# GameStop Corp.

Case Study Part 2: Data Collection and Briefing Report

Dev D. Rabadia (NF1005560),

Mark C. Guanizo (NF1003833),

Miko L. Tan (NF1008647)

Master of Data Analytics, University of Niagara Falls Canada

DAMO-511-7: Winter 2025 Data Analytics Case Study 2

Professor Touraj Banirostam

March 12, 2025

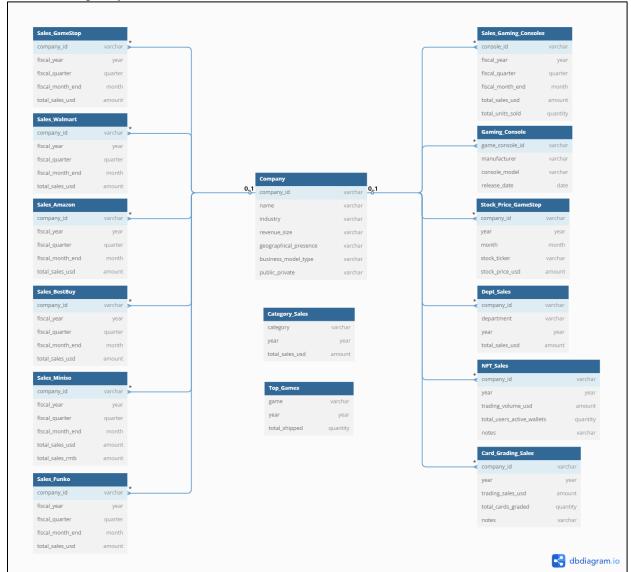
#### Introduction

In this research paper, the group aims to analyze the Marketing and Sales Operations of GameStop, applying the PESTEL framework to understand the key external factors influencing the company's performance. Specifically, the group will focus on the Economic, Social, and Technological factors, as these are particularly relevant to GameStop's operations in the retail gaming industry. The objective of this analysis is to provide insights into how these external forces shape the company's marketing and sales strategies and their potential impact on overall business performance.

#### **Data Collection and Challenges**

The group encountered several challenges in acquiring relevant data for analysis. While the group had access to GameStop's financial statements, obtaining comprehensive data for the PESTEL analysis proved to be difficult. The group explored various sources, including publicly available datasets, industry reports, and governmental databases. However, many of the datasets found were either incomplete, outdated, or misaligned with the group's research objectives.

To bridge these gaps, the group employed web scraping techniques to extract data from corporate websites, industry reports and other relevant online sources. Additionally, the group leveraged AI tools for additional inputs and insights, and then manually encoded the dataset to ensure consistency and usability. The group then evaluated the relevance of the collected data sources to ensure that only valuable information will be incorporated into the analysis. This collaborative effort enabled the group to establish a well-structured dataset that aligns with the study's objectives.



**Table 1:** Output of Data Collection

#### **Dataset Table and Variable Analysis**

Upon collecting the data, the team performed Exploratory Data Analysis (EDA) and Data Preprocessing using Python. The detailed steps and results have been fully documented in the *gamestop-data-collection.pdf*. The list of tables below summarizes key details about the variables in each dataset table, which will be used later in Power BI. These variables were selected to evaluate operational risks and opportunities in GameStop's business model, including:

- Industry trends in retail and gaming.
- Financial performance over time.
- The impact of NFTs and collectibles on revenue streams.
- Competitive landscape with other retailers.

 Table 2-1: Sales Table (Fact Table)

Variable Name	Source	Level of Measurement	Rationale for Selection
Company ID	Manually Defined Key	Nominal	It is a unique identifier for each company, used to distinguish between different companies.
Fiscal Year	Financial Reports	Ordinal	Represents the year in which the company's fiscal period falls, typically in a sequential order.
Fiscal Quarter	Financial Reports	Ordinal	Represents a quarter within the fiscal year, which is sequential in nature (Q1, Q2, Q3, Q4).
Fiscal Month End	Financial Reports	Ordinal/Interval	Represents the last day of each month in the fiscal year, typically sequential and ordered.
Total Sales (USD)	Financial Reports	Ratio	Measures the total revenue in USD, with an absolute zero and meaningful ratios.
Total Units Sold	Product Website	Ratio	Measures the number of units sold, with an absolute zero and meaningful ratios (if applicable).

