

Info 474 A2 EDA

Yichao Wang

4/26/2019

Data

This data set contains a list of video games with sales greater than 100,000 copies along with critic and user ratings (from Kaggle.com). The fields include name, platform, year of release, publisher, sales in different region, rating scores, and etc.

EDA

Initial Question

How many video games are produced each year from 2000 to 2016?

From Figure 1, we can see there are only 354 video games produced in 2000. The number has reached to 1550 in 2009 but it decreased to 618 in 2016. It seems there is a decline in video game industry.

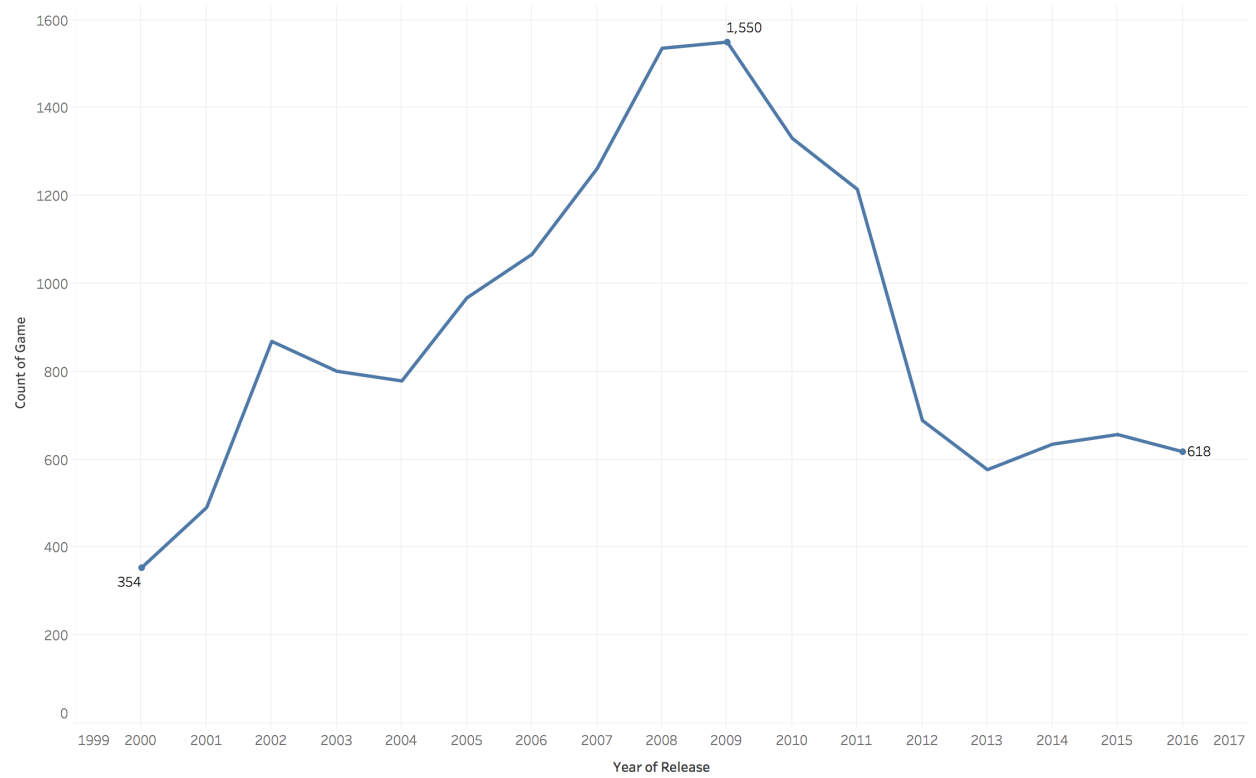


Figure 1: Count of game produced each year

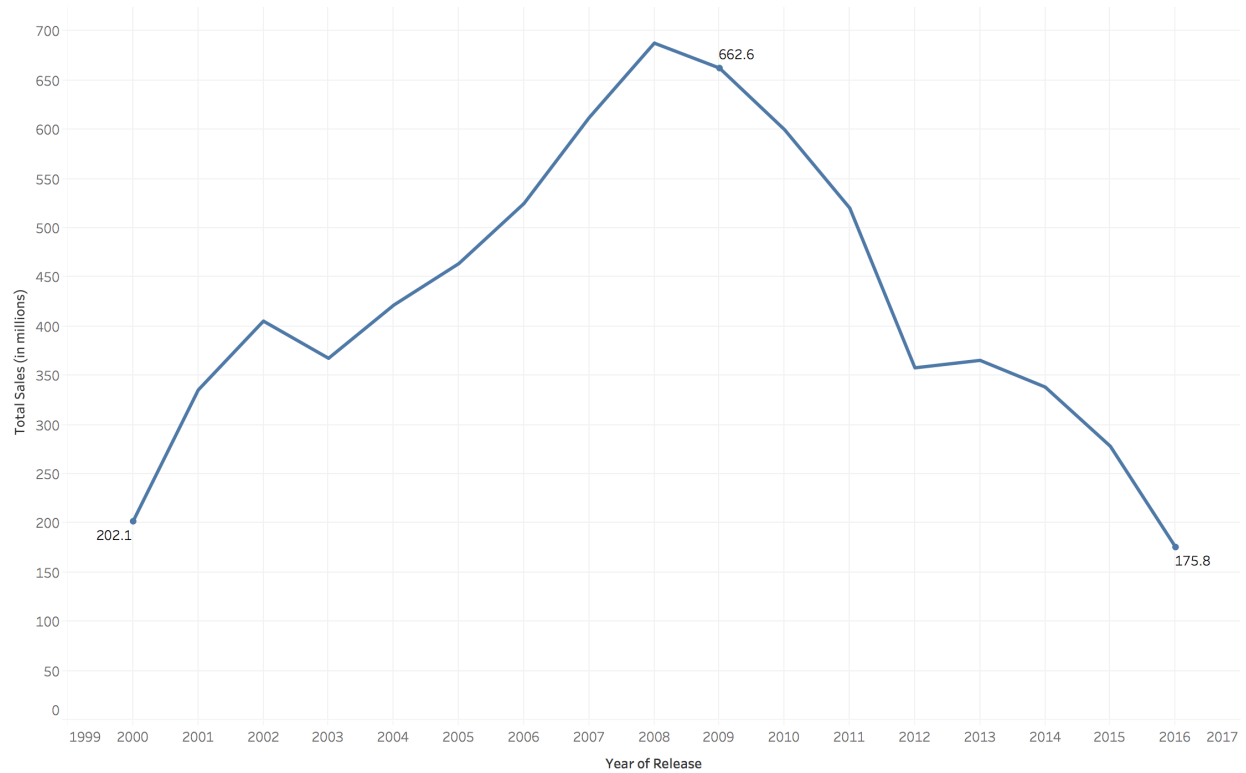


Figure 2: Total sales each year

Question 2

Are the total sales also decreased in 2016?

I sum all the sales from different regions to get total sales over the world. Similar as video game production, the total sales increased as production increased and decreased as production decreased (Figure 2).

