Spring 2023

Trends in Young Adult Literature: A Quantitative Approach to Characterizing the New York Times Young Adult Bestsellers of 2020, 2021, and 2022

Anna Wehmeier Giol
Portland State University

Trends in Young Adult Literature:

A Quantitative Approach to Characterizing the *New York Times* Young Adult Bestsellers of 2020, 2021, and 2022

Anna Wehmeier Giol

Master of Arts in Book Publishing

Portland State University

Spring 2023

Abstract

Understanding how a book becomes a bestseller is one of the biggest mysteries in the publishing industry. This research paper focuses on YA bestsellers and uses a quantitative methodology to try to solve the mystery. Its main objectives are to (1) review the current trends in YA literature and (2) find patterns and similarities in the text of YA bestsellers. By building a dataset and analyzing the 175 books featured in the "Young Adult Hardcover" bestseller list of the *New York Times* for 2020, 2021, and 2022, the study shows that YA bestsellers share characteristics and patterns, such as the page count or the point of view of the stories, and urges to expand on this research and continue exploring and understanding bestsellers.

1. Introduction

The publishing industry is one of the world's oldest industries, and it has evolved throughout time in line with society and new technologies. Research around this field has emerged in recent years and has tried to analyze and explain different aspects of publishing books, such as the disruption of artificial intelligence in internal processes,[1] the diversity of authors and stories,[2] and banned books,[3] among others. However, one of the biggest unanswered questions in this industry continues to be what makes a book a bestseller.

Previous research tries to answer this complex question from different perspectives, such as content,[4] marketing,[5] and context.[6] And most of it chooses to focus on adult literature and tackle the question with qualitative methods,[7] while fewer try quantitative methods.[8]

In this research paper, I present a quantitative study of Young Adult (YA) bestsellers. The main objectives are (1) to review the current trends in YA literature and (2) to find patterns and similarities in the text of YA bestsellers.

YA literature is an understudied field that is often overlooked by scholars but has great potential for mass reach because it is read by both young adults and adults.[9] In order to identify YA

---

[1] Gould Finch and Frankfurter Buchmesse, "The Future Impact of Artificial Intelligence on the Publishing Industry," *Livre Blanc*, 2019.

[2] Kendra Fix, "Diversity in Children's Literature," *Education: Student Scholarship & Creative Works* 17 (2019).

[3] Yvonne Vissing and Melissa Juchniewicz, "Children's Book Banning, Censorship and Human Rights," in *Globalisation, Values Education and Teaching Democracy*, ed. Joseph Zajda, Pamela Hallam, and John Whitehouse, Globalisation, Comparative Education and Policy Research (Cham: Springer International Publishing, 2023), 181–201, https://doi.org/10.1007/978-3-031-15896-4_12.

[4] Giovana D. da Silva et al., "Using Full-Text Content to Characterize and Identify Best Seller Books" (arXiv, October 5, 2022), http://arxiv.org/abs/2210.02334.

[5] Michael Wright Johnson, "Bestsellers beyond Bestsellers: The Success of a Good Story," *Online Journal of Communication and Media Technologies* 4, no. 4 (October 15, 2014), https://doi.org/10.29333/ojcmt/2483.

[6] John Sutherland, *Bestsellers: A Very Short Introduction* (Oxford, United Kingdom: Oxford University Press, 2007), http://ebookcentral.proquest.com/lib/psu/detail.action?docID=415119.

[7] Qualitative methods for the study of bestsellers are used in *The Bestseller Code: Anatomy of the Blockbuster Novel*; "Top or Flop: Characteristics of Bestsellers;" *Bestellers: A Very Short Introduction;* "Bestsellers beyond Bestseller;" and *Writing Bestsellers*.

[8] Quantitative methods for the study of bestsellers are used in "Succes with Style: Using Writing Style to Predict the Success of Novels;" "Using Full-Text Content to Characterize and Identify Best Seller Books;" ad "Success in Books: A Big Data Approach to Bestsellers."

[9] Caroline Kitchener, "Why So Many Adults Read Young-Adult Literature," The Atlantic, December 1, 2017, https://www.theatlantic.com/entertainment/archive/2017/12/why-so-many-adults-are-love-young-adult-literature/547334/.

bestsellers, however, I first needed to define the term *bestseller*. Because of the influence and importance of the *New York Times* (*NYT*) bestseller lists in previous research and the book publishing industry, I decided to rely on these lists to gather my data. And I specifically pulled the data from the "Young Adult Hardcover" list. Consequently, the term *bestseller* is understood in this research paper as the following:

> A bestseller is a book that has reached one of the ten positions in the *New York Times* bestseller list for "Young Adult Hardcover" for at least one week.

Because I wanted to identify trends in current YA bestsellers, this research paper focuses only on the lists for 2020, 2021, and 2022. And the analysis is divided into two parts that tackle both of the main objectives: an analysis of the lists and an analysis of the titles and the first chapters. In order to find patterns and similarities, I decided to only focus on titles and first chapters because they are some of the first elements that readers encounter when picking up a book and what might lead them to read it.[10]

By using quantitative methods to analyze the data, an underused method in research on the publishing industry, I want to find patterns in the stories and answer questions. Are the books that reach the *NYT* bestseller list similar? Do they have patterns in common? Can potential bestsellers be predicted?

This research paper is an attempt to answer all of these questions and provide a clear vision of what YA fiction bestsellers currently look like. All my analysis is done from the starting point of the existing research on bestsellers and YA literature. My ultimate goal is to paint a picture of the landscape of YA bestsellers and find patterns in the books that make the *NYT* bestseller list. I also want to create a tool for publishers and authors to determine a book's potential to become a bestseller. My research question is, therefore, the following:

> What are the characteristics of a contemporary bestselling Young Adult fiction book in the US? And how can publishers use this information to create or identify potential future bestsellers?

