# What do you like to do in your Free time?

# What do you like to do in your Free time?

Exercising, reading, meeting friends, listening to music, watching movies, cooking, studying, sleeping, going on trips, visiting restaurants, and **playing games** 



Dec 13th, 2022



START



1

#### **Data & Business Questions**

Data Cleaning in Excel,
Rationale of Data and
Business Questions

#### ERD & Data Upload & Cleaning

Entity Relationship Diagram,
Data Upload in MySQL, bit of a
data cleaning



#### Assign 'id', PK and FK in MySQL Answering Q by Queries

Making PK & FK relationship in MySQL Answering business questions by MySQL queries



### Game Sales Dataset

First look of game dataset, **16,720** entries and **15** columns Here, we consider facts and dimensions to convert and manage data into normal form.

4	Α	В	С	D	E	F	G	Н	1	J	K	L	М	N	0	Р
1	Name	Platform	Year_	Genre	Publisher	NA_Sales	EU_	JP_	Other	Global	Critic_Scor	Critic_Cou	User_Scor	User_Cour	Developer	Rating
2	Wii Sports	Wii	2006	Sports	Nintendo	41.36	29	4	8.45	82.53	76	51	8	322	Nintendo	E
3	Super Mario	NES	1985	Platform	Nintendo	29.08	4	7	0.77	40.24	NULL	NULL	NULL	NULL	NULL	NULL
4	Mario Kart W	Wii	2008	Racing	Nintendo	15.68	13	4	3.29	35.52	82	73	8.3	709	Nintendo	E
5	Wii Sports Re	Wii	2009	Sports	Nintendo	15.61	11	3	2.95	32.77	80	73	8	192	Nintendo	E
6	Pokemon Red	GB	1996	Role-Playir	Nintendo	11.27	9	10	1	31.37	NULL	NULL	NULL	NULL	NULL	NULL
7	Tetris	GB	1989	Puzzle	Nintendo	23.2	2	4	0.58	30.26	NULL	NULL	NULL	NULL	NULL	NULL
8	New Super N	DS	2006	Platform	Nintendo	11.28	9	7	2.88	29.8	89	65	8.5	431	Nintendo	E
9	Wii Play	Wii	2006	Misc	Nintendo	13.96	9	3	2.84	28.92	58	41	6.6	129	Nintendo	E
10	New Super N	Wii	2009	Platform	Nintendo	14.44	7	5	2.24	28.32	87	80	8.4	594	Nintendo	E
11	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	1	0	0.47	28.31	NULL	NULL	NULL	NULL	NULL	NULL
12	Nintendogs	DS	2005	Simulation	Nintendo	9.05	11	2	2.74	24.67	NULL	NULL	NULL	NULL	NULL	NULL
13	Mario Kart D	DS	2005	Racing	Nintendo	9.71	7	4	1.9	23.21	91	64	8.6	464	Nintendo	E

### Game Sales Dataset

the important information

Game info Sales critic & user score Developer

4	Α	В	С	D	Е	F	G	Н	1	J	K		L		М	N	0	Р
1	Name	Platform	Year_	Genre	Publisher	NA_Sales	EU_	JP_	Other	Global_	Critic_	Scor	Critic_Co	ou U	Jser_Scor	User_Cou	r Developer	Rating
2	Wii Sports	Wii	2006	Sports	Nintendo	41.36	29	4	8.45	82.53		76	5	51	8	322	Nintendo	E
3	Super Mario	NES	1985	Platform	Nintendo	29.08	4	7	0.77	40.24	NULL		NULL	N	IULL	NULL	NULL	NULL
4	Mario Kart W	Wii	2008	Racing	Nintendo	15.68	13	4	3.29	35.52		82	7	73	8.3	709	Nintendo	E
5	Wii Sports Re	Wii	2009	Sports	Nintendo	15.61	11	3	2.95	32.77		80	7	73	8	192	Nintendo	E
6	Pokemon Red	GB	1996	Role-Playir	Nintendo	11.27	9	10	1	31.37	NULL		NULL	N	IULL	NULL	NULL	NULL
7	Tetris	GB	1989	Puzzle	Nintendo	23.2	2	4	0.58	30.26	NULL		NULL	N	IULL	NULL	NULL	NULL
8	New Super N	DS	2006	Platform	Nintendo	11.28	9	7	2.88	29.8		89	(	55	8.5	431	Nintendo	E
9	Wii Play	Wii	2006	Misc	Nintendo	13.96	9	3	2.84	28.92		58	4	11	6.6	129	Nintendo	E
10	New Super N	Wii	2009	Platform	Nintendo	14.44	7	5	2.24	28.32		87	8	30	8.4	594	Nintendo	E
11	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	1	0	0.47	28.31	NULL		NULL	N	IULL	NULL	NULL	NULL
12	Nintendogs	DS	2005	Simulation	Nintendo	9.05	11	2	2.74	24.67	NULL		NULL	N	IULL	NULL	NULL	NULL
13	Mario Kart D	DS	2005	Racing	Nintendo	9.71	7	4	1.9	23.21		91	(	54	8.6	464	Nintendo	E

Ouestion One

Which gaming platform has the largest number of games available?

Question Two

Which gaming platform has the highest total sales?

Question Three

What are the average sales figures for each gaming platform?

Question Four

Which game genre has the highest total sales average?

Question Five

Comparison of the top 10 games by critic score and the top 10 games by user score, along with their sales figures.



