

# Amazon Best Selling Books Analysis using Python



Written By

***Nnenna Njoku***

# INTRODUCTION

In the early days, Amazon functioned exclusively as a bookstore—a fact that might surprise those familiar only with its current status as a retail giant. Nevertheless, many avid readers turn to Amazon to discover their next literary journey.

Amazon provides a wide selection of books for its customers. Books are never judged by their cover, but everyone should also be able to rate books (not just journalists or editors). Amazon enables customers to rate books by giving them scores between 1-5 and a detailed review text to help others decide whether to purchase a particular book or not.

In this project, I will be presenting insights derived from the analysis of Amazon's best-selling book between 2009 and 2019.

Due to the restriction on Amazon website scraping, I downloaded the dataset from [Github](#)

The analysis was done using [Python Jupyter Notebook](#).

## About Dataset

Amazon best-selling book for years 2009 to 2019. The dataset has 8 columns.

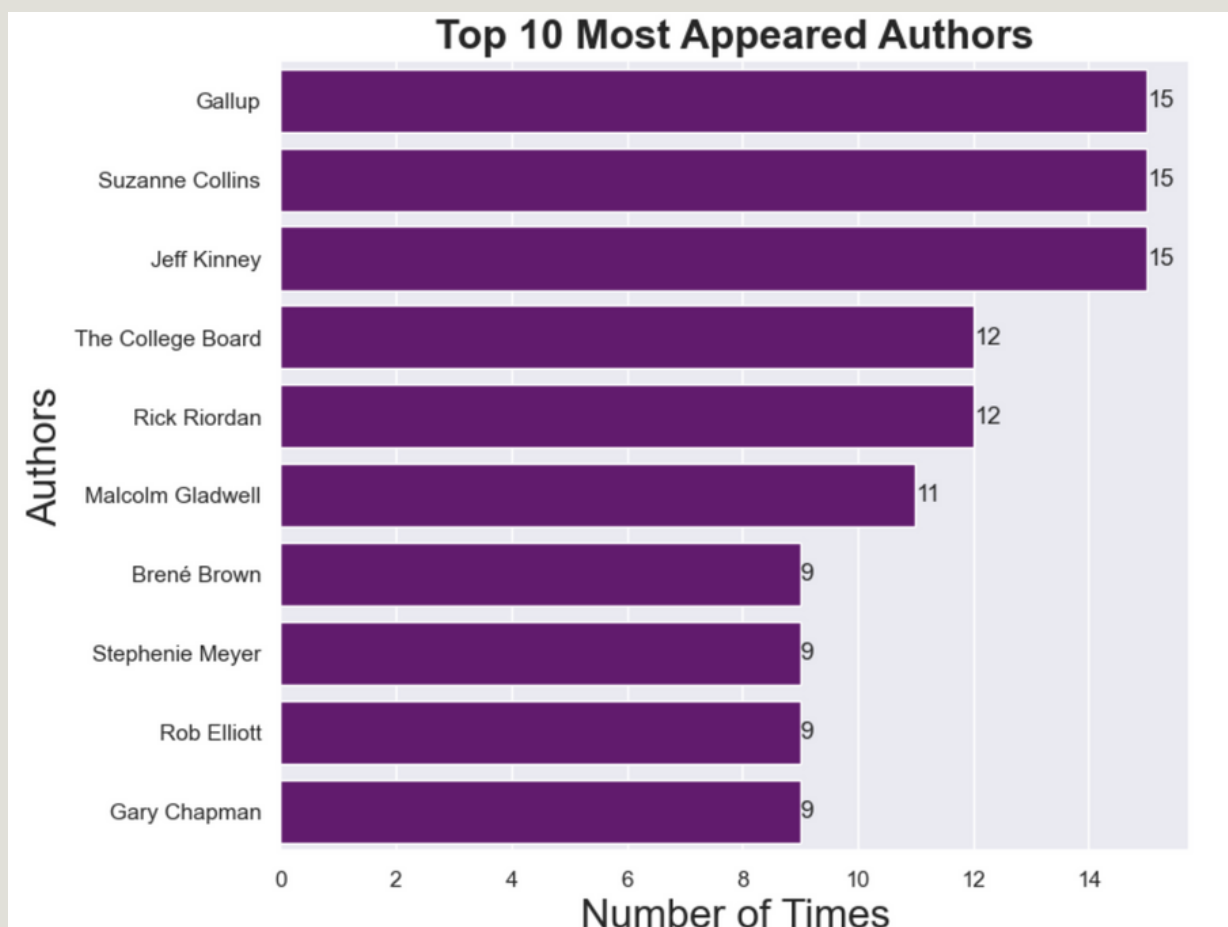
- 1.Price
- 2.Rank
- 3.Title
- 4.Number of Reviews
- 5.Ratings
- 6.Author
- 7.Cover Type
- 8.year

