Final Compiled Report

Authors:
Aayush Malde
Henry McLane
James Sweat
Daniel Bonilla

Abstract:

For our project, we decided to investigate and analyze the video game industry, by acquiring data on the top 150 best selling games. The data displayed the games' names, platforms, genre, publisher, year released, and sales. With this we were able to gather useful trends and characteristics. First, we were able to identify that the year 2010 had the most selling games with 14 games sold out of the 150 games. We identified that the most successful platform was Xbox360 with 22 games sold. Through filtering, we discovered that the most popular shooting game after the year 2000 was Call of Duty Modern Warfare 3. The best selling Playstation 2 game was Grand Theft Auto San Andreas. The most popular Nintendo game before the year 2000 was Super Mario Bros released in 1985 for the NES. We determined that Nintendo was the most successful company because they have the most popular games in a wide variety of genres. We discovered that there is an inverse relationship between the total sales and the amount of games sold, as well as an inverse relationship between the rank and the total sales. We were able to determine that the USA sales, Japan Sales, And EU Sales and Other Sales, were our only relevant predictors for the Total Sales. We were also able to predict whether a game was released after the year 2000 based on the USA Sales, Japan Sales, and Other Sales. Through hypothesis testing, we determined that Nintendo games sales take up between 55.4% and 55.5% of the total sales. Through a two-sample inference test, we discovered that Activision had a greater proportion than Nintendo in USA Sales, meaning that despite Nintendo's overall success, Activision is more successful in the United States.

Introduction:

While conversing about what the subject of our data should be, we started to discuss our own personal interests. One common interest our group had was our enjoyment for video games. Knowing this, we set out to search for a dataset. Our data contains a list of popular video games with sales greater than 100,000 copies. This was perfect because all our group members were heavily engaged within the video game scene prior to discussion and this dataset included many of the familiar game titles we love the most. In addition, the publisher of the dataset provided us with records of the data stating that, "2 records were dropped due to incomplete information." I believe that these 2 records will not affect the conducting of this data investigation especially when considering there are 16,598 records. Some of the relevant variables our group will utilize:

Rank - Ranking of overall sales
Name - The game's name
Platform - Platform on which game can be played (i.e. PC,PS4, etc.)
Year - Year of the game's release
Genre - Genre of the game
Publisher - Publisher of the game
NA_Sales - Sales in North America (in millions)
EU_Sales - Sales in Europe (in millions)
JP_Sales - Sales in Japan (in millions)
Other_Sales - Sales in the rest of the world (in millions)
Global_Sales - Total worldwide sales (in millions).

A Description of Reasoning Behind Wrangling:

We found the standard deviation of these key numeric variables: Year, USA Sales, Europe Sales, Japan Sales, Total Sales, and Other Sales[Appendix #1-3]. These summaries can reveal the prevalence of video games in a given year. Also, in terms of sales, these summaries can reveal the ranges of sales within the dataset. Additionally, we found the frequency tables for Name[4-12], Platform[13], Genre[13], and Publisher[14] and two-way tables comparing Publisher and Genre[15], Platform and Genre[16], and Publisher and Rank[17]. These portray the relevancy of a given variable within the dataset. We also did some simple subsetting of the dataset. We included Popular Modern Shooters[18], Best Selling PS2 Games[19], and Old Nintendo Classics[20]. In terms of data cleaning, we made some simple changes to the format of the dataset. We changed the names of the "NA_Sales, EU_Sales, JP_Sales, Global_Sales" vectors to "USA_Sales, Europe_Sales, Japan_Sales, Total_Sales." We removed those previous rows from the dataset entirely because they were no longer needed. We made the name change to make the column names more comprehensible. Also, during cleaning, we mutated the "Rank" variable to "Ranking" while setting the games with rank <= 100 to "Top 100", rank <= 500 to "Top 500", rank <= 1000 to "Top 1000" and anything under as "Not Worthy." Adding this new ranking provides a secondary ranking system to reference the games' relevance in the dataset. We also replaced all the values of 0 to "NA," so the variables' summaries were precise. Finally, we grouped by the Year[21] and Platform[22] variables to see which year and platform were the most common in our dataset.

A Description of Findings After Wrangling:

After wrangling, we found that the United States marketplace is much more lucrative than Europe and Japan. Also, we were surprised by the older years included in the dataset. The most popular platform found was the PS3 and the XBOX 360 with 21 and 22 games, respectively. Xbox is also the most popular platform for shooter games. We observed that the most popular genre within the dataset was again shooter games. In terms of publishers, Nintendo is very dominant. Nintendo has the most games listed with 63, the most platform games listed, the most role-playing games and this publisher occupies the entirety of the top ten games listed. Our first subset is popular modern shooters, which shows the games ranked higher than 50 in the shooter genres after 2000, with total sales of over 13,370,000. From this, we can see the popularity of each game within the subset and the popularity within the total subset, which could settle arguments on which game is ultimately the best. From this subset, we can observe that Activision is the dominant publisher of shooter video games. Another subset we made was old Nintendo classics, which The Games published by Nintendo before the year 2000 with total sales of over 13,370,000. We can observe older Nintendo games' relevancy concerning the other games within the data set from this subset.

A Description of Our Graphs:

For this part of the project, we were curious about finding several visual trends about our video game data. We used a word cloud, bar graph, box plot, scatterplot, and bubble graph so that we can properly see what these trends mean.

Starting with the word cloud[23], we saw in our dataset that there were a lot more video game companies covered than we anticipated and wanted to see how some of the lesser known companies stood out against the giants of today, for example Nintendo. After looking at this word cloud, we knew that Activision and Take-Two Interactive would have the biggest presence. After all, Take-Two Interactive owns a lot of other popular companies that play a big influence in the video game market. However, we were surprised to see that there were multiple “Nintendos”

in varying sizes. We don't know whether it was an error in our coding or if it's because our range wasn't big enough to include only one "Nintendo," which would cause the code to create smaller clones to compensate for its limited size.

For the bar graph[24], we wanted to see a visualization of the frequency of video game companies that had large/small sales. From this, we saw unsurprisingly that most video games made under $1 million and that there would be a negative trend as you progress along the x axis. However, we were surprised to see that there was one outlier game that sold over $80 million, completely separate from the rest of the pack. That game was Wii Sports, which completely overtook the original Super Mario Bros game which sold around $40 million. Wii Sports was a lot more popular than we originally thought.