# 2. ERD, Data Cleaning & Data Upload

Entity Relationship Diagram (ERD),
Bit of a data cleaning in Excel,
Data Upload in MySQL

## Game Table

MAIN TABLE columns A:D & P





4	Α	В	C	D	E	F		G	Н	1	J	K	L	М	N	0	Р
1	Name	Platform	Year_	Genre	Publisher	NA_Sal	es	EU_	JP_	Other_	Global	Critic_Sco	Critic_Cou	User_Scor	User_Cour	Developer	Rating
2	Wii Sports	Wii	2006	Sports	Nintendo	41.	36	29	4	8.45	82.53	76	51	8	322	Nintendo	E
3	Super Mario	NES	1985	Platform	Nintendo	29.	08	4	7	0.77	40.24	NULL	NULL	NULL	NULL	NULL	NULL
4	Mario Kart W	Wii	2008	Racing	Nintendo	15.	68	13	4	3.29	35.52	82	73	8.3	709	Nintendo	E
5	Wii Sports Re	Wii	2009	Sports	Nintendo	15.	61	11	3	2.95	32.77	80	73	8	192	Nintendo	E
6	Pokemon Red	GB	1996	Role-Playir	Nintendo	11.	27	9	10	1	31.37	NULL	NULL	NULL	NULL	NULL	NULL
7	Tetris	GB	1989	Puzzle	Nintendo	23	3.2	2	4	0.58	30.26	NULL	NULL	NULL	NULL	NULL	NULL
8	New Super N	DS	2006	Platform	Nintendo	11.	28	9	7	2.88	29.8	89	65	8.5	431	Nintendo	E
9	Wii Play	Wii	2006	Misc	Nintendo	13.	96	9	3	2.84	28.92	58	41	6.6	129	Nintendo	E
10	New Super N	Wii	2009	Platform	Nintendo	14.	44	7	5	2.24	28.32	87	80	8.4	594	Nintendo	E
11	Duck Hunt	NES	1984	Shooter	Nintendo	26.	93	1	0	0.47	28.31	NULL	NULL	NULL	NULL	NULL	NULL
12	Nintendogs	DS	2005	Simulation	Nintendo	9.	05	11	2	2.74	24.67	NULL	NULL	NULL	NULL	NULL	NULL
13	Mario Kart D	DS	2005	Racing	Nintendo	9.	71	7	4	1.9	23.21	91	64	8.6	464	Nintendo	E

