

EZ gaming Publisher Market Entry

Names:

Mohamed Aziz Srairi & Rayen Fares

Project Title:

Video Games Industry Analysis

Business Problem:

EZ gaming is a new **video game** publisher that wants to understand the industry before entering the market and producing Video Games. We want to analyze the market, who is the market leader? What are the global sales? Sales per region? What do gamers prefer the most? On which console should we adopt our game?

Project Description:

This project will involve collecting and analyzing video game data from all gaming publishers. The data will be integrated & cleaned using Talend, transformed, and visualized using Power Bi. Reports, dashboards, and visualizations will be created to help EZ gaming management understand the market. EZ gaming will be able to use the insights gained from the analysis to choose the best genre of game, depending on the region.

Datasets Used:

- Video Games names dataset (Kaggle *csv file*)
- Video Games sales dataset (Kaggle *.txt file*)
- Console sales (kaggle *csv file*)
- Video Games genre preferences (Web scrapping from Wikipedia *csv file*)

Project Summary:

1. Business Understanding
2. Data Collection
3. Data Integration & cleaning (Talend)
4. Data Visualisation (Power BI)
5. Insights & Analysis
6. Conclusion

Project Details:

1. Business Understanding

The video game industry is a rapidly growing and highly profitable business. The industry is driven by a combination of factors, including advancements in technology, the increasing popularity of gaming as a form of entertainment. The largest segment of the industry is PC and consoles. The major players in the console market are Sony, Microsoft, and Nintendo. As the industry continues to evolve and expand, it is expected to see continued growth in the coming years.

2. Data Collection

After understanding the video game industry, the first step is to start by collecting different datasets about our needs (Video games sales per region, Video Games Genres Preferences, Consoles Sales ...) So, we started by using Google Dataset Search to look for our datasets using the exact keywords, and then we looked for the best dataset that fit our needs in different sources (Kaggle, Wikipedia, Data.world, Statistica)

- a. What did we find?
 - i. Video Games names dataset (in Kaggle as a **csv file**)
 - ii. Video Games sales dataset (in Kaggle as **.txt file**)
 - iii. Console type, sales & Manufacturer (in kaggle csv file transformed into **MySQL database**)
 - iv. Video Games genre preferences (**Web scraping** from Wikipedia using python code and transformed into a **csv file**)

3. Data Integration & cleaning (Talend)

- a. We Started first by cleaning the web scraped file as it contains different names that belong to the same genre using the average.
- b. We uploaded the different files into Talend: database connection and the remaining inputDelimitedFiles.

c. Database connection

Connexion à la base de données

Mettre à jour la connexion à une base de données - Etape 2/2

Vous devez cliquer sur le bouton Vérifier afin de vérifier les paramètres de la base de données.

Type de BD: MySQL

Version de la base de données: MySQL 5

Chemin de caractères de connexion: jdbc:mysql://localhost:3306/console?useUnicode=true

Identifiant: root

Mot de passe: root

Serveur: localhost

Port: 3306

Base de données: console

Paramètres supplémentaires: useUnicode=true

Tester la connexion

Exporter en tant que contexte Revenir au contexte précédent

Comment installer un pilote

< Back Next > Finish Cancel

- d. **JOB1:** Join csv and text file
 - i. *(Check Photo 2 join_csv_and_txt)*
- e. **JOB2:** Join csv and web scraping file
 - i. *(Check Photo3 join_csv_web_scrapping)*
- f. **JOB3:** Join the database
 - i. *(Check join_database)*
- g. **JOB4:** Tsortrow (sort file by name)
 - i. *(Check Photo 5 sort_by_name)*
- h. **JOB5:** Remove empty rows
 - i. *(Check Photo 6 remove_rows_with_empty_sales)*

4. Data Visualisation (Power BI) and Analysis

- a. Once finished, we imported the final output file (as csv) into our POWER BI program.
- b. We changed the names of some columns to make them suitable
 - i. Name -> Name of video games
 - ii. Year -> Release year of video game
 - iii. women percentage -> women percentage per genre
- c. We added a **measure** that sums the sales of different regions(NA,EU,JP and Others) under the name of **Global Sales**.
- d. We started visualizing our data in different visuals (Pie chart, line chart, 100% stacked bar chart, Table, Treemap, Donut Chart, cards, Clustered column chart ...)

e. Insights About Video games, publishers

i. Video Game Sales per region

- 1. NA has the highest Sales for all the video games
- 2. Wii sports is the most sold game with **\$82,74M** in global sales
- 3. GTA V is the most sold game in the last decade (2006-2016) with **\$55,92**

ii. Video Games and their global sales per year

- 1. 2002, 2008, 2009, and 2015 represent the peaks of VG sales

2. Video games and their global sales are positively correlated as global sales increase when the number of video games increases
3. After conducting research we found that in 2000 ps2 console was released and it resulted in a peak of sales in 2002, the same in 2006 when the ps3 was released, 2 years after that the number of sales increased again, the same thing applies in 2013 when ps4 was released and yield to a peak in 2015.

iii. Video Game Sales per Genre

1. Most genre preferred by the gamers community was Action (it represents 20% of the total sales)
2. NA dominates the sales in most of the genre types, except for Role playing where Japan has the highest sales.

iv. Video Game Genre Distribution

v. Video Games Preferences per gender

1. The most genre preferred by men is Action (98%)
2. The most genre preferred by women is Simulation(69%)

vi. Number of video Games produced by publisher

1. EA is the highest publisher producing video games

vii. Most produced genre by publisher

1. Most genre produced by publishers is Action

viii. Global Sales per publisher

1. Although EA sports has the highest number of video games produced, Nintendo has the highest global sales, we can conclude that Nintendo games are more expensive than EA sports games.
2. EA is the second in the global sales

f. Insights About console sales

i. percentage of console sales

1. PC is the highest console sold with **39.3%** of the total console sales, followed by PS2 with **13.72%**

ii. Sum of console sales per manufacturer

1. The top 4 Manufacturers of consoles are:
 - a. Computer companies

- b. Sony
- c. Nintendo
- d. Microsoft

iii. Console Type Distribution

1. 68% of consoles are HOME, which means can be played through controllers (PCs, Xbox, PS2 ...) and the rest can be played handheld like (Nintendo DS, PSP)

5. Conclusion

- **Nintendo and EA** are the market leaders in the production and sales of video games.
- Most Video games genre preferred by gamers is **Action**.
- **Women** prefer Simulation games
- **Japan** prefers Role-playing video Games.
- Most used console is **PC (Personal Computer)**
- The main segment interested in video Games is the North American **(NA)** people followed by Europeans **(EU)**
- The release of a new console is followed by an increase in video game sales.

6. Suggestions:

- If we are going to produce a video game, we should adapt it for Personal computers and may adapt it for other consoles, especially for the newly released consoles.
- If we choose a role-playing game genre, it should be marketed mainly in Japan.
- If we want to attract women to our segment (target audience), we can choose a Simulation game as it's the most preferred game genre for women
- The safest and best game genre choice is Action
- Our games should be mainly marketed in North America
- The optimal time of game release is when a new console is released and of course, the game should be adapted to it!