---

[10] Jodie Archer and Matthew L. Jockers, *The Bestseller Code: Anatomy of the Blockbuster Novel*, First (New York: St. Martin's Press, 2016).

2. Literature Review

The most extensive study on bestsellers to date is *The Bestseller Code*, conducted by Jodie Archer and Matthew L. Jockers. Their premise is that "the novels that hit the *New York Times* bestseller lists are not random," and their goal is to identify patterns, feed them to a machine, and have the machine identify potential bestsellers.[11] They approach this hypothesis with qualitative methods and extract information about style, themes, emotional highs and lows, characters, settings, and linguistic data.

They use the *NYT* bestseller list because it is the most popular and prestigious list and, as Wilkins and Bennett put it, "it has become the gold standard."[12] The *NYT* list will also be used in this research and is used in many other studies. However, it is important to be aware that, since its conception in 1931, the newspaper has become "more mysterious about how the list [is] compiled"[13] and has been accused of not being transparent about the ways they collect their data.[14]

A few years after the publication of *The Bestseller Code*, in 2020, Jockers launched "Authors A.I.," an algorithm to identify potential bestsellers that improves the "original algorithm."[15] The publication of Archer and Jocker's book was also followed by a wave of studies of bestseller books that approached the question from different perspectives. Most of them focused on qualitative methods and adult fiction, such as "Top or Flop: Characteristics of Bestseller," which focuses on identification characteristics in bestsellers;[16] "Bestsellers beyond Bestsellers: The Success of a Good Story," which analyzes *The Da Vinci Code* and *Harry Potter*;[17] and *Writing Bestsellers: Love, Money, and Creative Practice*, which analyzes the relationship between art and market.[18] But others decided to innovate and approach the question of bestsellers with quantitative methods, such as "Success with Style: Using Writing Style to Predict the Success of

---

[11] Archer and Jockers.

[12] Kim Wilkins and Lisa Bennett, *Writing Bestsellers*, Elements in Publishing and Book Culture (Cambridge: Cambridge University Press, 2021), doi:10.1017/9781108663724.

[13] Laura J Miller, "The Best-Seller List as Marketing Tool and Historical Fiction," *Book History* 3, no. 1 (2000): 286–304, https://doi.org/10.1353/bh.2000.0012.

[14] Pamela Koehne-Drube, "The New York Times Bestseller Lists Explained," *Novlr*, September 16, 2022, https://www.novlr.org/the-reading-room/the-new-york-times-bestseller-lists-explained.

[15] "Beyond 'The Bestseller Code,'" Authors A.I., accessed January 29, 2023, https://authors.ai/beyond-the-bestseller-code/.

[16] Sabine Albers, "Top or Flop: Characteristics of Bestsellers," in *Stylistics and Social Cognition*, PALA Papers (Leiden, The Netherlands: Brill, 2007), 205–15, https://doi.org/10.1163/9789401200646_014.

[17] Johnson, "Bestsellers beyond Bestsellers."

[18] Wilkins and Bennett, *Writing Bestsellers*.

Novels," which claims to be the "first quantitative study of bestsellers";[19] and "Using Full-Text Content to Characterize and Identify Best Seller Books," which carries an extensive analysis of the texts in their sample.[20]

Both quantitative and qualitative research has been key in determining how to identify potential bestsellers. Some of the characteristics that have been identified are the following:

- There are similar linguistic patterns in bestsellers and linguistic differences between fiction and nonfiction.[21]
- Fiction is generally more popular than nonfiction. However, the genres that reach the bestseller list more often are general fiction and memoir.[22]
- "Female and male authors are equally represented on the fiction bestseller list."[23]
- Titles are relevant in determining the success of a book.[24]
- The opening scene is key to making a book successful. The scene has to be emotional, touching, authentic, believable, and easy to read.[25]
- Authors of bestseller books choose simple sentences and active characters to tell their stories.[26]
- The success of a book is not only determined by its content but also by visual and external elements, such as "cultural moments," "personal and political taste," and "national and international zeitgeists." And these elements are difficult to detect by machines.[27]

3. Methodology: dataset collection and analysis

The main objectives of this research paper are to understand the trends in current YA bestseller books and to find patterns and characteristics in their titles and first chapters. Therefore,

---

[19] Vikas Ganjigunte Ashok, Song Feng, and Yejin Choi, "Success with Style: Using Writing Style to Predict the Success of Novels," *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 2013, 1753–64.

[20] da Silva et al., "Using Full-Text Content to Characterize and Identify Best Seller Books."

[21] Ashok, Feng, and Choi, "Success with Style: Using Writing Style to Predict the Success of Novels."

[22] Burcu Yucesoy et al., "Success in Books: A Big Data Approach to Bestsellers," *EPJ Data Science* 7, no. 1 (December 2018): 7, https://doi.org/10.1140/epjds/s13688-018-0135-y.

[23] Yucesoy et al.

[24] Archer and Jockers, *The Bestseller Code: Anatomy of the Blockbuster Novel*.

[25] Albers, "Top or Flop: Characteristics of Bestsellers"; Archer and Jockers, *The Bestseller Code: Anatomy of the Blockbuster Novel*.

[26] Archer and Jockers, *The Bestseller Code: Anatomy of the Blockbuster Novel*.

[27] Wilkins and Bennett, *Writing Bestsellers*.

different methodologies are followed: the first part of the research is a descriptive approach to the list, and the second part follows a quantitative methodology.

The first step to accomplishing the objectives was to collect a database of YA books that appeared on each week of the *NYT* bestseller lists for 2020, 2021, and 2022. The following information was collected in a spreadsheet: year, week, rank in the list, title, author, publisher, and the number of weeks in the list.[28] The organization of this dataset was based on the "NYT Hardcover Fiction Bestsellers: List" database.[29]

There was no selection process in this first step. All books were added to the dataset, resulting in a total of 1,560 books (because each week has 10 books and the total number of weeks for the three years is 156). The next step was to create a final list in which books were only listed once, removing repetitions. This process resulted in a list of 175 books, of which 8 were nonfiction and 167 were fiction.