## Fact & Dimension Tables

**FACT TABLE** 

**DIMENSION TABLES** 

F:J sales

critic\_ review

user\_ review

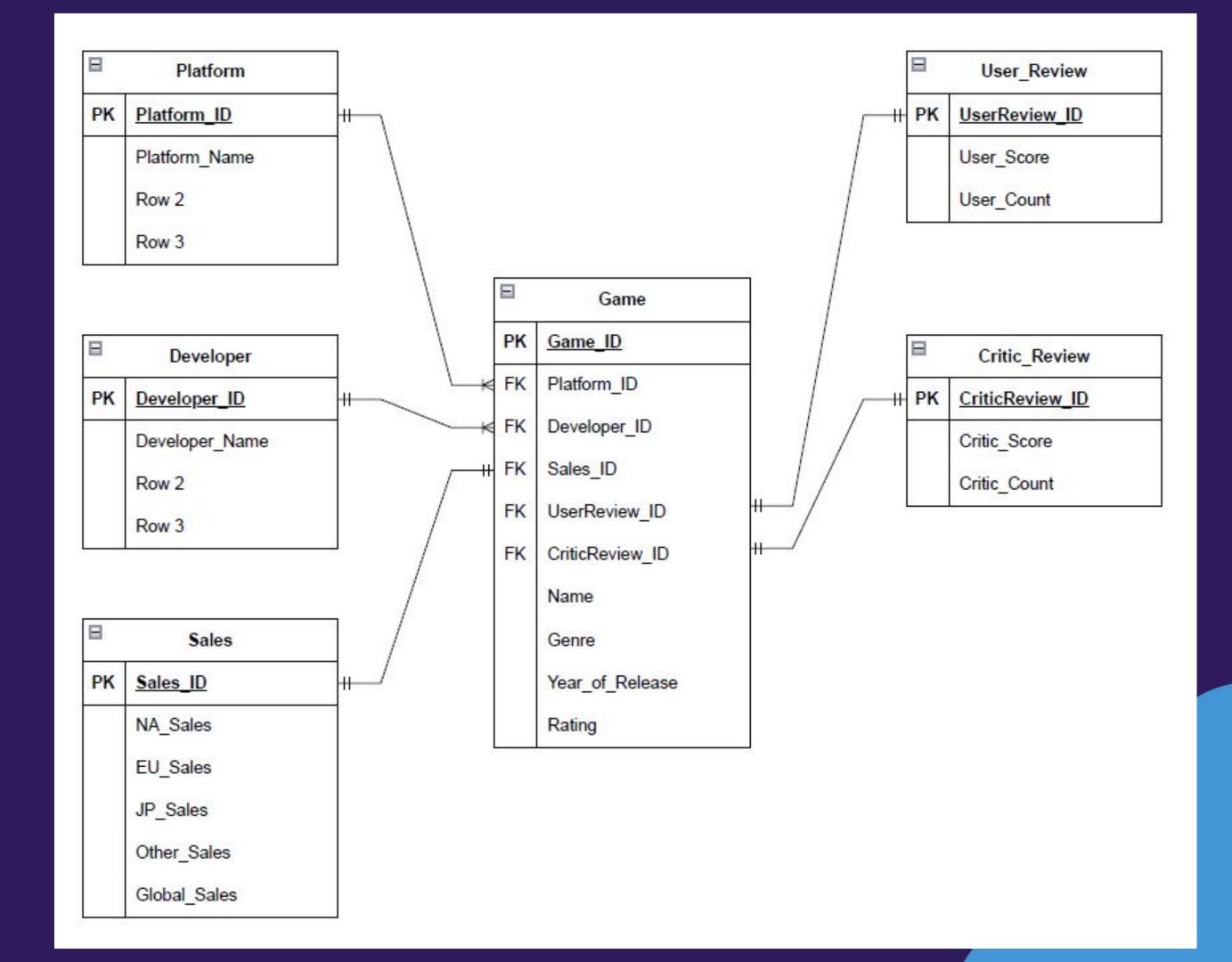
Developer

1	Α	В	С	D	Е	F	G	Н	1	J	K		L	М		N		0	Р
1	Name	Platform	Year_	Genre	Publisher	NA_Sales	EU_	JP_	Other	Global_	Critic_	Scor	Critic_Cou	User_	Score	User_	Cour	Developer	Rating
2	Wii Sports	Wii	2006	Sports	Nintendo	41.36	29	4	8.45	82.53		76	51		8		322	Nintendo	E
3	Super Mario	NES	1985	Platform	Nintendo	29.08	4	7	0.77	40.24	NULL		NULL	NULL		NULL		NULL	NULL
4	Mario Kart W	Wii	2008	Racing	Nintendo	15.68	13	4	3.29	35.52		82	73		8.3		709	Nintendo	E
5	Wii Sports Re	Wii	2009	Sports	Nintendo	15.61	11	3	2.95	32.77		80	73		8		192	Nintendo	E
6	Pokemon Red	GB	1996	Role-Playir	Nintendo	11.27	9	10	1	31.37	NULL		NULL	NULL		NULL		NULL	NULL
7	Tetris	GB	1989	Puzzle	Nintendo	23.2	2	4	0.58	30.26	NULL		NULL	NULL		NULL		NULL	NULL
8	New Super N	DS	2006	Platform	Nintendo	11.28	9	7	2.88	29.8		89	65		8.5		431	Nintendo	E
9	Wii Play	Wii	2006	Misc	Nintendo	13.96	9	3	2.84	28.92		58	41		6.6		129	Nintendo	E
10	New Super N	Wii	2009	Platform	Nintendo	14.44	7	5	2.24	28.32		87	80		8.4		594	Nintendo	E
11	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	1	0	0.47	28.31	NULL		NULL	NULL		NULL		NULL	NULL
12	Nintendogs	DS	2005	Simulation	Nintendo	9.05	11	2	2.74	24.67	NULL		NULL	NULL		NULL		NULL	NULL
13	Mario Kart D	DS	2005	Racing	Nintendo	9.71	7	4	1.9	23.21		91	64		8.6		464	Nintendo	E

# ERD V1.0

**Entity Relationship Diagram** 





# Data Cleaning in Excel









#### Correct 'name'

Remove blank, remove strange names starting with .hack, etc., and modify data included with ??

#### Fill with 'NULL'

in 'Year\_of\_Release'
N/A to NULL
Critic\_score, Critic\_count,
User\_score, user\_count,
Developer, and Rating
were all changed
from Blank to NULL.

#### Divide into tables

Put into dimensions to Critic\_review,
Developer,
Platform, Sales, and
User\_review tables

# Unique Platform & Developer

Add Developer\_id
Add Platform information
which are released\_year
(of platform), Company
refer to Wikipedia

# Adding platform information

After Data cleaning in Excel and after uploading data into MySQL 11,491 entries and 17 columns

Find additional information about the platform and add it manually in Excel

"Super Nintendo" redirects here. For Universal Studios Japan & Hollywood attraction, see Super Nintendo World.

The Super Nintendo Entertainment System, commonly shortened to Super Nintendo, [b] Super NES or SNES, [c] is a 16-bit home video game console developed by Nintendo that was released in 1990 in Japan and South Korea, [16] 1991 in North America, 1992 in Europe and Oceania, and 1993 in South America. In Japan, it is called the Super Famicom (SFC). [d] In South Korea, it is called the Super Comboy [e] and was distributed by Hyundai Electronics. [17] The system was released in Brazil on August 30, 1993, [16][18] by Playtronic. Although each version is essentially the same, several forms of regional lockout prevent cartridges for one version from being used in other versions.

The Super NES is Nintendo's second programmable home console, following the Nintendo Entertainment System (NES). The console introduced advanced graphics and sound capabilities compared with other systems at the time. It was designed to accommodate the ongoing development of a variety of enhancement chips integrated into game cartridges to be competitive into the next generation.

The Super NES received largely positive reviews and was a global success, becoming the bestselling

#### Super Nintendo Entertainment System



# Adding platform information

Before After

#### manually

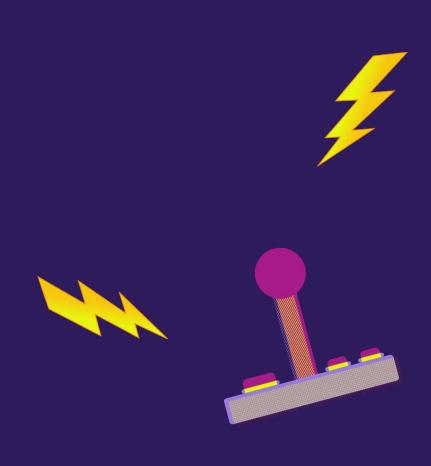
A	Α
1	Platform
2	Wii
3	NES
4	GB
5	DS
6	Xbox360
7	PS3
8	PS2
9	SNES
10	GBA
11	PS4
12	3DS
13	N64
14	PS
15	XB
16	PC