**Table 2-2:** *Stock Price Table (Dimension Table)* 

Variable Name	Source	Level of Measurement	Rationale for Selection
Company ID	Manually Defined Key	Nominal	It is a unique identifier for each company, used to differentiate companies.
Year	Stock Market Data	Ordinal	Represents the calendar year, with an inherent order (e.g., 2023, 2024).
Month	Stock Market Data	Ordinal/Interval	Represents the months each year, ordered sequentially (January to December).
Stock Ticker	Stock Market Data	Nominal	A unique identifier symbol for the stock, used to distinguish one company from another.
Stock Price (USD)	Stock Market Data	Ratio	Represents the value of a company's stock in USD, with a true zero point and meaningful ratios.

 Table 2-3: Game Console Table (Dimension Table)

Variable Name	Source	Level of Measurement	Rationale for Selection
Game Console ID	Manually Defined Key	Nominal	It is a unique identifier for each game console, used to distinguish between different consoles.
Manufacturer Name	Product Website	Nominal	Represents the name of the manufacturer (e.g., Sony, Microsoft), used to categorize consoles.
Console Model Name	Product Website	Nominal	Identifies the specific model of the console (e.g., PlayStation 5, Xbox Series X), used to differentiate consoles.
Release Date	Product Website / Industry Reports	Interval/Ratio	Represents the date the console was released, with meaningful intervals (e.g., months or years).

**Table 2-4:** Category Sales Table (Dimension Table)

Variable Name	Source	Level of Measurement	Rationale for Selection
Category Name	Industry Reports	Nominal	It represents the classification of products which helps to categorize data.
Year	Industry Reports	Ordinal/Interval	Represents the calendar year, with an inherent order and meaningful intervals between years.
Total Sales (USD)	Industry Reports	Ratio	Measures total sales (e.g., in USD), with a true zero and meaningful ratios for comparison.

**Table 2-5:** Dept Sales Table (Dimension Table)

Variable Name	Source	Level of Measurement	Rationale for Selection
Company ID	Manually Defined Key	Nominal	It uniquely identifies each company, helping to differentiate between companies.
Department	Financial Reports	Nominal	Categorizes the type of products sold (Hardware, Software, Collectibles), helping to distinguish sales across different departments.
Year	Financial Reports	Ordinal/Interval	Represents the calendar year, ordered in a meaningful sequence (e.g., 2023, 2024).
Total Sales (USD)	Financial Reports	Ratio	Represents total sales revenue in USD, with a true zero point and meaningful ratios.

**Table 2-6:** *Top Games Table (Dimension Table)* 

Variable Name	Source	Level of Measurement	Rationale for Selection
Game Name	Industry Reports	Nominal	Represents the specific name of the game, used to identify and distinguish between games.
Year	Industry Reports	Ordinal/Interval	Represents the calendar year, ordered in a meaningful sequence (e.g., 2023, 2024).
Total Shipped	Industry Reports	Ratio	Represents the total number of game units shipped, with a true zero point and meaningful ratios.

**Table 2-7:** NFT Sales Table (Dimension Table)

Variable Name	Source	Level of	Rationale for Selection
		Measurement	
Company ID	Manually Defined	Nominal	It uniquely identifies each company, helping
	Key		to differentiate between companies.
Year	Trading Platform /	Ordinal/Interval	Represents the calendar year, ordered in a
	Industry Reports		meaningful sequence (e.g., 2023, 2024).
Trading Volume	Trading Platform /	Ratio	Represents the total trading volume in USD,
(USD)	Industry Reports		with a true zero point and meaningful ratios.
Total Users	Trading Platform /	Ratio	Represents the total number of users
	Industry Reports		involved, with a true zero and meaningful
			ratios.
Notes	Industry Reports	Nominal/Ordinal	Contains additional information, often text-
			based or ordered data that may describe
			specific conditions or observations.

 Table 2-8: Card Grading Sales Table (Dimension Table)

Variable Name	Source	Level of Measurement	Rationale for Selection
Company ID	Manually Defined Key	Nominal	A unique identifier for each company, used to distinguish between different companies.
Year	Company Website / Industry Reports	Ordinal/Interval	Represents the calendar year, ordered in a meaningful sequence (e.g., 2023, 2024).
Total Sales (USD)	Company Website / Industry Reports	Ratio	Represents the total sales revenue in USD, with a true zero point and meaningful ratios.
Total Cards Graded	Company Website / Industry Reports	Ratio	Represents the total number of cards graded, with a true zero point and meaningful ratios.
Notes	Company Website / Industry Reports	Nominal/Ordinal	Contains additional information, often text- based or ordered data that may describe specific conditions or observations.

