

EPISODE 57: SPIDERMAN – EXPLORING AMAZON WITH DATA GUY

James Blatch: Hello. Before we get going with this podcast, and it's a good one ... It's an intriguing one actually, so definitely keep listening. But this is James to tell you that if you're listening to the podcast at the time of its release, end of March, beginning of April 2017, then Mark Dawson's Ads for Authors course is open. We opened it once last year.

This is a rare event. We opened it for 10 or 11 days last year. It's open for enrollment right now. You can go to selfpublishingformula.com/ads17, that's A-D-S-1-7, ads17. What can I say about the course? Well, I suppose what I should say is that the best thing that I ever get to do in Self Publishing Formula is to visit some of the students who've taken it in the past.

Speaker 2: Oh, the impact that Mark's course has made on my writing career, well basically is that it's given me a career.

Speaker 3: I was just a freelance writer. I was spending all my time writing resumes and cover letters for people, and now I am writing fiction. I'm supporting my family on fiction and we're actually making more money than we ever had with my passion.

Speaker 4: Mark has opened my eyes to a lot of things. I was using Facebook ads pretty quickly and I found all the sudden I started getting subscribers, and then I started getting sales.

Speaker 5: I decided to put a box set together. The sales that I get from that box set I can almost exclusively credit Mark Dawson for all of them.

Speaker 6: From day one when I first typed a word of the book I said all I want to do is support my family by doing something I'm passionate about. What this course enabled me to do in a matter of months was exactly that.

James Blatch: There you go. They say it better than I can. The course is open now. If you think it's for you and you think it's going to make a difference to your career, now is the time to get on board. If you're not sure if it's for you and it definitely isn't for everybody, you should be in a certain place with your writing to commence on this course, you can still pop along to the URL that I gave.

If you're watching this on videos on the screen as well, selfpublishingformula.com/ads17, we even have a live chat. Mark and I and John and Katherine and Alexandra and Susan and the team we've got together will be there answering questions for you.

We're not going to be hard sell. We're not hard sell people. If you say, "This is my situation, is the course right for me," and if the answer's no we'll tell you that as well.

We do offer a 30 day money back guarantee. I am sounding a bit sales-y now, aren't I? But that's always quite good advice from us is basically have a look at the course for a few days and have a really good dig around.

You can even implement the strategies, and if you think it's not for you, within a month you get all your money back no questions asked. Anyway, that's it. It's a rare thing for us and it's the one big commercial thing we do. It pays for the podcast and everything else, so it is important to us, but it could be important to you if you're in the right place to utilize the Ads For Authors course. Now is your chance to get in. Okay. Let's get on with this podcast.

Narrator: Two writers, one just starting out, the other a best seller. Join James Blatch and Mark Dawson and their amazing guests as they discuss

how you can make a living telling stories. There's never been a better time to be a writer.

James Blatch: Hello and welcome to the Self Publishing Formula podcast with Mark Dawson, also starring James Blatch.

Mark Dawson: Hello.

James Blatch: That sounds like it's been gone through everyone's agents and legal wranglings. Is that the gaffer tape? The gaffer tape ruins the magic, doesn't it?

Mark Dawson: For listeners, as you always forget-

James Blatch: I'm a visual man.

Mark Dawson: ... there is some gaffer tape on the desk we've just used to secure John Dyer.

James Blatch: Yes.

Mark Dawson: I'll just remove that to hide the evidence.

James Blatch: We have an increasing number of people who watch it on YouTube, but yeah, many thousands listen only on audio. We should always remember that. This is first and foremost a podcast but you can watch us in glorious color, high definition, on our YouTube channel as well. I frequently make references to what's around us. For people who can't see it we're actually in a sauna, just got very small towels wrapped around us. Haven't we? John Dyer's actually completely naked, just reclining next to us.

Mark Dawson: Wrapped in tape.

James Blatch: Wrapped in tape now. If that doesn't get people rushing to our YouTube channel I don't know what well.

Mark Dawson: Contacting the police.

James Blatch: Or contacting the police. Okay.

We have a really interesting interview today, actually quite an interview that will hopefully have you more enthused than ever about the work that you do and the dream that we all have being in the indie author space. It's about really the business of indie book publishing, where it is, where it is versus traditional publishing and publishing in general, but particularly indie publishing.