Additional fields were added to the resulting 175 books in order to be able to draw conclusions about current trends. These fields included publication date, genre, and page count.[30] The information for these specifications was pulled from Edelweiss by looking only at the hardcover format of the books.[31]

This first dataset was built to tackle the objective of describing current trends in YA. In order to address the second objective and identify patterns, extra steps were taken to modify the dataset. The quantitative analysis of titles and first chapters was conducted only in fiction YA bestsellers which required establishing a different initial dataset. Therefore, of the 175 books, only 167 were considered. And the next step was to collect the full text of the first chapter of each book. During the text collection process, further restrictions were added to the dataset, and the final list for the linguistic analysis excluded books that fit into one or more of the following patterns:
- The first chapter is not available online.

---

[28] The complette spreadsheet is available here.
[29] This dataset created by Jordan Pruett includes all the fiction titles that appeared on the *NYT* bestseller list between 1931 and 2020. It has 60,000 entries organized by year, week, rank, title_id, title, and author. And it can be found on the following link: https://view.data.post45.org/nytfull.
[30] The complete spreadsheet is available here.
[31] Edelweiss is a platform used by publishers to market their books and by reviewers to discover new titles (https://www.edelweissplus.com/).

- The first chapter is available but is not replicable.
- The book is not written in prose (e.g. poems, emails).
- The book is part of a series and is not the first one.
- The book is an adaptation of a show or movie.

Consequently, the process resulted in a total of 107 books. And the following step was to identify the linguistic features in order to find patterns.[32] To create a list of features, I relied on previous research on bestsellers, which had identified characteristics of these types of books. As a result, the features annotated for each title in this research paper are the following:
- word count
- common words
- word categories (verbs, pronouns, proper nouns, numerals)
- sentence structure
- punctuation
- other observations (words pointing to time and setting, words with a mysterious and bright tone)

And the features annotated for the first chapters are the following:
- word count (for the first chapter, the first paragraph, and the first sentence)
- sentence length (annotated as average words per sentence)
- point of view (first person, third person) and narrator (single narrator, multiple narrators)
- contractions
- punctuation (periods, commas, colons, semicolons, question marks, exclamation marks, ellipsis, and em dashes)
- dialogue
- word categories (nouns, verbs, adjectives, adverbs, pronouns, prepositions, articles, determiners, conjunctions, and interjections)

These features were collected using the following interfaces:
- Voyant tool: to determine the most common words in titles and the average words per sentence in each first chapter.

---

[32] The complete spreadsheet with annotated features for titles and first chapters is available here.

- Claws (UCREL's English part-of-speech tagger): to identify the word categories of the titles and first chapters. For the first chapters, I annotated the number of times each category appeared in the text.

Collecting this data ultimately allowed me to establish the image of the current YA literature landscape, provide conclusions about the linguistic patterns and similarities of bestsellers, and build a document with the general characteristics of current YA fiction bestsellers.

4. The list: an exploration of trends

As I mentioned previously, I will focus only on the "Young Adult Hardcover" bestseller list in the *NYT*. The *NYT* bestseller list "is considered the most important bestseller list"[33] and is "published online every Wednesday at 7 p.m. Eastern."[34] They have specific staff devoted to collecting data and building the lists. And they disclose that the only data that impacts which books appear on the lists is sales data reported by various stores across the country.

If the bestseller lists are built by considering sales across the country, analyzing a bestseller list should represent what the United States population is reading. Therefore, conclusions could be made about specific trends in the genre and length of the book, among others.

In this part of the research paper, I will be giving an overlook of what the lists look like for 2020, 2021, and 2022. I will be providing statistics and graphs for the following elements: genre, publication date, page count, author, and publisher.

4.1. Genre

YA literature stands out from other categories of literature because of its flexibility. This flexibility originates in the fact that YA offers the possibility of having genre crossovers, meaning that its books are not strictly classified into genres but all shelved together as YA. As Fitzsimmons and Wilson state, "young adult literature is not a single genre but a publishing category that

---

[33] Tucker Max, "How Bestseller Lists Actually Work -- And How To Get On Them," *Entrepreneur* (blog), August 30, 2016, https://www.entrepreneur.com/leadership/how-bestseller-lists-actually-work-and-how-to-get-on-them/280520.

[34] "How the New York Times Best-Seller Lists Come Together - The New York Times," accessed March 7, 2023, https://www.nytimes.com/2020/09/23/insider/best-seller-list-process.html.

encompasses multiple genres."[35] Some examples of genre crossovers in books from the dataset are *A Heart So Fierce and Broken* by Brigid Kemmerer, which is classified as fantasy and romance; and *Anatomy* by Dana Schwartz, which is fantasy, historical fiction, and romance.

For the sake of clarity, I classified fiction books into six genres following Cart's examination: speculative fiction (science fiction and fantasy), historical fiction, romance, horror, mystery/suspense, and humor.[36] I also added contemporary realism as a seventh genre because there were multiple books in the list that did not fit into any of the other genres and had Coming of Age, Social Themes, and LGBTQ+ as their BISAC codes. In addition to the genre classification, I also decided to analyze books that had BISAC codes pointing to wizards & witches, dystopia, paranormal romance, vampires, and werewolves because I wanted to observe if YA is still influenced by *Harry Potter*, *The Hunger Games*, and *Twilight*, which are the three book series that have shaped YA literature as we know it today.[37]

As can be seen in Figure 1, bestseller YA books are predominantly fiction (with 167 books out of 175). Figure 2 shows that most fiction books are speculative fiction, while the least popular genres are historical fiction and humor. By looking closer at speculative fiction books, fantasy clearly dominates over science fiction (Figure 3). The top three genres in YA are the following:
1. Fantasy
2. Romance
3. Mystery/suspense

