Video Game Sales Practice Case Study

Citation: Gregory Smith. (2022). Video Games Sales [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.5898311>

This dataset contains a list of video games with sales greater than 100,000 copies. It was generated by a scrape of [vgchartz.com](http://www.vgchartz.com/).

Platform Key: **2600** - Atari 2600 **GB** - Game Boy **GBA** - Game Boy Advance **GC** - GameCube

**GEN** - Sega Genesis **N64** - Nintendo **NES** - Nintendo Entertainment System

**PS** - PlayStation **PS2** - PlayStation 2 **PS3**  - PlayStation 3 **PS4** - PlayStation 4

**PSP** - PlayStation Portable **SNES** - Super Nintendo Entertainment System

**X360** - XBox 360 **XB** - XBox **XOne** - XBox One

Main Areas/ Continents: **na** - North America **eu** - Europe **jp** - Japan

Sales - **Sales are per 100,000**

ex: 1 sale = 100,000 sales OR .5 sale = 50,000 sales

Stats:

* Data spans 1980 to 2016 (36 years)
* 17 total genres
* 500 games

Cleaning Steps:

1. Reduced the data to the top 500 ranked games
2. Reviewed for missing/incomplete data and filled it in
   1. Genre, Year
      1. Friend Collection is actually Tomodachi Collection
3. Standardized formatting to be uniform
   1. i.e. Operating System 2600 is not a number it is text

Limitations:

* This dataset does include all country sales and instead groups them as “other\_sales”, making it difficult to verify the validity of the sales
* I limited the the dataset to the top 500 sales, which limited the year range to 2016

Questions:

1. **Which area consumes/ buys the most video games?**
2. **What are the 5 most popular genres in each listed region (North America, Europe, Japan)?**
3. **What is the most popular game in North America, Europe and Japan by sales?**
4. **What is the most popular game by sales?**
5. **What are the top 5 consoles?**
6. **Which console produced the most sales?**
7. **Total video game sales per year?**
8. **How many video games were sold per genre?**

**SQL Data Results**

1. **Which area consumes/ buys the most video games?**

I used the SUM function to look at the total sales for each listed region, instead of identifying the top consumer of video games. I rounded the total\_na\_sales as it was a trailing float.

SELECT

ROUND(SUM(na\_sales),2) AS total\_na\_sales,

SUM(eu\_sales) AS total\_eu\_sales,

SUM(jp\_sales) AS total\_jp\_sales

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`;



**Answer: North America bought the most video games total**

1. **What are the 5 most popular genres in each listed region (North America, Europe, Japan)?**

I used the SUM and GROUP BY functions on each regions sales to find the total sales for each genre listed

**North America**

SELECT

Genre,

ROUND(SUM(na\_sales),2) AS na\_genre\_sales,

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

GROUP BY

Genre

ORDER BY

na\_genre\_sales DESC

LIMIT 5;



**Europe**

SELECT

Genre,

ROUND(SUM(eu\_sales),2) AS eu\_genre\_sales,

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

GROUP BY

Genre

ORDER BY

eu\_genre\_sales DESC

LIMIT 5;



**Japan**

SELECT

Genre,

ROUND(SUM(jp\_sales),2) AS jp\_genre\_sales,

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

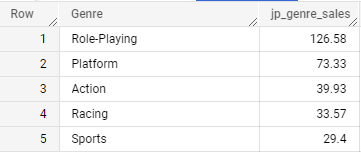
GROUP BY

Genre

ORDER BY

jp\_genre\_sales DESC

LIMIT 5;



**Answer: North America Europe Japan**

1. **Shooter Action Role-Playing**
2. **Action Sports Platform**
3. **Platform Shooter Action**
4. **Sports Platform Racing**
5. **Role-Playing Role-Playing Sports**

**3. What is the most popular game in North America, Europe and Japan by sales?**

I created a subquery to find the name of the most popular game.

**North America**

SELECT

Name,

na\_sales

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

WHERE

na\_sales IN (

SELECT MAX(na\_sales)

FROM **`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`**

**);**

****

**Europe**

SELECT

Name,

eu\_sales

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

WHERE

eu\_sales IN (

SELECT MAX(eu\_sales)

FROM `portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

);

****

**Japan**

**SELECT**

Name,

jp\_sales

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

WHERE

jp\_sales IN (

SELECT MAX(jp\_sales)

FROM `portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

);

****

**Answer:**

**North America: Wii Sports**

**Europe: Wii Sports**

**Japan: Pokemon Red/Pokemon Blue**

**4. What is the most popular game by sales?**

This is a variation of the subquery from question 3. The included column global sales is the sum of all 4 regions (North America, Europe, Japan, and Other). All that was left was to isolate the name of the most and least popular games.

**Most Popular**

SELECT

Name,

global\_sales

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

WHERE

global\_sales IN (

SELECT MAX (global\_sales)

FROM `portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

);



**Least Popular**

SELECT

Name,

global\_sales

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

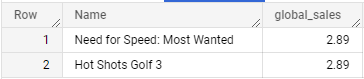
WHERE

global\_sales IN (

SELECT MIN (global\_sales)

FROM `portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

);



**Answer:**

**Most Popular: Least Popular:**

Wii Sports Need for Speed: Most Wanted

Hot Shots Golf 3

**5. What are the top 5 consoles?**

SELECT

Platform,

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

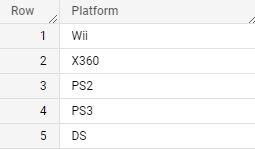
GROUP BY

Platform

ORDER BY

SUM(global\_sales) DESC

LIMIT 5;



**6. Which console produced the most sales?**

SELECT

Platform,

ROUND(SUM(global\_sales),2) as total\_sales

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

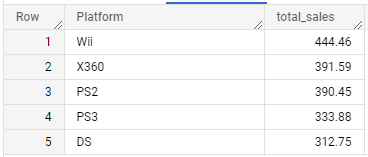
GROUP BY

Platform

ORDER BY

total\_sales DESC

LIMIT 5;



**Answer: Wii with 44,446,000 sales**

**7. Total video game sales per year?**

SELECT

Year,

ROUND(SUM(global\_sales),2) AS total\_sales

FROM

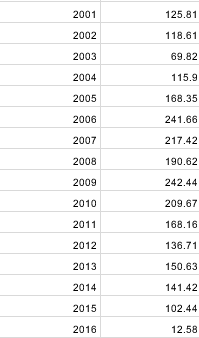
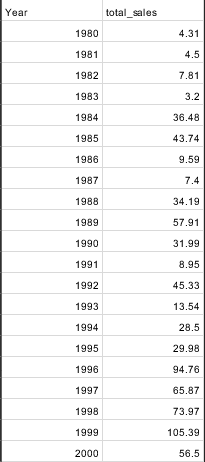
`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

GROUP BY

YEAR

ORDER BY

Year;



**8. How many video games were sold per genre?**

SELECT

Genre,

COUNT(Name) AS num\_sold

FROM

`portfolio-projects-365917.sql\_case\_studies.video\_game\_sales`

GROUP BY

Genre

ORDER BY

num\_sold;

****