# What should we be looking for?

What are the things we want to find? To Unravel what types of books were bought frequently, between 2009 and 2019. We were able to derived the insights below:

- Top 10 Authors - Based on the number of times their names appeared in the best-selling listings from 2009 to 2019
- The Top 10 Books - Based on the number of times the book appeared as a best-selling book from 2009 to 2019
- Top 10 Authors with the highest User Ratings
- Top 10 Authors with Highest Reviews
- Top 10 Best Selling Books by Number of Reviews
- Top Rated Authors from 2009 to 2019
- Top rated Books from 2009 to 2019
- Number of books by cover type
- Book by Top Average price
- Relationship between rating and price

## Best selling authors



**Figure 1: Best Selling Authors**

Gallup, Suzanne Collins, and Jeff Kinney are the top authors based on the number of times their books appeared from 2009 to 2019 based on figure 1 above.

## Top selling Books



Figure 2: Top selling Books

From figure 2 above, StrengthsFinder 2.0 is the top best selling book with total count of 11, followed by The Official SAT Study Guide with total count of 9, The Gifts of Imperfection: Let Go of Who You Think You're Supposed to Be and Embrace Who You Are with total count of 7.

## Authors with the Highest Ratings

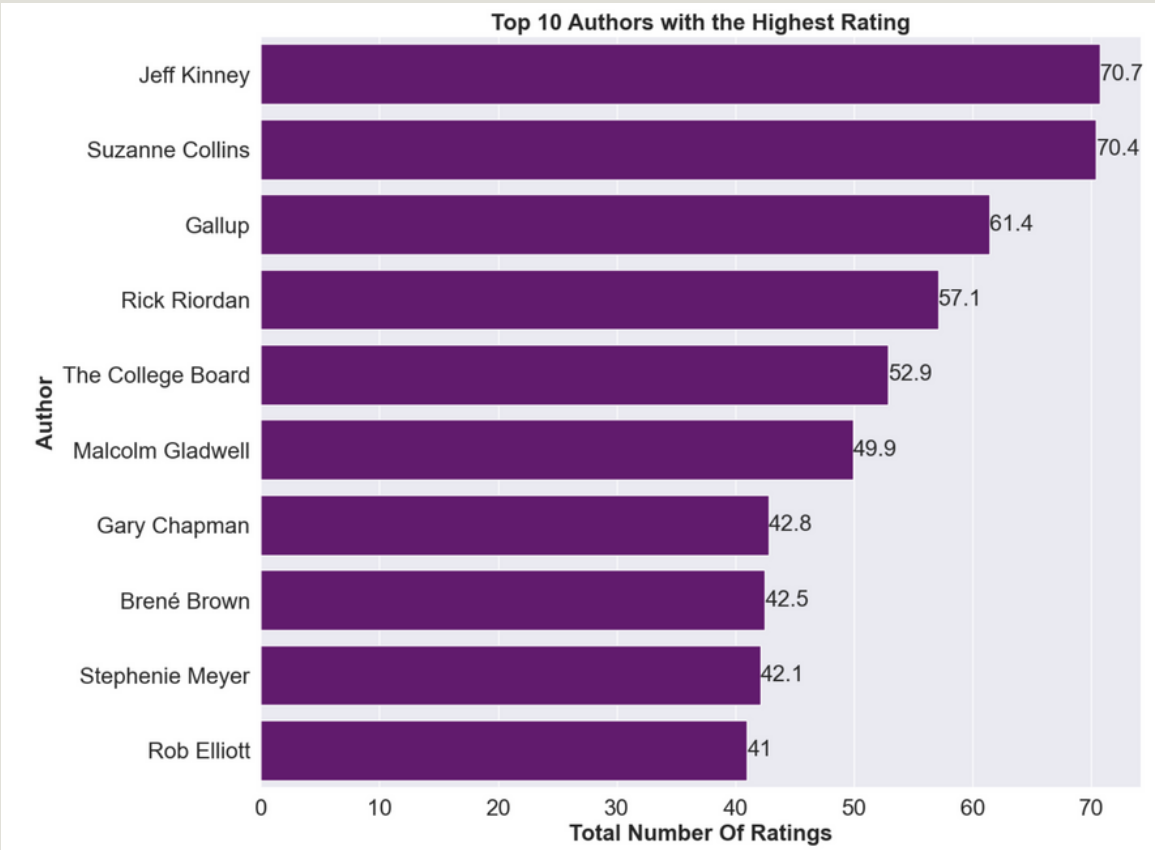


Figure 3: Authors with the Highest Rating

From Figure 3 above, we can see that Jeff Kinney has the highest rating with 70.7, followed by Suzanne Collins with 70.4 and Gallup with 61.4 ratings

