

# GameCo

## Video Game Sales Analysis



### DATA DICTIONARY

By: Jess Duong  
January 2026

# Table of Contents

GameCo Video Game Sales Analysis .....	3
Project Information .....	3
Dataset Overview .....	3
Description: .....	3
Variable Definition .....	4
Data Limitations and Notes .....	5
Calculated Fields .....	5
Key Metrics Summary .....	6

# GameCo Video Game Sales Analysis

## Project Information

Project	GameCo Video game Sales Descriptive Analysis
Analyst	Jess Duong
Date	January 2026
Data Source.	VGChartz

## Dataset Overview

<b>Dataset Name</b>	Vgsales_Clean
<b>Total Records</b>	16,598 video games
<b>Time Period</b>	1980-2016

## Description:

This dataset contains sales data for video games across multiple platforms, publishers, and genres. Sales figures are broken down by region (North America, Europe, Japan, and Other) and aggregated globally.

## Variable Definition

Column Name	Description	Data Type	Example
Rank	Ranking position based on global sales (1=highest selling)	Integer	1, 50, 100
Name	Title of the video game	Text	Wii Sports
Platform	Gaming platform/console the game was released on	Text	Wii, PS4, X360
Year	Year the game was released	Integer	2006, 2015
Genre	Category/Type of game	Text	Action, Sports
Publisher	Company that published the game	Text	Nintendo, EA
NA_Sales	Sales in North America (in millions of units)	Decimal	41.9
EU_Sales	Sales in Europe (in millions of units)	Decimal	29.4
JP_Sales	Sales in Japan (in millions of units)	Decimal	3.7
Other_Sales	Sales in other regions (in millions of units)	Decimal	8.6
Global_Sales	Total worldwide sales (sum of all regional sales, in millions of units)	Decimal	82.6

### Genre Categories

The following genres were included in the dataset: Action, Sports, Shooter, Role-Playing, Platform, Racing, Puzzle, Misc, Fighting, Simulation, Adventure, Strategy

## Data Limitations and Notes

1. Incomplete 2016 Data: The dataset appears to cut off in early 2016, so year-over-year comparisons involving 2016 should be interpreted with caution.
2. Publishing Naming Variations: Some publishers may appear under different names (e.g. “Sony Computer Entertainment” vs “Sony”). Data cleaning was required to consolidate these.
3. Character Encoding: Some game titles contain special characters that may display incorrectly (e.g. Pokémon appears with encoding issues).
4. Sales Units: All sales figures are in millions of units sold, not revenue
5. Franchise Identification: Franchise groupings (e.g. “Mario” games) were identified using wildcard matching on game names, which may not capture all related titles.

## Calculated Fields

The following fields were calculated during analysis:

Calculated Field	Formula/Description
Regional Percentage	$(\text{Region\_Sales} / \text{Global\_Sales}) \times 100$ : Shows what percentage of a game or genre’s sales come from which region
Year-over-Year Market Share	Sum of regional sales per year/Sum of global sales per year $\times 100$ : Tracks how much each region’s share of the market changes over time.
Publisher Total Sales	SUMIF on Global_Sales where Publisher matches: Aggregate all sales for a given publisher
Franchise Sales	SUMIF with wildcards on Name Field (e.g. “Mario”: Groups sales by franchise using pattern matching

## Key Metrics Summary

Metric	Value
Total Global Sales	8,941.52 million units
North America Share	49.1% (4,392.95M)
Europe Share	27.2% (2,435.03M)
Japan Share	14.4% (1,291.30M)
Other Regions Share	9.2% (822.24M)
Top Publisher	Nintendo (1.787M units)
Top Franchise	Super Mario Bros (556M units)