out launch, in many charts. (July, 2017). Retrieved from

<https://www.marketwatch.com/story/the-nintendo-switch-launch-in-many-charts-2017-03-27>