Analysis of a One-Sixth Scale  
Action Figure Collection

Leonard Armstrong  
IST-662 Scripting for Data Science  
February 7, 2021

# Introduction

Figure 1- A 1964 G.I. Joe Action Soldier

American children born in the 1960s experienced the advent of the action figure with a toy named *G.I. Joe*®, conceived by licensing agent Stan Weston and brought to life in 1964 by Hasbro executive Don Levine. To combat Hasbro executives’ concerns about whether or not boys would play with dolls, Levine insisted that his team refer to G.I. Joe as an “*action soldier*” or “*action figure*.” The phrase “*Action Soldier*” can be seen on the original G.I. Joe box shown in Figure 1. Anyone who referred to G.I. Joe as a “doll” risked losing their job at Hasbro. Thus, to counter the cultural biases of the era, the “action figure” was born.

The original G.I. Joe was an articulated[[1]](#footnote-1) figure slightly under 12-inches tall, making him one-sixth scale compared to the average male’s height. G.I. Joe sales’ early success led to the production of one-sixth scale figures by other companies such as Captain Action, Marx Knights, and Marx Best of the West, just to name a few. One-sixth scale figures continued production into the mid-1970s with the G.I. Joe brand leading the way. However, in the 1970s, in response to lower sales correlated to the Vietnam War, G.I. Joe evolved from a military figure to an adventurer with the G.I. Joe Adventure Team.

One-sixth scale action figures faded away in the late 70s to be replaced by smaller, less articulated figures such as Kenner’s Star Wars line of 3 ¾-inch toys. A key driver of this change was the rising price of petroleum-based plastics. Even G.I. Joe reinvented himself in 1982 with the 3 ¾-inch *“A Real American Hero”* line of toys, a distinctly different and more expansive concept from the original everyman G.I. Joe. With the end of G.I. Joe’s Adventure Team line, one-sixth scale action figures were very sparse from the late 1970s through the entire decade of the 1980s.[[2]](#footnote-2)

However, the one-sixth scale figure returned in 1992 with G.I. Joe once again blazing the trail. Hasbro introduced one-sixth scale versions of some of the more popular characters of the G.I. Joe: A Real American Hero line, initially available only as 3 ¾-inch figures. Their timing was impeccable. New one-sixth scale G.I. Joe figures caught the attention of young fathers and mothers that remembered the toy from their youth and could now share their experiences with their children. Alternatively, young professionals with new *disposable income* could start collecting the figures they remembered from days past.

The analysis provided herein looks at one such collection of 1:6 scale action figures, collected from 1996 to the present.

# The Data Set

The action figure data set under analysis is a custom database of the owner’s collection, affectionately named, ***Joebase***. The source data is in a multi-table relational database, described in the entity-relationship diagram shown in Figure 2.

Diagram

Description automatically generated

Figure 2- Action Figure Database ER Diagram

The relational model was denormalized into a single, wide table structure, written to a single comma-separated values (CSV) file.[[3]](#footnote-3) The primary unit of the table is an action\_figure record. The one-to-many relationship between an action figure and its genres is translated with one-hot encoding resulting in 30 new binary fields added to each action figure record. This research uses a snapshot of the database made on February 9, 2021, 6:10P EST.[[4]](#footnote-4)

## Data Elements

The table in Figure 3 lists all the fields in the single, denormalized input table used in this analysis. The CSV input file is generated from a single T-SQL query against the database shown in Figure 2.[[5]](#footnote-5)