For the boxplot[25] and scatterplot[26], we had similar results. Both graphs displayed the outlier previously mentioned, except we organized by the rank each game possessed, ranging from Top10 to Top150. For the boxplot we only looked at Japan Sales, as we knew that Japan plays a significant role in the video game industry and were curious to see how it compared with the rest of the world. The results lined up similarly with the Total Sales. However, thanks to the boxplot, it displayed a more accurate visual of the average sales in terms of ranking. We expected the boxplot of the Top10 ranking to have the biggest box but were surprised to see that it was the Top50 box instead.

Finally, the scatter plot helps us visualize a real trend in video game sales, showing the seeming plateau among the Top100 and Top500 rankings and the Top10 variables having wildly diverse values, making the outlier Wii Sports really stand out. Although we thought that the regression line would curve upwards a lot more because of this outlier. Perhaps because there were simply too many similar values in the previous ranking categories. As for the 3-category bubble graphs[27+28] we were also able to look which decade the top selling games were released in, and the distribution of publishers. Most of the Top10 games were from the 2010s, which highlights how much influence video games play in our modern society. The distribution of publishers had a similar result to the word cloud, with Nintendo overtaking the competition.

A Description of Our Regression:

For our data, we were curious about the relationships between the various Sales variables. We found these by writing multiple and logistic regression codes. Firstly, with the multiple regression, we wanted to see how increasing one of these variables impacts the others. We did this by making the Total Sales variable the dependent variable, and saw that every single other Sales variable, from USA to Japan, almost all had the same predicting value of 1(aka "slope" if this was a singular regression). We don't know whether this was due to an error in our code, but it's more likely because of the fact that there already is an inherent relation between the Total Sales and the smaller Sales values; Total = USA+EU+JP+Other. They can't be used to make predictions on a preset formula. But it's weird how EU sales is slightly less than 1, despite all of them having the same p value as well.

As for the logistic regression, we wanted to figure out how different video game sales from different countries can predict whether a game was released after 2000. We found here that the EU sales variable got eliminated as the worst predictor for the After2000 variable. It's hard to say whether it was a coincidence that the variable that stood out in the multiple regression code also stood out here. Again, we don't know why, but it's interesting to see here. After interpolation and extrapolation of sample values, we found the following: The first set of

sales entries for a video game has a 57% chance of it being released after 2000, whereas there is a 100% for the second set of entries. This indicates how video games became a lot more popular after the year 2000, which is why there's a low threshold for video game sales if it got released before 2000.

One and Two Sample Inference:

When considering our data, we became curious about the proportions between Nintendo/Activision game sales and total sales. We found these proportions using one and two sample inferences.

Our previous assumption was that Nintendo game sales were much higher than any other publisher within our data. To prove this true, we decided to conduct a one-sample inference on the sum of Nintendo Sales and the sum of total sales to determine the proportion of Nintendo sales in the dataset. We measure this proportion using a 95 percent confidence interval. After conducting the test we concluded an interpretation. We stated that we are 95% confident that the true percentage of Nintendo games sales within total sales is between 55.4% and 55.5%. This means that Nintendo games sales take up between 55.4% and 55.5% which is above half the dataset's sales. These findings further prove the successfulness of Nintendo gaming.

Secondly, we conducted a two-sample inference test on the difference of proportions of the sum of Nintendo sales and the sum of Activision sales over total sales. Before testing these proportions, the idea of Nintendo being a worldwide dominant publisher is still present. However after testing, the data revealed an interpretation that was not expected. We again measured this proportion difference with a 95 percent confidence interval. We stated that we are 95% confident that the proportion of Activision sales in the USA is between 0.03241773 and 0.03257421 greater than that of Nintendo USA sales. These findings reveal that although Nintendo is a very dominant company, Activision has an edge over Nintendo in the US.

Conclusion:

For our project, we wanted to answer the above questions and analyze some of the key factors that determined the overall sales of the top 150 best selling video games. While we gained valuable insight, there were quite a few limitations that limited our research. The most prominent example was that some of the entries in the regional sales columns had NAs. Because we did not want to lose out on other valuable information from that row, we decided to wrangle the data by replacing the NAs with 0s. However, this meant that analysis on regional sales were not as accurate as analysis on total sales, and so for the most part we only did data analysis with regards to the total sales.

In addition, there were not many factors that we had to consider. Unfortunately, the only factors we could really use to predict sales were Platform, Year, Genre, and Publisher. Currently our research is useful because investors would be able to determine which game/company they should invest in. Companies would also be able to accurately predict the sales of a video game, and can also optimize the type of games, platform, and publisher to maximize profits. While effective now, our research can be continued to improve the accuracy of sales predictions. One way is to improve the data scraping process such that there would be no NAs in our original data set. We could also update our data set to include 2020 sales as our data set is not as modern as it could be. Another way is to have more predictors and answer their research questions, a few examples being:

Date of Release- When in the year is the best time to release a game? Near the holidays? Summer where there is less competition?

Money spent on Advertisements - Where should the advertising money be spent on to maximize regional sales as well as total sales?

Packaging with console - Are sales higher when a game is bundled with a console or separate and available on multiple platforms?

References:

Smith, Gregory. “Video Game Sales.” *Kaggle.* Accessed 7 Dec. 2020. https://www.kaggle.com/gregorut/videogamesales

Renato DinhaniRenato Dinhani 28.6k4646 gold badges124124 silver badges190190 bronze badges, et al. “How Do I Replace NA Values with Zeros in an R Dataframe?” *Stack Overflow*, 1 Jan. 1961, stackoverflow.com/questions/8161836/how-do-i-replace-na-values-with-zeros-in-an-r-dataframe.

“Reordering Data Frame Columns in R.” *STHDA*, www.sthda.com/english/wiki/reordering-data-frame-columns-in-r.

Appendix:

1.

```
#year
summary(vg_data$Year)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1982    2001    2007    2005    2011    2015
```

```
sd(vg_data$Year)
```

```
## [1] 7.391951
```

```
#usa
summary(vg_data$USA_Sales)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##    70000  3010000  4285000  5708867  6577500 41490000
```

```
sd(vg_data$USA_Sales)
```

```
## [1] 5188504
```

```
#europe
summary(vg_data$Europe_Sales)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##        0  1972500  2800000  3543267  4015000 29020000
```

```
sd(vg_data$Europe_Sales)
```

```
## [1] 3072281
```

```
#japan
summary(vg_data$Japan_Sales)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##        0   130000   850000  1587200  2667500 10220000
```

```
sd(vg_data$Japan_Sales)
```

```
## [1] 1861846
```

2.

```
#other countries
summary(vg_data$Other_Sales)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##        0   475000   780000  1103333  1252500 10570000
```

```
sd(vg_data$Other_Sales)
```

```
## [1] 1330852
```

```
#total
summary(vg_data$Total_Sales)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##  5840000  6902500  8930000 11942400 13370000 82740000
```

```
sd(vg_data$Total_Sales)
```

```
## [1] 8948990
```

3.