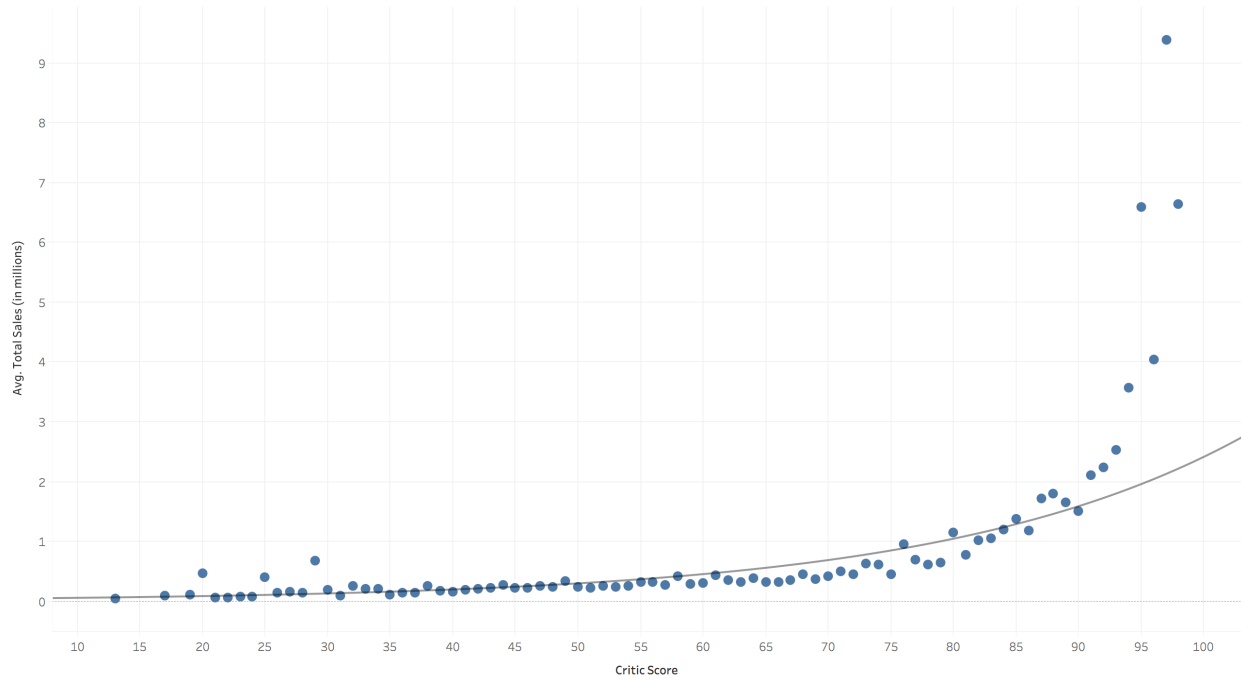


Figure 3: Critic rating score VS average total sales

Question 3

Is there a correlation between average total sales and rating scores?

After excluding some null values, from Figure 3, we can see there is an exponential correlation between critic rating score and average total sales.

From Figure 4, we can see there is an positive linear correlation between critic rating score and average total sales.