That's the main focus of this because there's a quarterly report that comes out called The Author Earnings Report. You can go to authorearnings.com and download them. It's very high profile. We're going to hear a little bit about the history of it in a moment.

But the primary way it is put together is by a single individual who is known only as Data Guy. He's well known in the industry for this. He's got his own motif, which is a spider. The reason for that is this concept of a spider that crawls across the websites, the various websites such as Amazon, and scrapes off the data that's available to you and turns it into meaningful charts and graphs and stats that we can look at. They're very honest. It's data, so they can't manipulate it. It's just how it's presented to you caused a huge moment in the industry probably 18 months ago when indie publishing overtook traditional publishing on e-readers.

Of course, for the big five publishers there was all sorts of recriminatory type talk from them. Then within a quarter there was a slight slowdown of that growth and everyone was talking about that. We go through all of this in the interview. What's really interesting for us in the self-publishing formula community is how Data Guy became known to us.

Mark Dawson: Yes, and therein lies a tale. So come back with me to six months ago and I was in Dublin Airport. I just had been in Dublin for two days with Amazon to talk at an event for them. I was going back so I just got a taxi in the airport and I had something to eat, checked my phone, and there was an email from someone who I shan't name for obvious reasons. He was a student of our advertising course and we had a little discussion about some things that he discovered. I told him some things I thought would help him sell a few more books and I think it had gone quite well. I emailed off and then he emailed back and said, "Oh, by the way, I'm Data Guy."

I was like, "Holy," expletive. We had a little chat and we've spoken since. He is a very clever guy and I was immediately terribly ashamed because part of the course involves some spreadsheets that I put together. Now, I quite like data, but I would describe myself as a very happy amateur.

When you see the spreadsheets and the data that Data Guy deals with it makes my efforts look positively like I'm a cave man compared to a very sophisticated data analyst like he is. So there was that, but he was very kind. He didn't criticize my spreadsheets too much.

We are delighted to get him on the podcast. I don't think he's been on one before. He has spoken at some events and we'll talk about that later, but it's a real pleasure to get him on. The work he's doing with Hugh Howey on the authorearnings.com website is really important and it has started to ruffle some feathers in the traditional space as well.

James Blatch: Definitely has. It's put in black and white what's happening in the industry and people can be in denial about it if they want. But what these guys have done, Hugh and Data Guy have put together, is irrefutable evidence of the transition that's happening in publishing.

Now, we do have a giveaway. I'm going to mention it before the interview as well. If you don't have time to listen to the whole interview you can actually go to selfpublishingformula.com/download59. Is that right?

Mark Dawson: Yes.

James Blatch: Download 59. What you're going to get is a really beautifully put together presentation from Data Guy that he's used as a keynote on one occasion, but he's allowed it to go to us to give away to our listeners. It fleshes out in very visually engaging ways all the data that they've come across and all the key findings. He talks about this in some detail and a bit about the background of it all in the interview. We'll give that URL out at the end as well.

Okay, let's listen to the mysterious Data Guy.

Data Guy: I actually came to data analysis for this industry rather indirectly. I was an author and an early indie author. Actually, let's go even further back than that.

I was working a video game industry as the CTO of a social and mobile video game company, a fairly large one, 400 people. At the time the Apple iPhone came out we had a need to understand what our competitors were doing.

The data which we used to get in the video game industry from looking at the number of CDs shipped and the sales reports from places, Game Stop, et cetera, that data disappeared because the bulk of purchases had moved into the digital realm. They were app store downloads on iPhone and on Google Play.

What we had to do was solve the problem of getting access to that data, and so I developed a technical map and we put it into production at a fairly large scale, of scraping Apple's video game app store categories and later

the Google Play ones, and then using that to break down our own sales and the sales of our competitors to look at which genres we should target, what kinds of games, what the sales were trending like for various different segments of the game industry so we could basically figure out what games to put into production.

The way in which we converted rankings to sales was, and anyone who's familiar with Author Earnings will be smiling at this point, was by taking a large catalog of products that we had ourselves, close to 100 of them at various points along the sales ranking curves. We actually had several number one games on Apple, and so we knew what number one on Apple was worth, number 10, number 20, number 30, et cetera.

We were able to use that to build the calibration curve and then predict and project what our competitors were doing. When I became an indie author I noticed that the book industry had the exact same problem, because I as an author trying to figure out whether to accept a publishing contract that had come my way after I was selling well as an indie or not, I just didn't have the data I needed to make that decision.