| Column[[6]](#footnote-6) | Type | Description |
| --- | --- | --- |
| ProductId,  product\_id | Integer | Unique identifier of a product. A product is an atomic unit of sales. A product typically contains one figure, but in some cases, it may include multiple figures. |
| FigureId,  figure\_id | Integer | Unique identifier of a single action figure. An action figure is a single 1:6-scale plastic human form. |
| Manufacturer,  manufacturer | String | Name of the company that manufactured the action figure. |
| Product,  product\_name | String | Name of the product as sold. |
| Release Year,  year | Integer | The year that the manufacturer released the action figure, which is not necessarily the year the action figure was added to the collection. However, almost all figures produced after 1995 were acquired within a year of their release. |
| Product Description,  product\_descr | String | A longer description of the product. |
| Product Type | String | In all cases, this field will contain the value “1:6 figure”. This field is in the database for expansion to additional collectibles other than one-sixth scale action figures. |
| Purchased From,  seller | String | Retailer or another source that sold or gifted the product. |
| purchase\_price,  price | Float | The price paid for the product. |
| exclusive\_to\_retailer\_id | Integer | Unique identifier of a retailer that sold the product as an exclusive, if applicable. |
| af\_descr | String | A longer description of the action figure. This description often repeats information in product\_descr or is left null when the parent product contains only a single action figure. |
| likeness | String | The real person whose likeness is represented by the action figure. |
| storage\_location | String | Indicator of where the action is figure is either stored or displayed in the owner’s collection. |
| Genres (30 fields):  Adventure, Air Force, Armor, Army, Astronaut, Avengers, Celebrity, Civilian, Coast Guard, Comics, DC Comics, Fashion, Fire Fighter, Foreign, Horror, Knight, Marines, Martial Arts, Marvel Comics, Navy, Police, RAH/Cobra, Sci-Fi, Sports, Spy, TV/Film, Warrior, Western, World Leader, X-Men | Binary | Thirty, one-hot encoded genres that classify action figures. Multiple genres often categorize each action figure. For example, a figure of Robert Downey Jr.’s *Iron Man* is simultaneously part of the “**TV/Film**,” “**Comics,**” and “**Marvel Comics**” genres. |

Figure 3- Fields in the Input Data File

## Data Cleansing and Shaping

After reading the input file, the data is cleansed. First, the renaming of select data fields takes place. The renaming simplifies some attribute names and gives many names the structure of a single snake case string without whitespace.

After the renaming step, any data records containing a null figure\_id are removed.[[7]](#footnote-7) Then, null year values were set to 0 to signify a special case of an unknown purchase year. After assuring that all figure\_id and year rows had values, the cleansing process converts these two fields from float64 to int64.[[8]](#footnote-8)

Finally, a new field, Half Decade, is derived from the year field and added to the data set. For example, the years 1995 through 1999, inclusive, are mapped to the half-decade starting in 1995. Likewise, 2000-2004 are mapped to the half-decade beginning in 2000, and so on. The addition of this field helps to simplify the question detailed in **Section 3.1**.

## Basic Descriptive Statistics of the Data Set

The base data set, immediately after intake, contains 497 records with 43 attributes per record. After cleansing, the data set consists of 491 records with 44 attributes per record.

### Manufacturer and Seller/Retailer

Figure 4 shows descriptive statistic counts for the Manufacturer and Seller attributes. The lower volume of recorded sellers (retailers and other sellers) results from this information not being recorded for many early purchases.

|  |  |  |
| --- | --- | --- |
|  | Manufacturer | Seller |
| Count Non-null | 484 | 262 |
| Unique Values | 66 | 48 |

Figure 4- Counts of Non-null and Unique Manufacturers and Sellers

### Year of Release

Years-of-release of figures in the collection span from 1964 (vintage G.I. Joe figures) to figures released in 2020.[[9]](#footnote-9) Figure 5 shows a distribution of figure counts across these years. The collection has over 60 action figures released in 2001. Action figures released over the past decade (2011-2020) come in at more modest numbers of ten or less per year.

