

Increasing
Book Sales and

Visibility with
Keywords



KADAXIS

Increasing Book Sales and Visibility with Keywords

A research report by



KADAXIS

version 1 | February 8, 2017

The Hypothesis	2
Research Approach	2
What are Keywords?	3
How Keywords Affect Book Sales	3
The Best Source of Keywords	4
Keyword Timing and the Virtuous Cycle	5
Research Results	6
Dover Publications	7
Kaplan	10
Yale University Press	13
Andrews McMeel	14
Conclusions	15

The Hypothesis

Firebrand Technologies and Kadaxis began a research project in July 2016 to study the impact of targeted keywords on the sale of books. The hypothesis of this ongoing research project is that by analyzing the natural language and common phrases that readers use in describing a book in reviews and other online social media, we can determine the best possible keywords that will aid in discovery, increase visibility, and ultimately improve sales.

This keyword analysis process can be aided by machine learning, and can be performed at scale across large lists of titles and across many retailers. Using audience analysis we can create high value keywords that are better targeted for readers, which will improve conversion rates and visibility.

Research Approach

To test this hypothesis, Firebrand and Kadaxis teamed up with four publishers in July 2016: Andrews McMeel, Dover, Kaplan, and Yale University Press. We also added Abrams and University of Nebraska Press to this research project toward the end of that year. Through these partnerships we have tried to create a cross-section of publisher types to test how audience analysis can have the broadest impact on the marketplace. These are the companies and the people that we know like to experiment and like to learn.

We also wanted to select titles that would be considered in a steady state, books that were relatively predictable in volume. Each publisher chose twenty titles from the backlist even some titles from the deep backlist. We looked at sales and activity retrospectively for two years, and we are planning on collecting at least one full year in forward-looking results.

Our research has been primarily focused on Amazon. Amazon is currently the only major book retailer that accepts keywords that are sent in ONIX metadata, and the presence of keywords are actually utilized in Amazon's search algorithms.

During the course of this research project we have been ingesting the publishers' sales data across retailers and other sources, as well as Amazon metrics when they were are available. This data is being imported into the Firebrand Research Labs platform and analyzed to track trends.

What are Keywords?

Before we show the results of the research, it is informative to discuss keywords in more detail. The textbook definition of keywords is a little bit dry, but in essence it is this:

Keywords are words or phrases used to describe books. Keywords are assigned to metadata and they supplement other metadata like the title, author, and description. Keywords are added to an ONIX file and sent from publisher to retailer, and then indexed for search.

This definition is accurate, but it completely leaves out the motivation—why do we actually use keywords and what are the benefits behind using them?

On the surface, keywords are just a metadata, but used properly they can be a powerful discovery mechanism to capture a reader's experience with a book in a way that facilitates sharing that experience with others. Creating keywords is an exercise in studying reader psychology and linguistics, requiring empathy and insight into how people communicate about books with each other. If you're able to think and talk like a reader, you're more likely to reach them. If you reach them, then you can sell your books to them.

How Keywords Affect Book Sales

Most books are sold online today, most books are sold on Amazon, and most people find them through search.**[citation needed]** So, if you can improve a book's visibility in search on Amazon, you can improve its likelihood of selling more copies. Effective keywords directly increase the exposure of the book to potential consumers, giving you more chances to sell that book. The math on this is simple:

If a book shows up a thousand times in search results, then you add keywords and the shows up two thousand times, you have increased the number of chances to sell that book by one

thousand.

The more data a publisher supplies to any retailer, providing it is good quality data, the easier it is to match their books to more search queries. This is true for Amazon even more so than some other retailers. If you don't give Amazon keywords, they'll match your book to search queries based on the standard metadata: categories, description, title, and so on. However, if you give them fifty keywords, they'll look at that standard metadata, *as well* as trying to match your book to each of those fifty keywords, plus the derivative versions of them.

Therein lies the big opportunity for publishers. Most of the books for sale on Amazon right now don't have keywords assigned to them. In 2016 Kadaxis and Bowker [ran a study](#) that looked at 115,000 different publishers and their keyword usage. The study found that 15% of those publishers had at least one book with keywords. However, of the books that did have the keywords, most were completely ineffective, including redundant data like the format of the product (audiobook, ebook, etc.), the title, and the author.

What this means is that publishers who add keywords to titles today have a significant advantage in visibility over publishers who don't. This is a testament to the fact that keywords have become more important and beneficial in the last few years, and most publishers just haven't caught up with their backlists yet. As time goes by and more publishers catch on, the improvements that we're witnessing today are going to be less significant. Once there's more prevalent coverage of baseline keywords across publishers' backlists, then the focus will be on tuning those keywords, which will be a lot harder to do.

It is also important to remember that keywords at Amazon represent a low risk. If you accidentally say the wrong thing (for example, inadvertently adding some offensive text or misspelling a keyword) you will not be penalized by Amazon's search system because the keywords are pretty much invisible to the reader. All you need to do is re-send the keywords to correct the mistake.

The Best Source of Keywords

In many circles the editorial and marketing teams are considered to be the best source of keywords for a book. However, these groups often have a more sophisticated perspective of their books, describing books in terms that are highly appropriate for jacket copy but aren't always how readers themselves would describe them. For example, an editor might use these keywords to describe a mystery suspense novel: "noir atmosphere," "urban settings," "world-weary protagonists," "harsh realism," "psychological depth," "wistful poetry," "atmospheric," "urban landscape," "violent portrait."