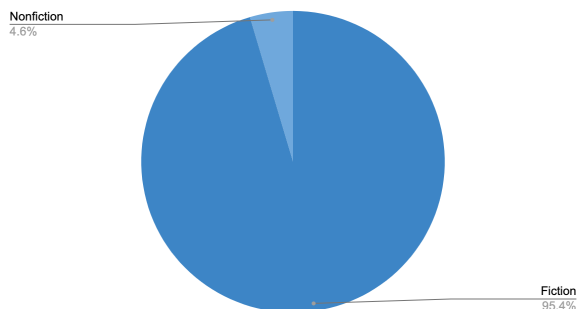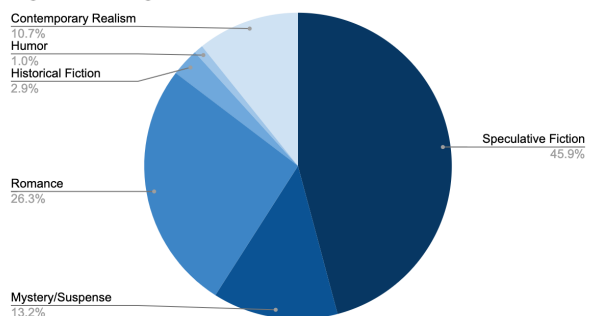Figure 1. Young Adult Literature

Nonfiction 4.6%
Fiction 95.4%

Figure 2. Young Adult Fiction

Contemporary Realism 10.7%
Humor 1.0%
Historical Fiction 2.9%
Romance 26.3%
Mystery/Suspense 13.2%
Speculative Fiction 45.9%

---

[35] Rebekah Fitzsimmons and Casey Alane Wilson, *Beyond the Blockbusters: Themes and Trends in Contemporary Young Adult Fiction* (Univ. Press of Mississippi, 2020).
[36] Michael Cart, "Genre on the Agenda," in *Young Adult Literature: From Romance to Realism*, Fourth (Chicago: ALA Neal-Schuman, 2022), 75–84.
[37] Fitzsimmons and Wilson, *Beyond the Blockbusters*.

Figure 3. Speculative Fiction

Both
2.1%
Science Fiction
9.6%

Fantasy
88.3%
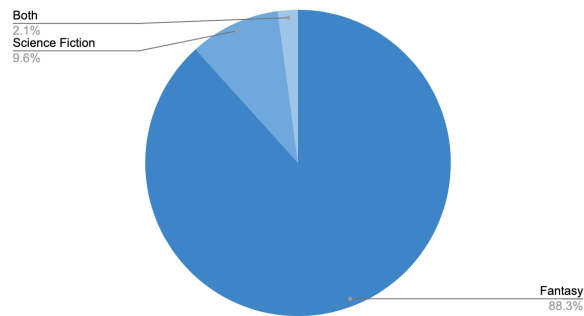
As for tags that point to the influence of *Harry Potter*, *The Hunger Games*, and *Twilight*, there are 25 fiction books with one or more of the tags for wizards & witches, dystopia, paranormal romance, vampires, and werewolves (Figure 4). This shows that those blockbuster books still have some influence on today's YA books, but they are not the most popular and significant trend.

Figure 4. Young Adult Fiction Trends

Werewolves
3.6%

Paranormal Romance
28.6%

Wizards & Witches
39.3%

Vampires
10.7%

Dystopia
17.9%
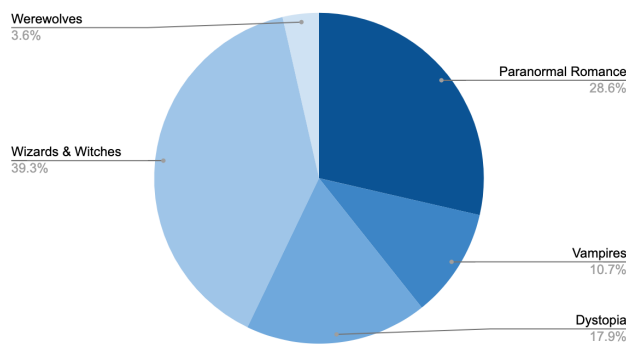
4.2. Publication date

The publication date of the books in the bestseller lists was studied to observe the number of backlist and frontlist titles populating the list. In this research paper, frontlist and backlist titles will be understood as the following:[38]

> A frontlist title is a book that has been published in the last 12 months.
> A backlist title is a book that has been published for more than 12 months.

---

[38] Karl Berglund and Ann Steiner, "Is Backlist the New Frontlist?: Large-Scale Data Analysis of Bestseller Book Consumption in Streaming Services," *Logos* 32, no. 1 (May 25, 2021): 7–24, https://doi.org/10.1163/18784712-03104006.

Traditionally, frontlist titles have occupied the top positions in bestseller lists and have been the main sales drivers. However, since 2020, the industry is seeing a new phenomenon in which backlist titles "[have] become more important."[39] Is that phenomenon true for YA?

My analysis suggests not only that there are more frontlist than backlist titles in the list but also that backlist titles have seen a significant drop in representation on the lists (Figure 5). The year 2020 has the most backlist titles out of the three years, with backlist titles appearing in 181 of the 520 spots in the 52 weeks of the year (which represents only a 34 percent of backlist tiles). There was a significant drop in backlist titles in 2022, a year in which they only occupied 98 spots (18 percent).