Chart, histogram

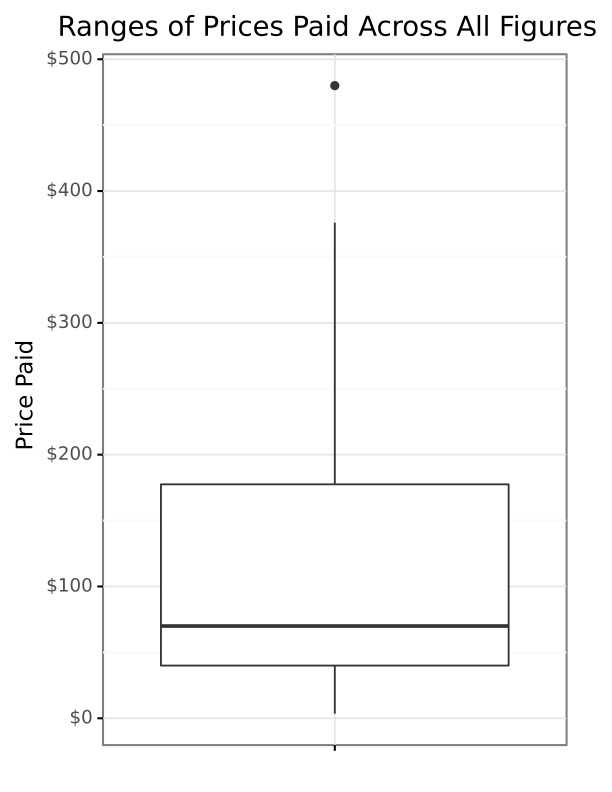
Description automatically generated

Figure 5- Distribution of Figures by Year of Release

The chart shows that the collection contains a small volume of figures from the “vintage” era of one-sixth scale figures, 1964-1976; however, most action figures in the collection were produced in 1996 and after. In fact, the collection started with the purchase of a single reproduction, G.I. Joe Action Soldier figure that was packaged and sold along with a book titled *G.I. Joe: The Story Behind the Legend[[10]](#footnote-10)* as a “Masterpiece Edition” boxed set.

### Prices

There are 166 figures with non-zero prices recorded with a mean price of $110.31. Figure 6’s boxplot and associated quartile statistics show the range of prices across all action figures in Joebase.

Table

Description automatically generated

Figure 6- Range of Action Figure Prices

# Analysis

Analyses of the Joebase data set targeted the following two questions:

* What are the average prices paid, per year of release, for a figure?
* How has the collection focus shifted over the years?

## Question 1 - What are the average prices paid, per year or release, for a figure?

To understand price differences over time, the data set was grouped by year, eliminating figures with unrecorded prices and figures whose production year is unrecorded or unknown. Figure 7 shows the rise in the mean purchase price of action figures in the collection produced over the 26 years from 1995 to 2020. One can see a trend of prices rising at a superlinear rate.

Chart, line chart

Description automatically generated

Figure 7- Mean Figure Price Per Year of Release

However, looking at mean averages across time only tells part of the story. Figure 8 shows a more detailed picture of price ranges during the period. While one still sees the superlinear price rise, the use of box-and-whisker plots reveals the central tendency of action figure prices (in this case, median instead of the mean) and the growing price variations between higher end and lower end figures. Looking at the interquartile ranges (IQR) for action figures produced from 1998 through 2008, the spread is relatively small, with all years having an IQR of less than $20.00.[[11]](#footnote-11) Then, starting in 2009, the IQRs are much larger, excepting 2010. During that period, the IQR range is between $30.00 in 2014 and $106.25 in 2019.

In summary, Figure 8 allows one to visualize both general price increases of figures over the years and the growth span of action figure price ranges.

Chart, bar chart

Description automatically generated

Figure 8- Range of Prices Paid Per Production Year

## Question 2 - How has the collection focus shifted over the years?

To understand “*shift in focus,*” the term “*focus*” must be defined. The definition of “focus” in Joebase is the genres that categorize each action figure. The collection is classified by 30 genres ranging, alphabetically, from “Adventure” to “X-Men.” The table in Figure 9 shows the 30 genres and the volume of figures currently classified by each genre.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Genre | Volume | Genre | Volume | Genre | Volume |
| TV/Film | 128 | World Leader | 18 | Warrior | 6 |
| Comics | 95 | Spy | 13 | Fire Fighter | 6 |
| Army | 64 | Civilian | 13 | Western | 6 |
| Foreign | 50 | Marvel Comics | 11 | DC Comics | 5 |
| Adventure | 43 | Astronaut | 10 | Fashion | 5 |
| Sci-Fi | 36 | Sports | 9 | Armor | 4 |
| Horror | 34 | Martial Arts | 9 | Knight | 2 |
| Navy | 31 | Celebrity | 8 | X-Men | 1 |
| Air Force | 27 | RAH/Cobra | 8 | Avengers | 1 |
| Marines | 24 | Police | 7 | Coast Guard | 1 |

Figure 9- All Genres, Sorted by Volume of Figures in Each Genre

As shown in Figure 2’s ER diagram, there is a many-to-many relationship between action figures and genre. Each genre may contain multiple action figures, as Figure 9 shows, but also each action figure may be classified by many genres. For example, the recent glut of action figures related to Marvel and DC comic book movies (e.g., *Iron Man*, *The Dark Knight*) are in the TV/Film genre, the Comic genre, and either the Marvel Comics or DC Comics genres. Thus, the sum of values in Figure 9 is 675, a volume higher than the number of figures in the collection currently.[[12]](#footnote-12)

The treemap shown in Figure 10 helps to visualize each genre's relative size.

