Codebook
Video Game Sales Analysis Codebook
By: Miranda Gemme-Ellis

Dataset name:	vg_codebook
Dataset size:	2.7 Mb
Column count:	16
Row count:	16,327
Updated date:	2024-10-23

Description:

The objective of this project is to analyze video game sales worldwide, focusing on factors such as genre, game ranking, publisher, platform, and regional sales (including North America, Japan, and other regions). The goal is to analyze gaming industry trends, assess video games' performance across different localities, and understand the influence of different platforms and publishers on sales. Analyzing this data can help stakeholders, such as publishers and developers, to strategize marketing more effectively.

Column Attributes:

1	**Column name:**	**Rank**
	Column description:	Ranking of overall sales
	Source information:	Video Game Sales
	Data type:	Integer
	Unique non-missing value count:	16,327
	Missing value count:	0

Min	Mean	Median	Max	SD
1.00	8292.87	8295.00	16600.00	4792.67

2	**Column name:**	**Name**
	Column description:	The video games name
	Source information:	Video Game Sales
	Data type:	Factor
	Unique non-missing value count:	11,360
	Missing value count:	0

Categories with Smallest Values	Frequency	Categories with Largest Values	Frequency
¡Shin Chan Flipa en colores!	1	The LEGO Movie Videogame	8
.hack: Sekai no Mukou ni + Versus	1	FIFA 14	9
.hack//G.U. Vol.1//Rebirth	1	LEGO Marvel Super Heroes	9
.hack//G.U. Vol.2//Reminisce	1	Ratatouille	9
.hack//G.U. Vol.2//Reminisce (jp sales)	1	Need for Speed: Most Wanted	12

3	Column name:	Platform
	Column description:	Platform of the video games release (i.e. PC,PS4, etc.)
	Source information:	Video Game Sales
	Data type:	Factor
	Unique non-missing value count:	31
	Missing value count:	0

Categories with Smallest Values	Frequency	Categories with Largest Values	Frequency
GG	1	X360	1235
PCFX	1	Wii	1290
TG16	2	PS3	1304
3DO	3	PS2	2127
SCD	6	DS	2133

4	Column name:	Year
	Column description:	Year of video game's release
	Source information:	Video Game Sales
	Data type:	Numeric
	Unique non-missing value count:	39

Missing value count:		0		

Min	Mean	Median	Max	SD
1980.00	2006.41	2007.00	2020.00	5.83

5	**Column name:**	**Genre**

Column description:	Genre of video game
Source information:	Video Game Sales
Data type:	Factor
Unique non-missing value count:	12
Missing value count:	0

Categories with Smallest Values	Frequency	Categories with Largest Values	Frequency
Puzzle	571	Shooter	1282
Strategy	671	Role-Playing	1471
Fighting	836	Misc	1710
Simulation	851	Sports	2304
Platform	876	Action	3253

6	**Column name:**	**Publisher**

Column description:	Publisher of video game
Source information:	Video Game Sales
Data type:	Factor
Unique non-missing value count:	577
Missing value count:	0

Categories with Smallest Values	Frequency	Categories with Largest Values	Frequency
2D Boy	1	Konami Digital Entertainment	823
49Games	1	Ubisoft	918
989 Sports	1	Namco Bandai Games	928
Abylight	1	Activision	966

| Activision Blizzard | 1 | Electronic Arts | 1339 |

7	**Column name:**	**NA_Sales**

Column description:	Sales in North America (millions)
Source information:	Video Game Sales
Data type:	Numeric
Unique non-missing value count:	408
Missing value count:	0

Min	Mean	Median	Max	SD
0.00	0.27	0.08	41.49	0.82

8	**Column name:**	**EU_Sales**

Column description:	Sales in Europe (millions)
Source information:	Video Game Sales
Data type:	Numeric
Unique non-missing value count:	305
Missing value count:	0

Min	Mean	Median	Max	SD
0.00	0.15	0.02	29.02	0.51

9	**Column name:**	**JP_Sales**

Column description:	Sales in Japan (millions)
Source information:	Video Game Sales
Data type:	Numeric
Unique non-missing value count:	244
Missing value count:	0

Min	Mean	Median	Max	SD
0.00	0.08	0.00	10.22	0.31

10	**Column name:**	**Other_Sales**

Column description:	Sales in the rest of the world (millions)			
Source information:	Video Game Sales			
Data type:	Numeric			
Unique non-missing value count:	157			
Missing value count:	0			

Min	Mean	Median	Max	SD
0.00	0.05	0.01	10.57	0.19

11	Column name:	Global_Sales

Column description:	Total worldwide sales
Source information:	Video Game Sales
Data type:	Numeric
Unique non-missing value count:	621
Missing value count:	0

Min	Mean	Median	Max	SD
0.01	0.54	0.17	82.74	1.57

12	Column name:	NA_Sales_Prop

Column description:	The proportion of a game's total global sales that were made in North America.
Source information:	Video Game Sales
Data type:	Numeric
Unique non-missing value count:	2,952
Missing value count:	0

Min	Mean	Median	Max	SD
0.00	0.45	0.50	1.00	0.34

13	Column name:	EU_Sales_Prop

Column description:	The proportion of a game's total global sales that were made in Europe.

Source information:	Video Game Sales
Data type:	Numeric
Unique non-missing value count:	2,703
Missing value count:	0

Min	Mean	Median	Max	SD
0.00	0.23	0.20	1.00	0.25

14	Column name:	JP_Sales_Prop

Column description:	The proportion of a game's total global sales that were made in Japan.
Source information:	Video Game Sales
Data type:	Numeric
Unique non-missing value count:	1,746
Missing value count:	0

Min	Mean	Median	Max	SD
0.00	0.24	0.00	1.00	0.40

15	Column name:	Other_Sales_Prop

Column description:	The proportion of a game's total global sales that were made in other regions.
Source information:	Video Game Sales
Data type:	Numeric
Unique non-missing value count:	1,963
Missing value count:	0

Min	Mean	Median	Max	SD
0.00	0.06	0.06	1.00	0.08

16	Column name:	Decade

| Column description: | Categorizes the release year of each game into decade-based time periods. |
| Source information: | Video Game Sales |

Data type:		Factor	
Unique non-missing value count:		4	
Missing value count:		0	

Categories	Frequency	Cumulative Frequency	Percent
1980s	221	221	1.35
1990s	2,102	2,323	12.87
2000s	10,118	12,441	61.97
2010s	3,886	16,327	23.80