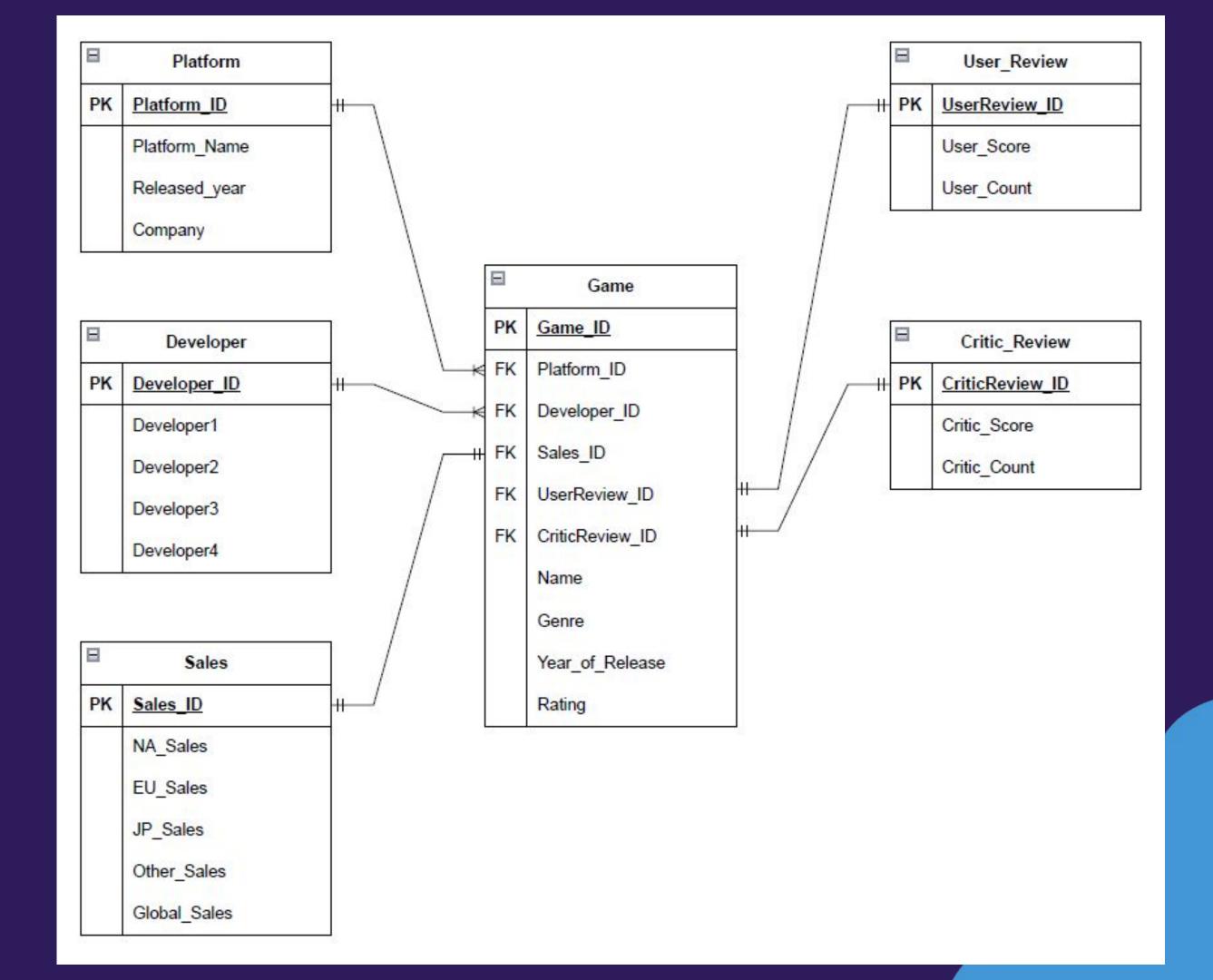


	Α	В	С
1	Platform	Released_year	Company
2	Wii	2006	Nintendo
3	NES	1985	Nintendo
4	GB	1989	Nintendo
5	DS	2004	Nintendo
6	Xbox360	2005	Microsoft
7	PS3	2006	Sony
8	PS2	2000	Sony
9	SNES	1991	Nintendo
10	GBA	2001	Nintendo
11	PS4	2013	Sony
12	3DS	2010	Nintendo
13	N64	1996	Nintendo
14	PS	1994	Sony
15	XB	2001	Microsoft
16	PC	1998	NEC

