

APPENDIX A: METHODOLOGY

The *BookStats* project, a joint venture of the Association of American Publishers and the Book Industry Study Group, encompasses a complete accounting of all publisher sales, in both print and electronic formats, across the entire spectrum of books and book-related content in the marketplace today. For tracking purposes, *BookStats* divides the marketplace into a “data cube.” Each axis of the cube segments the market in one of three ways: Format, Category, or Channel.

BookStats Volume Four uses the same basic methodology originally created for the inaugural release of the study in 2011. The methodology has undergone rigorous review and testing—both internally to the *BookStats* team, and externally with select industry experts, resulting in subtle changes in how data is collected and parsed.

For a full list of definitions used, please see www.BookStats.org.

COLLECTING DATA

Survey Process — Source data for the *BookStats* project was derived from an extensive survey of U.S. publishers. The survey requested sales data (net units and dollars, less returns) for both 2012 and 2013 against all segments within the *BookStats* data cube. If a publisher could not provide this level of detail, or was unwilling, they were asked to complete a much simpler survey that asked for total net sales over the survey period, expressed in both dollars and units.

All active individuals or entities (“publishers”) with an active ISBN in R. R. Bowker’s Books In Print® database were contacted and asked to participate in the survey. This list included all publishers with an active ISBN between 2000 and 2013.

The total number of active publishers covered by this study is 71,759. The total number of publishers covered by *BookStats Volume Three* (featuring 2008-2012 data) was 59,719.

Three tactics were used to reach the entire universe of publishers, encouraging them to participate in the survey:

1. Staff members from the Association of American Publishers (AAP) and the Book Industry Study Group (BISG) sent direct requests for information to all major publishers.

2. Nielsen conducted a major email campaign, targeting more than 36,000 small and midsize publishers.
3. Major Distributors (e.g. Perseus, Independent Publishers Group, etc.) were contacted directly to secure the large list of publishers distributed by such companies for inclusion in the *BookStats* data.

For Volume Four, 2013 source data was collected from 1,616 publishers. Of these 1,616 publishers, 673 also submitted or resubmitted source data for 2012. Again, all data provided represented publishers’ net unit and dollar sales, less returns.

Publisher Banding According to Size—To ensure all market revenue segments were fairly represented in the source data received, all 1,616 publishers were grouped according to annual revenue (i.e., size) in the following way:

Publisher Size	Minimum Sample Size	Number Submitted
Large (Sales over \$100M)	15	24
Medium (Sales from \$5M--\$99.9M)	70	76
Small (Sales from \$500K-- \$5M)	127	196
Very Small (Sales under \$500K)	383	1320
TOTAL		1616

Participating publishers were asked to provide both 2012 and 2013 sales information. While the primary purpose was to provide new 2013 sales data, *BookStats* also recognized the need to collect 2012 data. Additionally, publishers who provided 2012 data were also allowed to restate their net sales if they had made adjustments after their submission for last year’s volume.

Enriched Data Reporting—*BookStats* also includes a trade-focused “Category Deep-Dive” enhancement for the Online Data Dashboard. This more granular view of the data is based on survey responses from publishers who provided detailed categories of sales reporting beyond the top-level super categories. These categories are reflective of the second-level BISAC subject headings for the general trade market. A four-part criterion was used to determine which categories would be open for analysis in the Online Data Dashboard in order to assure both accuracy and confidentiality. In order to be represented in the Dashboard, categories must:

- Include more than five publishers
- Have no single publisher represent >45% of market
- Have a source data ISBN count of >50% of total ISBNs
- A category value > \$300,000 in source data

For those categories not meeting these requirements a process was undertaken with the assistance of industry category experts to ‘fold’ them into broader categories wherever possible. A detailed list of these deep-dive categories can be found in Appendix A.

Submission Results — Overall source data was provided from 1,616 publishers. These publishers, ranged from zero dollars to billions of dollars in annual net sales. All data provided represents publishers’ net sales and revenue, less returns. While many publishers provided sales data directly, *BookStats* also received substantial data from distributors on behalf of their clients—with some choosing to remove identities of individual publishers or providing a single sales value and number of publishers.

It is also important to note that for the K-12 and Higher Education markets, a significant portion of the data (especially from the largest publishers in this segment) is derived from sales numbers collected as a part of the AAP annual collection process for those market vertices. The decision to take a ‘single feed’ for both the *BookStats* and AAP efforts was endorsed by both publishing committees in an effort to begin the process of bringing both data sets into alignment in the reporting process. The *BookStats* management team will work to achieve this goal, especially in the effort of aligning definitions of accounting for certain sales (i.e. export sales) in order to bring the two reports into greater harmony.