4.

```
#frequency
table(vg_data$Name)
```

```
##
##                      Animal Crossing: New Leaf
##                                              1
##                    Animal Crossing: Wild World
##                                              1
##                           Assassin's Creed III
##                                              1
##                                  Battlefield 3
##                                              2
##                              Big Brain Academy
##                                              1
##    Brain Age 2: More Training in Minutes a Day
##                                              1
##   Brain Age: Train Your Brain in Minutes a Day
##                                              1
##                 Call of Duty 4: Modern Warfare
##                                              2
##                 Call of Duty: Advanced Warfare
##                                              1
##                        Call of Duty: Black Ops
##                                              2
##                      Call of Duty: Black Ops 3
##                                              2
##                     Call of Duty: Black Ops II
##                                              2
##                           Call of Duty: Ghosts
##                                              2
##                 Call of Duty: Modern Warfare 2
##                 Call of Duty: Modern Warfare 3
##                                              2
##                     Call of Duty: World at War
##                                              1
##                                Crash Bandicoot
##                                              1
##         Crash Bandicoot 2: Cortex Strikes Back
##                                              1
##                      Crash Bandicoot 3: Warped
##                                              1
##                            Donkey Kong Country
##                                              1
##                   Donkey Kong Country Returns
##                                              1
## Dragon Quest IX: Sentinels of the Starry Skies
##                                              1
##                                         Driver
##                                              1
##                                      Duck Hunt
##                                              1
##                                      Fallout 4
##                                              1
##                                        FIFA 12
##                                              1
##                                        FIFA 14
##                                              1
##                                        FIFA 15
##                                              1
##                                        FIFA 16
##                                              1
##                                 FIFA Soccer 13
##                                              1
```

5.

```
##                             Final Fantasy VII
##                                             1
##                            Final Fantasy VIII
##                                             1
##                               Final Fantasy X
##                                             1
##                             Final Fantasy XII
##                                             1
##                                  Gears of War
##                                             1
##                                Gears of War 2
##                                             1
##                                Gears of War 3
##                                             1
##                                 GoldenEye 007
##                                             1
##                                  Gran Turismo
##                                             1
##                                Gran Turismo 2
##                                             1
##                           Gran Turismo 3: A-Spec
##                                             1
##                                Gran Turismo 4
##                                             1
##                                Gran Turismo 5
##                                             1
##                            Grand Theft Auto III
##                                             1
##                             Grand Theft Auto IV
##                                             2
##                              Grand Theft Auto V
##                                             3
##          Grand Theft Auto: Liberty City Stories
##                                             1
##                    Grand Theft Auto: San Andreas
##                                             1
##                     Grand Theft Auto: Vice City
##                                             1
##                                         Halo 2
##                                             1
##                                         Halo 3
##                                             1
##                                    Halo 3: ODST
##                                             1
##                                         Halo 4
##                                             1
##                           Halo: Combat Evolved
##                                             1
##                                    Halo: Reach
##                                             1
##                                     Just Dance
##                                             1
##                                   Just Dance 2
##                                             1
##                                   Just Dance 3
##                                             1
##                                   Just Dance 4
##                                             1
##                             Kinect Adventures!
##                                             1
##                                  Kinect Sports
##                                             1
```

```
##                                Kingdom Hearts
##                                             1
##                                LittleBigPlanet
##                                             1
##             Mario & Sonic at the Olympic Games
##                                             1
##                                  Mario Kart 64
##                                             1
##                                   Mario Kart 7
##                                             1
##                                   Mario Kart 8
##                                             1
##                                  Mario Kart DS
##                                             1
##                                 Mario Kart Wii
##                                             1
##                     Mario Kart: Double Dash!!
##                                             1
##                                  Mario Party 8
##                                             1
##                                 Mario Party DS
##                                             1
##                     Medal of Honor: Frontline
##                                             1
##                                Metal Gear Solid
##                                             1
##           Metal Gear Solid 2: Sons of Liberty
##                                             1
##      Metal Gear Solid 4: Guns of the Patriots
##                                             1
```

8.

```
##                                     Minecraft
##                                             1
##                    Need for Speed Underground
##                                             1
##                  Need for Speed Underground 2
##                                             1
##                         New Super Mario Bros.
##                                             1
##                       New Super Mario Bros. 2
##                                             1
##                    New Super Mario Bros. Wii
##                                             1
##                                   Nintendogs
##                                             1
##                                     Pac-Man
##                                             1
##                     PokÃ©mon Crystal Version
##                                             1
##                     PokÃ©mon Emerald Version
##                                             1
##                    PokÃ©mon Platinum Version
##                                             1
##       PokÃ©mon Yellow: Special Pikachu Edition
##                                             1
##              Pokemon Black 2/Pokemon White 2
##                                             1
##                  Pokemon Black/Pokemon White
##                                             1
##                 Pokemon Diamond/Pokemon Pearl
##                                             1
```

9.

```
##              Pokemon FireRed/Pokemon LeafGreen
##                                              1
##                    Pokemon Gold/Pokemon Silver
##                                              1
##           Pokemon HeartGold/Pokemon SoulSilver
##                                              1
##      Pokemon Omega Ruby/Pokemon Alpha Sapphire
##                                              1
##                      Pokemon Red/Pokemon Blue
##                                              1
##                  Pokemon Ruby/Pokemon Sapphire
##                                              1
##                             Pokemon X/Pokemon Y
##                                              1
##                             Red Dead Redemption
##                                              2
##                           Sonic the Hedgehog 2
##                                              1
##                  Star Wars Battlefront (2015)
##                                              1
##          Street Fighter II: The World Warrior
##                                              1
##                             Super Mario 3D Land
##                                              1
##                                  Super Mario 64
##                                              2
##                           Super Mario All-Stars
##                                              1
##                              Super Mario Bros.
```