Figure 5. Backlist and Frontlist Titles
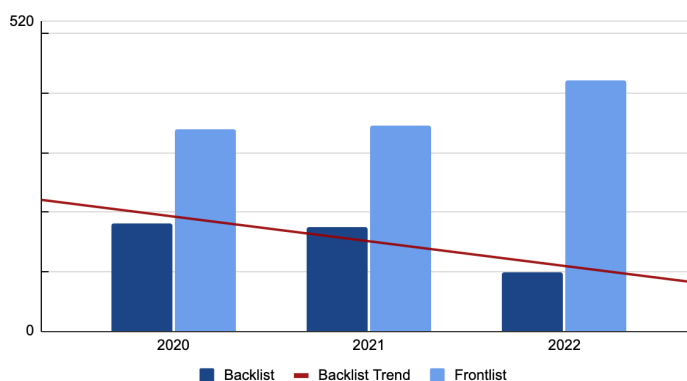
Although the appearance of backlist titles on the lists is in decline, there are five blacklist titles that are worth highlighting because they were published in 2017 and are still on the lists for 2020, 2021, and 2022. These books are *One Of Us Is Lying*, *The Hate U Give, Stamped*, *Those Violent Delights*, and *The Cousins*. Because of the success they are having, they have the potential to become the next blockbuster YA titles that mark a generation of readers and a trend in the market. That is, they might become the next *Harry Potter*, *The Hunger Games*, or *Twilight*.

4.3. Page count

Research conducted by Wordsrated in 2022 pointed out that bestsellers are becoming shorter over time, with most books being fewer than 400 pages.[40]

---

[39] Berglund and Steiner.
[40] WordsRated, "Bestselling Books Have Never Been Shorter [Study of 3,444 NYT Bestselling Titles]," June 20, 2022, https://wordsrated.com/bestselling-books-have-never-been-shorter/.

Figure 6 shows that it is true that most of the YA books in the lists have less than 400 pages (with 90 books having between 100 and 399 pages). However, it is not possible to make generalizations and affirm that that is the trend because the rest of the sample, which includes 85 books, has more than 400 pages. Therefore, my observations differ from the Wordsrated results: the trend in YA bestsellers is to be between 300 and 500 pages, with the 300-to-399-page range being the most popular one.

Figure 6. Page count in YA bestseller
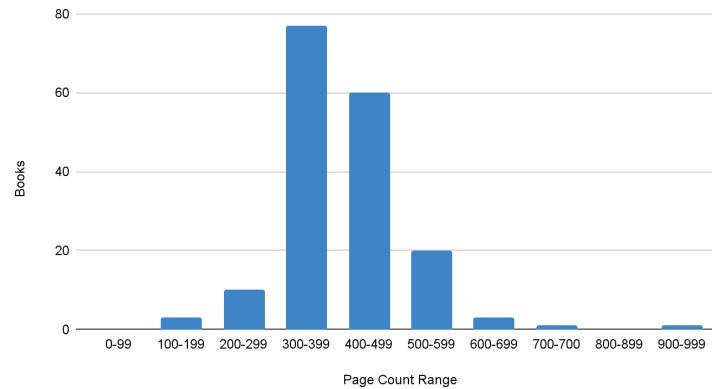
4.4. Authors and publishers

The last fields considered for the analysis of trends are authors and publishers. By including this, I wanted to provide a list of authors and publishers that are currently controlling the YA landscape.

4.4.1. Authors

There is a total of 143 authors represented in the 2020, 2021, and 2022 lists. The analysis suggests that the three authors with more books on them are the following:

1. Karen M. McManus
2. Tiffany D. Jackson
3. Angie Thomas

It is important to point out, though, that Karen M. McManus is the only author out of the three that has five individually written books on the lists. She also has the book that has been on the lists the longest, that is, 155 weeks out of the 156 weeks that represent the three years. This book is *One Of Us Is Lying*.

Although there are other authors with multiple books on the lists, such as Brigid Kemmerer with four books and Holly Jackson with three books, most of the authors (100 out of 143) only have one book on the list (Figure 7). This is, therefore, the general trend.

Figure 7. Authors

Five books
2.1%
Three books
7.0%

Two books
20.3%

One book
69.9%
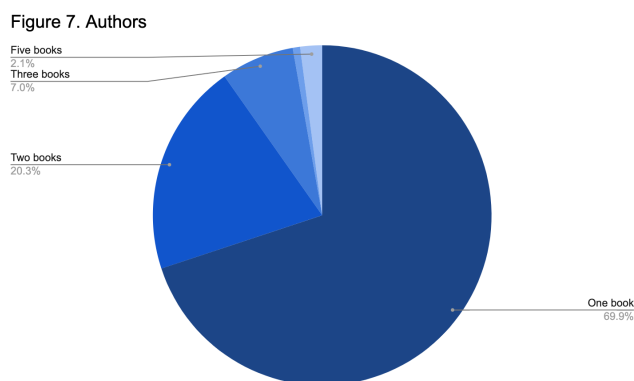
Coauthoring books is also a trend worth mentioning. My analysis shows that only 20 out of the 175 books have been written by two or more authors. These 20 books occupy a total of 214 positions throughout the three years, with 113 positions in 2020, 65 in 2021, and 36 in 2022. Therefore, coauthoring is a decreasing trend that only represents a small number of bestseller books. However, many successful authors are participating in it, which points to a possible revival of the trend in the future. Some authors that have written a book collaboratively are Adam Silvera, Cassandra Clare, Jason Reynolds, and Angie Thomas.

4.4.2. Publishers

The *NYT* bestseller list has been criticized for being biased and favoring major publishing houses. It is sometimes said that the *NYT* "won't recognize a book that doesn't come from one of the big New York publishing houses."[41]

My observations are that there are a total of 48 publishers and imprints represented on the list: 34 imprints of the Big Five, 9 imprints of medium-sized publishers (Disney Publishing Companies, ABRAMS, Scholastic, Bloomsbury), and 5 smaller or independent publishers. The publishers and imprints that have the most books on the list are the following:

1. Delacorte (imprint of Penguin Random House)
2. Wednesday (imprint of Macmillan)
3. HarperTeen (imprint of HarperCollins)

---

[41] Max, "How Bestseller Lists Actually Work -- And How To Get On Them."