 Table 2-9: Company Table (Dimension Table)

Variable Name	Source	Level of	Rationale for Selection
		Measurement	
Company ID	Manually Defined	Nominal	A unique identifier for each company, used to
	Key		distinguish between different companies.
Company or	Company Website /	Nominal	Identifies the specific company or console,
Console Name	Product Website		serving as a unique label for the entity.
Industry	Company Website /	Nominal	Represents the sector the company operates
	Industry Reports		in (e.g., Tech, Entertainment), used to
			categorize companies.
Revenue Size	Company Website /	Ordinal	Represents the revenue size of the company
	Industry Reports		(small, medium, large), ordered in terms of
			size categories.
Geographical	Company Website /	Ordinal/Nominal	Indicates whether the company operates
Presence	Industry Reports		locally, regionally, or globally, either as
			categories or ordered levels.
Business Model	Company Website /	Nominal	Represents the business model of the
Type	Industry Reports		company (e.g., B2B, B2C, subscription-
			based), used for classification.

Public or Private	Company Website /	Nominal	Distinguishes whether a company is publicly
	Industry Reports		traded or privately held, an important
			classification.

## **Descriptive Statistics**

Table 3-1: Sales Table (Fact Table) Statistics, Missing Values & Outliers

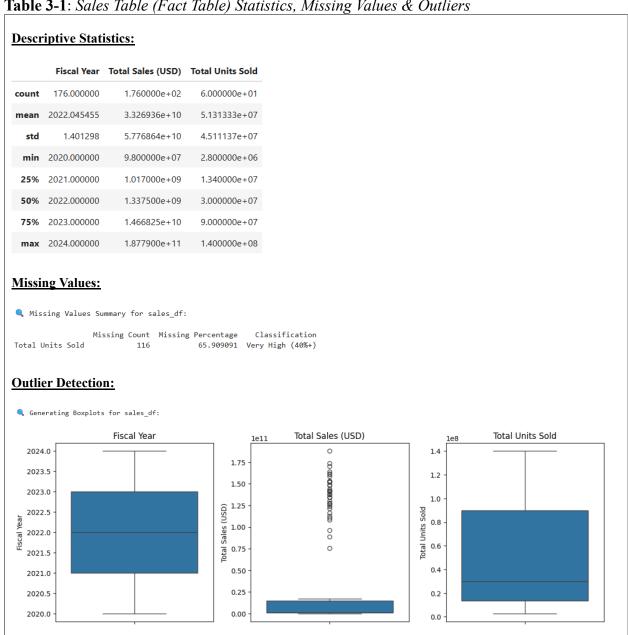


Table 3-2: Stock Price Table (Dimension Table) Statistics, Missing Values & Outliers

# **Descriptive Statistics:**

	Year	Stock Price (USD)
count	60.000000	60.000000
mean	2022.000000	34.876500
std	1.426148	35.711445
min	2020.000000	3.250000
25%	2021.000000	22.350000
50%	2022.000000	22.350000
75%	2023.000000	22.350000
max	2024.000000	200.000000

## **Missing Values:**

- Missing Values Summary for stock\_price\_df:
- ✓ No missing values in stock\_price\_df!

#### **Outlier Detection:**

Generating Boxplots for stock\_price\_df:

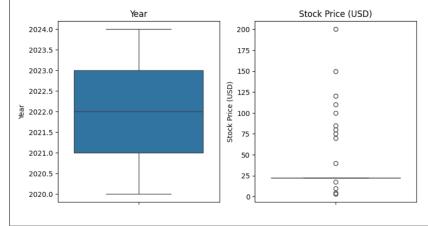


Table 3-3: Game Console Table (Dimension Table) Statistics, Missing Values & Outliers

#### **Descriptive Statistics:**

	Release Date
count	6
mean	2021-06-18 20:00:00
min	2019-08-01 00:00:00
25%	2020-11-10 00:00:00
50%	2020-11-11 00:00:00
75%	2021-07-17 12:00:00
max	2024-11-07 00:00:00

## **Missing Values:**

- Missing Values Summary for game\_console\_df:
- ✓ No missing values in game\_console\_df!

