Case Study: Raw & Relational: Exploring Game Sales with SQL

Objective

To simulate a full dashboard experience using only SQL, no Power BI, and no Excel visuals. The aim was to uncover insights from video game sales data by designing a normalized schema and writing clean, business-focused queries.

Background

Most dashboards rely heavily on BI tools. But in this project, I wanted to test what happens when we strip everything back to logic and relationships, can we still get meaningful, visual-style insights?

Using raw sales data from 1,500 video games, I cleaned the dataset in Excel to prepare it for import, then created a relational database in MySQL. From there, I built SQL queries that mimic the behavior of KPI tiles, bar charts, and heatmaps, with no BI tools involved.

Tools Used

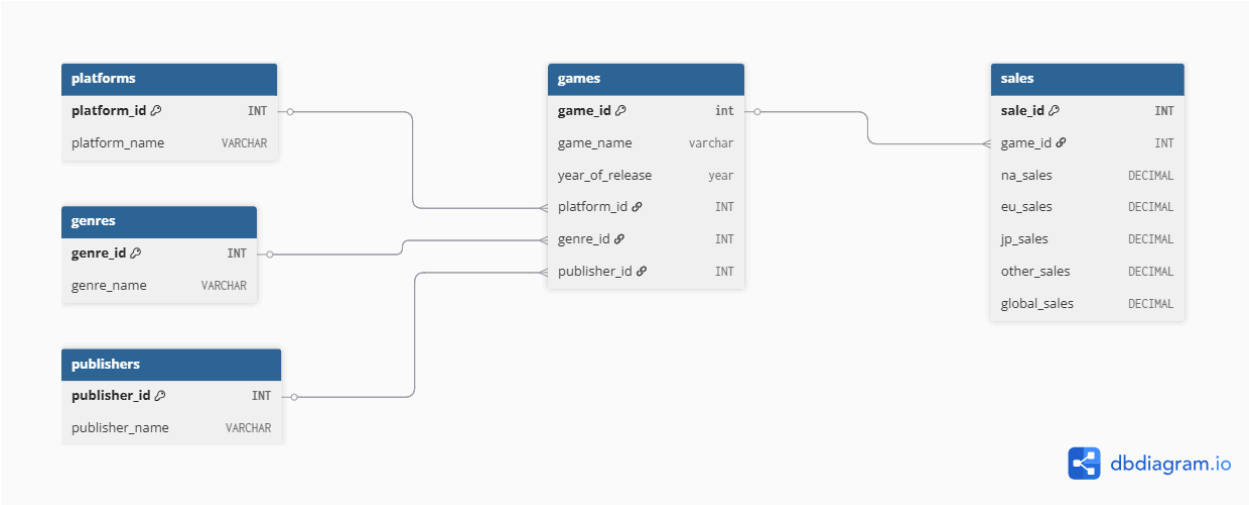
- Excel- data cleaning
- MySQL- schema design, SQL querying, logic building
- dbdiagram.io- ERD and schema planning
- GitHub- documentation and publishing

Schema Design

The flat dataset was normalized into 5 main tables:

- Games: game details (title, platform, year, etc.)
- Genres: distinct genre types
- Publishers: publisher information
- Platforms: console/platform types
- Sales: sales split by region, and global

All foreign keys were manually created to enforce referential integrity.



“Entity Relationship Diagram (ERD) designed manually in dbdiagram.io based on the normalized schema.”

Key Business Questions

- Which platforms and publishers drive the highest global sales?
- Which game genres perform the best and worst?
- How have global sales evolved over the years?
- What sales distribution exists are there more niche games or blockbusters?

Insights Uncovered

- **Top Platform: Xbox 360**
  With over 10.8K units in global sales, Xbox 360 led platform performance. It consistently outperformed others across both NA and EU regions, pointing to strong market penetration and game library depth.

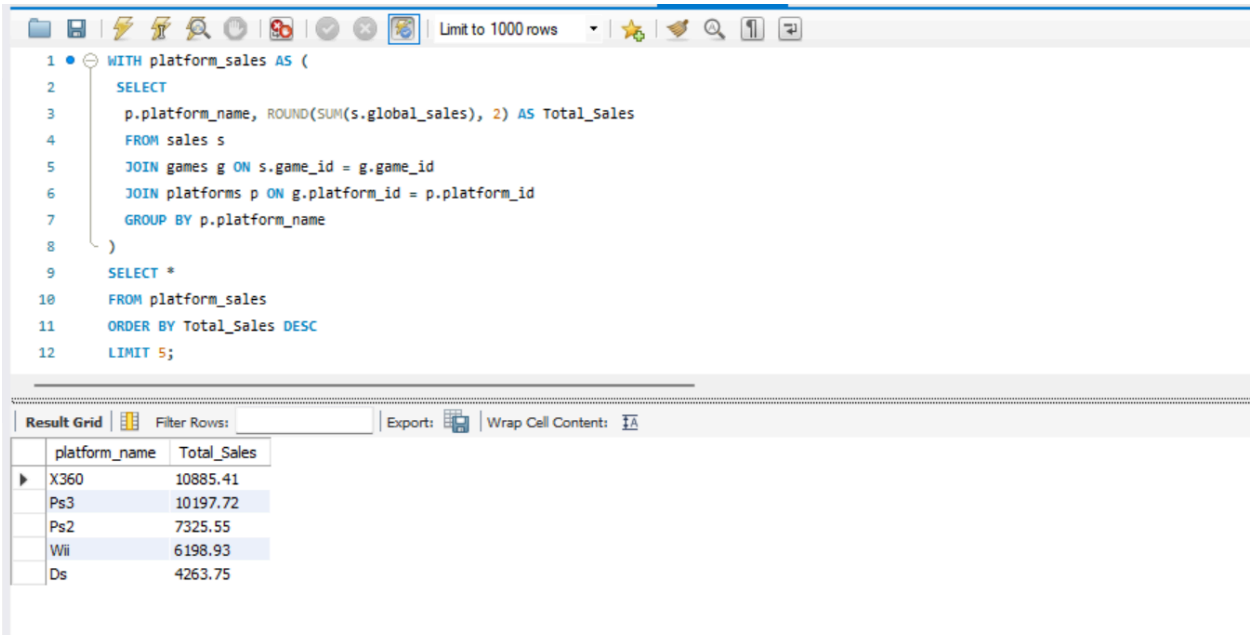
- **Top Publisher: Nintendo**
  Nintendo topped the charts with 14.4K in total global sales, significantly ahead of other publishers. Its dominance highlights the power of flagship franchises and long-term brand loyalty.