While these are all accurate and beautifully crafted phrases, readers of the genre are more likely to search for less sophisticated terms like "bad guys," "FBI," "action-packed," "surprise ending," and "courtroom drama." These keywords provided by the reader are very different from the more sophisticated keywords provided by the marketing department, and they provide

insights into how readers think about a book.

Understanding the reader is synonymous with identifying your book's target market, or target audience. From a pure marketing perspective understanding the target audience involves understanding the book contents, analyzing who the reader is, developing personas, researching demographics, identifying customer intent, understanding where readers buy, and so on.

These are all very useful and practical book marketing research activities, but they do not necessarily impact connect with the reader as well as just listening to what the reader is saying about the book. The language most readers use to describe a book is also the language most consumers are likely to use in searching for a book. This means that the most efficient method of scaling discovery, and thus the best way to connect with potential consumers, is to highlight the topics readers of the book find most appealing and express them in that language.

Check back in this report later for information on the keyword creation process.

Keyword Timing and the Virtuous Cycle

It can take some time for keywords to affect sales, so the general rule is that the sooner you add keywords to the title, and the more often you refine those keywords, the better the results will be. This refining process allows you to proactively capture future trends. Capitalizing on the initial spike of traffic for a specific topic leads to significantly higher returns than being reactive and updating your keywords when everyone else is updating theirs.

This also ties into the concept of the Virtuous Cycle (Figure 8). The Virtuous Cycle refers to how Amazon's algorithms react to book sales and increases in book. If a book does well on Amazon, they give it a lot of extra love. For example, if you assign twenty-five keywords to a title and ten of those keywords rank high in search results, then the title will see a spike in sales. This also leads to an increase in the ranking for significantly more keywords. Amazon's system sees these ranking increases and the title will start to show up on the "frequently bought together" lists, bestseller lists, and similar recommendations.

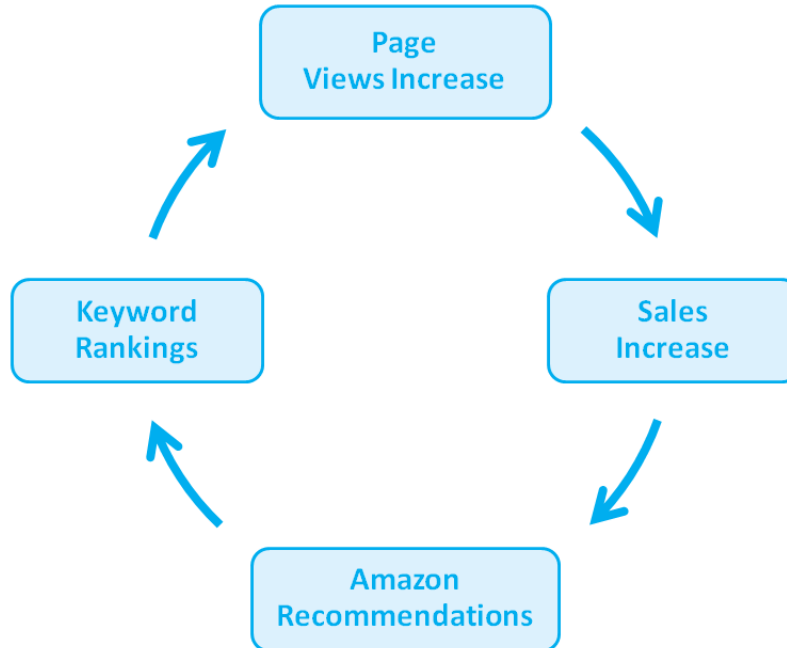


Figure 8. The Virtuous Cycle

This Virtuous Cycle leads to more visibility, which leads to more sales, and back around the circle. Amazon knows that books that sell well will typically be more likely to sell well again in the future, so it uses the virtuous cycle to encourage those sales.

This strategy can also tie into marketing campaigns. If you are planning to do a paid marketing campaign with Amazon, it can be beneficial to update the keywords right before the campaign starts so that you can increase the number of sales for the book, giving Amazon the ability to then leverage more of that data during the campaign.

Now, keywords are powerful, but their effect in the market is usually very slow. You are not likely to see overnight changes. The best approach is to update the keywords for a large number of titles, wait for six to twelve months, and review from there.

Research Results

Now that you understand the approach we took during this research project to audience analysis and the creation and application of keywords, we would like to share our findings.

As was mentioned above, very few publishers are currently adding keywords to their titles, and most of the keywords that are being added are quite ineffective. This current state of affairs represents a unique opportunity for early adopters to increase the sales ranking and positioning of their titles in the marketplace.

Our research unequivocally shows that targeted keywords work. This single piece of

metadata—keywords derived from an analysis of reader language—had a measurable impact on visibility in the marketplace and on sales. While not every title in the study was impacted in a measurable way, the impact on some titles was extraordinary, and we saw no downside to applying keywords to titles that were not impacted.

Instead of just providing you with data, we would like to tell the story of some key example titles from each of the publishers who participated in this research project. Each publisher submitted twenty titles for the project, and the sections below provide a snapshot of our key findings.

Dover Publications

Dover Publications saw some of the most impressive increases in sales during the testing period. Figure 9 shows the year-over-year increase in unit sales for all twenty of their initial test titles. Dover has a large backlist with consistent sales, and none of those titles had keywords, so their selection of test titles was a great fit for the program.