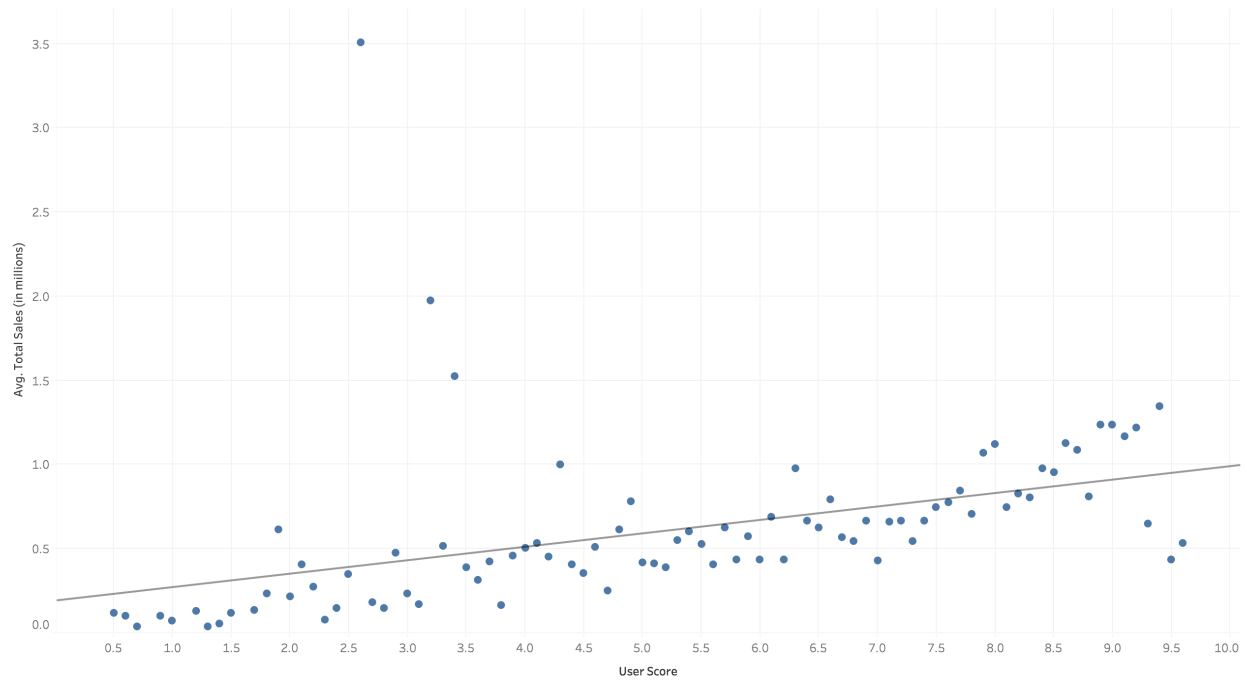


Figure 4: User rating score VS average total sales

Question 4

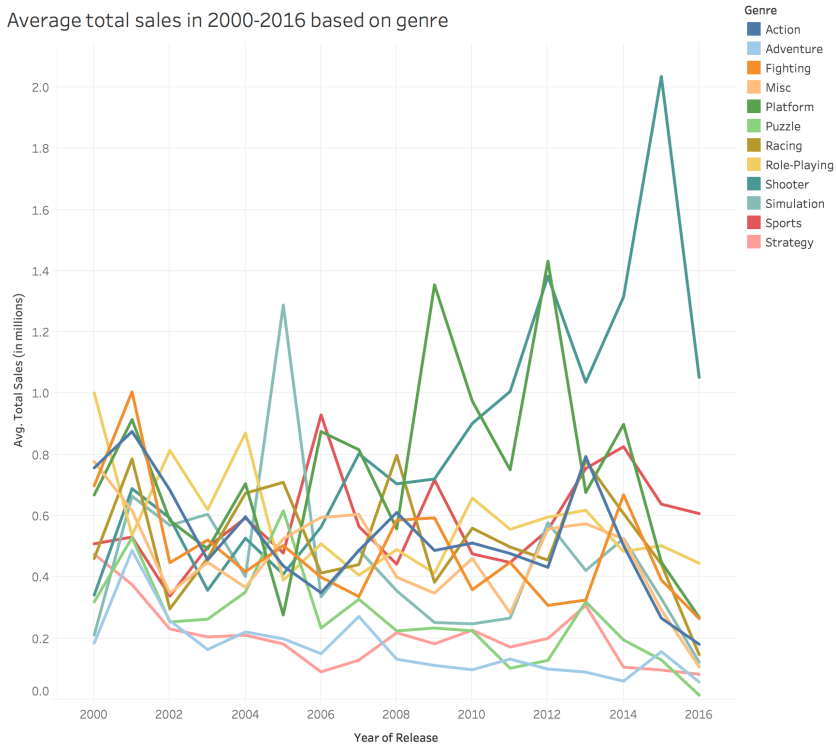
How does genre, platform, and the ESRB ratings affect average total sales in years?

From Figure 5, we can see shooter genre start dominate the market after 2008. In 2016, the top 3 (most average total sales) genre are shooter, sports, and role-playing.