## Authors with the Highest Reviews

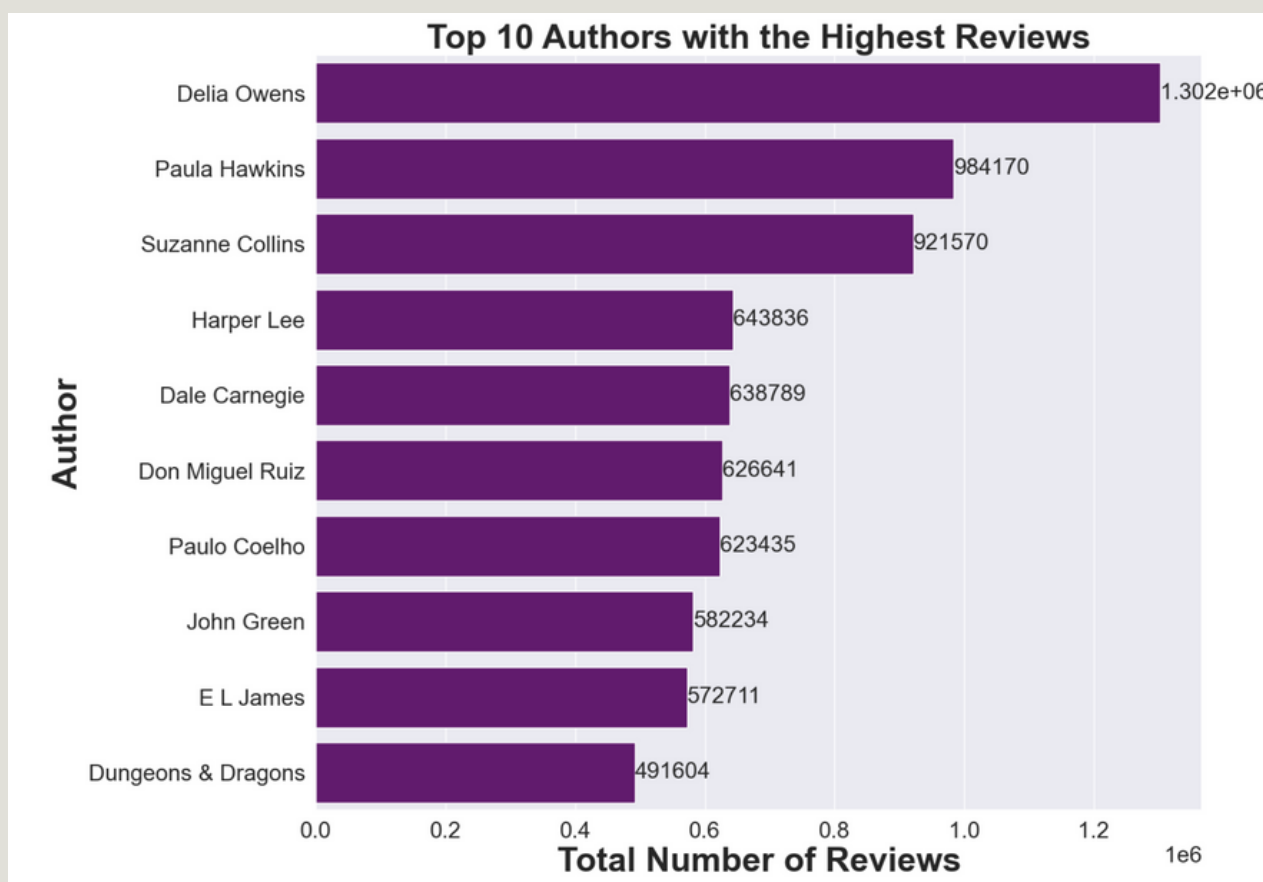


Figure 4: Authors with the Highest Reviews

Figure 4 shows that Delia Owens has the highest reviews with 1,301,999 followed by Paula Hawkins with 984,170, Suzanne Collins with 921,570 reviews

## Best Selling Books by Number of Reviews

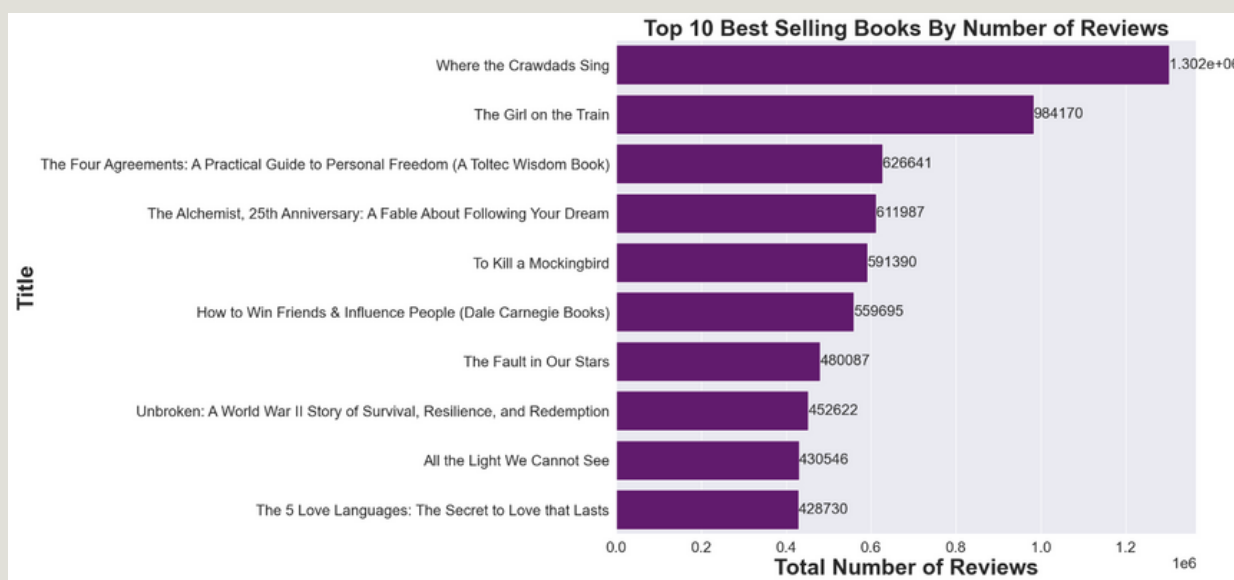


Figure 5: Best selling Books by number of reviews



From the chart above in Figure 5, we can see that the best-selling book by number of reviews is *Where the Crawdads Sing* with 1,301,999 followed by *The Girl on the Train* with 984,170 and *The Four Agreements: A Practical Guide to Personal Freedom* (A Toltec Wisdom Book) with 626,641 reviews.

From the top 10 authors analyzed earlier, the best 3 authors with the highest reviews were Delia Owens, Paula Hawkins, and Suzanne Collins while in this analysis for the best-selling book by the number of reviews, the authors are Delia Owens, Paula Hawkins, and Don Miguel Ruiz

## Top Rated Authors and Books from 2009 to 2019

	author	ratings
0	Patrick Thorpe	4.9
1	Rush Limbaugh	4.9
2	J. Kenji López-Alt	4.9
3	Shel Silverstein	4.9
4	Dr. Seuss	4.9
5	Margaret Wise Brown	4.9
6	Sherri Duskey Rinker	4.9
7	Andrea Beaty	4.9
8	Robert Munsch	4.9
9	Emily Winfield Martin	4.9