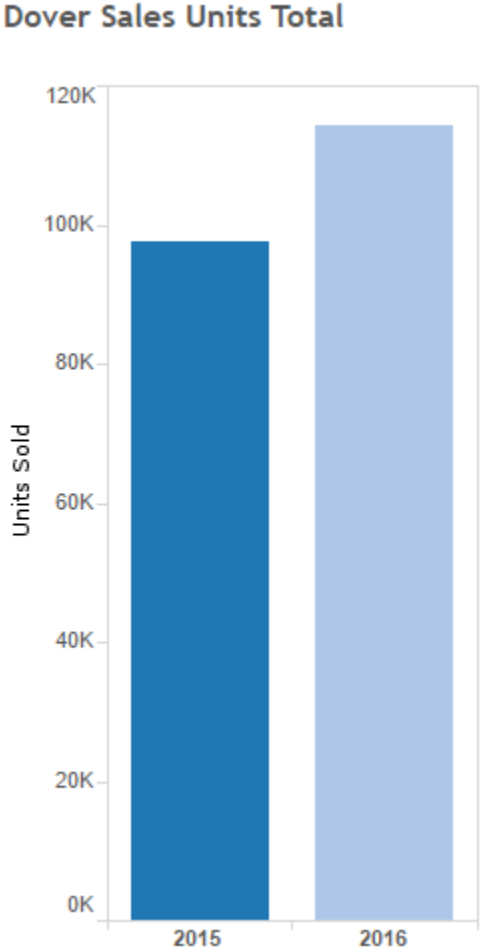


Figure 9. Dover unit sales year-over-year

Narrowing down the results to look at just a few titles (Figure 10) is even more interesting. *The*

Technical Manual and Dictionary of Classical Ballet, the title published in 1967 which we discussed earlier, stands out. We placed Kadaxis' suggested keywords into Title Management, and sent them out through the Eloquence on Demand service in ONIX to Amazon. No other metadata on the title was updated. Figure 11 provides the difference in sales for this title over the period of August through November in 2014, 2015, and 2016, showing a large jump in sales in the months following the addition of the keywords and an overall sales increase of 25% over the testing period. Don't forget, that title was published in 1967. It is a steady backlist title with no marketing campaigns or other influencing elements. This spike in sales was only attributable to the addition of the suggested keywords.

