datalab

Video games are big business: the global gaming market is projected to be worth more than $300 billion by 2027 according to Mordor Intelligence. With so much money at stake, the major game publishers are hugely incentivized to create the next big hit. But are games getting better, or has the golden age of video games already passed?

In this project, you'll analyze video game critic and user scores as well as sales data for the top 400 video games released since 1977. You'll search for a golden age of video games by identifying release years that users and critics liked best, and you'll explore the business side of gaming by looking at game sales data.

Your search will involve joining datasets and comparing results with set theory. You'll also filter, group, and order data. Make sure you brush up on these skills before trying this project! The database contains two tables. Each table has been limited to 400 rows for this project, but you can find the complete dataset with over 13,000 games on Kaggle.

`game_sales` table

Column	Definition	Data Type
name	Name of the video game	varchar
platform	Gaming platform	varchar
publisher	Game publisher	varchar
developer	Game developer	varchar
games_sold	Number of copies sold (millions)	float
year	Release year	int

`reviews` table

Column	Definition	Data Type
name	Name of the video game	varchar
critic_score	Critic score according to Metacritic	float
user_score	User score according to Metacritic	float

`users_avg_year_rating` table

Column	Definition	Data Type
year	Release year of the games reviewed	int
num_games	Number of games released that year	int
avg_user_score	Average score of all the games ratings for the year	float

`critics_avg_year_rating` table

Column	Definition	Data Type
year	Release year of the games reviewed	`int`
num_games	Number of games released that year	`int`
avg_critic_score	Average score of all the games ratings for the year	`float`

🐘 Projects Data   DataFrame as `best_selling_games`

```sql
-- best_selling_games
SELECT *
FROM game_sales
ORDER BY games_sold DESC
LIMIT 10;
```

	name	platform	publisher
0	Wii Sports for Wii	Wii	Nintendo
1	Super Mario Bros. for NES	NES	Nintendo
2	Counter-Strike: Global Offensive for PC	PC	Valve
3	Mario Kart Wii for Wii	Wii	Nintendo
4	PLAYERUNKNOWN'S BATTLEGROUNDS for PC	PC	PUBG Corporation
5	Minecraft for PC	PC	Mojang
6	Wii Sports Resort for Wii	Wii	Nintendo
7	Pokemon Red / Green / Blue Version for GB	GB	Nintendo
8	New Super Mario Bros. for DS	DS	Nintendo
9	New Super Mario Bros. Wii for Wii	Wii	Nintendo

10 rows ⬇

🐘 Projects Data   DataFrame as `critics_top_ten_years`

```sql
-- critics_top_ten_years
SELECT g.year, COUNT(g.name) AS num_games, ROUND(AVG(r.critic_score),2) AS avg_critic_score
FROM game_sales g
INNER JOIN reviews r
ON g.name = r.name
GROUP BY g.year
HAVING COUNT(g.name) >= 4
ORDER BY avg_critic_score DESC
LIMIT 10;
```

	year	num_games	
0	1998		1
1	2004		1
2	2002		
3	1999		1
4	2001		1
5	2011		2
6	2016		1
7	2013		1
8	2008		2
9	2017		1

10 rows ⬇

🐘 Projects Data   DataFrame as `golden_yea`

```
-- golden_years
SELECT u.year, u.num_games, c.avg_critic_score, u.avg_user_score, c.avg_critic_score -
u.avg_user_score AS diff
FROM critics_avg_year_rating c
INNER JOIN users_avg_year_rating u
ON c.year = u.year
WHERE c.avg_critic_score > 9 OR u.avg_user_score > 9
ORDER BY year ASC
```

	year	num_games	avg_critic_score
0	1997	8	
1	1998	10	
2	2004	11	
3	2008	20	
4	2009	20	
5	2010	23	