#### **Outlier Detection:**

- Generating Boxplots for game\_console\_df:
- ${\color{red} \underline{ {\bf \Lambda}}}$  No numerical columns in game\_console\_df for boxplot visualization.

Table 3-4: Category Sales Table (Dimension Table) Statistics, Missing Values & Outliers

#### **Descriptive Statistics:** Year Total Sales (USD) 20.000000 2.000000e+01 count 4.048000e+07 mean 2022.000000 4.205354e+07 std 1.450953 7.400000e+06 min 2020.000000 **25%** 2021.000000 8.875000e+06 **50%** 2022.000000 2.405000e+07

#### **Missing Values:**

**75%** 2023.000000

max 2024.000000

Missing Values Summary for category\_sales\_df:

5.070000e+07

1.500000e+08

No missing values in category\_sales\_df!

## **Outlier Detection:**

2020.0

Generating Boxplots for category\_sales\_df:

Total Sales (USD) Year 1e8 2024.0 1.4 2023.5 1.2 2023.0 Total Sales (USD) 2022.5 2022.0 2021.5 2021.0 2020.5

0.2

Table 3-5: Dept Sales Table (Dimension Table) Statistics, Missing Values & Outliers

#### **Descriptive Statistics:** Year Total Sales (USD) 15.00000 1.500000e+01 count mean 2022.00000 1.878453e+09 1.46385 9.258897e+08 min 2020.00000 5.964000e+08 **25%** 2021.00000 9.878000e+08 **50%** 2022.00000 1.842900e+09 **75%** 2023.00000 2.753350e+09 max 2024.00000 3.140000e+09 **Missing Values:** Missing Values Summary for dept\_sales\_df: ✓ No missing values in dept\_sales\_df! **Outlier Detection:** Generating Boxplots for dept\_sales\_df: Total Sales (USD) Year 2024.0 3.0 2023.5 2023.0 2.5 Total Sales (USD) 2022.5 2022.0 2021.5

1.0

2021.0

2020.5

Table 3-6: Top Games Table (Dimension Table) Statistics, Missing Values & Outliers

# **Descriptive Statistics:** Year Total Shipped 15.00000 1.500000e+01 count mean 2022.00000 1.574867e+07 1.46385 1.097849e+07 std 2020.00000 3.910000e+06 25% 2021.00000 9.750000e+06 2022.00000 1.289000e+07 75% 2023.00000 1.853000e+07 max 2024.00000 4.744000e+07 **Missing Values:** Missing Values Summary for top\_games\_df: ✓ No missing values in top\_games\_df! **Outlier Detection:** Generating Boxplots for top\_games\_df: Total Shipped 2024.0 2023.5 4 2023.0 Total Shipped 2022.5 2022.0 2021.5 2021.0 2020.5 1 2020.0

Table 3-7: NFT Sales Table (Dimension Table) Statistics, Missing Values & Outliers

#### **Descriptive Statistics:** Year Trading Volume (USD) Total Users (Active Wallets) 18.000000 1.300000e+01 1.400000e+01 count 1.032214e+08 mean 2022.166667 5.392070e+12 1.676527e+08 std 1.424574 9.882791e+12 min 2020.000000 1.000000e+08 1.250000e+05 **25%** 2021.000000 2.100000e+09 5.112500e+05 **50%** 2022.000000 2.300000e+10 2.830000e+07 **75%** 2023.000000 5.000000e+12 8.500000e+07

#### **Missing Values:**

max 2024.000000

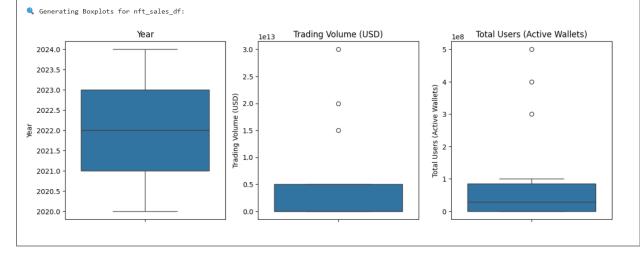
Missing Values Summary for nft\_sales\_df:

3.000000e+13

Missing Count Missing Percentage Classification
Trading Volume (USD) 5 27.777778 High (20-40%)
Total Users (Active Wallets) 4 22.22222 High (20-40%)

5.000000e+08

#### **Outlier Detection:**



**Table 3-8:** Card Grading Sales Table (Dimension Table) Statistics, Missing Values & Outliers

#### **Descriptive Statistics:** Year Trading Sales (USD) Total Cards Graded 4.000000e+00 4.000000 3.0 count mean 2022.500000 400000000.0 3.381660e+06 1.290994 100000000.0 4.414606e+06 std min 2021.000000 300000000.0 9.870920e+05 **25%** 2021.750000 350000000.0 1.146773e+06 **50%** 2022.500000 400000000.0 1.269774e+06 **75%** 2023.250000 450000000.0 3.504661e+06 **max** 2024.000000 500000000.0 1.000000e+07

#### **Missing Values:**

Missing Values Summary for card\_grading\_sales\_df:

Missing Count Missing Percentage Classification
Trading Sales (USD) 1 25.0 High (20-40%)

#### **Outlier Detection:**

Generating Boxplots for card\_grading\_sales\_df:

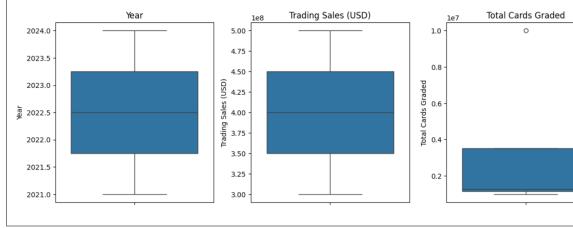
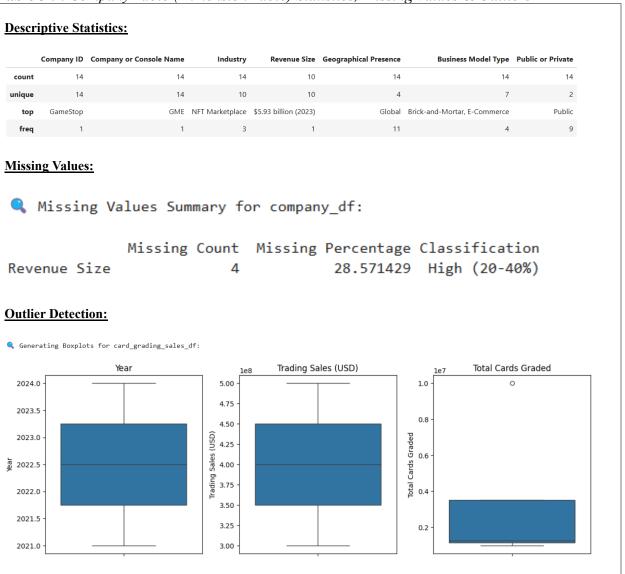


Table 3-9: Company Table (Dimension Table) Statistics, Missing Values & Outliers



## **Handling Missing Data**

**Table 4-1:** Sales Table (Fact Table) Handling of Missing Data

# Python Code: # Sales Table - Total Units Sold - Default missing values to 0 because this feature is only applicable to Gaming Consoles sales\_df.loc[:, "Total Units Sold"] = sales\_df["Total Units Sold"].fillna(0) Rationale: The missing Total Units Sold was defaulted to 0 because this feature is only applicable to Gaming Consoles. The missingness only happens to records that are not related to Gaming Consoles.

#### **Table 4-2:** NFT Sales Table (Dimension Table) Handling of Missing Data

#### **Python Code:** # NFT Sales Table - Trading Volume (USD) and Total Users - Forward fill missing data for 2024, retain 0 for 2020 # Sort the dataframe by 'Company ID' and 'Year' to ensure proper grouping and chronological order nft\_sales\_df = nft\_sales\_df.sort\_values(by=['Company ID', 'Year']) columns\_to\_impute = ['Trading Volume (USD)', 'Total Users (Active Wallets)'] # Handle 2020 missing data (set to 0 for each company independently) for col in columns\_to\_impute: # Set the 2020 missing values to 0 for each company independently nft\_sales\_df.loc[(nft\_sales\_df['Year'] == 2020) & (nft\_sales\_df[col].isnull()), col] = 0 # Forward fill missing values within each company for company\_id, company\_data in nft\_sales\_df.groupby('Company ID'): for col in columns\_to\_impute: # Apply forward fill directly company\_data[col] = company\_data[col].ffill() # Update the original dataframe with forward-filled values for each company nft\_sales\_df.update(company\_data) # Print the updated dataframe print(nft\_sales\_df)

#### Rationale:

The missing Trading Volume (USD) and Total Users was imputed using forward fill while retaining 0 for the 1<sup>st</sup> date which is 2020. Forward fill was utilized since the missing data follows a time-based or sequential pattern.

