

Video Game Popularity Data Project

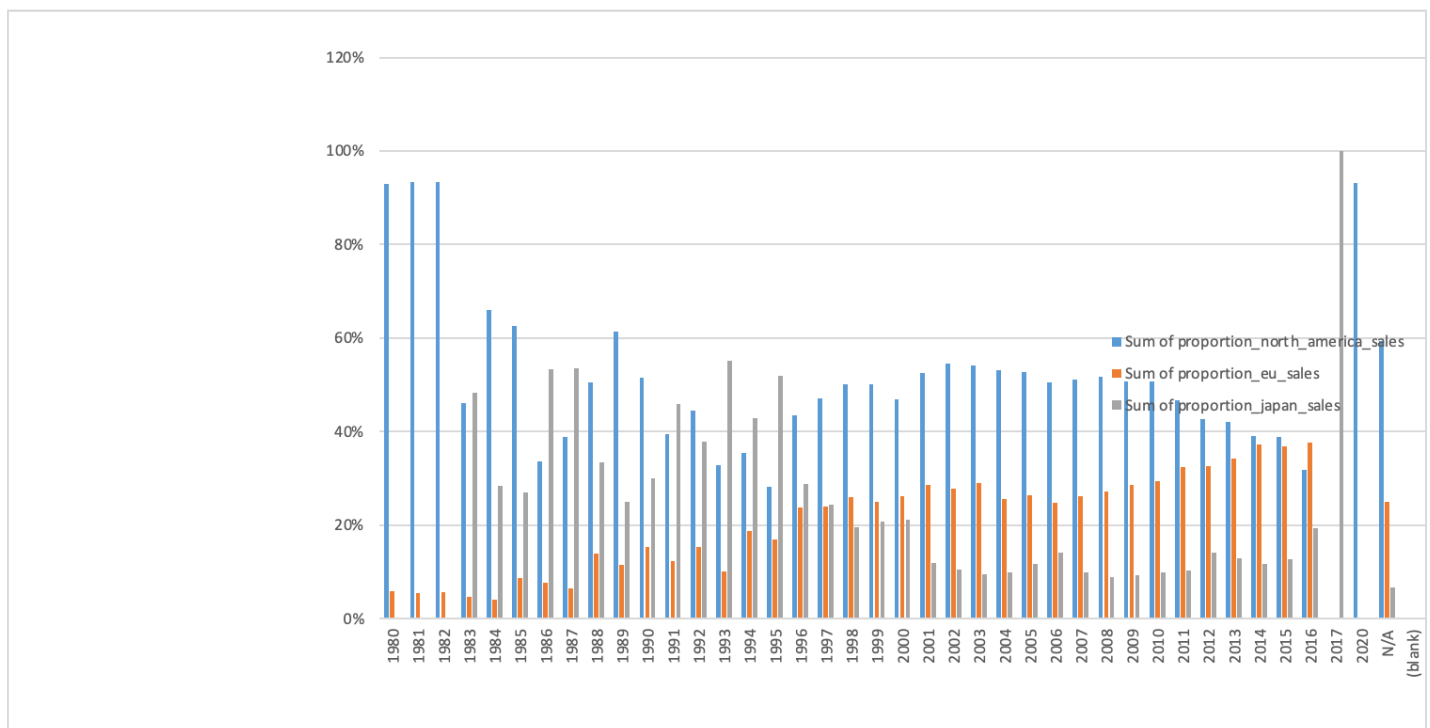
This analysis started with data cleaning step in which duplicates and some rows that didn't have any info for any of the variables were removed, the average for the variable "Other_Sales" was imputed (576 replacements have been made) and some of the columns were adjusted to fit the text and data.