Dover Sales Units YOY Highlights

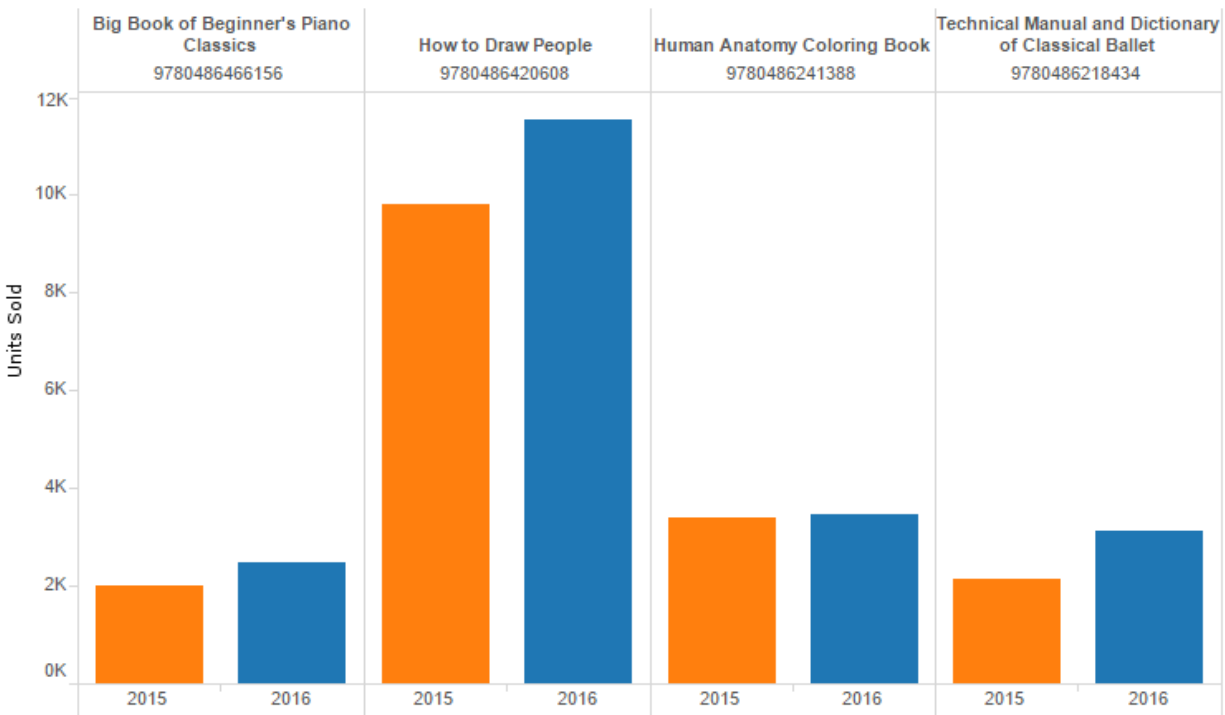


Figure 10. Specific Dover titles

Sales Units YOY Ballet

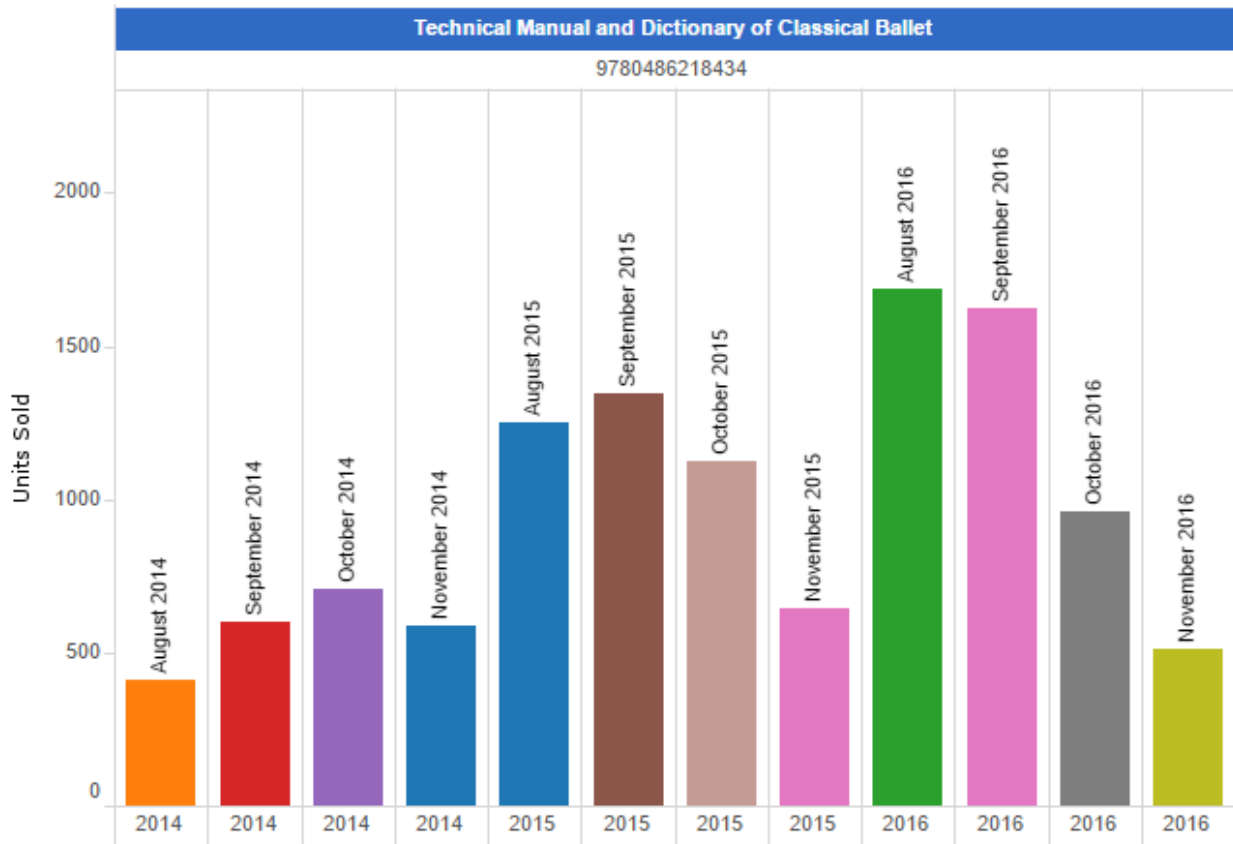


Figure 11. Year-over-year unit sales for the *Ballet* book, Aug-Dec, 2014-2016

Another example from Figure 10 is *The Human Anatomy Coloring Book*. At the Digital Book World conference in January 2017, Jonathan Stolper from Nielsen showed how adult coloring book sales dropped dramatically between 2015 and 2016 (see page 15 of his presentation [here](#)). Despite this overall trend, *The Human Anatomy Coloring Book* stayed steady and other coloring books in Dover's list actually *increased* in sales during this testing period in 2016, when the overall market was not just down, but dramatically down.

In October 2016, following the success of the first few months of the study period, Dover increased the number of titles in the keyword program to one thousand. As can be seen in Figure 12, just one week after the release of this new set of keywords, the sales on 11 of Dover's titles went up 54% year-over-year; within 4 weeks sales had increased an average of 70%. Those same 11 titles had been down 11% year-over-year in the seven weeks leading up to their inclusion in the program and no other marketing was done to target those titles specifically.