# ERD v2.0

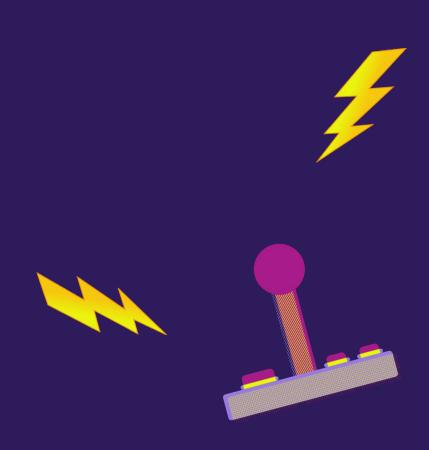
Entity Relationship Diagram

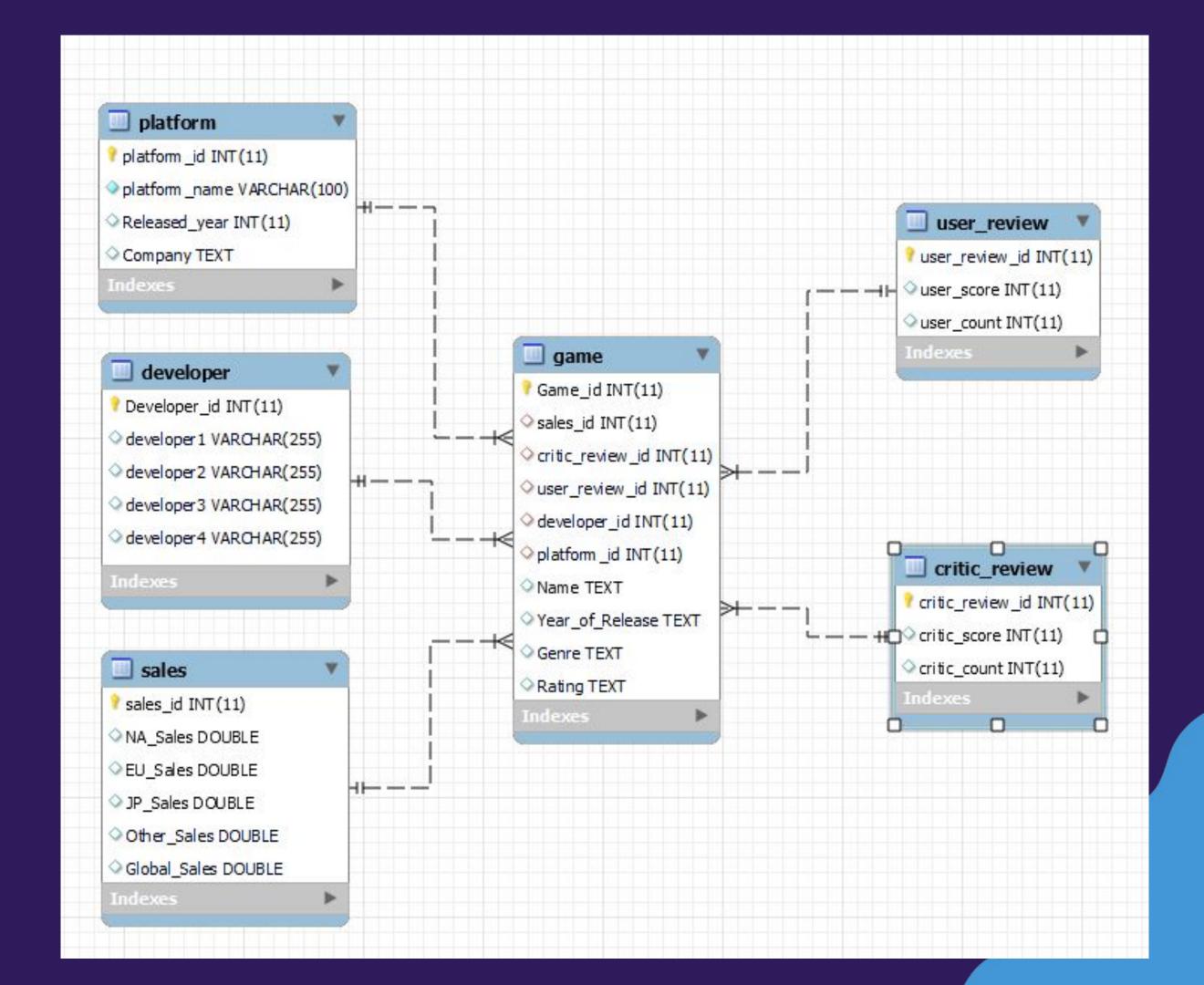




# Final ERD

**Entity Relationship Diagram** 





# MySQL Tables

		Game		
1	Α	В	С	D
1	Name	Year_of_Release	Genre	Rating
2	Wii Sports	2006	Sports	E
3	Super Mario Bros.	1985	Platform	NULL
4	Mario Kart Wii	2008	Racing	E
5	Wii Sports Resort	2009	Sports	E
6	Pokemon Red/Pokemon	1996	Role-Playing	NULL
7	Tetris	1989	Puzzle	NULL
8	New Super Mario Bros.	2006	Platform	E
9	Wii Play	2006	Misc	E
10	New Super Mario Bros.	2009	Platform	E
11	Duck Hunt	1984	Shooter	NULL
12	Nintendogs	2005	Simulation	NULL
13	Mario Kart DS	2005	Racing	E
14	Pokemon Gold/Pokemo	1999	Role-Playing	NULL
15	Wii Fit	2007	Sports	E
16	Kinect Adventures!	2010	Misc	E

2009 Sports

E

17 Wii Fit Plus

#### Critic\_review

Α

1	Critic_Score	Critic_Count
2	76	51
3	NULL	NULL
4	82	73
5	80	73
6	NULL	NULL
7	NULL	NULL
8	89	65
9	58	41
10	87	80
11	NULL	NULL
12	NULL	NULL
13	91	64
14	NULL	NULL
15	80	63
16	61	45
17	80	33