4. Margaret K. McElderry (imprint of Simon & Schuster)
5. Feiwel & Friends (imprint of Macmillan)
6. Simon & Schuster

These are all imprints of the Big Five, which confirms that the current general trend is still favoring the biggest publishers. The Big Five continue to dominate the list and make true the statement that most of the books in the list "come from one of the big New York publishing houses."[42] However, other publishers, such as Scholastic, Disney, Amulet, Sourcebooks Fire, and Candlewick, are slowly gaining positions in the list. It will be interesting to continue observing the shifts in the future.

5. Analysis of titles and first chapters

The second part of this research paper focuses on the content of YA fiction bestsellers and analyzes the titles and first chapters of the books using a quantitative methodology. As mentioned in the methodology section, only 107 titles are considered for this analysis.

5.1. Titles

As Jockers and Archer point out in *The Bestseller Code*, titles are essential because they are one of the first things consumers see when looking for a book.[43] By analyzing them, I wanted to observe patterns and trends among titles in different genres and from different publishing houses.

In examining the 107 titles, it was found that there are no relevant distinctions between titles belonging to different genres. The dataset has 99 titles that have between one and five words, with three and four words (52 titles) being the most common word counts for titles (Figure 8). It is also important to point out that there is not a lot of word repetition among all the titles. The most popular words, excluding stopwords, are "girl" (5), "witch" (3), and "kingdom" (3).[44]
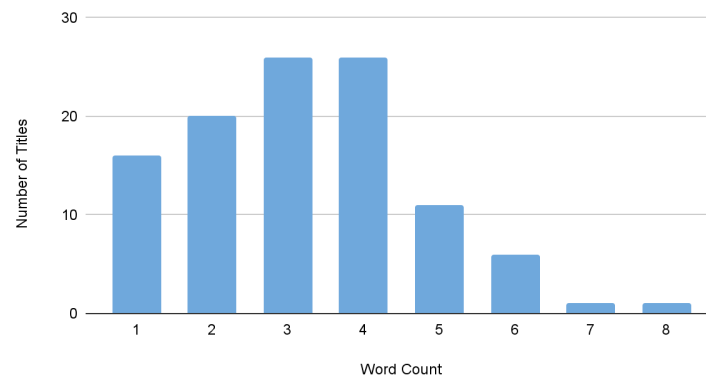
---

[42] Max.
[43] Archer and Jockers, *The Bestseller Code: Anatomy of the Blockbuster Novel*.
[44] "Voyant Tools," accessed April 5, 2023,
https://voyant-tools.org/?corpus=2e5a110cf2bcfb5a43c4b90b9fcde31e.

Figure 8. Word Count in Titles

There is a clear trend for titles to capture a mysterious and dark tone, with 39 titles including words such as "die," "murder," "lie," and "hate"; in contrast, titles that capture a bright and light atmosphere are fewer, with only 11 titles including words such as "love," "life," and "delightful." There is a minor trend (with only 9 titles) to include words in the royalty realm, such as "kingdom" or "queen," and references to elements, such as "silver," "gold," and "iron." A more significant trend is to include specifications of setting and time, with 20 titles out of 107 including words such as "boat," "house," "tonight," and "yesterday." Only 20 titles refer to a female or male figure, of which 13 are focused on a female figure.

My analysis also found that 21 out of the 107 titles start with "the." In comparison, only 7 titles start with "a." A verb is included in 27 titles, and 21 titles mention characters by either using their name or a pronoun. In contrast, including numbers, punctuation, and contractions is not popular, with only 12 titles doing so. Some trends observed were titles with the structure "noun+of+noun" (*The Fountains of Silence*, *Chain Of Gold*, *A Scatter Of Light*), one-word titles (*Legendborn*, *Blackout*, *Anatomy*), and titles with a noun modified by an adjective or noun (*Cemetery Boys*, *Instant Karma*). It was also interesting to see other similar title structures (Table 1).

Table 1. Structures in YA titles

Structure	Titles				
The + Noun	*The Beautiful*	*The Betrothed*	*The Cousins*	*The Agathas*	*The Luminaries*
Noun + Of + Noun Complement	*Wings Of Ebony*	*House Of Hollow*	*Ace Of Spades*		*Beasts Of Prey*
These/This + Adjective + Noun	*These Violent Delights*		*These Hollow Vows*	*This Woven Kingdom*	
The + Noun + Of + Noun Complement	*The Fountains Of Silence*		*The Kingdom Of Back*	*The Nature of Witches*	
A + Noun + Of + Noun Complement	*A Song Of Wraiths and Ruin*		*A Scatter of Light*		
A + Adjective + Girl's Guide To + Noun Complement	*A Good Girl's Guide To Murder*		*A Cuban Girl's Guide To Tea and Tomorrow*		
Once Upon A + Noun Phrase	*Once Upon A Broken Heart*		*Once Upon A K-Prom*		

Although these similar structures might point to trends, my sample is not large enough to be able to make generalizations and conclusions about current trends in the structure of titles. Therefore, it can only be considered a coincidence.

5.2. First Chapters

The first chapter of a book is what might incline readers to either read the entire text or leave the book unread. Albers identifies that "an emotionally touching opening scene" is key in bestsellers.[45] Moreover, other research studies, such as "Success with Style: Using Writing Style to Predict the Success of Novels," also choose to explore the beginning of stories by taking into consideration "1000 sentences from the beginning of each book."[46]

My analysis indicates that the first chapters run between 1,000 and 4,000 words, with 82 out of the 107 books having word counts between those numbers. The most common sentences are between 11 and 13 words long, with 63 books with an average between those numbers. As for

---

[45] Albers, "Top or Flop: Characteristics of Bestsellers."
[46] Ashok, Feng, and Choi, "Success with Style: Using Writing Style to Predict the Success of Novels."

the point of view, YA fiction bestsellers are mainly written in the first person, with 76 out of 107 books. And most of them have a single narrator. However, it is notable that 26 books in the sample use multiple voices and narrators to tell the story by alternating the narrating character in each chapter. This technique allows readers to see the story from multiple points of view. It is a trend that is not currently the most common but that is represented across the three years (with 10 books in 2020, 10 books in 2021, and 12 books in 2022) and, therefore, should be followed and analyzed in the future.