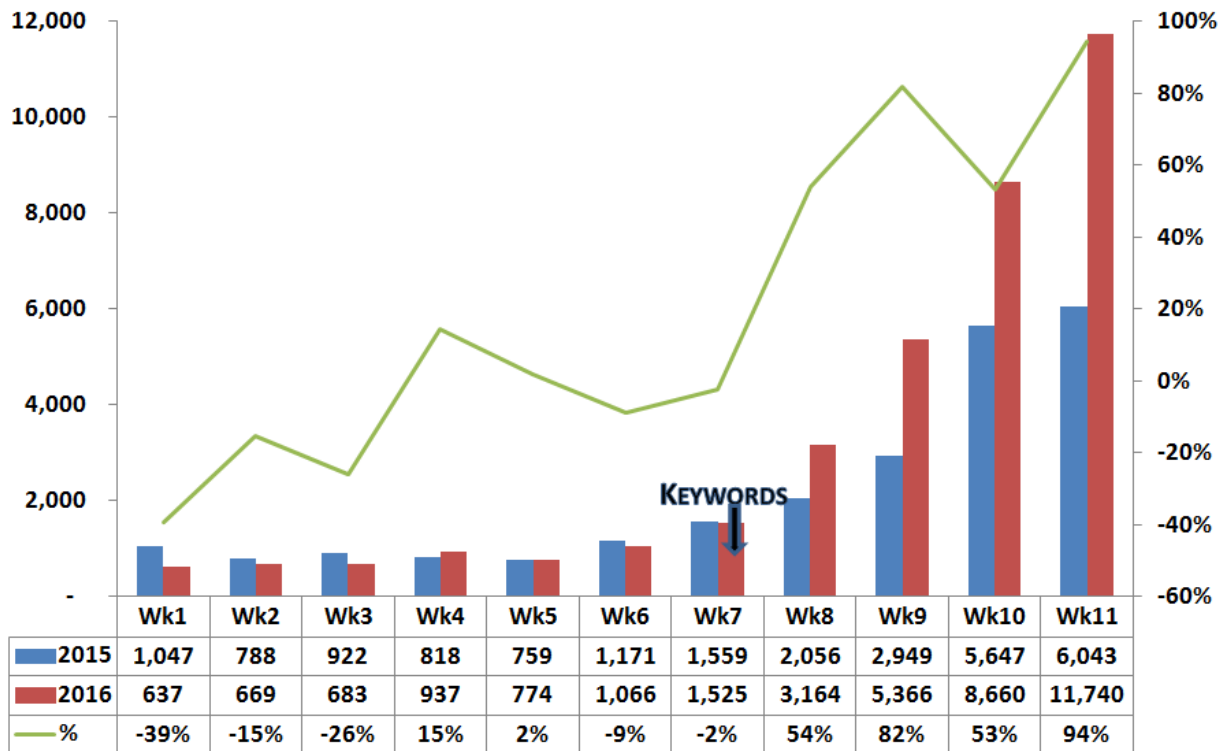


Figure 12. Dover sales after keywords were added to 1,000 titles

Upon looking deeper, Dover noticed that three of those eleven titles were not even in the keywords program at all. They are related to some of the titles in the program, and Dover had those relationships defined in their ONIX feed, but they themselves were never assigned keywords or adjusted in any other way. That is the Virtuous Cycle in action.

Sales on those one thousand titles have increased so much that Dover has had trouble keeping many of the books in stock.

Kaplan

Kaplan publishes educational titles, especially prep books for various tests like the LSAT and GMAT. However, *The Anatomy Coloring Book* is an interesting example because it is not tied to any specific test or testing date like other Kaplan titles so it is not affected by seasonal sales changes. *The Anatomy Coloring Book* already had some keywords assigned to it, but the recommended keywords were added to the ONIX file in August 2016. The title saw an immediate spike in pageviews and sales on Amazon (Figure 13).

