

# Group Project: Part 1 (RDBMS) MCSD 1123/SCSP5023 (BIG DATA MANAGEMENT) Section 1

Ahmad Khaidir Amir bin Rodzman (A18CS0014) Putera Muhammad Syabil Bin Sarianto (A18CS0235) LI YEXIN(MCS211046)

### 1) Selected Data

The data that has been selected for this project is the Video Game Sales in 2019 dataset from Kaggle. There are 55,792 records in the dataset as of April 12th, 2019. It was generated by a scrape of vgchartz.com.

Dataset Link - https://www.kaggle.com/datasets/ashaheedq/video-games-sales-2019

# 2) Data Scenario

This is a Game Shop company and in 2019 we need to create a report based on history to increase sales for the shop in future. Game sales data have been collected since the establishment of the company, and use the following analysis to determine which developers, platforms, and publishers we should work with more closely, as well as what types of games we should focus on to increase the sales and profit for upcoming years and what niche or genre developer need to focus on. How does the market trend every year affects the sales and how does user opinion on the game that they buy.

#### 2.1 Column Definition

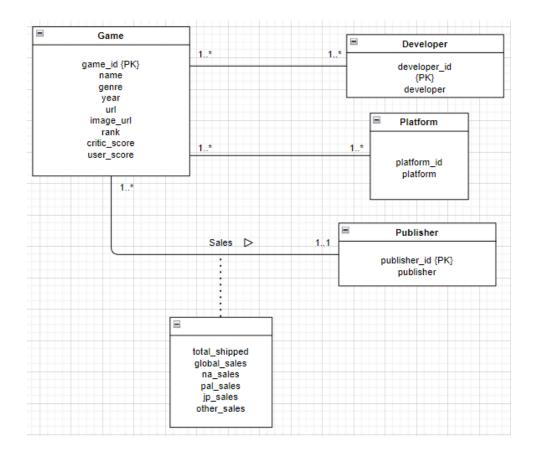
- Rank Ranking of overall sales
- Name Name of the game
- Platform Platform of the game (i.e. PC, PS4, XOne, etc.)
- Genre Genre of the game
- ESRB Rating ESRB Rating of the game
- Publisher Publisher of the game
- Developer Developer of the game
- Critic Score Critic score of the game from 10
- User Score Users score the game from 10
- Total Shipped Total shipped copies of the game
- Global Sales Total worldwide sales (in millions)
- NA Sales Sales in North America (in millions)
- PAL Sales Sales in Europe (in millions)
- JP Sales Sales in Japan (in millions)
- Other Sales Sales in the rest of the world (in millions)
- Year Year of release of the game

#### 2.2 Possible Questions for Analysis

- 1) What are the games that have the highest user score?
- 2) Which game and publisher produce the games that have the highest user score?
- 3) Which game that is produced by Nintendo that has the highest user score
- 4) Which developer developed the rank 55th game?
- 5) Which platform has the top 20 highest critic scores?
- 6) Which platform is used between 1999 to 2005
- 7) Which publisher and game has the top 10 highest global sales?
- 8) Which top 5 genre contribute the most sales In Japan.
- 9) Which top 5 publisher contribute to the most sales in Europe?
- 10) What is the top 10 genre with the highest average of critic score?

## 3) Possible Relation

This section cover on what is possible relation or table that could exist in the Game Store Database and how they can relate to each other.



PK - gameID

FK - developerID reference Developer

FK - platformID reference Platform

FK - publisherID reference Publisher

FK - gameID reference Sales

Developer(developerID,developer)

PK - developerID

Platform(platformID, platform)

PK - platform ID

Publisher(publisherID, publisher)

PK - publisherID

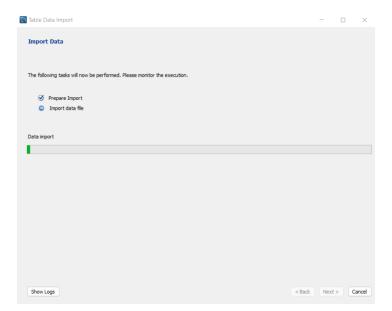
Sales(total\_shipped,global\_sales,na\_sales,pal\_sales,jp\_sales,other\_sales)