Moving on to the text, my analysis suggests that there are more question marks than exclamation marks, more periods than commas, and more colons than semicolons. Overall, however, it can be said that bestseller YA first chapters have more question marks, periods, and commas, than exclamation marks, semicolons, and colons. The texts also include ellipsis and em dashes, which are mainly used "as a way to indicate unfinished thought."[47]

I also observed dialogue and contractions because *The Bestseller Code* indicates that they are used to show a conversational tone and authentic voice.[48] My analysis shows that contractions represent between 2 and 5 percent of the total word count of the chapter and that each chapter has an average of between 10 and 40 instances of dialogue. This shows that these elements are not a big part of YA bestsellers. However, it would be necessary to compare bestsellers to non-bestsellers to be able to extract better conclusions.

The last feature observed was word categories. Previous research pointed out that nouns and verbs are more common than adjectives and adverbs in bestseller books.[49] My research suggests a similar outcome. In YA fiction bestsellers, there are more verbs than nouns and more adverbs than adjectives. Overall, however, there is a lower proportion of adjectives and adverbs than nouns and verbs. This indicates that authors prefer more concise sentences.[50]

6. Results

The results from the dataset collection and the data analysis indicate that YA bestsellers are predominantly fiction, usually fantasy, although romance and mystery/suspense are also

---

[47] Archer and Jockers, *The Bestseller Code: Anatomy of the Blockbuster Novel.*
[48] Archer and Jockers.
[49] Archer and Jockers.
[50] Archer and Jockers.

prominent. Trends in subject matter indicate a distancing from wizards, dystopias, and paranormal romances, that is, from those themes portrayed in *Harry Potter*, *The Hunger Games*, and *Twilight*. In contrast, current YA books are shifting toward the representation of LGBTQ+ characters and themes and of social issues, such as friendship, racism, and depression. This reflects the flexibility of YA to follow trends and adapt to the needs of its readers, which are currently directing their attention to social movements and diverse voices.[51]

The *NYT* bestseller list continues to include mostly books published by the Big Five, although medium-sized and independent publishers are slowly finding a space in the lists. More than half of the authors in the list only have one book throughout the three years included in this research. This indicates that getting a spot is a difficult task. While most of the books are written by a single author, which is clearly the general and more popular trend, there is a slight inclination to writing books collaboratively that should be observed in the future. As for the length of the books, YA bestsellers have a consistent length of between 300 and 500 pages. The majority of books use prose as their medium to tell the story, but there seems to be a new wave of trying out different techniques. For instance, the dataset includes books that use poetry and emails to tell the story. This is an interesting twist in narrating stories, and although it is still not relevant enough in bestseller books, it might have a bigger impact in the future as authors continue to innovate and change their styles.

Moving on to the writing style, the most common YA bestseller book titles have between one and five words, lean toward a mysterious and dark tone, and include verbs and references to characters (by using pronouns or their names). The stories are predominantly written in the first person and portray only one point of view, which is usually the protagonist. However, the portrayal of multiple voices by changing the narrator every chapter is a technique that might gain importance in the future. As for specific features, YA bestsellers are written with more question marks, periods, commas, verbs, and nouns, than exclamation marks, semicolons, colons, adjectives, and adverbs. And the author's writing style also reflects a conversational tone by including contractions, ellipsis, em dashes, and dialogue. However, these latter features should not be used excessively but only represent around 10 percent of the text. The first chapter should be between 1,000 and 4,000 words and have sentences with an average of 12 words.

---

[51] Jacqueline Bach, "Young Adult Boom," in *The Encyclopedia of Contemporary American Fiction 1980–2020* (John Wiley & Sons, Ltd, 2022), 1–10, https://doi.org/10.1002/9781119431732.ecaf0270.

The results of my analysis indicate, therefore, that current YA titles should have the characteristics included in Table 2, which are the more common among YA fiction bestseller books, to have a greater chance of hitting the *NYT* bestseller list.

Table 2. Characteristics of a YA bestseller

Title:	three- or four-word title that includes verbs, characters (name or pronoun), and mysterious words (e.g., blood, violence, lie, night)
Genre:	fantasy, romance, or mystery/suspense
Page count:	300–500 pages
First chapter word count:	1,000–4,000 words
Point of view:	first person
Writing style:	question marks, periods, commas, nouns, and verbs; dialogue and contractions

7. Conclusions

The study of bestsellers continues to be a fascinating field with many open possibilities to explore and discover. In this research paper, I presented a quantitative study of Young Adult bestsellers to show the current trends and patterns in this category of literature. My results show that YA fiction bestsellers in the *NYT* bestseller list have a similar structure and writing style, which implies that potential bestsellers could be identified before publication. And the methodology used could be applied to further study the text of bestsellers.

This research, however, has limitations that restrict the extraction of conclusions. By only being able to analyze bestsellers for 2020, 2021, and 2022 from a single list, I can present conclusions for the *NYT* bestseller list, but this does not indicate that the characteristics observed apply to YA overall (bestsellers and even non-bestsellers). Moreover, it is also important to reinforce the idea that the text is not the only element that drives a book to become a bestseller. There are other elements, such as marketing or the cultural moment, that can also have a big influence on the success of a book.[52]

---

[52] Wilkins and Bennett, *Writing Bestsellers*.