**Figure 6: Top-rated Authors**

We can see from Figure 6 above that the top 10 rated authors are Patrick Thorpe, Rush Limbaugh, J.Kenji Lopez-Alt, Shel Silverstein, Dr. Seuss, Margaret Wise Brown, Sherri Duskey Rinker, Andre Beaty, Robert Munsch, Emily Winfield Martin

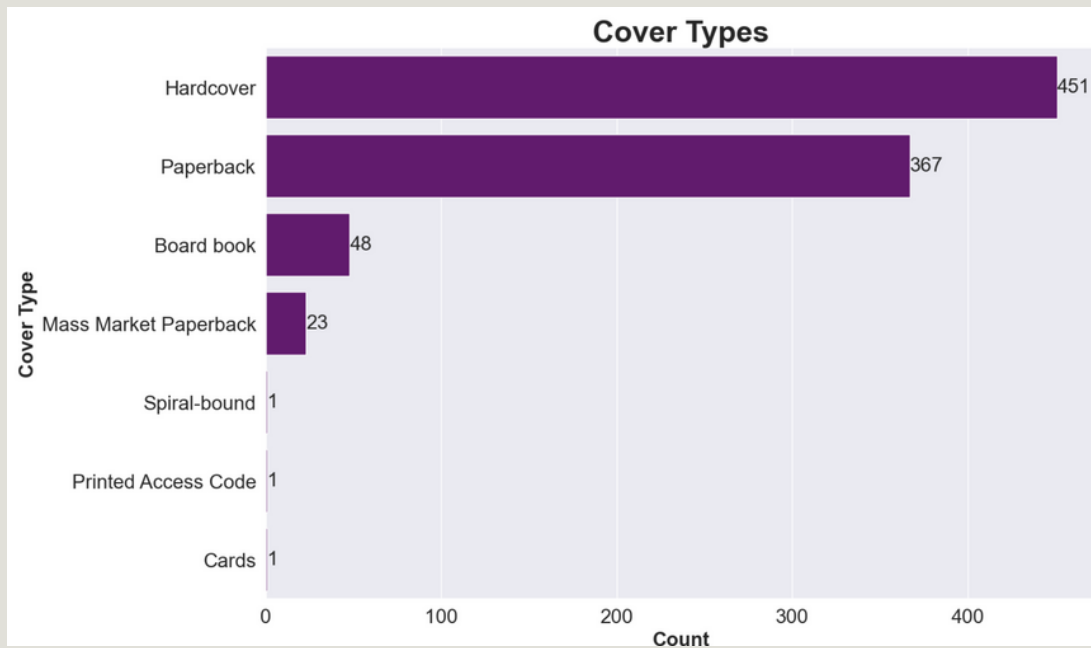
	author	title	ratings
0	Dungeons & Dragons	Xanathar's Guide to Everything (Dungeons & Dra...	4.9
1	Eric Carle	The Very Hungry Caterpillar	4.9
2	Emily Winfield Martin	The Wonderful Things You Will Be	4.9
3	J. Kenji López-Alt	The Food Lab: Better Home Cooking Through Science	4.9
4	J.K. Rowling	Harry Potter and the Sorcerer's Stone: The Ill...	4.9
5	Patrick Thorpe	The Legend of Zelda: Hyrule Historia	4.9
6	Bill Martin Jr.	Brown Bear, Brown Bear, What Do You See?	4.9
7	Bill Martin Jr.	Chicka Chicka Boom Boom (Board Book)	4.9
8	Dungeons & Dragons	Love You Forever	4.9
9	Dungeons & Dragons	D&D Player's Handbook (Dungeons & Dragons Core...	4.9

**Figure 7: Top-rated Books**

From Figure 7 above, we can see the authors that have the top rating and the title of their books.

We can see that the authors of top-rated books are different from the top-rated authors.