10.

```
##                                              1
##                            Super Mario Bros. 2
##                                              1
##                            Super Mario Bros. 3
##                                              1
##                             Super Mario Galaxy
##                                              1
##                           Super Mario Galaxy 2
##                                              1
##                                Super Mario Kart
##                                              1
##                                Super Mario Land
##                                              1
##             Super Mario Land 2: 6 Golden Coins
##                                              1
##                           Super Mario Sunshine
##                                              1
##                              Super Mario World
##                                              1
##                        Super Smash Bros. Brawl
##                                              1
##            Super Smash Bros. for Wii U and 3DS
##                                              1
##                        Super Smash Bros. Melee
##                                              1
##                                      Tekken 3
##                                              1
##                                      Tetris
##                                              1
##                    The Elder Scrolls V: Skyrim
```

11.

```
##                                              2
```

```
##                               The Last of Us
##                                            1
##                           The Legend of Zelda
##                                            1
##             The Legend of Zelda: Ocarina of Time
##                                            1
##           The Legend of Zelda: Twilight Princess
##                                            1
##                                    The Sims 3
##                                            1
##                      Uncharted 2: Among Thieves
##                                            1
##                  Uncharted 3: Drake's Deception
##                                            1
##                                       Wii Fit
##                                            1
##                                  Wii Fit Plus
##                                            1
##                                     Wii Party
##                                            1
##                                      Wii Play
##                                            1
##                                    Wii Sports
##                                            1
##                               Wii Sports Resort
##                                            1
##                               World of Warcraft
##                                            1
##                                   Zumba Fitness
##                                            1
```

12.

```
table(vg_data$Platform)
```

```
##
## 2600  3DS   DS   GB  GBA   GC  GEN  N64  NES   PC   PS  PS2  PS3  PS4  PSP SNES
##    1    7   15    7    3    3    1    4    5    2   10   12   21    7    1    5
##  Wii WiiU X360   XB XOne
##   20    1   22    2    1
```

```
table(vg_data$Genre)
```

```
##
##       Action    Adventure     Fighting         Misc     Platform       Puzzle
##           24            1            5           12           22            3
##       Racing Role-Playing      Shooter   Simulation       Sports
##           14           23           31            4           11
```

13.

14.

```
table(vg_data$Publisher)
```

```
##
##                          505 Games                           Activision
##                                  1                                   17
##                              Atari                      Bethesda Softworks
##                                  1                                    3
##                            Capcom                        Electronic Arts
##                                  1                                   12
##                     GT Interactive             Konami Digital Entertainment
##                                  1                                    3
##              Microsoft Game Studios                              Nintendo
##                                 12                                   63
##                               Sega              Sony Computer Entertainment
##                                  2                                   15
## Sony Computer Entertainment Europe                          Square Enix
##                                  1                                    1
##                          SquareSoft                    Take-Two Interactive
##                                  1                                   11
##                             Ubisoft
##                                  5
```

15.

```
#two way tables
vg_data%>%
  select(Publisher,Genre)%>%
  table()
```

```
##                                 Genre
## Publisher                        Action Adventure Fighting Misc Platform
##   505 Games                           0         0        0    0        0
##   Activision                          0         0        0    0        0
##   Atari                               0         0        0    0        0
##   Bethesda Softworks                  0         0        0    0        0
##   Capcom                              0         0        1    0        0
##   Electronic Arts                     1         0        0    0        0
##   GT Interactive                      1         0        0    0        0
##   Konami Digital Entertainment        3         0        0    0        0
##   Microsoft Game Studios              0         0        0    2        0
##   Nintendo                            4         1        3    6       17
##   Sega                                0         0        0    0        1
##   Sony Computer Entertainment         2         0        1    0        4
##   Sony Computer Entertainment Europe  1         0        0    0        0
##   Square Enix                         0         0        0    0        0
##   SquareSoft                          0         0        0    0        0
##   Take-Two Interactive               11         0        0    0        0
##   Ubisoft                             1         0        0    4        0
##                                 Genre
## Publisher                        Puzzle Racing Role-Playing Shooter
##   505 Games                           0      0            0       0
##   Activision                          0      0            1      16
##   Atari                               1      0            0       0
##   Bethesda Softworks                  0      0            3       0
##   Capcom                              0      0            0       0
##   Electronic Arts                     0      2            0       4
##   GT Interactive                      0      0            0       0
##   Konami Digital Entertainment        0      0            0       0
##   Microsoft Game Studios              0      0            0       9
##   Nintendo                            2      7           14       2
##   Sega                                0      0            0       0
##   Sony Computer Entertainment         0      5            3       0
##   Sony Computer Entertainment Europe  0      0            0       0
##   Square Enix                         0      0            1       0
##   SquareSoft                          0      0            1       0
##   Take-Two Interactive                0      0            0       0
##   Ubisoft                             0      0            0       0
##                                 Genre
## Publisher                        Simulation Sports
##   505 Games                             0      1
##   Activision                            0      0
##   Atari                                 0      0
##   Bethesda Softworks                    0      0
##   Capcom                                0      0
##   Electronic Arts                       1      4
##   GT Interactive                        0      0
##   Konami Digital Entertainment          0      0
##   Microsoft Game Studios                0      1
##   Nintendo                              3      4
##   Sega                                  0      1
##   Sony Computer Entertainment           0      0
##   Sony Computer Entertainment Europe    0      0
##   Square Enix                           0      0
##   SquareSoft                            0      0
##   Take-Two Interactive                  0      0
##   Ubisoft                               0      0
```

16.

```
vg_data%>%
  select(Platform,Genre)%>%
  table()
```

```
##          Genre
## Platform Action Adventure Fighting Misc Platform Puzzle Racing Role-Playing
##     2600      0         0        0    0        0      1      0            0
##     3DS       0         0        1    0        2      0      1            2
##     DS        1         0        0    3        2      1      1            5
##     GB        0         1        0    0        1      1      0            4
##     GBA       0         0        0    0        0      0      0            3
##     GC        0         0        1    0        1      0      1            0
##     GEN       0         0        0    0        1      0      0            0
##     N64       1         0        0    0        1      0      1            0
##     NES       1         0        0    0        3      0      0            0
##     PC        0         0        0    0        0      0      0            1
##     PS        2         0        1    0        3      0      2            2
##     PS2       4         0        0    0        0      0      4            3
##     PS3       9         0        0    0        1      0      1            1
##     PS4       1         0        0    0        0      0      0            1
##     PSP       1         0        0    0        0      0      0            0
##     SNES      0         0        1    0        3      0      1            0
##     Wii       1         0        1    7        4      0      1            0
##     WiiU      0         0        0    0        0      0      1            0
##     X360      3         0        0    2        0      0      0            1
##     XB        0         0        0    0        0      0      0            0
##     XOne      0         0        0    0        0      0      0            0
```

```
## Platform Shooter Simulation Sports
##     2600       0          0      0
##     3DS        0          1      0
##     DS         0          2      0
##     GB         0          0      0
##     GBA        0          0      0
##     GC         0          0      0
##     GEN        0          0      0
##     N64        1          0      0
##     NES        1          0      0
##     PC         0          1      0
##     PS         0          0      0
##     PS2        1          0      0
##     PS3        7          0      2
##     PS4        3          0      2
##     PSP        0          0      0
##     SNES       0          0      0
##     Wii        0          0      6
##     WiiU       0          0      0
##     X360      15          0      1
##     XB         2          0      0
##     XOne       1          0      0
```