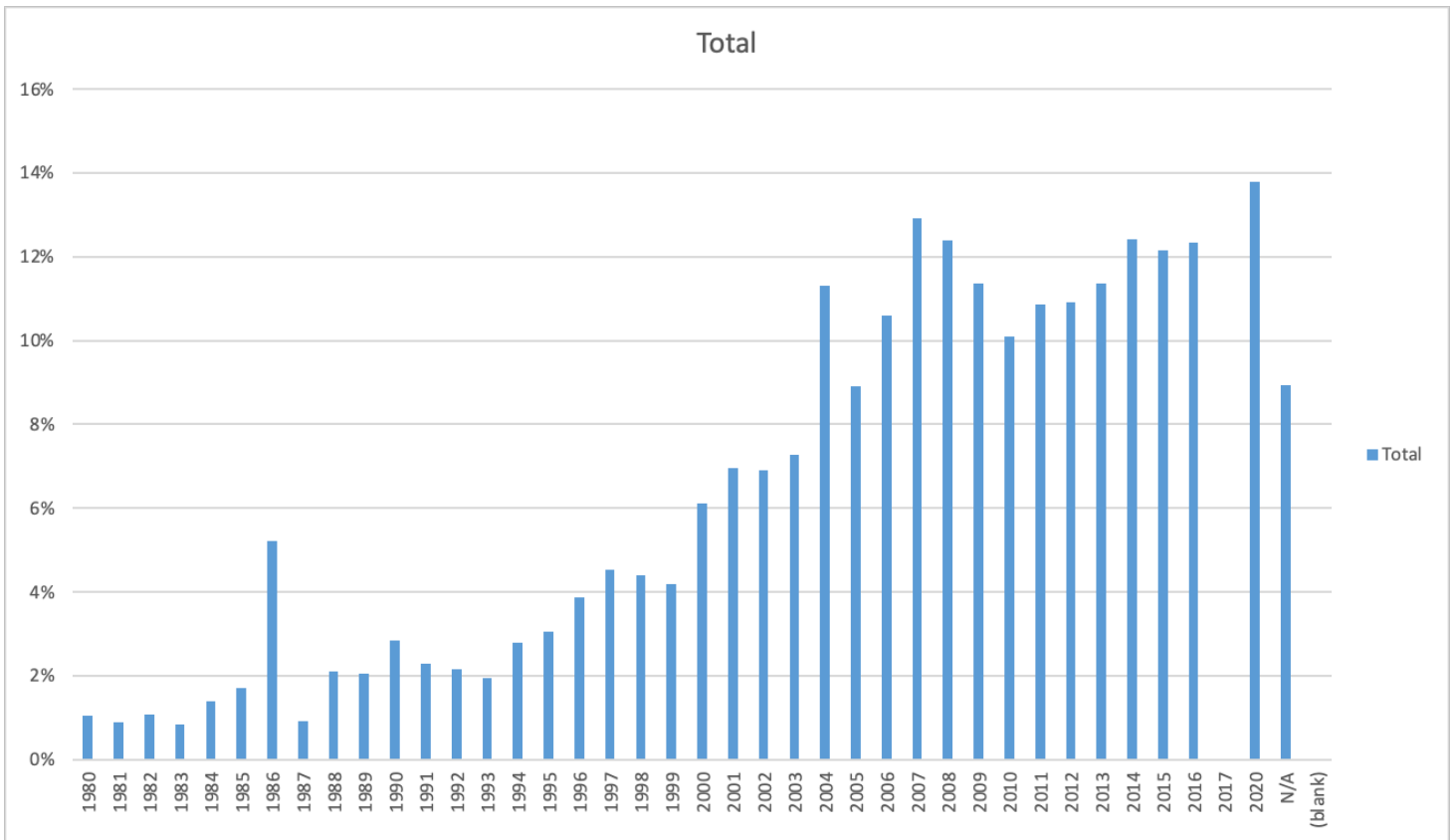
After that, the EDA was conducted and the mean, median and mode were calculated for all the five sales variables and outliers identified. Pivot Tables were run to analyze the sales in North America, Europe, Japan and Other_Sales variables.

Although North America sold the most video games, the distribution of the units of games sold in all the regions by years showed that there was a downward trend in sales in North America.

Further, some N/As for the Year variable for each region were identified and kept just to notice what's the percentage to the global sales. The highest proportion for the N/A was occupied by North America. This aligned with the expectation that games will continue to do well in sales in the next years despite the downward trend. However, a further analysis would be necessary for the Other_Sales variable since its upward trend reveals that a new market can emerge for those unspecified regions.

Then, four other variables for each region were created to see how much is the proportion of sales to the global sales. Also a separate graphic for Sum_of_Proportion_Other_Sales was run to get a closer look at the ascending trend sales for this variable (second graph below). For the Final Project the first graphic below was re-run for all four sales variables and the year 2017, 2020 and the blanks have been removed.





I chose to display the Sales by Regions over the Years visualization as the first one in my Final Presentation because it will inform the expectation that the video games sales for various geographic regions stayed the same over time. This will allow the audience to see the trends for all the four regions right away. The visualization is easy to follow and the graphic looks cleaner and more accessible.