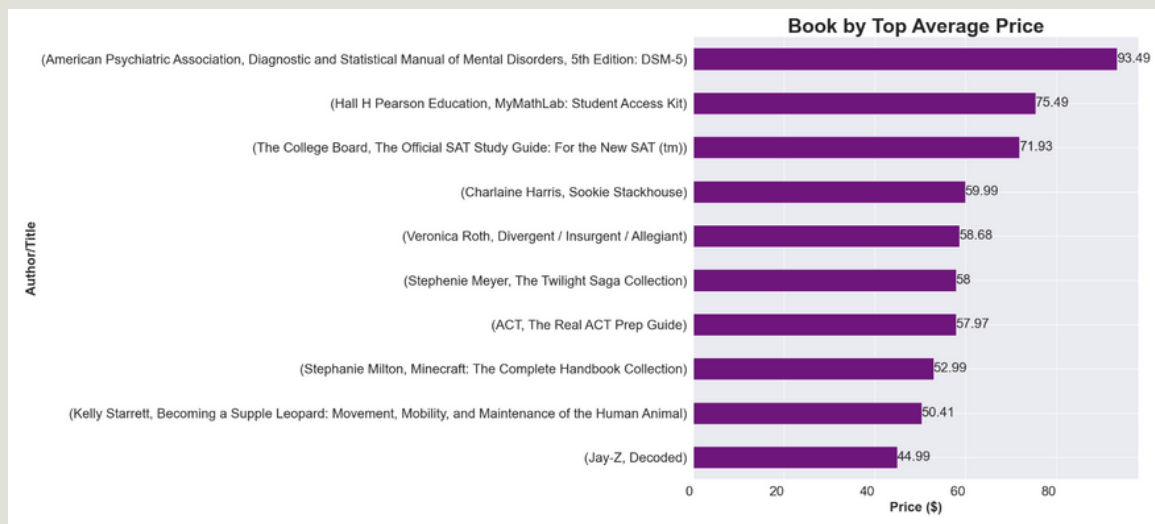
# Number of Books by Cover Type



**Figure 8: Cover Types**

In the total number of best-selling books, the majority of the books, a total of 451 books had a hardcover type, followed by 367 books that had paperback, and 48 had Board book.

# Book by Top Average price



**Figure 9: Book by Average Price**

Book by top average price is Diagnostic and Statistical Manual of Mental Disorders, 5th Edition: DSM-5, authored by American Psychiatric Association with the average price of 93.49, followed by MyMathLab: Student Access Kit, authored by Hall H Pearson Education with 75.49, and The Official SAT Study Guide: For the New SAT (tm), authored by The College Board with average price of 71.93

# Relationship Between Rating and Price

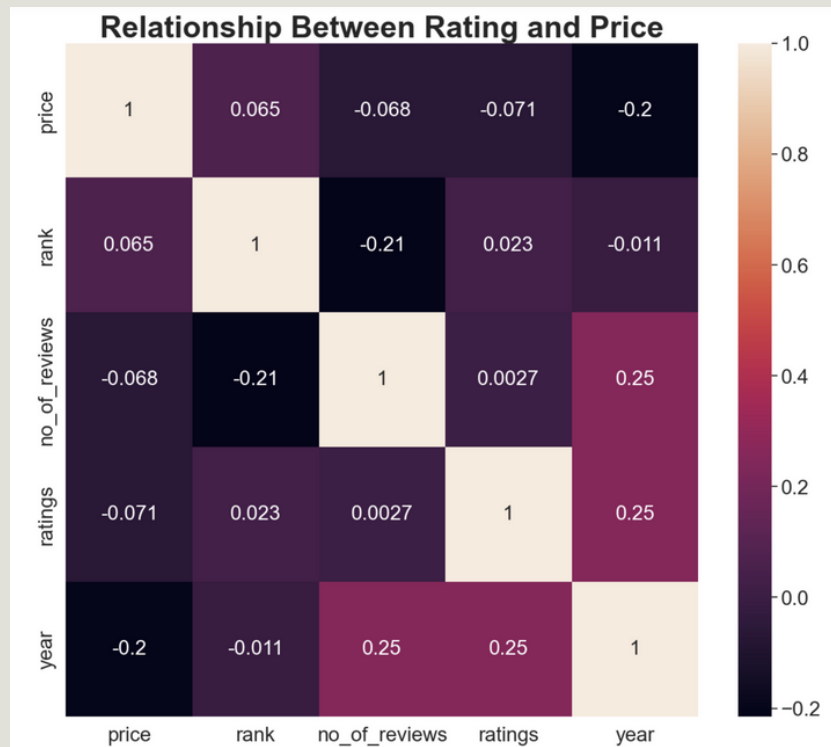


Figure 10: Relationship between Rating and Price

There is no relationship between rating and price.