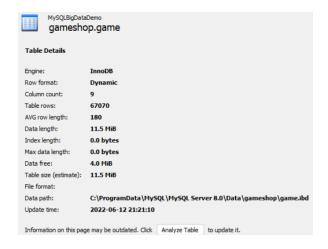
FK - publisherID reference Publisher

FK - gameID reference Game

# 4) SQL Queries

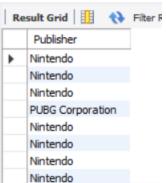
#### 4.1 Create Table and Import Data from csv files

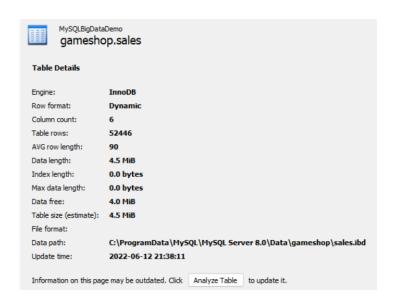




gameID	Name	Genre	year	url	image_url
32736	Gundemonium Recollection	Shooter	2011	http://www.vgchartz.com/game/44911/gunde	/games/boxart/default.jpg
32737	GunForce	Shooter	1992	http://www.vgchartz.com/game/13499/gunforc	/games/boxart/full_1195892AmericaFrontccc.jpg
32738	Gunlok	Strategy	2001	http://www.vgchartz.com/game/28335/gunlok/	/games/boxart/full_8339571AmericaFrontccc.jpg
32739	GunLord	Action	2001	http://www.vgchartz.com/game/49896/gunlord	/games/boxart/default.jpg
32740	Gunma's Ambition	Misc	2014	http://www.vgchartz.com/game/83434/gunmas	/games/boxart/full_3678161AmericaFrontccc.png
NULL	NULL	NULL	NULL	NULL	NULL







Total_Shipped	Global_Sales	NA_Sales	PAL_Sales	JP_Sales	Other_Sales
0	0.05	0.01	0	0.04	0
0	0.05	0.05	0	0	0
0	0.05	0	0	0.05	0
0	0.05	0.05	0	0	0
0	0.05	0.03	0.01	0	0
0	0.05	0.05	0	0	0
0	0.05	0	0	0.05	0
0	0.05	0	0	0.05	0



Platform
GB
Wii
Wii
SNES
PS3
PS4
DS
PC



Platform
GB
Wii
Wii
SNES
PS3
PS4
DS
PC

#### 4.2 Primary Key Creation

a) From developer entity

alter table Developer

ADD COLUMN developerID INT not null auto\_increment Primary Key;

	Developer	developerID
•	Nintendo EAD	1
	Nintendo EAD	2
	Nintendo EAD	3
	PUBG Corporation	4
	Nintendo EAD	5
	Game Freak	6
	Nintendo EAD	7
	Bullet Proof Software	8

b) From publisher entity

alter table Publisher

ADD COLUMN publisherID INT not null auto\_increment Primary Key;

	Publisher	publisherID
•	Nintendo	1
	Nintendo	2
	Nintendo	3
	PUBG Corporation	4
	Nintendo	5
	Nintendo	6
	Nintendo	7
	Nintendo	8

c) From platform entity

alter table platform

ADD COLUMN platformID INT not null auto\_increment Primary Key;

	Platform	platformID
•	Wii	1
	NES	2
	Wii	3
	PC	4
	Wii	5
	GB	6
	DS	7
	CD	0

#### 4.3 Adding foreign key column and inserting value

#### a) From game entity:

alter table game add column publisherID int null; update game set publisherID = gameID;

alter table game add column platformID int null; update game set platformID = gameID;

alter table game add column developerID int null; update game set developerID = gameID;



#### b) From sales entity:

alter table sales

ADD COLUMN dummyPrimary INT not null auto\_increment Primary Key;

alter table sales add column gameID int null; alter table sales add column publisherID int null;

update sales set gameID = dummyPrimary; update sales set publisherID = dummyPrimary;

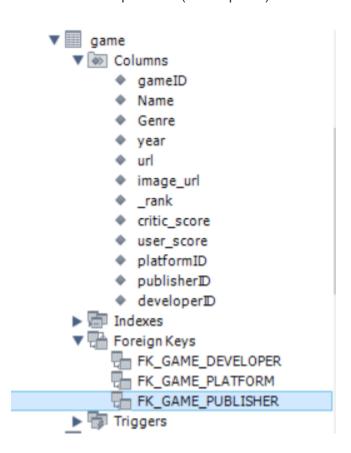
alter table sales
DROP COLUMN dummyPrimary;