From Figure 6, there are many platforms have been eliminated. Some of the platforms have been replaced by their new generations (e.g. PS4 replaced PS3). In 2016, the top 4 platform are Wii, PS4, 3DS, and XOne. The console game has dominate the market.

From Figure 7, we can see M rating game has dominate the market after 2008. In 2016, it is M rating game has the most average sales.

Average total sales in 2000-2016 based on genre



Average total sales in 2016 by

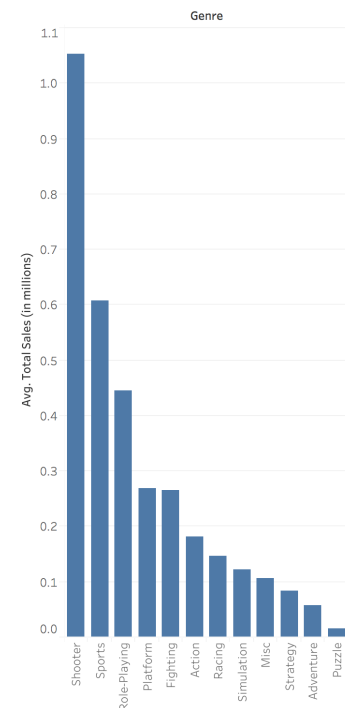
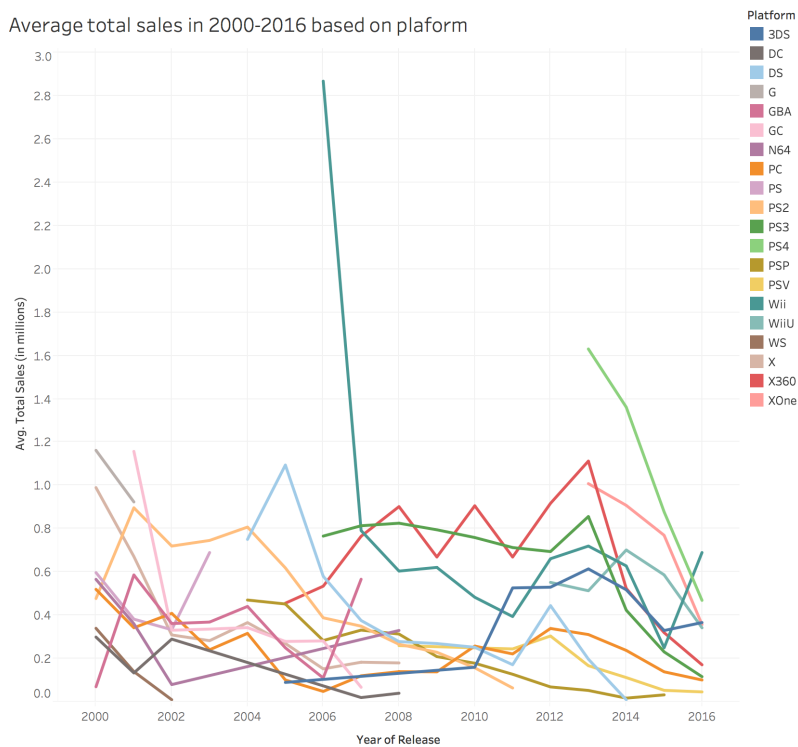