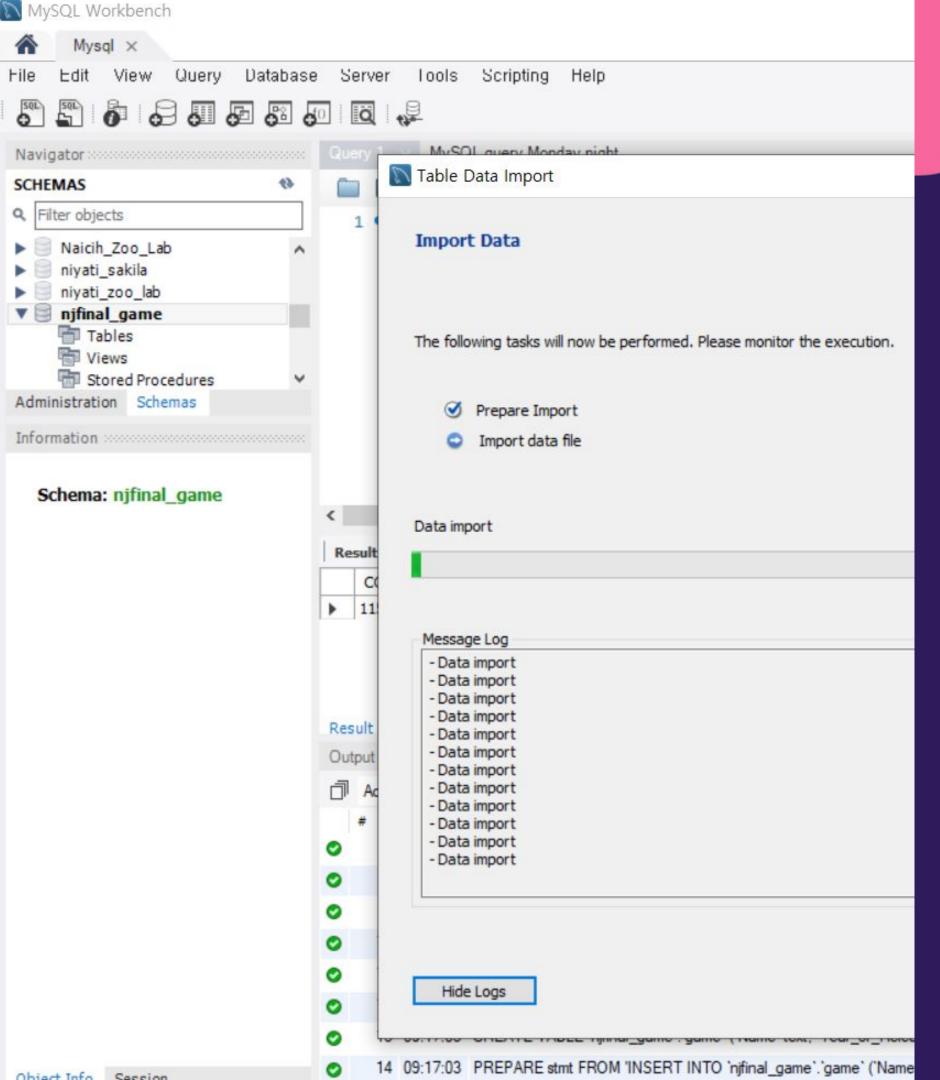
#### User\_review

В

1	User_Score	User_Count
2	8	322
3	NULL	NULL
4	8.3	709
5	8	192
6	NULL	NULL
7	NULL	NULL
8	8.5	431
9	6.6	129
10	8.4	594
11	NULL	NULL
12	NULL	NULL
13	8.6	464
14	NULL	NULL
15	7.7	146
16	6.3	106
17	7.4	52

# MySQL Tables

	and the second	•	Sale	es				Developer			Platform	
N	Α	В	С	D	E	4	Α	В	4	Α	В	С
1	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	1	Deve	Developer_name	1	Platform	Released_yea	Company
2	41.36	28.96	3.77	8.45	82.53	2	1	Nintendo	2	Wii	200	Nintendo
3	29.08	3.58	6.81	0.77	40.24	3	2	Good Science Studio	3	NES	198	Nintendo
4	15.68	12.76	3.79	3.29	35.52	4	3	Rockstar North	4	GB	1989	Nintendo
5	15.61	10.93	3.28	2.95	32.77	5	4	Polyphony Digital	5	DS	2004	Nintendo
6	11.27	8.89	10.22	1	31.37	6	5	Infinity Ward, Sledgeham	6	Xbox360	200	Microsoft
7	23.2	2.26	4.22	0.58	30.26	7	6	Treyarch	7	PS3	200	Sony
8	11.28	9.14	6.5	2.88	29.8	8	7	Infinity Ward	8	PS2	2000	Sony
9	13.96	9.18	2.93	2.84	28.92	9	8	DMA Design	9	SNES	199	Nintendo
10	14.44	6.94	4.7	2.24	28.32	10	9	Game Arts	10	GBA	200	Nintendo
11	26.93	0.63	0.28	0.47	28.31	11	10	Retro Studios, Entertainm	11	PS4	2013	Sony
12	9.05	10.95	1.93	2.74	24.67	12	11	Bungie Software, Bungie	12	3DS	201	Nintendo
13	9.71	7.47	4.13	1.9	23.21	13	12	Ubisoft	13	N64	199	Nintendo
14	9	6.18	7.2	0.71	23.1	14	13	Bungie	14	PS	199	Sony
15	8.92	8.03	3.6	2.15	22.7	15	14	SquareSoft	15	XB	200	Microsoft
16	15	4.89	0.24	1.69	21.81	16	15	343 Industries	16	PC	199	NEC
17	9.01	8.49	2.53	1.77	21.79	17	16	Hudson Soft	17	2600	197	Atari, Inc.





# Data Upload into MySQL

Upload 6 Divided Excel CSV files each. Perhaps because of the large amount of data, it takes more than 2 hours just to upload one 'game' table.