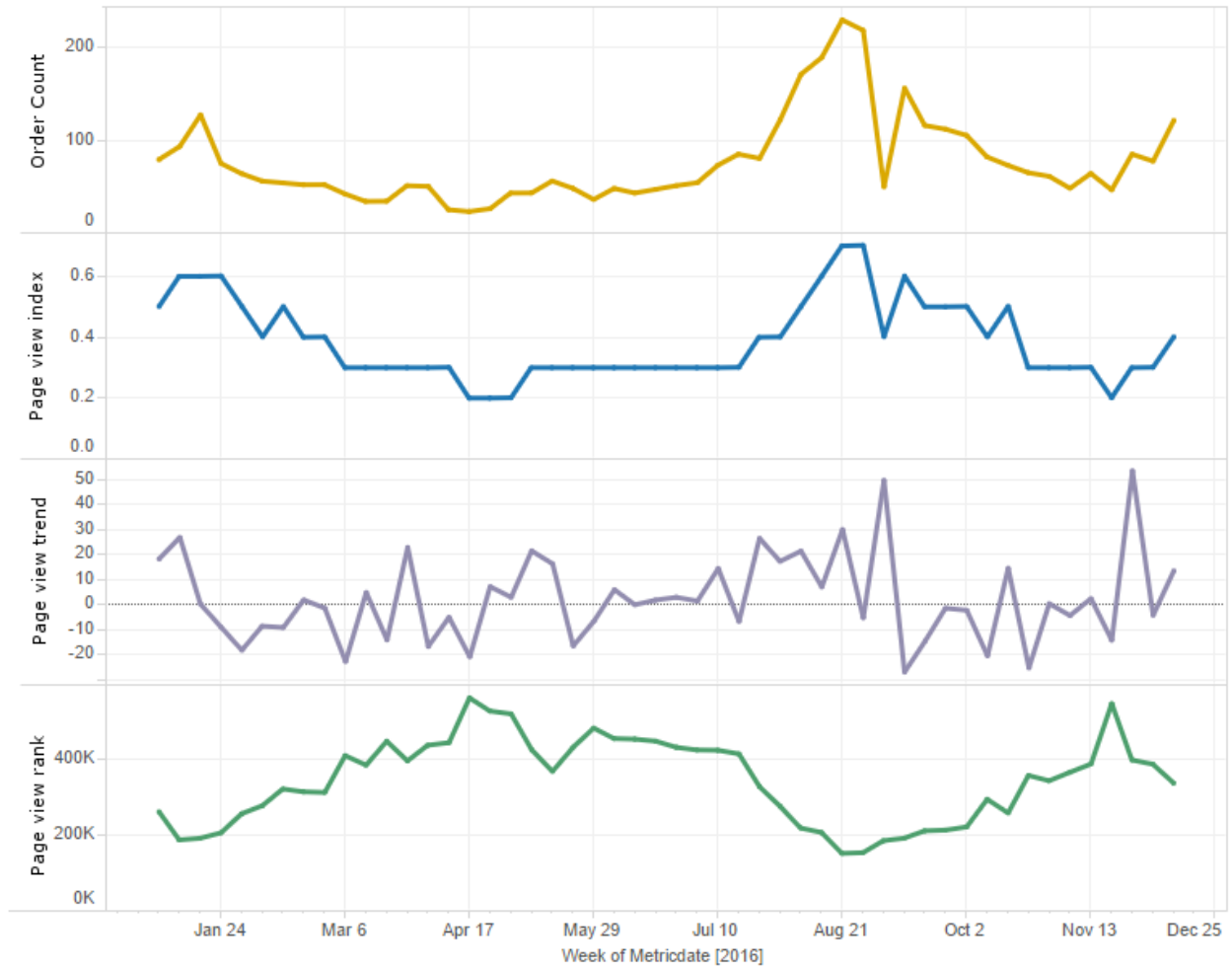


Figure 13. *The Anatomy Coloring Book* sales and page views

Kaplan also included some of their test prep books in the study, such as the *GMAT Math Workbook*. This title started without any keywords, and it also saw a bump in sales and pageviews after the keywords were added. In addition to the titles that were given keyword changes, Kaplan also decided to track the numbers on some control titles that were not adjusted. The control title for the *GMAT Math Workbook* was the *GMAT 800*, which is a similar title that was published at around the same time. As Figure 6 shows, there were tangible results from adding the keywords; while the *GMAT Math Workbook* increased in sales, the *GMAT 800* title stayed flat (Figure 14).

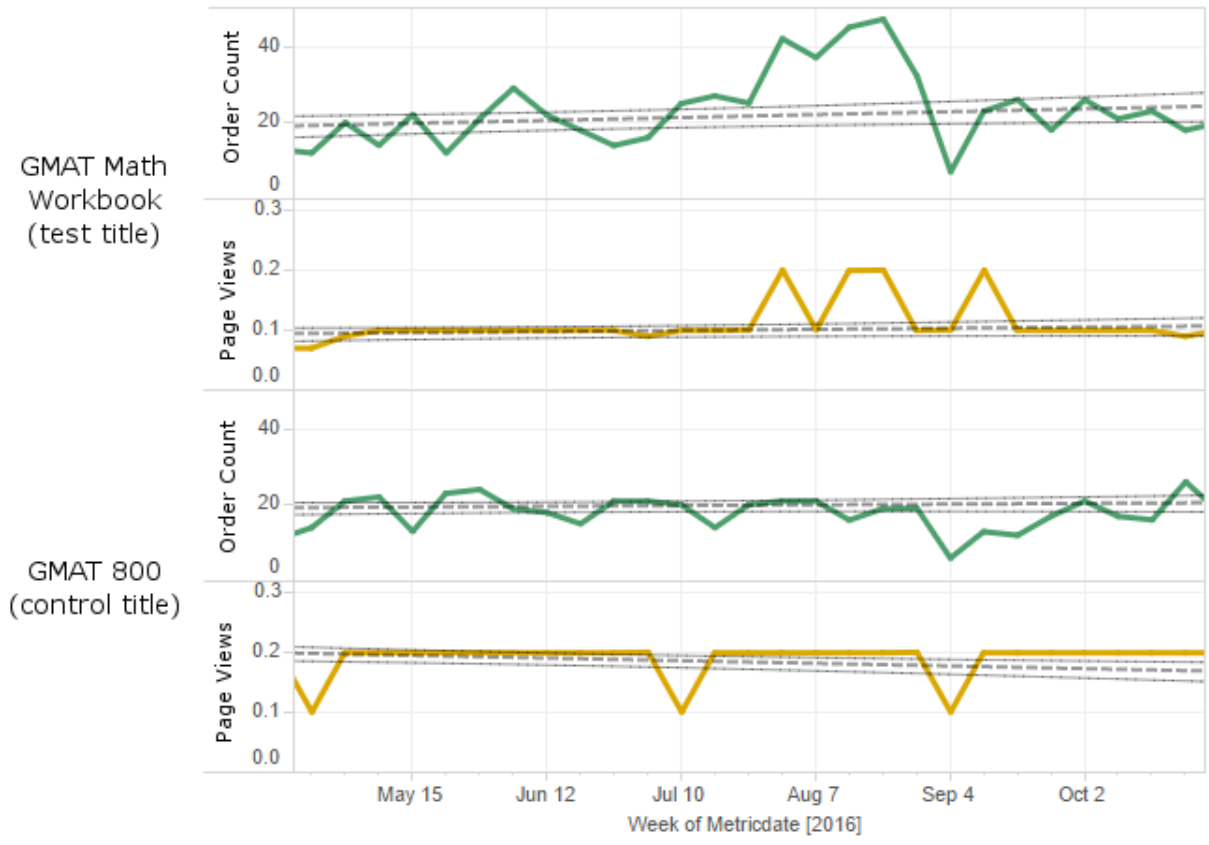


Figure 14. Comparison between *GMAT Math Workbook* (test title) and *GMAT 800* (control title)

The *NCLEX-RN Drug Guide* also saw some differences compared to its control title. However, another NCLEX title in the program, *NCLEX-RN Premier 2016*, did not change from its expected numbers. This underscores our earlier point that while there are no downsides to adding keywords to a title, there are also no guarantees that adding them will result in an increase in sales. The *NCLEX-RN Drug Guide* did see that increase—not just once, but twice, both times after the keywords were adjusted in the ONIX feed (Figure 15).

