Faizan Virani

Professor Binowski

Data Science Programming

7 May 2024

I performed data analysis on video game sales data, which included information on games’ sales, release year, genre, platform, and rank. The questions I was looking to answer were as follows:

1. What video game genre is the most popular?
2. What video game platform had the most games?
3. What is the average percent of total sales which are from the US?
4. What year had the most popular games released?
5. How do Japanese sales compare to European sales on average?

I used a jupyter notebook, the pandas library, the seaborn library, the numpy library, and the matplotlib library to create graphs and do calculations.

I discovered that the most popular platform for games from 2000 - 2020 was the DS. This was interesting to me because I expected newer consoles to eclipse old ones in popularity, but in hindsight, the extra time that older platforms had over newer ones can explain the difference. It can also be that less games are being developed now due to higher standards and longer production times.

If I were presenting this to a client, I would recommend creating an action-based game because those seem to perform the best. I would like to continue to analyze video game sales with the same data points in the future. I would also like to analyze the attributes of games which become the most popular, to find out what makes the ideal game.