- **Lowest-Performing Genre: Strategy**
  Strategy games generated under 500 in global sales, making it the weakest performer. This suggests either limited mainstream appeal or niche targeting, especially in fast-paced markets.

- **Sales Trend Over Time**
  Global video game sales peaked between 2008-2010, coinciding with console booms and hit releases. Post-2013, a steady decline occurred, possibly due to digital transformation, mobile gaming shift, or market saturation.

- **Blockbuster Titles (20M+ in Sales)**
  Only 189 games crossed the 20M+ sales mark, proving how rare it is to achieve true blockbuster status. The majority of titles remained in the “niche” or “moderate hit” categories, reinforcing how top performers skew overall averages.

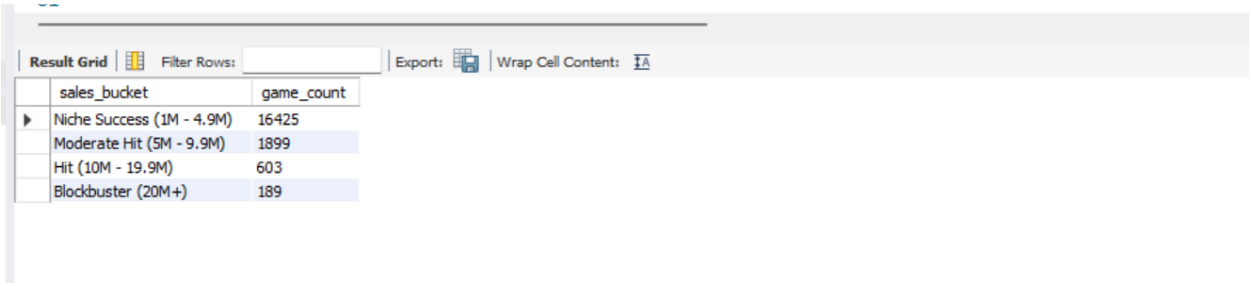
Sample Query: Top 5 platforms by Global Sales



Simulated dashboard tile using SQL CTE logic

Sales Buckets by Game Volume

```sql
WITH sales_buckets AS (
 SELECT g.game_name,
  CASE
   WHEN s.global_sales >= 20 THEN 'Blockbuster (20M+)'
   WHEN s.global_sales >= 10 THEN 'Hit (10M - 19.9M)'
   WHEN s.global_sales >= 5 THEN 'Moderate Hit (5M - 9.9M)'
   WHEN s.global_sales >= 1 THEN 'Niche Success (1M - 4.9M)'
   ELSE 'Low Impact (< 1M)'
END AS sales_bucket
FROM sales s
JOIN games g ON s.game_id = g.game_id
)
SELECT sales_bucket, COUNT(*) AS game_count
FROM sales_buckets
GROUP BY sales_bucket
ORDER BY game_count DESC;
```



sales_bucket	game_count
Niche Success (1M - 4.9M)	16425
Moderate Hit (5M - 9.9M)	1899
Hit (10M - 19.9M)	603
Blockbuster (20M+)	189

Sales bucket segmentation using CASE WHEN

Why It Matters

Most dashboards begin with drag-and-drop. This one began with a question:
“Can I build insights from scratch, using nothing but SQL?”
No visuals, just raw data, logic, and a clear business goal.

This project was my way of proving that insight isn't about aesthetics, it's about architecture. The structure of relationships, the weight of logic, and the flow of queries. I wanted to see if SQL alone could tell the story and it did.
Behind every dashboard is someone asking better questions, and I'm learning to be that someone.

What I'd Explore Next

If this were a real-world consulting case, I'd explore:

- Regional pricing to evaluate *revenue*, not just unit sales
- Game ratings to compare *quality vs sales*
- Launch timing to spot seasonal patterns (e.g., Q4 releases)
- Window functions for more advanced segment ranking (by genre, publisher, region)

Explore the Project

- GitHub Repo
- Portfolio