17.

```
vg_data%>%
  select(Publisher,Ranking)%>%
  table()
```

```
##                                 Ranking
## Publisher                        Top_10 Top_100 Top_150 Top_50
##   505 Games                           0       0       1      0
##   Activision                          0       6       3      8
##   Atari                               0       1       0      0
##   Bethesda Softworks                  0       1       2      0
##   Capcom                              0       0       1      0
##   Electronic Arts                     0       5       7      0
##   GT Interactive                      0       0       1      0
##   Konami Digital Entertainment        0       0       3      0
##   Microsoft Game Studios              0       4       6      2
##   Nintendo                           10      20      11     22
##   Sega                                0       1       1      0
##   Sony Computer Entertainment         0       6       7      2
##   Sony Computer Entertainment Europe  0       0       1      0
##   Square Enix                         0       0       1      0
##   SquareSoft                          0       1       0      0
##   Take-Two Interactive                0       3       2      6
##   Ubisoft                             0       2       3      0
```

18.

```
#(b) Subsetting (or filtering) the data to summarize key variables. Think about interesting questions you can investigate th
rough filtering the data.

PopularModernShooters<-vg_data%>%
  filter(Year>2000,Rank<=50,Genre=="Shooter")
head(PopularModernShooters)
```

```
##   Rank                         Name Platform Year   Genre  Publisher
## 1   30 Call of Duty: Modern Warfare 3     X360 2011 Shooter Activision
## 2   32        Call of Duty: Black Ops     X360 2010 Shooter Activision
## 3   34      Call of Duty: Black Ops 3      PS4 2015 Shooter Activision
## 4   35     Call of Duty: Black Ops II      PS3 2012 Shooter Activision
## 5   36     Call of Duty: Black Ops II     X360 2012 Shooter Activision
## 6   37 Call of Duty: Modern Warfare 2     X360 2009 Shooter Activision
##   USA_Sales Europe_Sales Japan_Sales Other_Sales Total_Sales Ranking
## 1   9030000      4280000      130000     1320000    14760000  Top_50
## 2   9670000      3730000      110000     1130000    14640000  Top_50
## 3   5770000      5810000      350000     2310000    14240000  Top_50
## 4   4990000      5880000      650000     2520000    14030000  Top_50
## 5   8250000      4300000       70000     1120000    13730000  Top_50
## 6   8520000      3630000       80000     1290000    13510000  Top_50
```

19.

```
BestSellingPS2Games<-vg_data%>%
  filter(Platform=="PS2",Total_Sales>13370000)
head(BestSellingPS2Games)
```

```
##   Rank                         Name Platform Year  Genre
## 1   18 Grand Theft Auto: San Andreas      PS2 2004 Action
## 2   25   Grand Theft Auto: Vice City      PS2 2002 Action
## 3   29        Gran Turismo 3: A-Spec      PS2 2001 Racing
##                    Publisher USA_Sales Europe_Sales Japan_Sales Other_Sales
## 1        Take-Two Interactive   9430000       400000      410000    10570000
## 2        Take-Two Interactive   8410000      5490000      470000     1780000
## 3 Sony Computer Entertainment   6850000      5090000     1870000     1160000
##   Total_Sales Ranking
## 1    20810000  Top_50
## 2    16150000  Top_50
## 3    14980000  Top_50
```

```
OldNintendoClassics<-vg_data%>%
  filter(Publisher=="Nintendo",Year<2000,Total_Sales>13370000)
head(OldNintendoClassics,10)
```

```
##   Rank                                  Name Platform Year        Genre
## 1    2                   Super Mario Bros.      NES 1985     Platform
## 2    5               Pokemon Red/Pokemon Blue       GB 1996 Role-Playing
## 3    6                                Tetris       GB 1989       Puzzle
## 4   10                             Duck Hunt      NES 1984      Shooter
## 5   13           Pokemon Gold/Pokemon Silver       GB 1999 Role-Playing
## 6   19                     Super Mario World     SNES 1990     Platform
## 7   22                      Super Mario Land       GB 1989     Platform
## 8   23                   Super Mario Bros. 3      NES 1988     Platform
## 9   31 PokÃ©mon Yellow: Special Pikachu Edition       GB 1998 Role-Playing
##   Publisher USA_Sales Europe_Sales Japan_Sales Other_Sales Total_Sales Ranking
## 1  Nintendo  29080000      3580000     6810000      770000    40240000  Top_10
## 2  Nintendo  11270000      8890000    10220000     1000000    31370000  Top_10
## 3  Nintendo  23200000      2260000     4220000      580000    30260000  Top_10
## 4  Nintendo  26930000       630000      280000      470000    28310000  Top_10
## 5  Nintendo   9000000      6180000     7200000      710000    23100000  Top_50
## 6  Nintendo  12780000      3750000     3540000      550000    20610000  Top_50
## 7  Nintendo  10830000      2710000     4180000      420000    18140000  Top_50
## 8  Nintendo   9540000      3440000     3840000      460000    17280000  Top_50
## 9  Nintendo   5890000      5040000     3120000      590000    14640000  Top_50
```