## Table 4-3: Card Grading Sales Table (Dimension Table) Handling of Missing Data

```
Python Code:
# Card Grading Sales Table - Trading Sales - Forward fill missing data for 2024, retain 0 for 2020
 # Sort the dataframe by 'Company ID' and 'Year' to ensure proper grouping and chronological order
card_grading_sales_df = card_grading_sales_df.sort_values(by=['Company ID', 'Year'])
 # Define the columns to impute
columns_to_impute = ['Trading Sales (USD)']
 # Handle 2020 missing data (set to 0 for each company independently)
for col in columns to impute:
    # Set the 2020 missing values to 0 for each company independently
    card_grading_sales_df.loc[(card_grading_sales_df['Year'] == 2020) & (card_grading_sales_df[col].isnull()), col] = 0
 # Forward fill missing values within each company
for company id, company data in card grading sales df.groupby('Company ID'):
    for col in columns to impute:
        # Apply forward fill directly
        company_data[col] = company_data[col].ffill()
        # Update the original dataframe with forward-filled values for each company
        card_grading_sales_df.update(company_data)
 # Print the updated dataframe
print(card_grading_sales_df)
```

#### **Rationale:**

The missing Trading Sales (USD) was imputed using forward fill while retaining 0 for the 1<sup>st</sup> date which is 2020. Forward fill was utilized since the missing data follows a time-based or sequential pattern.

**Table 4-4:** Company Table (Dimension Table) Handling of Missing Data

```
Python Code:
# Company Table - Revenue Size - Tag missing values as Not Available
company_df.loc[:, "Revenue Size"] = company_df["Revenue Size"].fillna("Not Available")
```

#### Rationale:

The missing Revenue Size was imputed with a "Not Available" value to signify data that data was unavailable at the time of capture.

## **Handling Outliers**

For the Sales Table, which includes **Total Sales (USD)**, no action was taken, as the data represents historical sales from multiple companies and may naturally contain outliers due to variability in sales volumes. These outliers will be identified and visualized during the analysis.

Similarly, for the Category Sales Table, which also reflects **Total Sales (USD)**, no action was taken as it represents historical data from different companies. Outliers are expected, and their impact will be analyzed and visualized.

In the case of NFT Sales Table, specifically **Trading Volume (USD)** and **Total Users**, no action was taken since these are historical data points that may include outliers due to market fluctuations. These outliers will be visualized and considered during the analysis.

Lastly, for Card Grading Sales Table, which tracks **Total Cards Graded**, no action was taken as this historical data may include potential outliers caused by variations in grading volumes across different companies. These outliers will also be visualized and analyzed.

#### Conclusion

In conclusion, this research paper provides a comprehensive analysis of GameStop's Marketing and Sales Operations through the lens of the PESTEL framework, focusing on the Economic, Social, and Technological factors. Despite challenges in data collection, the team utilized web scraping, AI tools, and manual encoding to build a robust and consistent dataset.

This completed dataset, now fully prepared, is ready for further analysis and visualization in Power BI. The insights derived from this research will enhance the understanding of how external factors influence GameStop's marketing and sales strategies, contributing to a deeper analysis of the company's business performance.

#### References

Best Buy. (n.d.). Quarterly results. Retrieved from

https://investors.bestbuy.com/financials/quarterly-results/default.aspx.

Blur. (n.d.). *Blur blog*. Retrieved from https://blur.io/blog.

Card Ladder. (n.d.). Sports card market analysis. Retrieved from https://www.cardladder.com.

Chainalysis. (n.d.). *Blockchain analytics and insights*. Retrieved from https://www.chainalysis.com.

Coindesk. (n.d.). *NFT market insights*. Retrieved from https://www.coindesk.com.