EXTRAPOLATING THE SIZE OF THE MARKET

Publisher Banding according to Publishing Type — R. R. Bowker provided the BISAC subject headings assigned to all titles published by active U.S. publishers. Using this data, it was possible to categorize each publisher based on the kind of titles that publisher produced. Peer Groups were created based on the BISAC subject-heading most used by a particular publisher (as shown in chart below) as well as other common characteristics.

	SMALL	MEDIUM	LARGE	VERY LARGE	TOTAL
ACADEMIC	15	5	19	13	52
ARTS	951	197	12	1	1,161
BISAC UNKNOWN	89	47	14		150
BUSINESS	301	71	9		381
EDUCATION	245	79	14	2	340
LAW	70	47	7		124
MEDICAL	146	36	8		190
NON-TRADITIONAL		2	13	28	43
OTHER	543	176	40	1	760
RELIGION	567	184	26	1	778
SMALL PUBLISHER	61,105				61,105
ADULT FICTION	1,095	224	42	6	1,367
TECHNICAL AND SCIENTIFIC	744	269	64	8	1,085
GENERAL	1,259	397	75	3	1,734
JUVENILE	336	174	62	4	576
TRADE NON-FICTION	1,369	301	37		1,707
UNIVERSITY PRESS	103	67	36		206
TOTAL	68,938	2,276	478	67	71,759

Deriving Revenue Factors — Once all 1,616 submitting publishers had been assigned a Peer Group, the following two-step process occurred:

1. Sales data from the 1,616 publishers was linked to known ISBN counts, and a per-unit lifetime revenue factor was calculated for each Peer Group.
2. All revenue factors were ranked within each Peer Group and the middle (median) value for per-unit lifetime revenue was found. The median was used in order to exclude outliers at both ends of the spectrum. Other statistical measures, such as mean, standard deviation, and skew, were used to test the reasonableness of this value.

A summary of these revenue factors is shown below.

	SMALL	MEDIUM	LARGE	VERY LARGE
ACADEMIC	3,000	39,000		
ARTS		5,000		
BISAC UNKNOWN		13,900		
BUSINESS		13,900	20,800	30,000
EDUCATION				
LAW		20,000	31,000	60,000
MEDICAL				
NON-TRADITIONAL		350		
OTHER				
RELIGION		13,900	20,800	31,000
SMALL PUBLISHER				
TECHNICAL AND SCIENTIFIC		13,900	20,800	31,000
TRADE - GENERAL				
TRADE - JUVENILE				
TRADE - NONFICTION		13,900	29,100	60,000
TRADE - ADULT FICTION				
UNIVERSITY PRESS	5,300			

Important to highlight here is the strong correlation between the number of ISBN holdings and revenue. In other words, when comparing similar publishers, those with more ISBNs typically have greater revenue than those with fewer ISBNs. To further illustrate this point using the chart above, the average median revenue for an adult fiction title produced by a medium, large, and very large publisher is \$13,900, \$29,100, and \$60,000, respectively.

Extrapolating the Total Market Size — By recombining the above matrix with the ISBN data taken from R. R. Bowker’s Books In Print, the estimated revenue and category for each title published within a specific Peer Group was calculated. The sum total of revenue for all titles in print becomes the overall estimate for the total size of the book publishing industry in the United States: \$27.0 billion in 2013. The source data was then used to reapportion the total sales estimate according to the axes on the data cube.

CAVEATS

As with any project that seeks to calculate size based on a complex and varied data set such as the publishing industry, it is important to point out the various caveats and limitations in the data, whether it is related to the collection process, the calculation of the numbers, or other contributing factors. To fully understand the numbers represented in *BookStats Volume Three* report, the following factors should be taken into consideration:

Net Units for K-12 and Higher Education — Due to the complexities associated with the nature of bundled products within the Higher Education and K-12 categories, which obscure true unit sales figures, textbook publishers were not able to furnish net unit sales. As a result, a fixed, average net unit price based on ancillary market research into average prices and annual fluctuations was used in these cases. In the case of Higher Education, this was an ANUP of \$76.13 in 2013. In the case of K-12, this was an ANUP of \$42.40 in 2013. Unit data for education markets is available in the Online Data Dashboard, but users are encouraged to treat that data as a broad or rough estimate.

