**Term Project Write-Up**

Dean Winters

DSC530 Exploratory Data Analysis

Dr. Shankar Parajulee

November 20, 2021

Outcome of Your EDA:

The most jarring piece of information I’ve found from this analysis is how many titles are developed by various companies that are relatively unsuccessful when compared to major well-known titles. At least 95% of titles don’t sell more than 500,000 units. Companies like Nintendo and Electronic Arts produce so many successful games, but these companies also have a much higher cost of running business when compared to small indie-developers. Selling 500,000 units for a one-person-show would be an enormous achievement, but for a AAA development company that would likely be a major failure.

I didn’t expect North America to have so many sales compared to the other regions, but it’s possible that the sales of the past 5 years have differed from what’s taken place in the data observed. Nintendo really knocked it out of the park with Wii Sports.

What do you feel was missed during the analysis?

I feel like there could be a lot more to explore if I had more quantitative data to compare this data with. If I could compare the total populations for different regions and the buying rates, then I’d be able to determine the percentage of buyers in a population. Maybe North America does sell better than the other regions, or maybe the other regions just have less people in general.

Were there any variables you felt could have helped in the analysis?

Yes, region population would have helped. Demographics of the purchasers would be tremendously helpful. The total number of platforms sold would have been useful. Production time for each title would be interesting to look at. That could help answer the question of “Does spending more time developing the product lead to a more successful product?”

Were there any assumptions you made you felt were incorrect?

Yes, I honestly did not expect Nintendo to have so many titles that are in the top 50 best selling games from 2016 and prior. The top 15 best selling titles are all Nintendo, then various publishers take credit for a few of the top performing titles, but Nintendo has so many high-performing titles.

What challenges did you face, what did you not fully understand?

I faced challenges with merging my data correctly and making it usable for analysis. I need to practice more to improve this skill since it’s so crucial to data science. I also did not fully understand hypothesis testing in Python. I may need different data in order to perform the tests better but conducting a hypothesis test in Python is very different from conducting one by hand. Overall, the project forced me to use multiple pieces covered in the class and it was interesting to see how they all interact with data.