CoinGecko. (n.d.). *Crypto and NFT market overview*. Retrieved from https://www.coingecko.com.

Collectors Universe. (n.d.). *Sports card grading and authentication*. Retrieved from https://www.collectorsuniverse.com.

CoinMarketCap. (n.d.). NFT price tracking. Retrieved from https://coinmarketcap.com.

CryptoSlam. (n.d.). NFT market data. Retrieved from https://www.cryptoslam.io.

DevToDev. (2023, March). *Game market overview: Important reports published in March 2023*.

Retrieved from https://www.devtodev.com/resources/articles/game-market-overview-the-most-important-reports-published-in-march-2023.

Dune Analytics. (n.d.). OpenSea analysis. Retrieved from https://dune.com/siege/OpenSea.

- GameStop. (2024). Fourth quarter and fiscal year 2023 results. Retrieved from https://investor.gamestop.com/news-releases/news-release-details/gamestop-reports-fourth-quarter-and-fiscal-year-2023-results.
- GameStop. (2024). *Third quarter 2024 results*. Retrieved from https://investor.gamestop.com/news-releases/news-release-details/gamestop-discloses-third-quarter-2024-results.
- Macrotrends. (n.d.). *Amazon revenue 2010-2024*. Retrieved from https://www.macrotrends.net/stocks/charts/AMZN/amazon/revenue.
- Macrotrends. (n.d.). *Funko revenue 2010-2024*. Retrieved from https://www.macrotrends.net/stocks/charts/FNKO/funko/revenue.
- Macrotrends. (n.d.). *GameStop revenue 2010-2024*. Retrieved from https://www.macrotrends.net/stocks/charts/GME/gamestop/revenue.
- Macrotrends. (n.d.). *Walmart revenue 2010-2025*. Retrieved from https://www.macrotrends.net/stocks/charts/WMT/walmart/revenue.

Microsoft. (n.d.). Investor relations. Retrieved from https://www.microsoft.com/en-us/Investor.

Miniso. (n.d.). *Investor relations*. Retrieved from https://ir.miniso.com/.

Nintendo. (n.d.). *Investor relations*. Retrieved from https://www.nintendo.co.jp/ir/en/index.html.

NonFungible. (n.d.). NFT trends and marketplace data. Retrieved from https://nonfungible.com.

OpenSea. (n.d.). OpenSea blog. Retrieved from https://opensea.io/blog.

- PSA Card. (n.d.). *Professional sports authenticator (PSA)*. Retrieved from https://www.psacard.com.
- Sensor Tower. (n.d.). *State of mobile gaming 2024*. Retrieved from https://sensortower.com/blog/state-of-mobile-gaming-2024.

- Sony. (n.d.). *Investor relations*. Retrieved from https://www.sony.net/SonyInfo/IR/.
- Sports Collectors Daily. (n.d.). *Sports collectibles news and insights*. Retrieved from https://www.sportscollectorsdaily.com.
- Statista. (n.d.). *Global digital gaming revenue statistics*. Retrieved from https://www.statista.com/statistics/1344686/global-digital-gaming-revenue/.
- Statista. (n.d.). NFT market research and statistics. Retrieved from https://www.statista.com.
- The Block. (n.d.). NFT trends and analysis. Retrieved from https://www.theblock.co.
- The Guardian. (2024, September 11). *Sony reveals \$700 PlayStation 5 Pro console*. Retrieved from https://www.theguardian.com/games/article/2024/sep/11/sony-reveals-700-playstation-5-pro-console.
- Wikipedia. (n.d.). *Nintendo Switch*. Retrieved from https://en.wikipedia.org/wiki/Nintendo Switch.
- Wikipedia. (n.d.). *PlayStation 5*. Retrieved from https://en.wikipedia.org/wiki/PlayStation\_5.
- Wikipedia. (n.d.). *Xbox Series X and Series S*. Retrieved from https://en.wikipedia.org/wiki/Xbox Series X and Series S.
- WN Hub. (n.d.). *Gaming market analysis*. Retrieved from https://wnhub.io/news/other/item-46612.
- VGChartz. (n.d.). *Video game sales ranking*. Retrieved from https://www.vgchartz.com/games/games.php.
- Yahoo Finance. (n.d.). *GameStop historical stock data*. Retrieved from https://finance.yahoo.com/quote/GME/history/.