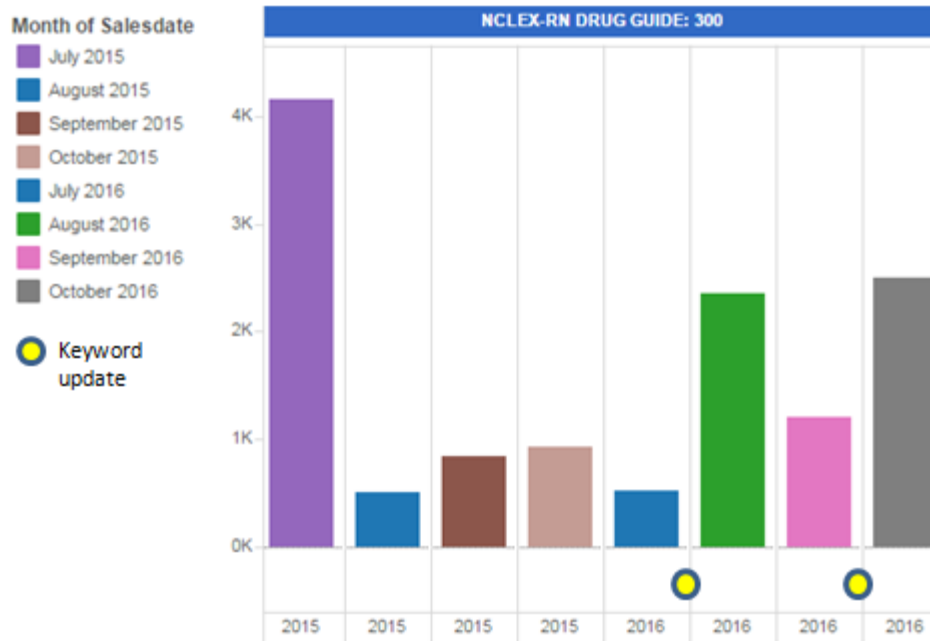


Figure 15. Unit sales of *NCLEX-RN Drug Guide* after keyword adjustments

Yale University Press

Yale University Press chose titles from their backlist that had under 5% variation over the course of four years excluding the first two years after their release. Despite this initial effort at choosing consistent titles for the test, it has been difficult to tell if the keywords made an appreciable difference in the sales of most of those titles. While many of the test titles continued to have steady sales over the test period, the company saw downward trends in their overall book sales over the same period. The question that remains is whether the steady sales of the test titles is the result of the positive impact of adding the keywords to those titles, or if that steadiness can be attributed to the consistency of those titles in general.

One title that did see an increase in sales was *How Rome Fell*, one of Yale's crossover titles that has both scholarly appeal and trade appeal. This title's use in college classes typically leads to higher sales at the beginning of the school year, and its trade appeal leads to more sales at the holidays. However, the recommended keywords were added to the title in July, and the sales started to increase immediately, outside of the normal school-related sales. Again, this could be attributed to an early school effect, but it does appear that the keywords had an immediate, if small, impact (Figure 16).

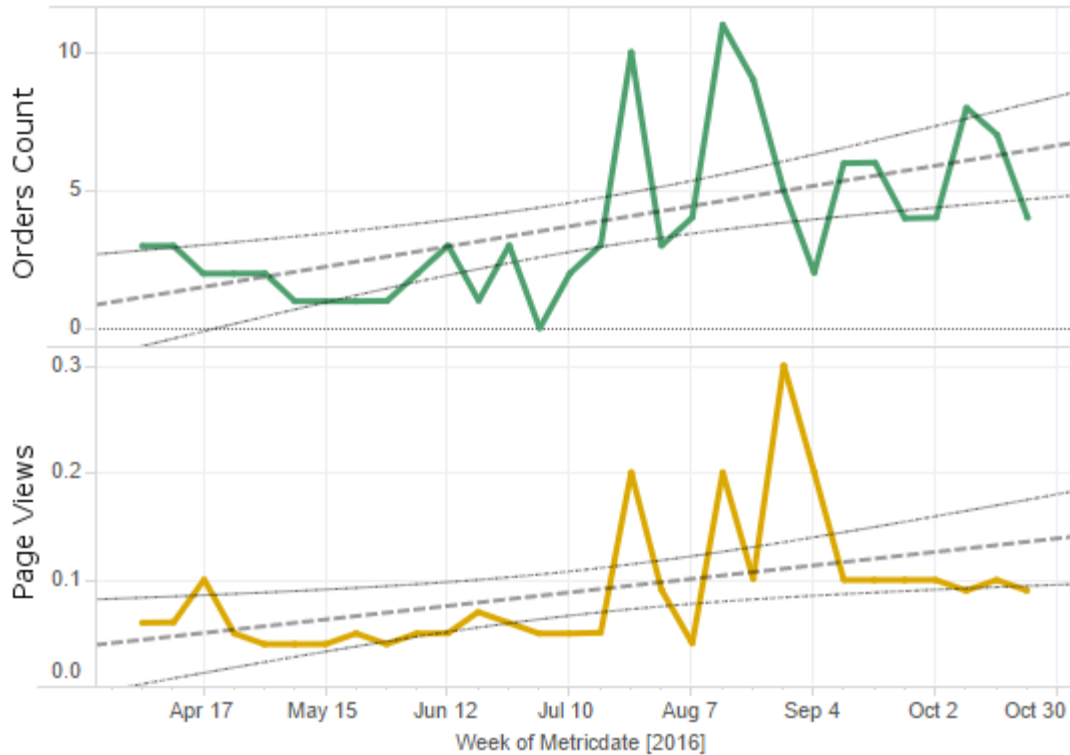


Figure 16. Sales of *How Rome Fell* after keywords were added

This title, as well as others in Yale’s test list that did well during the research period, provided Yale with the impetus to address keywords and other metadata quality in their titles in a more targeted way, focusing on marketing their titles in the trade differently than they might market them in the university press arena. It will be interesting to see how those titles fare over a longer period of time.

