# Dataset Name

VG Sales

# Description

Originally published on Zenodo (http://doi.org/10.5281/zenodo.2672927), this dataset is part of a final presentation titled “Analyzing the Global Video Game Sales.” It includes data on global video games sales from 1980 to 2015. There are 11 columns with 16,598 non-null values collected from Kaggle. The Python packages used to develop this dataset include Pandas, Zipfile, plotly, numpy, and copy. This dataset was normalized for upload into the GG! Repository.

#Variables

rank

A popularity rank based on the sales data contained within the dataset.

(numeric value)

title

Title of video game

(string)

platform

Platform the title is played on

(string)

year

Year of the title’s release

(numeric value: YYYY)

genre

The genre or genres associated with the title

(string)

publisher

Company responsible for the title’s publication

(string)

naSales

Sales data from North America

(in millions)

euSales

Sales data from Europe

(in millions)

jpSales

Sales data from Japan

(in millions)

otherSales

Sales data from other geographic regions

(in millions)

globalSales

Total global sales data

(in millions)

# Normalization Notes

Normalization included transforming variable names to suit the repository’s preferred naming conventions.

# Usage and Limitations

This dataset can be used to identify video game titles with the most recorded sales as of 2015.

It is possible to quickly identify titles with the most global sales through the “rank” variable. It is also possible to identify titles with more sales compared to others released on the same platform, in the same year, in the same genre, or by the same publisher.

While it is possible to identify titles with the most sales in a specific geographic region, this dataset only includes data specific to North America, Europe, and Japan. Other regions are rolled into “otherSales,” limiting the dataset’s usage for evaluating sales in Africa, Oceania, South America, or Asia at large.

# Authors and Acknowledgement

Jun Ying

# License

Attribution 4.0 International (CC BY 4.0)