## Relationship between the Price of Books and Years

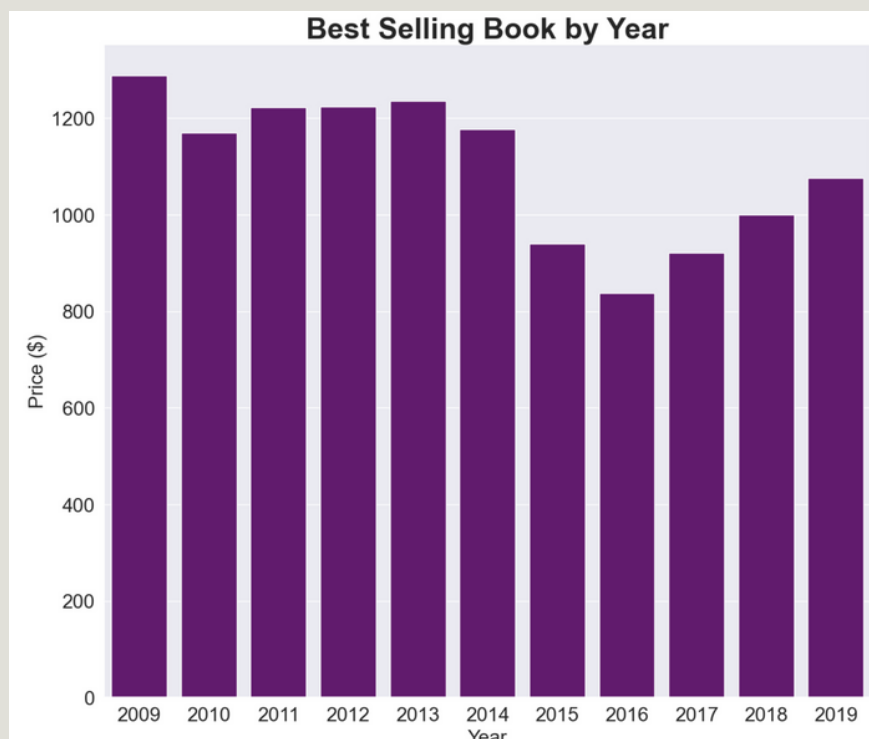


Figure 10: Relationship between the Price of Books and Years

We observed that 2009 had the highest best-selling book with the sum of 1288.12 followed by 2003 with 1236.07 and 2013 with 1224.43. The sum price of books decreased in the year 2016



# Conclusion

- Gallup, Suzanne Collins, and Jeff Kinney are the top authors based on the number of times their books appeared from 2009 to 2019.
- StrengthsFinder 2.0 is the top best-selling book, followed by The Official SAT Study Guide, The Gifts of Imperfection: Let Go of Who You Think You're Supposed to Be and Embrace Who You Are.
- Jeff Kinney has the highest rating with 70.7, followed by Suzanne Collins with 70.4 and Gallup with 61.4 ratings.
- Delia Owens has the highest reviews with 1,301,999 followed by Paula Hawkins with 984,170, and Suzanne Collins with 921,570 reviews.
- The best-selling book by number of reviews is Where the Crawdads Sing with 1,301,999 followed by The Girl on the Train with 984,170 and The Four Agreements: A Practical Guide to Personal Freedom (A Toltec Wisdom Book) with 626,641 reviews.
- The best 3 authors with the highest reviews were Delia Owens, Paula Hawkins, and Suzanne Collins while this analysis for the best-selling book by several reviews, the authors are Delia Owens, Paula Hawkins, and Don Miguel Ruiz.
- The top 10 rated authors are Patrick Thorpe, Rush Limbaugh, J.Kenji Lopez-Alt, etc
- Hardcover has the highest number of books with 451 followed by paperback 367 and Boardbook with 48.
- Book by top average price is the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition: DSM-5, authored by the American Psychiatric Association with an average price of 93.49, followed by MyMathLab: Student Access Kit, authored by Hall H Pearson Education with 75.49, and The Official SAT Study Guide: For the New SAT (tm), authored by The College Board with an average price of 71.93
- There is no relationship between rating and price.
- We observed that 2009 had the highest best-selling book with the sum of 1288.12 followed by 2003 with 1236.07 and 2013 with 1224.43. The sum price of books decreased in the year 2016