Chart, treemap chart

Description automatically generated

Figure 10- Treemap Comparing Relative Genre Sizes

While Figure 10 provides a static view of the genres that dominate the collection, it does not give any sense of how the collection's focus has shifted over time. For the historic view, only the top six genres in terms of volume are compared. These six genres each classify more than 35 figures or at least 7% of the total Joebase collection. Also, instead of viewing the changes year to year, a coarser scale of a half-decade is used. Each figure was given a half-decade attribute during the data cleansing stage using the following formula.

|  |
| --- |
| ***# Add a half-decade field as a string type. This works out better for graphing.***  half\_decade = fig\_data["year"] - fig\_data["year"].mod(5)  fig\_data["Half Decade"] = [str(x) for x in half\_decade] |

Figure 11 visualizes the changes in the size of the six top genres over time. Each bar represents 100% of the six genres for the half-decade labeled on the x-axis. The volumes of each genre from that half-decade are compared, relative to the other five genres. This chart is independent of the fact that the overall volume of action figures collected from 2000-2004 is far greater than the overall volume of figures collected during the 2015-2019 half-decade. It instead looks at the focus of genre sizes relative to each other during each half-decade.

Chart, bar chart

Description automatically generated

Figure 11- Relative Size of Top Genres Over Time

The chart shows that the Adventure, Army, and Foreign genres declined over time to a point where purchases in these genres stopped after 2015. Conversely, Comics, TV/Film, and Sci-Fi genres all grow during the 27 years from 1995 to 2021, inclusive. Adventure, Army, and Foreign figures are associated with the one-sixth scale G.I. Joe line started in 1995 and ended by Hasbro around the early 2010s. On the other side, the onset and growth of comic-based movies with artistic character designs relate to the rise of the Comics, TV/Film, and Sci-Fi genres. Producers such as Hot Toys, Sideshow, Medicom all have strong sales with high-quality products based on these movies.

# Conclusion

An analysis of Joebase, a structured database of one-sixth scale action figures, is presented. The analysis focuses on two questions: (1) What are the average prices paid, per year of release, for a figure? (2) How has the collection focus shifted over the years? With respect to prices, a superlinear increase in figures prices was shown as well as a widening gap in figure prices. The focus of figure purchases was also shown to shift from Army, Adventure, and Foreign figures to TV/Film, Comics, and Sci-Fi figures.

1. Movable joints and body parts allowing the figure to be posed. [↑](#footnote-ref-1)
2. Occasional exceptions exist, such as the 12-inch versions of Mego’s World’s Greatest Super-Heroes™ figures, introduced in 1979. [↑](#footnote-ref-2)
3. The denormalized structure maintained all significant information except for the series data which is not part of this analysis. [↑](#footnote-ref-3)
4. As the collection is still growing, the associated database is updated on a regular basis. [↑](#footnote-ref-4)
5. The database explorer tool, **DBeaver Enterprise**, was used to perform the query and write the results to an output CSV file. [↑](#footnote-ref-5)
6. In the cases where the **Column** description contains two values, one represents the name that is present in the CSV file’s header row and the other representing a column renaming performed during data cleansing. [↑](#footnote-ref-6)
7. One such row existed due to an erroneous record in the original RDBMS that has since been corrected. [↑](#footnote-ref-7)
8. Numpy’s int64 data type does not allow for null or NA values. Thus, since these fields contained null values in the original source CSV, these fields were read as float64 which does allow permit NA values. [↑](#footnote-ref-8)
9. As of this writing, no figures released in 2021 have been added to the collection although some are on order and may make their way into the collection shortly. [↑](#footnote-ref-9)
10. Levine, Don with John Michlig, *G.I. Joe: The Story Behind the Legend, An Illustrated History of America’s Greatest Fighting Man*. Chronicle Books, 1996. [↑](#footnote-ref-10)
11. Max IQR for this range is $17.51 in 2004. [↑](#footnote-ref-11)
12. But give it time… ☺ [↑](#footnote-ref-12)