# 3. Assign 'id', PK & FK Answering Business Questions

Assign 'id' in MySQL

Making PK & FK relationship in MySQL

Answering business questions by MySQL queries

# Add 'table\_id' in each table as Primary key

ALTER TABLE sales
ADD COLUMN sales\_id INT NOT
NULL auto\_increment Primary Key;

sales_id	NA_Sales	EU_Sales	JP_Sales	Other_Sales
1	41.36	28.96	3.77	8.45
2	29.08	3.58	6.81	0.77
3	15.68	12.76	3.79	3.29
4	15.61	10.93	3.28	2.95
5	11.27	8.89	10.22	1
6	23.2	2.26	4.22	0.58
7	11.28	9.14	6.5	2.88
8	13.96	9.18	2.93	2.84
9	14.44	6.94	4.7	2.24
10	26.93	0.63	0.28	0.47
11	9.05	10.95	1.93	2.74
12	9.71	7.47	4.13	1.9
13	9	6.18	7.2	0.71
14	8.92	8.03	3.6	2.15
15	15	4.89	0.24	1.69
16	9.01	8.49	2.53	1.77
17	7.02	9.09	0.98	3.96
18	9.43	0.4	0.41	10.57
19	12.78	3.75	3.54	0.55
20	4.74	9.2	4.16	2.04

# Add 'table\_id' in game table as Foreign key

ALTER TABLE game
ADD FOREIGN KEY (sales\_id)
REFERENCES sales(sales\_id)

Game_id	sales_id	critic_review_id
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15

# Multiple Developers into 1NF

#### SET

```
developer1 = SUBSTRING_INDEX(developer_name, ',', 1),
developer2 = CASE WHEN LENGTH(developer_name) -
LENGTH(REPLACE(developer_name, ',', ")) >= 1 THEN
```

SUBSTRING\_INDEX(SUBSTRING\_INDEX(developer\_nam e, ',', 2), ',', -1) ELSE NULL END,

(...) to developer4
Because MAX(comma\_count) is 3

#### Developer\_name

Epic Games

Pipeworks Software, Inc.

Rockstar San Diego

Retro Studios

Blizzard Entertainment

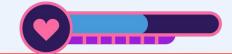
Reflections Interactive

Rare Ltd.

KCEJ

Kojima Productions

Square Enix



SCE/WWS, Media Molecule

#### Developer\_name

Epic Games

Pipeworks Software, Inc.

Rockstar San Diego

Retro Studios

Blizzard Entertainment

Reflections Interactive

Rare Ltd.

KCEJ

Kojima Productions

Square Enix

SCE/WWS, Media Molecule

Delete like 'Inc. after

into INF
Atomic

Developer_id	developer 1	developer2	developer3	developer4
1	Nintendo	NULL	NULL	NULL
2	Good Scienc	NULL	NULL	NULL
3	Rockstar North	NULL	NULL	NULL
4	Polyphony D	NULL	NULL	NULL
5	Infinity Ward	Sledgeha	NULL	NULL
6	Treyarch	NULL	NULL	NULL
7	Infinity Ward	NULL	NULL	NULL
8	DMA Design	NULL	NULL	NULL
9	Game Arts	NULL	HULL	NULL
10	Retro Studios	Entertain	NULL	NULL
11	Bungie Soft	Bungie	NULL	NULL
12	Ubisoft	NULL	NULL	NULL
13	Bungie	NULL	NULL	NULL
14	SquareSoft	NULL	NULL	NULL
15	343 Industries	NULL	NULL	NULL
		_	-	

Ouestion One

Which gaming platform has the largest number of games available?

Question Two

Which gaming platform has the highest total sales?

Question Three

What are the average sales figures for each gaming platform?

Question Four

Which game genre has the highest total sales average?

Question Five

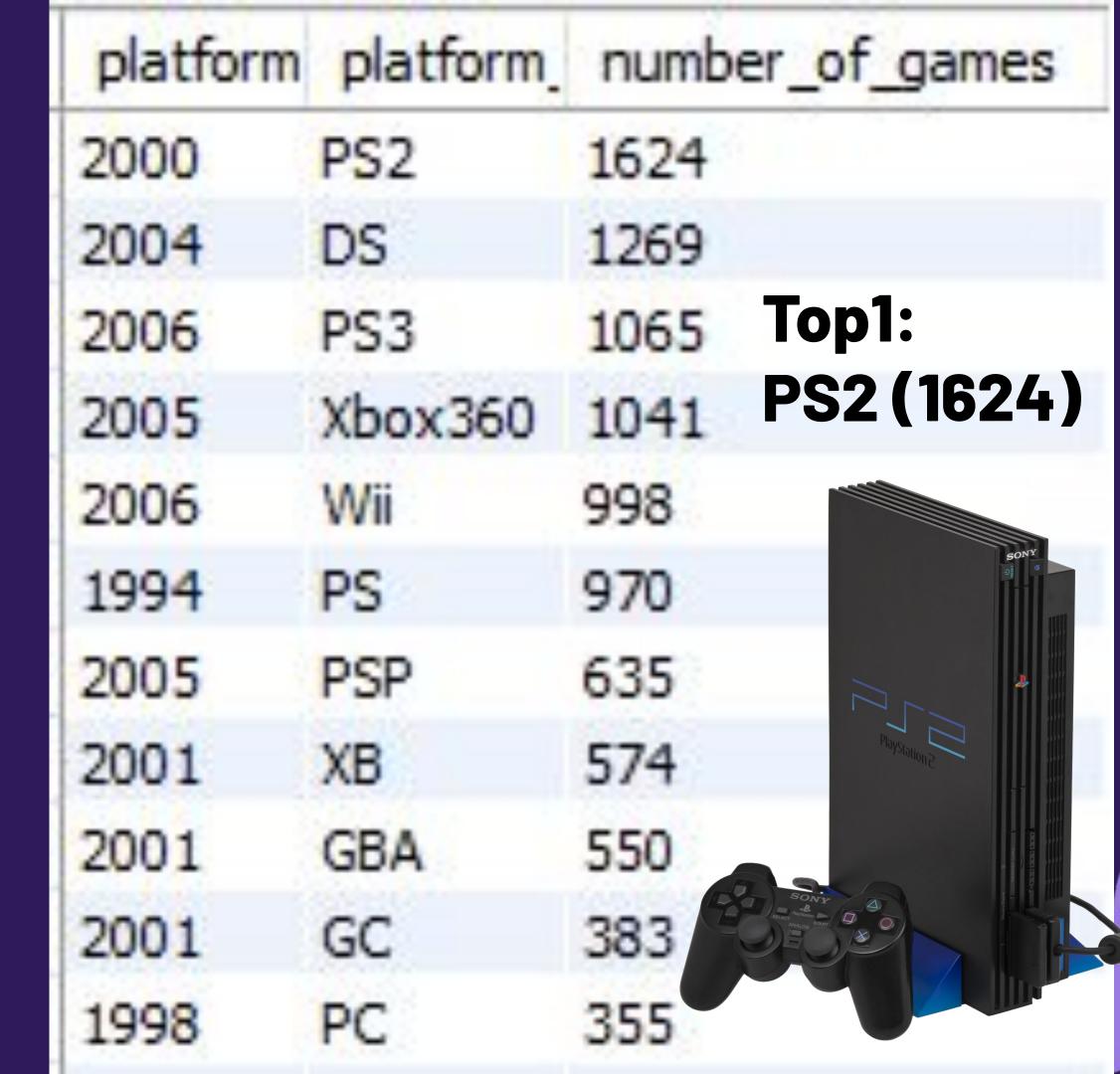
Comparison of the top 10 games by critic score and the top 10 games by user score, along with their sales figures.