Channel Reporting — The data in *BookStats* address to whom publishers sell their products; the data does not address where consumers purchase the publishers' products. Publishers sell into many intermediary channels and cannot always be certain of their products' final destinations—a publisher's sale to a wholesaler could wind up in a big-box store, a library, or an independent bookstore. In addition, sales from publishers to some specific companies are accounted for in multiple channels. For example, some brick-and-mortar retailers sell eBooks. However, because those books are sold through a website, those sales are accounted for in the "Online Retail" channel, and not in any of the brick-and-mortar retail channels. For this reason, what might appear in *BookStats* as market growth or decline in some channels might simply represent a shift where publishers are shipping their products. Finally, since the mass merchandiser, jobber, and wholesaler channels are so dependent on each other, it is difficult to provide an accurate analysis of them as separate entities. In short, while *BookStats* provides an excellent representation of how publishers interact with various channels, it should not be used to gauge the health or lack of health of any particular channel.

Mass Market Books — Mass Market is a specific format of book that typically fits into a specialized kind of display rack. Independent analysis of the industry suggests that the vast majority of Mass Market books are sold by publishers who already report to *BookStats*. As such, there was concern that apportioning the percent of Mass Market units sold by reporting publishers to the extrapolated whole would overestimate sales of that format in the study. With this in mind, the *BookStats* team conducted a thorough review of the Mass Market landscape and adjusted the methodology for extrapolating sales and units for this one particular format. Data collected through Nielsen's Books

& Consumers program was used to adjust the extrapolated result downward, to better reflect the reality in the marketplace.

The Unclassified Choice — The source data survey allowed publishers to categorize books as “Unclassified” on each axis of the data cube. Although this was supposed to be a category of last resort, allowances have to be made for those publishers that could not properly categorize their books, either because of systems limitations or an inability to extract data in a particular way. Therefore, in some cases, the unclassified choice is likely higher than it necessarily should be.

The Self-Publishing Market — While the self-publishing market continues to grow in terms of the number of books published, *BookStats* is limited in its ability to ascribe total value to this group, especially in the case where ISBNs are not utilized. While self-publisher outreach was attempted, response was relatively low. It can be concluded that this sector is underrepresented for these reasons.

DEEP DIVE CATEGORY LIST

General Category	Category Deep-Dive	Includes...
Adult Fiction	Action & Adventure	Technological, War & Military, Westerns
	Christian	
	Classics	
	Comics & Graphic Novels	
	Contemporary Women	
	Erotica	
	Fantasy	Fairy Tales, Folk Tales, Legends & Mythology, Science Fiction
	General	African American, Amish & Mennonite, Anthologies (Multiple Authors), Asian American, Classics, Coming of Age, Cultural Heritage, Dystopian, Family Life, Gay, Hispanic & Latino, Holidays, Humorous, Jewish, Legal, Lesbian, Mashups, Media Tie-In, Medical, Political, Religious, Sagas, Satire, Sea Stories, Sports, Urban Life
Historical	Alternative History	

Adult Non-Fiction	Horror	Ghost, Gothic, Occult & Supernatural, Visionary & Metaphysical
	Literary	
	Mystery & Detective	Crime
	Romance	
	Short Stories	
	Thrillers	Psychological
	Antiques & Collectibles	
	Architecture	
	Art	
	Biography & Autobiography	
	Body, Mind & Spirit	
	Business, Careers, Occupations	Business & Economics
	Comics & Graphic Novels	
	Computers, Technology & Engineering	
	Cooking	
	Crafts & Hobbies	
	Design	
	Drama	
	Education and Reference	Foreign Language Study, Language Arts & Discipline, Study Aids
	Family & Relationships	
	Games	
	Health & Fitness	
	History	
	House & Home	Gardening
	Humor	
	Law	
	Literary Criticism & Collections	
	Mathematics	
	Medical	
	Music	
	Nature	
	Performing Arts	
	Pets	
	Philosophy	
Photography		

