

T5 Boot camp

Data science

# Project proposal

## Games sales dataset

mohammad alajmi

Instructor:

Mejdal Al-Qahtani

## abstract:

A famous video games company has decided to make a new game in north America and Europe regions, as the client asked to study the last years data's as they want to make predictions about the market; the client specified a genre in general that is going to be worked on, the client game is action genre, the client does not mind if there are better genre's that is makes more profit, also would accept any suggestion about a better region to publish the game.

The client asked about what the bestselling regions in case they have plans to extend their sales, also what is the best platform for action games, the client asked what the best three publishers of the action games are.

As we about to answer these questions:

**1-** what are the bestselling regions in case they have plans to extend their sales?

And to choose the bestselling regions for action games.

**2-** what is the best platform for action games based on records of the last 5 years?

To help the client to recognize the propriety for action games.

**3-** the client asked what are the best three publishers of the action games?

To help the client choose the best publishers for the game genre.

**4-**Is there a better genre in terms of profit in terms of the last year?

To help the client choose the most profit game genre.

The main goal of this project:

The main goal of this project is to study the video games market any analyze it to help the client to have a good understanding about the market and the sales of the video games to help the company establish their own game.

## The Data set:

The used data set in this project is "Video\_Games\_Sales\_as\_at\_22\_Dec\_2016", it is about the sales of games in the latest years in the different regions all over the world, the dataset covers all the sales before 2016The data can be found on Kaggle website, the data has about 15 thousand row and about 11 columns. The columns describe a games sale within every region and the expected and actual sales.

The columns also describe the games itself such as genre, platform and the date when it released.

## The used tools:

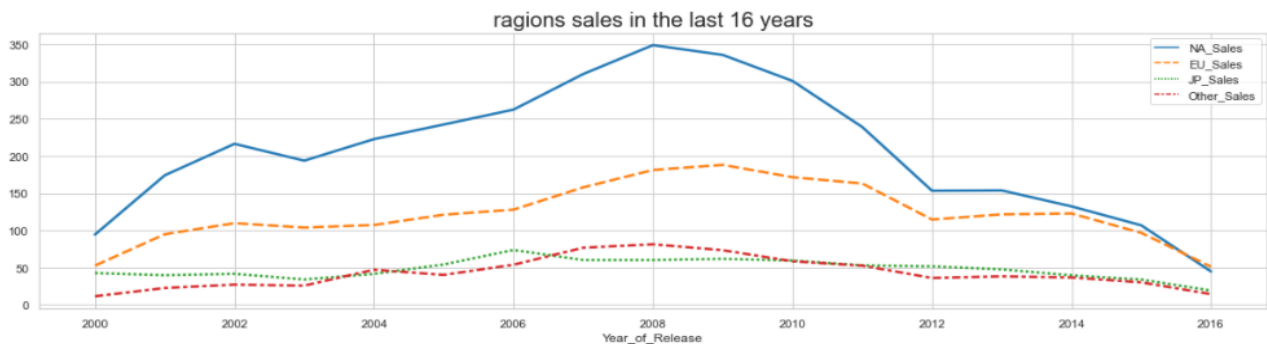
I have used different tools to achieve this goal such as:

- 1- Python programming language**
- 2- SQL programming language(optional)**
- 3- SQLALCHEMY library(optional)**

- 4- Numby library
- 5- Pandas' library
- 6- Matplotlib library

## Communication:

In figure 1 the used plot is called lineplot to represent the sales of every region in the last 16 years based on feature called "Year\_of\_Release".



the figure shows the Sales of each region as it follows:

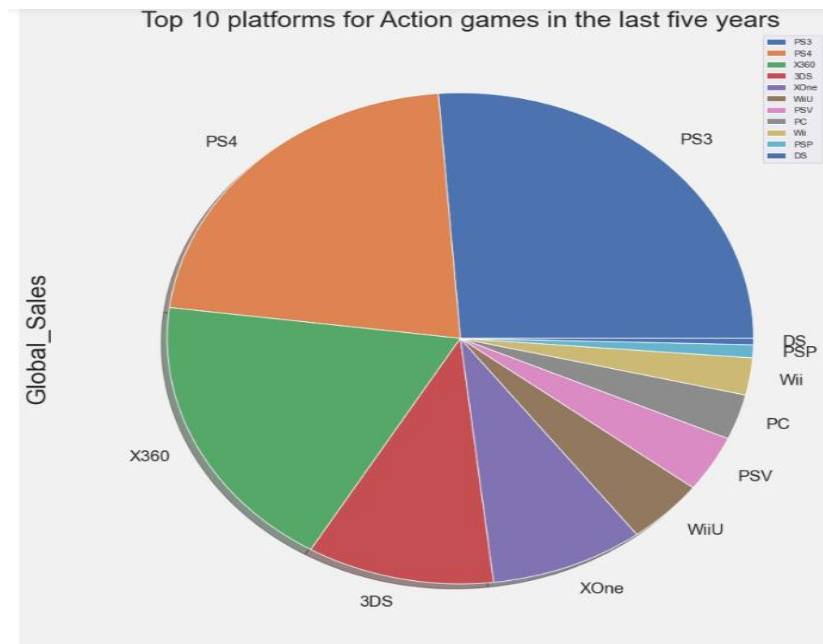
- 1- the North America in "Blue" line
- 2- the Europe in "orange" dashes
- 3- the Japan in "green" dots
- 4- the other sales in "red" dots and dashes

the result indicates that North America region is the most profitable region specially between the year 2006 to 2010 as it starts to downgrade in the last few years

Follows it in the second place "Europe" as it is the second most region in term of sales,"Europe" region has not changed much compared with North America region

"Japan" region takes the third place in Sales, in the last 4 years the region decreased significantly as other sales became above "Japan" region in terms of sales

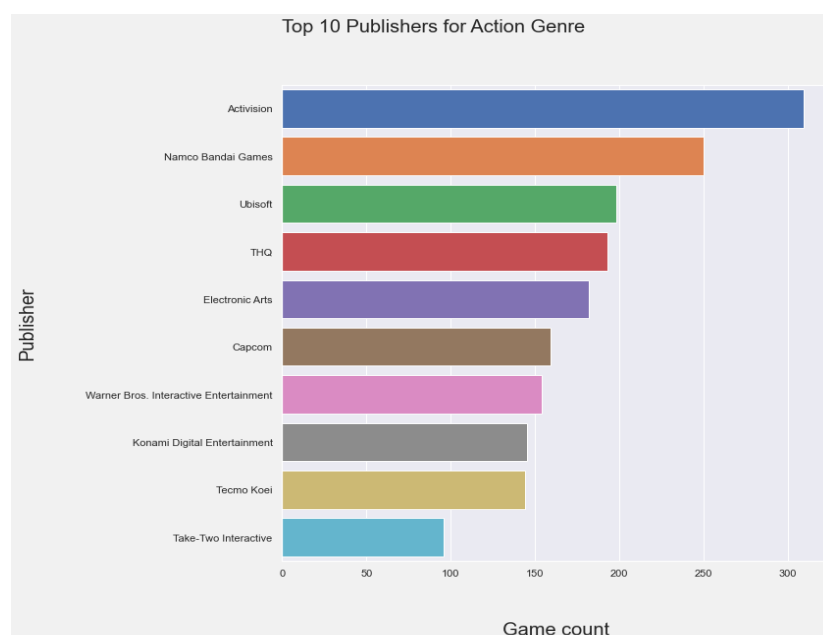
In figure 2 the used plot is called pie chart plot to represent the top 10 platforms for Action games in the last five years based on feature called "Global Sales".



the figure shows the top three platforms in term of profit as it follows:

- 1- the PS3 in "Blue" part was the most profitable platforms for Action Games
- 2- the PS4 in "orange" part was the second most profitable platforms for Action Games
- 3- the xbox360 in "green" part was the third most profitable platforms for Action Games

In figure 3 the used plot is called BarH to represent the top 10 publishers for Action games based on feature called "Publisher".