20.

```
#by groups

vg_data%>%
  group_by(Year)%>%
  summarize(Count=n())
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

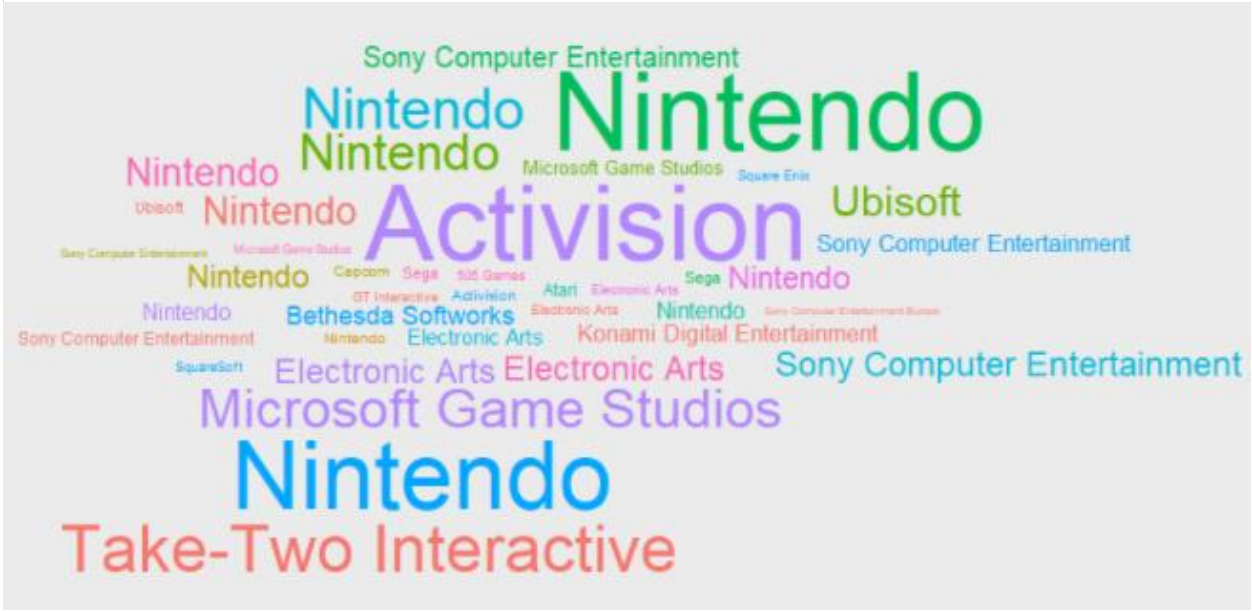
```
## # A tibble: 30 x 2
##     Year Count
##    <dbl> <int>
##  1  1982     1
##  2  1984     1
##  3  1985     1
##  4  1986     1
##  5  1988     2
##  6  1989     2
##  7  1990     1
##  8  1992     4
##  9  1993     1
## 10  1994     1
## # ... with 20 more rows
```

21.

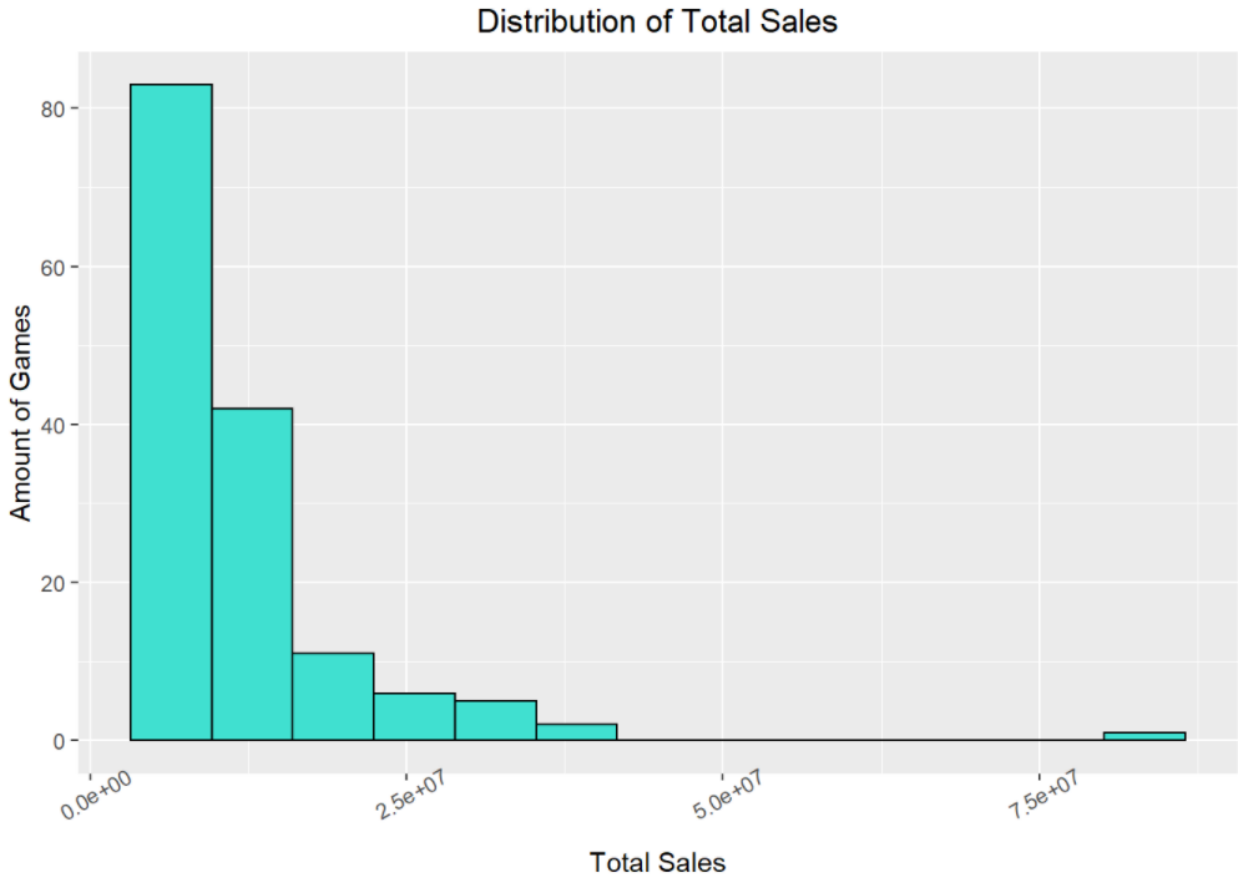
```
vg_data%>%
  group_by(Platform)%>%
  summarize(Count=n())
```

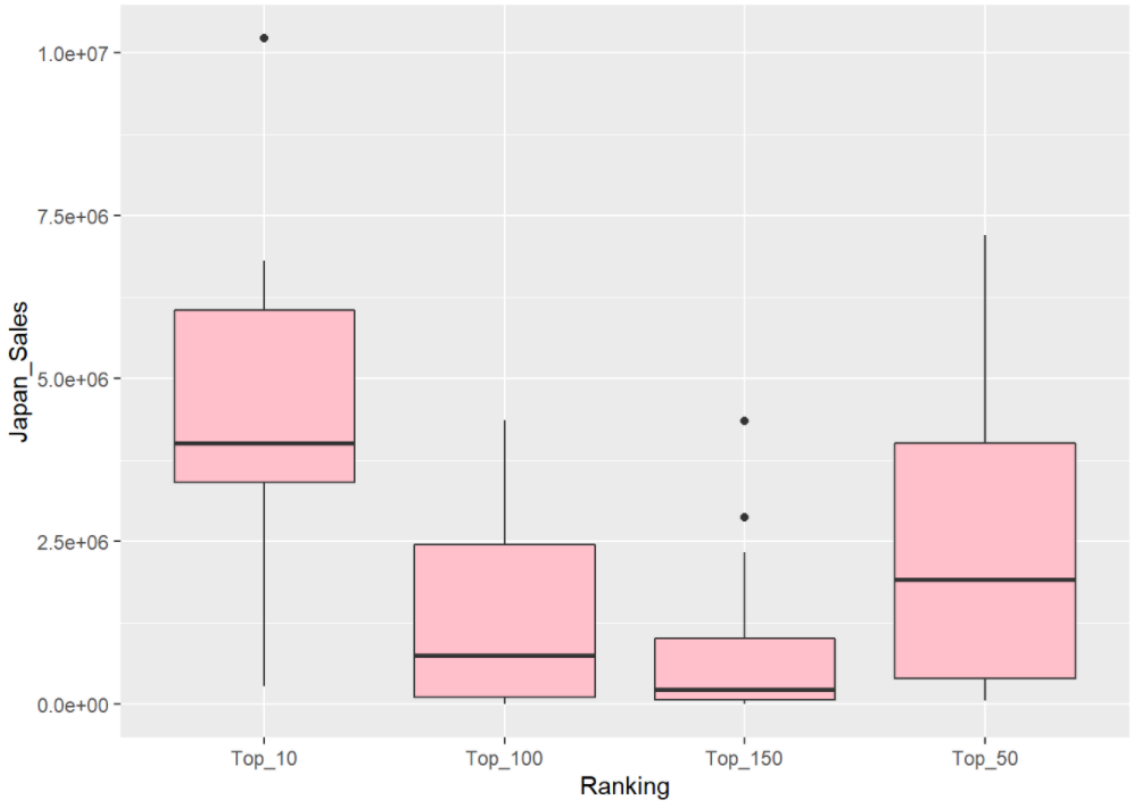
```
## `summarise()` ungrouping output (override with `.groups` argument)
```