#### 4.4 Adding foreign key to all entities

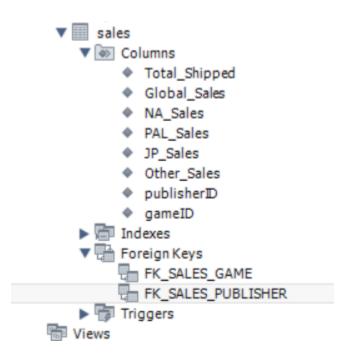
SET FOREIGN\_KEY\_CHECKS=0;

- a) From game entity:
  - To add foreign key to publisher
     ALTER TABLE game
     ADD CONSTRAINT FK\_GAME\_PUBLISHER
     FOREIGN KEY (publisherID)
     REFERENCES publisher(publisherID)
  - 2) To add foreign key to platform ALTER TABLE game ADD CONSTRAINT FK\_GAME\_PLATFORM FOREIGN KEY (platformID) REFERENCES platform(platformID)
  - 3) To add foreign key to developer ALTER TABLE game ADD CONSTRAINT FK\_GAME\_DEVELOPER FOREIGN KEY (developerID) REFERENCES publisher(developerID)



b) From sales entity:

- To add foreign key to publisher
   ALTER TABLE sales
   ADD CONSTRAINT FK\_SALES\_PUBLISHER
   FOREIGN KEY (publisherID)
   REFERENCES publisher(publisherID)
- 2) To add foreign key to publisher ALTER TABLE sales ADD CONSTRAINT FK\_SALES\_GAME FOREIGN KEY (gameID) REFERENCES game(gameID)



#### 4.5 SQL Queries for possible scenario analysis

1) What are the games that have the highest user score? select name as GAME\_NAME, user\_score as USER\_SCORE from game

where user score=(select Max(user score)from game);

	GAME_NAME	USER_SCORE
•	The Legend of Zelda: Breath of the Wild	10
	Final Fantasy X	10
	The Legend of Zelda: Ocarina of Time	10
	Red Dead Redemption	10
	Kirby Star Allies	10
	Dissidia: Final Fantasy	10
	Monster Hunter Tri	10
	The Legend of Zelda: Breath of the Wild	10
	Lunar: Silver Star Story Complete	10
	Okami HD	10
	Archibald's Adventures	10
	Astro Match3	10
	The Legend of Zelda Collector's Edition	10
	Viper Racing	10
	Blossom Tales: The Sleeping King	10
	Dust: An Elysian Tail	10
	Hollow Knight	10

2) Which game and publisher produce the games that have the highest user score?

select name as GAME\_NAME, user\_score as USER\_SCORE, publisher. Publisher as PUBLISHER\_NAME

from game inner join publisher on game.publisherID= publisher.publisherID where user\_score=(select Max(user\_score)from game)



#### 3) Which game that is produced by Nintendo that has the highest user score

select name as GAME\_NAME, user\_score as USER\_SCORE, publisher.Publisher as PUBLISHER NAME

from game inner join publisher on game.publisherID= publisher.publisherID where user\_score=(select Max(user\_score)from game) and publisher.Publisher="Nintendo"

	GAME_NAME	USER_SCORE	PUBLISHER_NAME
•	The Legend of Zelda: Ocarina of Time	10	Nintendo
	Red Dead Redemption	10	Nintendo
	Monster Hunter Tri	10	Nintendo

4) Which developer developed the rank 55th game?

Select Developer as DEVELOPER\_NAME, game.name as GAME\_NAME, game.\_rank as GAME\_RANK

from game inner join developer on game.developerID=developer.developerID where game. rank=(select\_rank from game where \_rank=55)

	DEVELOPER_NAME	GAME_NAME	GAME_RANK
•	Rockstar Games	New Super Mario Bros. 2	55

5) Which platform has the top 20 highest critic scores?