Andrews McMeel

Our final example is a title from Andrews McMeel called *Today I’ll Be a Princess* (Figure 17). This illustrated children’s book saw a spike in sales in May 2016 that was not related to the keywords project at all. As a matter of fact, that spike is still a mystery, as Andrews McMeel did not run any marketing programs for the book or do anything else to promote it that month.

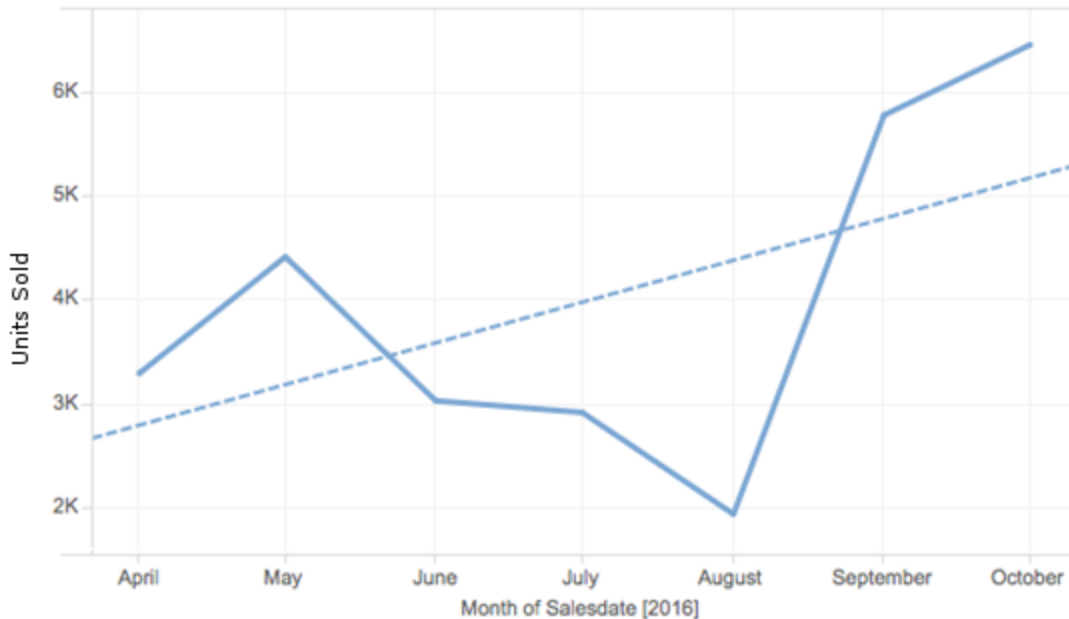


Figure 17. Sales of *Today I'll Be a Princess* after keywords were added

However, the sales tapered off again before the keywords were added in August, and there was an immediate effect, with sales jumping over 4,000 units in just one month. In addition, Andrews McMeel looked at examples of the Amazon carts of these book buyers and noticed that the book was being purchased along with princess dresses, wands, and tiaras.

This title is now ranked number 1 in some very good keyword searches, including “princess book”, which is a search phrase normally dominated by Disney’s princess books. It also serves as an example of the Virtuous Cycle effect keywords can have on books: an increase in visibility leads to an increase in sales, which leads to an increase in visibility, etc.

Conclusions

This research project has brought to light a couple of key conclusions that deserve specific attention.

1. Readers know the book best. The words readers use to describe a book in reviews and other online sources are the best words to use in the title’s keywords. This is, obviously, easiest to do with backlist titles, which have reviews that can be mined. For frontlist titles, the keyword recommendation process is a bit different. The common approach is to start with comparative titles chosen by the publisher, creating keywords from the reviews on those titles. However, we are also starting to do some tests with NetGalley to take advantage of pre-publication reviews to gather that audience language for frontlist titles.

2. Not all titles will benefit immediately. As we saw above, there are some titles that did not see an increase in sales over the initial test period. We will continue to track those titles and see

if the keywords affect them in the future. However, there will be some titles that just can't be impacted with this keyword approach. For example, we have seen the most powerful effects on non-fiction titles, not on fiction.

3. Keywords are not a short-term strategy. Yes, many titles in the test program saw an immediate effect after the keywords were added, but the goal, and the expectation, is that the addition of keywords will have a long-term effect and allow publishers to see steady increases in sales for backlist titles, not just spikes. Additionally, there is a lot of value in coming back to titles again in the future and re-running the analysis to take into consideration the additional reviews that will come with those increased sales. The changes may be minor in some cases, but that's an essential part of the reviewing process.

4. Keywords increase visibility in search. This is easily tracked by comparing the ranking of a book for certain keywords before the recommended keywords are added and after. This is also clear from the charts shown above, where pageviews for many of the test titles increased dramatically after keywords were added.

5. Better visibility leads to more sales. As was mentioned above, if you increase the number of times someone sees the title by 1,000 you have increased the number of opportunities to sell the book by 1,000. This is easily seen in the comparison between pageviews and sales in the examples given.

Thank you for taking the time to read this report. As was mentioned above, we expect to continue this research project until at least Fall 2017, and we will update this document with more results, conclusions, and interesting details along the way.

If you are interested in receiving an email whenever this report is updated, please fill out [this form](#).