Question One

Which gaming platform has the largest number of games available?

A: PS2 (1624),
DS (1269),
PS3 (1065)



Question One -1
Which gaming platform has the largest number of games available?

After 2012

A: PS4 (267), PS3 (222), XOne (171), 3DS (163) Xbox360 (139)

platform	platform_name	number_of_games
2013	PS4	267
2006	PS3	222
2013	XOne	171
2010	3DS	163
2005	Xbox360	<sup>139</sup> Top1
2012	PSV	<sup>116</sup> (2012~now)
1998	PC	<sup>93</sup> PS4 (267)
2012	WiiU	81
2005	PSP	
2006		
2004	DS	5

Question Two

Which gaming platform has the highest total sales in games?

A: PS2 (1231.88), Xbox360 (960.74), PS3 (927.61)



Top1: GAME BOY



What are the average sales figures for each gaming platform?

A: GB (2.74),
NES (2.61),
GEN (1.63)



platform	average_sales		
GB	2.74		
NES	2.61		
GEN	1.63		
PS4	1.16		
SNES	1.00		
XOne	0.92		
Xbox	0.92		
Wii	0.90		
PS3	0.87		
SCD	0.82		
N64	0.78		

Question Four

Which game genre has the highest total sales average?

A: Platform (1.18), Shooter (1.06), Role-playing (0.87)

genre_avg_sales		
1.18		
1.06		
0.87		
0.79		
0.71		
0.71		
0.70		
0.69		
0.67		
0.65		

## Platform Genre Game

A platformer (also called a platform game, and sometimes a jump 'n' run game) is a sub-genre of action video games in which the core objective is to move the player character between points in an environment.





Question Five

Comparison of the top 10 games by critic score and the top 10 games by user score, along with their sales figures.

game	u_score	game	c_score
Golden Sun The Lost	10	Grand Theft Auto IV	98
Harvest Moon Friend	10	Grand Theft Auto IV	98
MLB SlugFest Loaded	10	Tony Hawk's Pro S	98
Boktai The Sun is in Y	10	SoulCalibur	98
Cory in the House	10	Grand Theft Auto V	97
Super Puzzle Fighter II	10	Grand Theft Auto V	97
New Super Mario Bros.	9	Grand Theft Auto III	97
Mario Kart DS	9	Grand Theft Auto V	97
Grand Theft Auto Sa	9	Super Mario Galaxy	97
Grand Theft Auto Vic	9	Super Mario Galaxy 2	97

# Insights of MySQL Data Analysis

One

Importance of ERD for building a database



Two

Data Cleaning time takes up a large portion of all tasks.



After building data, it is easy to answer many questions.



Four

If analyzed with an expert with domain knowledge, it can be utilized more

business-wise.



# Thank you!

Do You Have Any Questions?



### References

RUSH KIRUBI. (2017). Video game sales with ratings. Kaggle. Retrieved from https://www.kaggle.com/datasets/rush4ratio/video-game-sales-with-ratings

Wikipedia. (n.d.). Nintendo DS. Retrieved from https://en.wikipedia.org/wiki/Nintendo\_DS

W3Schools. (n.d.). SQL Constraints. Retrieved from https://www.w3schools.com/sql/sql\_constraints.asp

CsharpTech. (n.d.). How to Allow Remote Access to MySql. YouTube. Retrieved from https://www.youtube.com/watch?v=XqADA64-d1c

TV Tropes. (n.d.). Miscellaneous Games. Retrieved from https://tvtropes.org/pmwiki/pmwiki.php/Main/MiscellaneousGames

Wikipedia. (n.d.). Platformer. Retrieved from https://en.wikipedia.org/wiki/Platformer