Juvenile Fiction	Poetry	
	Political Science	
	Psychology	
	Reference	General
	Science	
	Self-Help	
	Social Science	
	Travel & Transportation	
	True Crime	
	Action & Adventure	Westerns
	Animals	
	Art & Architecture	
	Classics	
	Comics & Graphic Novels	
	Concepts	Bedtime & Dreams
	Cooking	
	Education and Reference	
	Family	
	Fantasy	Fairy Tales & Folklore, Fantasy & Magic, Legends & Mythology, Paranormal
	General	Activity Books, Art & Architecture, Biographical, Books & Libraries, Boys & Men, Business, Careers, Occupations, Clothing & Dress, Computers, Girls & Women, Imagination & Play, Law & Crime, Nursery Rhymes, Performing Arts, Readers, Robots, Science & Technology, Steampunk, Toys, Dolls & Puppets, Media Tie-In, Religious
	Health & Daily Living	
	Historical	
	Holidays & Celebrations	
	Horror	Horror & Ghost Stories, Visionary & Metaphysical
	Humor	Humor, Humorous Stories
	Interactive Adventures	
Legends, Myths & Fables	Monsters	

Juvenile Non-Fiction	Lifestyles	Gay & Lesbian
	Love & Romance	
	Mystery & Detective	
	Nature & The Natural World	
	People & Places	Royalty
	School & Education	
	Short Stories	
	Social Issues	
	Sports & Recreation	
	Travel & Transportation	
	Animals	
	Art, Architecture & Photography	
	Biography & Autobiography	
	Concepts	
	Cooking	Cooking & Food
	Education & Reference	Books & Libraries, Business & Economics, Careers, Computers, Education, Foreign Language Study, Language Arts, Literary Criticism & Collections, Study Aids
	Family	
	General	Activity Books, Antiques & Collectibles, Boys & Men, Clothing & Dress, Comics & Graphic Novels, Crafts & Hobbies, Curiosities & Wonders, Drama, Games & Activities, Gardening, Girls & Women, House & Home, Humor, Law & Crime, Lifestyles, Media Tie-In, Performing Arts, Philosophy, Poetry, Social Issues, Social Science, Technology, Toys, Dolls & Puppets, Transportation, Travel
	Health & Daily Living	
	History	Adventure & Adventurers, Ethnic, Politics & Government
Holidays & Celebrations		
Mathematics		

Religion	Mystery & Detective	
	Nature	
	People & Places	
	School & Education	
	Sports & Recreation	
	Bibles	Biblical Commentary, Biblical Reference
	Biblical Studies	
	Buddhism	
	Christian Life	
	Christianity	Christian Ministry, Christian Theology, Christian Education
	Comparative Religion	
	Devotional	
	General	Antiques & Archeology, Atheism, Biblical Biography, Biblical Criticism & Interpretation, Biblical Meditations, Christian Church, Christian Rituals & Practice, Church & State, Clergy, Confucianism, Counseling, Cults, Deism, Discipleship, Eastern, Eschatology, Ethics, Ethnic & Tribal, Faith, Fundamentalism, Hinduism, Holidays, Institutions & Organizations, Islam, Jainism, Jewish Life, Leadership, Monasticism, Mysticism, Psychology of Religion, Religion & Science, Religion, Politics & State, Rituals & Practice, Sermons, Sexuality & Gender Studies, Shintoism, Stewardship, Taoism, Theology, Youth Ministries
	Judaism	
	Meditations, Prayer & Prayer Books	
	Philosophy	
	Reference	K-12
Spirituality		

APPENDIX B

BOOKSTATS VOLUME FOUR DATA DICTIONARY

Full list of terms, and their business definitions

Net Sales:

Gross dollar sales less returns dollars.

Billings to customers with all discounts deducted and exclusive of transportation charges. Does not include revenue from consignments except for books recorded as sold.

Does not include income from subsidiary rights or from non-publishing activities such as printing plants, sales representation, or fulfillment service operations. Subscription sales of all print and digital content recognize the full value of the subscription in the year the subscription started or was renewed.

Unit Sales:

Gross unit sales less units returned.

Category:

All trade books created for, and sold primarily as trade titles, were submitted at the top-level BISAC Subject Heading category for Trade publishers and as defined by the submitting publisher. A proprietary category structure supplied by the AAP was used for K-12 and Higher Education publishers. These detailed category views are available for an additional fee within the Category Deep-Dive enhancement in the Online Data Dashboard -- see www.bookstats.org for more details.

Note: If a publisher was not able to provide detailed category sales information, use of a higher categorization was requested which covers the following categories:

Adult Trade: Content marketed primarily for an adult or general audience. Segregated by the following subcategories:

Adult Fiction

Adult Non-Fiction

Juvenile Trade: Content marketed primarily for children, teens, and young adults. Segregated by the following subcategories:

Juvenile Fiction

Juvenile Non-Fiction

Religion: Religious content such as Bibles, as well as biblical studies titles, testaments, histories, spiritual titles, hymnals, and prayer books. Also includes religious non-fiction. Segregated by the following subcategories, if possible:

Bibles

Religion: All Religion Titles (excluding Bibles)

K-12: Textbooks, workbooks, review books, standardized tests, digital textbooks, and course material marketed primarily for the K-12 school market.