Figure 5: Genre VS average total sales in 2016

Average total sales in 2000-2016 based on platform



Average total sales in 2016 by platform

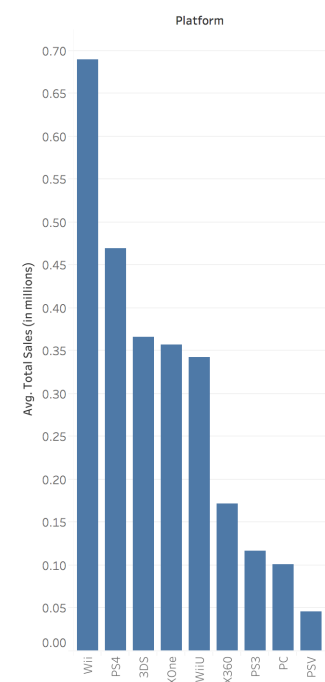
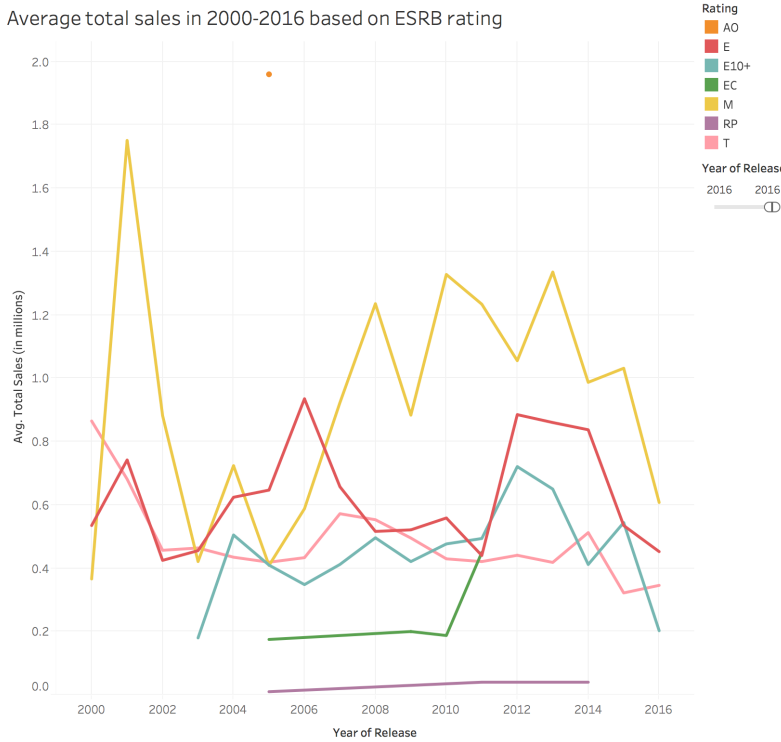


Figure 6: Platform VS average total sales in 2016

Average total sales in 2000-2016 based on ESRB rating



Average total sales in 2016 by ESRB rating

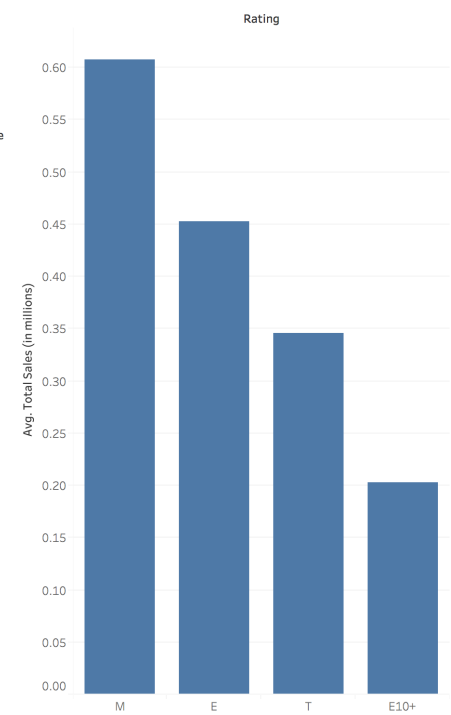


Figure 7: The ESRB rating VS average total sales in 2016

Final question

What kind of games can get most sales in 2016?

Since the game industry faced a decline after 2009, game company may want to know what kind of games can make more revenue now. After filtering the data based on previous question's result, I plot the chart for average sales in 2016 based on genre, platform, and rating. From Figure 8, we can see sports (with E rating) and shooter (with M and T rating) games on PS4 has the highest average total sales in 2016. Also, the game may need a high rating score from critic and user.