Even though this study is unable to provide generalizations for YA fiction, my findings can be used by publishers to help them identify potential bestsellers among the submitted manuscripts. The findings can also give them an idea about current popular genres and stories and, therefore, lead them to choose manuscripts that will probably be bought and read by many customers. That is, they will be able to find manuscripts that fit with current trends.

The study also paves the way for further studies on YA and adult bestsellers. It would be interesting to use quantitative methods to analyze a larger sample, which could include more years, the full text of books, non-bestseller books, nonfiction, different narration styles (such as prose, poetry, and emails), or even different bestseller lists. In addition, it would also be interesting to carry out research that combines qualitative and quantitative studies of the text of bestsellers and its surroundings (marketing, cultural moment, publicity, etc.). Finally, there is also the opportunity to take the analyses into practice and use artificial intelligence and machine learning to build a model that could identify potential bestsellers.

8. Bibliography

Albers, Sabine. "Top or Flop: Characteristics of Bestsellers." In *Stylistics and Social Cognition*, 205–15. PALA Papers. Leiden, The Netherlands: Brill, 2007. https://doi.org/10.1163/9789401200646_014.

Archer, Jodie, and Matthew L. Jockers. *The Bestseller Code: Anatomy of the Blockbuster Novel*. First. New York: St. Martin's Press, 2016.

Ashok, Vikas Ganjigunte, Song Feng, and Yejin Choi. "Success with Style: Using Writing Style to Predict the Success of Novels." *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 2013, 1753–64.

Authors A.I. "Beyond 'The Bestseller Code.'" Accessed January 29, 2023. https://authors.ai/beyond-the-bestseller-code/.

Bach, Jacqueline. "Young Adult Boom." In *The Encyclopedia of Contemporary American Fiction 1980–2020*, 1–10. John Wiley & Sons, Ltd, 2022. https://doi.org/10.1002/9781119431732.ecaf0270.

Berglund, Karl, and Ann Steiner. "Is Backlist the New Frontlist?: Large-Scale Data Analysis of Bestseller Book Consumption in Streaming Services." *Logos* 32, no. 1 (May 25, 2021): 7–24. https://doi.org/10.1163/18784712-03104006.

Cart, Michael. "Genre on the Agenda." In *Young Adult Literature: From Romance to Realism*, Fourth., 75–84. Chicago: ALA Neal-Schuman, 2022.

Finch, Gould and Frankfurter Buchmesse. "The Future Impact of Artificial Intelligence on the Publishing Industry." *Livre Blanc*, 2019.

Fitzsimmons, Rebekah, and Casey Alane Wilson. *Beyond the Blockbusters: Themes and Trends in Contemporary Young Adult Fiction*. Univ. Press of Mississippi, 2020.

Fix, Kendra. "Diversity in Children's Literature." *Education: Student Scholarship & Creative Works* 17 (2019).

"How the New York Times Best-Seller Lists Come Together - The New York Times." Accessed March 7, 2023. https://www.nytimes.com/2020/09/23/insider/best-seller-list-process.html.

Johnson, Michael Wright. "Bestsellers beyond Bestsellers: The Success of a Good Story." *Online Journal of Communication and Media Technologies* 4, no. 4 (October 15, 2014). https://doi.org/10.29333/ojcmt/2483.

Kitchener, Caroline. "Why So Many Adults Read Young-Adult Literature." The Atlantic, December 1, 2017. https://www.theatlantic.com/entertainment/archive/2017/12/why-so-many-adults-are-love-young-adult-literature/547334/.

Koehne-Drube, Pamela. "The New York Times Bestseller Lists Explained." *Novlr*, September 16, 2022. https://www.novlr.org/the-reading-room/the-new-york-times-bestseller-lists-explained.

Max, Tucker. "How Bestseller Lists Actually Work -- And How To Get On Them." *Entrepreneur* (blog), August 30, 2016. https://www.entrepreneur.com/leadership/how-bestseller-lists-actually-work-and-how-to-get-on-them/280520.

Miller, Laura J. "The Best-Seller List as Marketing Tool and Historical Fiction." *Book History* 3, no. 1 (2000): 286–304. https://doi.org/10.1353/bh.2000.0012.

Silva, Giovana D. da, Filipi N. Silva, Henrique F. de Arruda, Bárbara C. e Souza, Luciano da F. Costa, and Diego R. Amancio. "Using Full-Text Content to Characterize and Identify Best Seller Books." arXiv, October 5, 2022. http://arxiv.org/abs/2210.02334.

Sutherland, John. *Bestsellers: A Very Short Introduction*. Oxford, United Kingdom: Oxford University Press, 2007. http://ebookcentral.proquest.com/lib/psu/detail.action?docID=415119.

Vissing, Yvonne, and Melissa Juchniewicz. "Children's Book Banning, Censorship and Human Rights." In *Globalisation, Values Education and Teaching Democracy*, edited by Joseph Zajda, Pamela Hallam, and John Whitehouse, 181–201. Globalisation, Comparative Education and Policy Research. Cham: Springer International Publishing, 2023. https://doi.org/10.1007/978-3-031-15896-4_12.

"Voyant Tools." Accessed April 5, 2023. https://voyant-tools.org/?corpus=2e5a110cf2bcfb5a43c4b90b9fcde31e.

Wilkins, Kim, and Lisa Bennett. *Writing Bestsellers*. Elements in Publishing and Book Culture. Cambridge: Cambridge University Press, 2021. doi:10.1017/9781108663724.

WordsRated. "Bestselling Books Have Never Been Shorter [Study of 3,444 NYT Bestselling Titles]," June 20, 2022. https://wordsrated.com/bestselling-books-have-never-been-shorter/.

Yucesoy, Burcu, Xindi Wang, Junming Huang, and Albert-László Barabási. "Success in Books: A Big Data Approach to Bestsellers." *EPJ Data Science* 7, no. 1 (December 2018): 7. https://doi.org/10.1140/epjds/s13688-018-0135-y.