SELECT Platform as PLATFORM\_NAME, game.critic\_score as GAME\_CRITIC\_SCORE from game inner join platform on game.platformID= platform.platformID order by game.critic\_score desc limit 20

	PLATFORM_NAME	GAME_CRITIC_SCORE
•	Linux	10
	PC	10
	GB	10
	NES	10
	PS2	10
	PS3	10
	PC	10
	PS2	10
	PS4	10
	PC	10
	PC	10
	PS2	10
	X360	10
	PS	10
	NES	10
	SNES	10
	DS	10
	Wii	9.9
	PC	9.9
	PS2	9.9

6) Which platform is used between 1999 to 2005

SELECT Platform as PLATFORM\_NAME, year as GAME\_PRODUCED\_YEAR from platform inner join game on platform.platformID = game.platformID

# where year BETWEEN 1999 AND 2005 group by Platform order by year

	DI ATEODM NAME	CAME DOODLICED VEAD
	PLATFORM_NAME	GAME_PRODUCED_YEAR
	X360	1999
	N64	1999
	NGage	1999
•	DSi	1999
	Aco	1999
	PS4	1999
	GB	2000
	7800	2000
	ApII	2000
	GBC	2000
	AST	2000
	WinP	2000
	Ouya	2000
	Lynx	2000
	5200	2000
	PCE	2001
	XB	2001
	PS	2001
	Mob	2001
	Linux	2001
	PSV	2001
	SNES	2001
	14/14/	2002
Dag	rult QR 🗸	

7) Which publisher and game has the top 10 highest global sales.

SELECT Publisher as PUBLISHER\_NAME, game.name as GAME\_NAME, sales.global\_sales as GLOBAL\_SALES from ((game INNER JOIN publisher ON game.publisherID = publisher.publisherID) INNER JOIN sales ON game.gameID = sales.gameID)

order by sales.global\_sales desc LIMIT 10

	PUBLISHER_NAME	GAME_NAME	GLOBAL_SALES
	Electronic Arts	Dr. Mario	20.32
	Nintendo	Assassin's Creed III	19.39
	Nintendo	Uncharted: Drake's Fortune	16.15
•	Electronic Arts	Splatoon	15.86
	Activision	Batman: Arkham City	15.09
	Sony Computer Entertainment	The Lord of the Rings: The Two Towers	14.82
	LucasArts	Forza Motorsport 4	14.74
	Electronic Arts	Spider-Man: The Movie	13.94
	Nintendo	The Elder Scrolls IV: Oblivion	13.86
	Capcom	The Legend of Zelda: The Wind Waker	13.8

#### 8) Which top 5 genre contribute the most sales In Japan.

select genre as GAME\_GENRE, SUM(sales.jp\_sales) as JAPAN\_SALES\_SUM from game inner join sales on game.gameID= sales.gameID group by genre order by JAPAN\_SALES\_SUM desc limit 5

	GAME_GENRE	JAPAN_SALES_SUM
١	Action	116.95000000000026
	Sports	106.35000000000029
	Misc	77.21000000000053
	Role-Playing	71.89000000000019
	Shooter	68.66000000000003

#### 9) Which top 5 publisher contribute to the most sales in Europe

select Publisher as GAME\_PUBLISHER, SUM(sales.Pal\_sales) as PAL\_SALES\_SUM from publisher inner join sales on publisher.publisherID= sales.publisherID group by Publisher order by PAL\_SALES\_SUM desc limit 5

	GAME_PUBLISHER	PAL_SALES_SUM
١	Nintendo	282.6899999999992
	Electronic Arts	165.0099999999985
	Activision	127.14
	Ubisoft	100.5700000000003
	EA Sports	97.28

10) What is the top 10 genre with the highest average of critic score?

select genre as GENRE, AVG(critic\_score) as CRITIC\_SCORE\_AVERAGE from game

group by genre order by CRITIC\_SCORE\_AVERAGE desc limit 10

	GENRE	CRITIC_SCORE_AVERAGE
Þ	Role-Playing	1.4050977807075349
	Shooter	1.339467945922373
	Action-Adventure	1.2527960526315791
	Party	1.210666666666666
	Platform	1.196487663280116
	Fighting	1.1623021582733808
	Racing	1.0835643564356443
	MMO	1.0554054054054056
	Sandbox	1.02222222222221
	Music	1.0143589743589745