22.
```
## # A tibble: 21 x 2
##    Platform Count
##    <chr>    <int>
##  1 2600         1
##  2 3DS          7
##  3 DS          15
##  4 GB           7
##  5 GBA          3
##  6 GC           3
##  7 GEN          1
##  8 N64          4
##  9 NES          5
## 10 PC           2
## # ... with 11 more rows
```

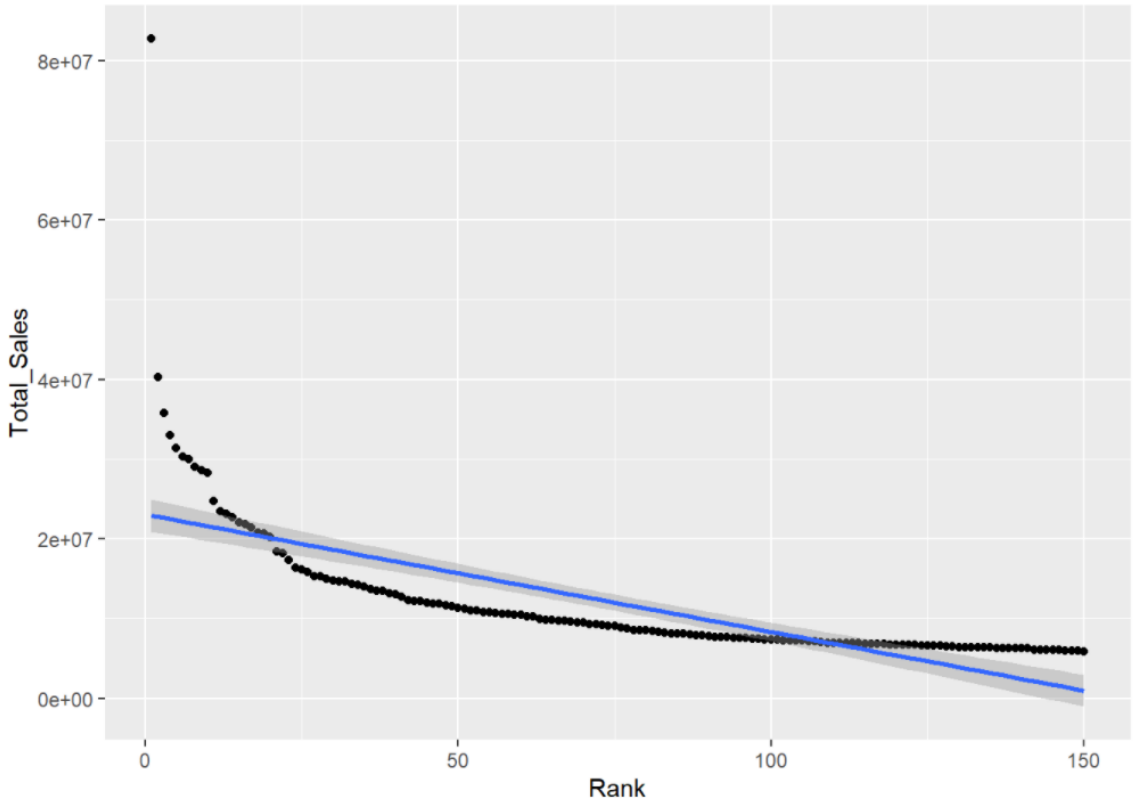
23.


24.

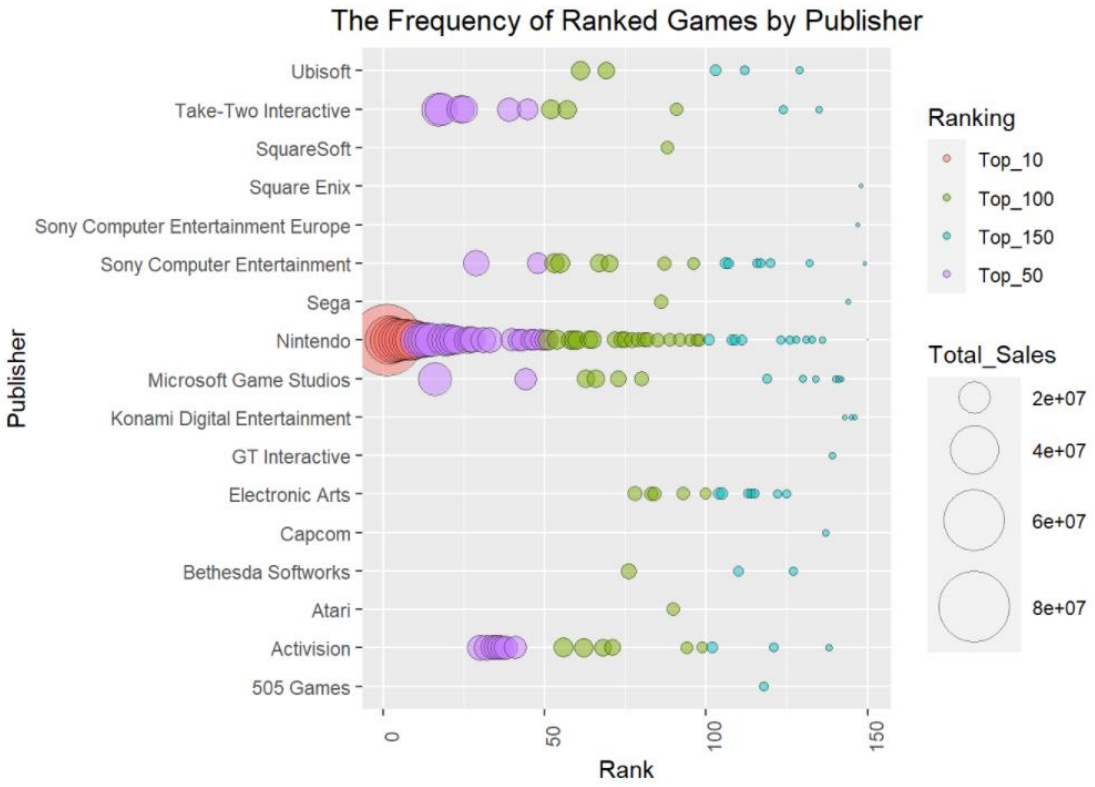

Distribution of Total Sales
Amount of Games
80
60
40
20
0
0.0e+00
2.5e+07
5.0e+07
7.5e+07
Total Sales


1.0e+07
7.5e+06
5.0e+06
2.5e+06
0.0e+00
Japan_Sales
Top_10
Top_100
Top_150
Top_50
Ranking

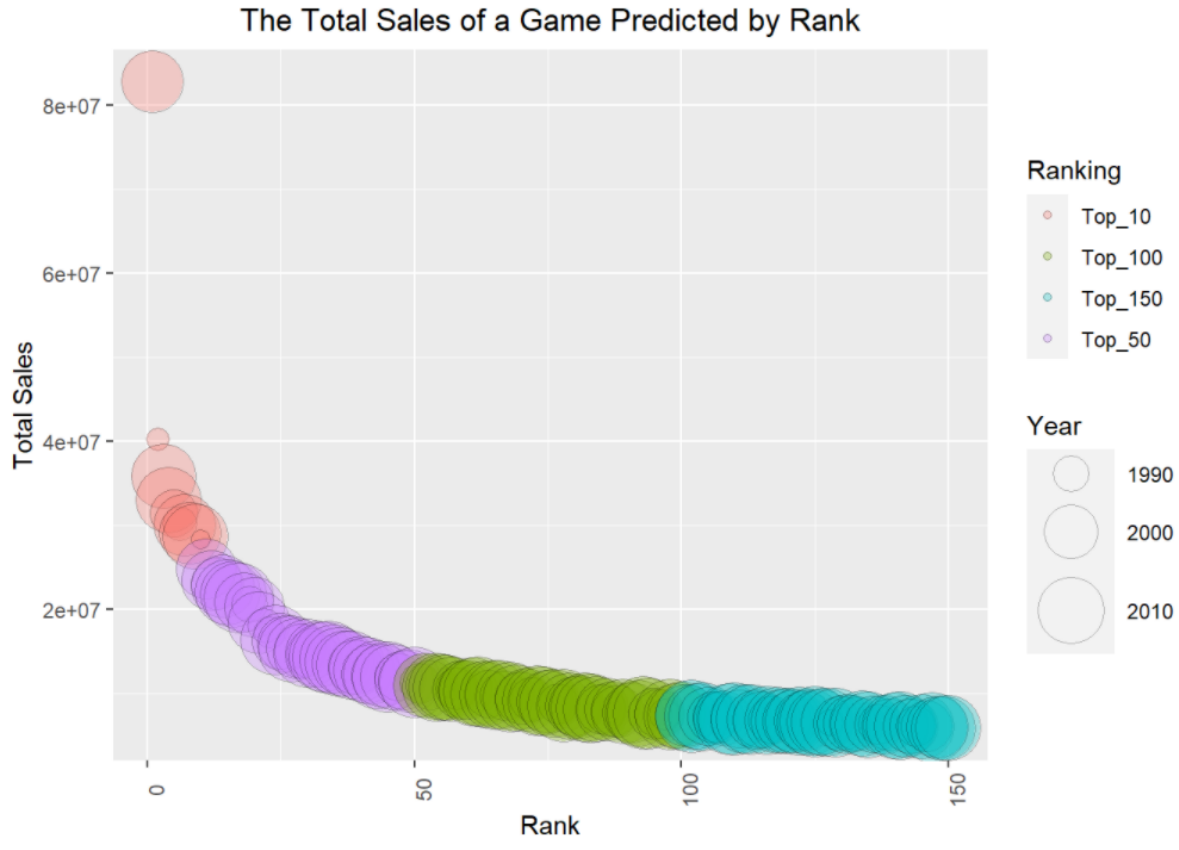
25.


8e+07
6e+07
4e+07
2e+07
0e+00
Total_Sales
0
50
100
150
Rank

26.


The Frequency of Ranked Games by Publisher
Ubisoft
Take-Two Interactive
SquareSoft
Square Enix
Sony Computer Entertainment Europe
Sony Computer Entertainment
Sega
Nintendo
Microsoft Game Studios
Konami Digital Entertainment
GT Interactive
Electronic Arts
Capcom
Bethesda Softworks
Atari
Activision
505 Games
Publisher
0
50
100
150
Rank
Ranking
Top_10
Top_100
Top_150
Top_50
Total_Sales
2e+07
4e+07
6e+07
8e+07

27.


The Total Sales of a Game Predicted by Rank
8e+07
6e+07
4e+07
2e+07
Total Sales
0
50
100
150
Rank
Ranking
Top_10
Top_100
Top_150
Top_50
Year
1990
2000
2010

28.