Higher Education: Textbooks, workbooks, review books, standardized tests, digital textbooks, and course material marketed primarily for the Higher Education market.

Professional: Content marketed for the Professional Publishing market; primarily Technical, Scientific, Law, Business, and Medical.

Scholarly: Content marketed for the Scholarly Publishing market; primarily books and journals which are used in the humanities, literature, and the arts.

Uncategorized: All other categories without a defined classification; to be used as a last resort.

Format

The form in which a book is delivered, in terms of binding type, media, or electronic method.

Physical: Includes all products that consist mostly (more than 50%) of print-on-paper materials or are sold on some physical medium other than paper, such as electronic products delivered on disk or as preloaded audio.

Hardcover: Printed books bound in hardcover.

Softcover: Printed books with flexible bindings, typically of better production quality, larger size, and higher price than rack-sized, Mass-Market Paperbacks. Includes digest format, spiral or comb-bound, or other specialty bound books that do not have hardcovers.

Textbooks: Printed books whose purpose is to educate and are used primarily by educational institutions.

Mass-Market Paperback: Rack-sized printed books with flexible bindings measuring c. 4½" by 7", as well as products in other physical formats intended for sale in what the industry has historically called the mass market, where books are most often defined as having strippable covers.

Audio - Physical: All audiobooks sold on physical media. Includes CD, MP3 CD, preloaded single use devices, CD ROM, and tape.

Physical - Other: All physical products without a defined classification above.

Non-Physical: Includes all products delivered in electronic format without any physical delivery mechanism.

Ebook: Digital content analogous to printed work. Includes direct and almost direct facsimiles of printed originals as well as e-original books as appropriate. These books may contain some hyperlinks, but if extensive conversion has taken place or the title is not book-like, it should be classified in one of the formats below.

Enhanced Ebooks and Paid Mobile Apps are also included within the general ebook format.

Digital Course Materials: Digital materials whose purpose is to educate and are used primarily by educational institutions.

Audio Download: Audiobooks delivered electronically.

Internet Products and Services: Book content delivered online. Includes all products and services that REQUIRE web access for use. For example, subscription models, homework managers, and portions of books sold separately.

Non-Physical - Other: All non-physical products without a defined classification above.

Bundled Product: All bundled items without a defined classification above because providing a defined classification was not possible or desired. Bundled Products may include any variety of print, electronic, or custom-created items.

All Other Formats: All other formats without a defined classification; to be used as a last resort.

Channel

The channel into which the publisher sold a book.

Retail Stores (Chains): Sales to all national and regional retail stores whose main business is selling books. Also see Online Retail below.

Retail Stores (Independents): Sales to independent retailers not affiliated with bookstore chains. Also see Online Retail below.

Retail Stores (College): Sales to on-site or off-site textbook bookstores that sell to a local university campus.

Retail Stores (Mass Merchants): Sales to large chain retailers including grocery stores, price clubs, warehouse stores, airport stores, and the like.

Retail Stores (Specialty): Sales to both chain and independent retailers whose principal business is not bookselling; includes gift stores, museum stores, dollar and discount stores, home improvement, cookery, and other similar stores.

Retail Stores (Unclassified): All other retail store sales without a defined classification above (i.e., Chains, Independents, Mass Merchants, or Specialty).

Online Retail: Sales to online retailers. For online retailers that also maintain physical Retail Stores, sales were apportioned in the Online Retail and Retail Stores categories, as appropriate and if possible.

Direct-to-Consumer: Sales directly from publisher to consumer; includes sales from publisher websites, telephone sales, sales at conferences, and sales to authors.

Institutional Sales: Sales to businesses, for-profit and not-for-profit groups, including libraries, schools, government agencies, and associations; includes premium sales.

Book Clubs and Fairs: Sales to companies that resell through clubs and fairs. Does not include income from subsidiary rights sales that do not involve the provision of physical books.

Export Sales: Direct export sales. Does not include sales to foreign subsidiaries.

Jobbers and Wholesalers: Sales to agents and companies that resell books to retailers, libraries, schools, and other institutions.

All Other Channels: All other sales without a defined channel were to be placed into the All Other Channels category.