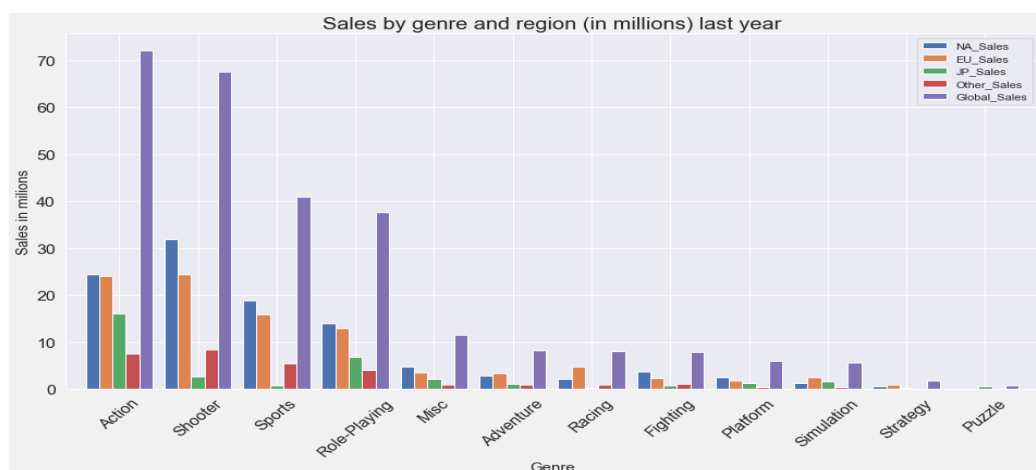
the figure shows the top three publishers in term of the number of publishes as it follows:

1- the Activision in "Blue" bar was the most number of publishes for Action Games

2- the Bandai Namco games in "orange" bar was the second most number of publishes for Action Games

3- the Ubisoft in "green" part was the third most number of publishes for Action Games

In figure 4 the used plot is called Bar chart to represent the sales by genre and region for Action games on last year based on feature called "Global\_Sales".



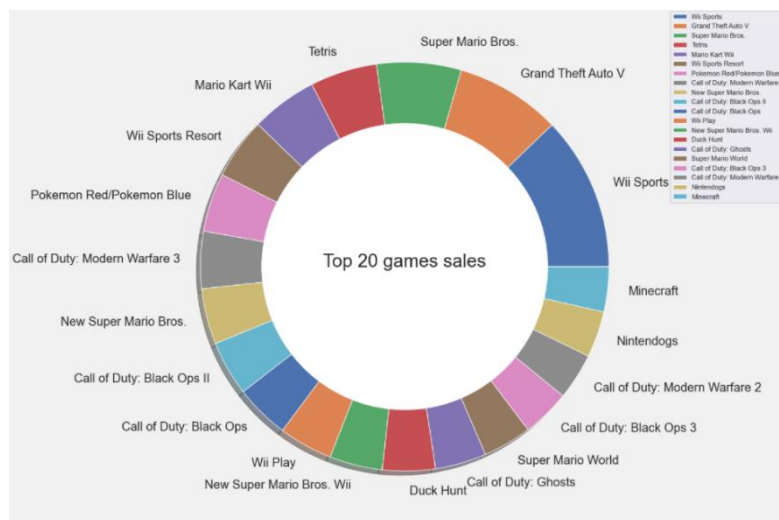
the figure shows the top three publishers in term of the number of publishes as it follows:

the Action Genre has the most sales followed by shooter games and in third place came sports games, the figure shows that North America and Europe has the most sales last year.

The demands on puzzle games and strategy games were the least.

The sales of shooter genre in Japan region was low unlike the other region sales which has an increase in sales which means that the shooter games in other regions has more popularity than Japan market.

In figure 5 the used plot is called circle pie (donuts) chart to represent the top 20 video games sales based on feature called "Global\_Sales".



the figure shows the top three games sales in term of the number of sales as it follows:

- 1- the Wii sports game in "Blue" part was the most game in term of sales.
- 2- the Grand theft auto game in "orange" part was the second most game in term of sales
- 3- the tetris in "green" part was the third most game in term of sales