Average total sales in 2016

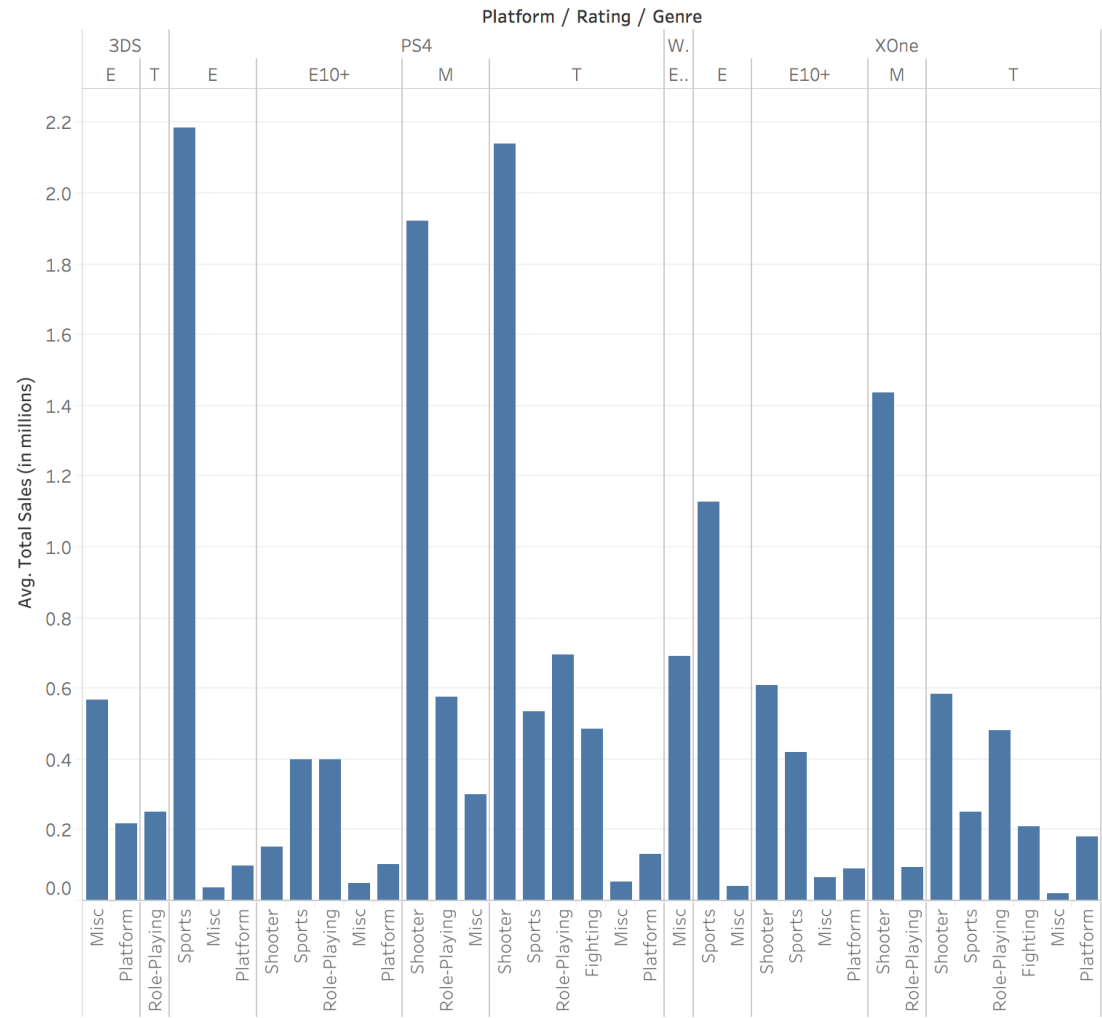


Figure 8: Average total sales in 2016 based on genre, platform, and rating