I dusted off the old expertise and did the very first Author Earnings data scraping. I was number one in science fiction for a very brief window on Amazon and Hugh and I had exchanged emails over it. So I reached out to him and I said, "Take a look at this data." He nearly fell out of his chair. He had the same reaction I had had, which is, "Wow." We had no idea that indies had already taken over so much of the online book sales market.

He said, "This is exactly the story I've been trying to tell. I cannot get the media interested in it." You know, whether it's Amanda Hawking or me or any of the authors who have done seven figures. They're missing the real story which is the vast middle class that's emerged for the first time that's paying some bills, paying their mortgages, sending their kids to college through self-publishing.

No one wants to talk about it because it's not sexy. So we put together the Author Earnings website and that was three years ago. Since then we've revisited on a quarterly basis the Amazon store. We've also hit the other major eBook stores, whether that's iBooks, Nook, Google Play, Kobo, and we've scraped those as well and then done analysis of the breakdowns of sales in each of those stores.

That's kind of in a nutshell the elevator pitch. Like most hobbies, for me it's gotten a bit out of control to the point where I'm now going to industry events and speaking and getting offers to present to more traditional publishing oriented audiences. It's kind of an interesting time because they realize that they can't get that data either through traditional means.

James Blatch: Let's talk about the data in a moment. I know from my journalist days that the devil is in the detail, it is such a great expression, and that within data is the story normally. But I do love it in those moments when somebody does some new research, opens in the spreadsheet, starts plotting things and goes, "Wow," when it's instantly revealed to you. That was one of those moments, and that was a big moment in the entire indie industry.

You should be proud of yourself for that work.

Data Guy: I basically really look to guys like Hugh, Joe Konrath, Barry Eisler for inspiration. What they introduced me to was the idea that we're all in this together, and that as authors we need to share information, share data. I came up in publishing reading their blogs and talking to them, and so this idea of giving something back was basically ingrained from day one into my author journey.

James Blatch: That's a trait of the industry, isn't it, which is a great thing of the non-traditional industry we should say at least.

Let's talk about the data a little bit then. As you say, they were easy to write off as outliers, people saying that they were having a lot of success in the industry. What you discovered is these tens of thousands, more than that, people around the world. Mark Dawson is one of them.

We have loads of our students now in that position where some of them are just making five figures a month and it's been transformative for them, but not necessarily in a multi-millionaire way, but in their lives.

This is what your data really uncovered, that there's a living to be made for writers and it's vibrant and it's exciting and it's definitely there.

Do you feel as excited about the possibilities and opportunities for indie writers now as you did when you first realized what sort of scale was going on here?

Data Guy: More than ever. The reason for that is over the last three years, aside from a recent setback which we saw in October which we'll loop back to and talk around, it's really been up and up and up as a sharer of a market that itself has been growing up and up and up in scale.

The overall opportunity as an indie has only grown. Now, at the same time we've also seen competition grow. There are now more titles chasing those sales, so not all authors are seeing success in a homogenous way. There are some who are doing well, some who are not doing as well, but that's always been the case. There is now more pie to slice up and divide up and share, though, and that's been a consistent trend.

The other reason I'm very excited about it is I'm seeing the international markets start to emerge for digital sales. One thing that is absolutely correlated in the data is digital sales and the growth of indie sales. The two go hand in hand.

James Blatch: Yeah, as I suppose you'd expect because most of us do look at the eBook publication as being our first port of call in an indie career. Though let's talk about the moments, then, because I remember it was probably third quarter 2015.

Was that the one where the indie overtook traditional and that really caused a stir when people looked at that graph on your Author Earnings report?

Data Guy: Yeah. I think that was a very, very significant turning point because I think that was the point at which the ... I don't know how to put this without being melodramatic. The wages of agency were being reaped fully by the traditional publishers.

The agency agreements prevented online retailers from discounting out of their own pockets traditionally published books from publishers who had signed those agency agreements. They rolled out between late 2014, I believe, and early 2015. Sorry, mid-2015. September 2015, the largest traditional publisher Penguin Random House went to on agency and the prices of their eBooks jumped up tremendously.

That was the point at which the indie share jumped and overtook the traditionally published share.

James Blatch: That was an amazing moment, I mean, despite the background to it. It was kind of coming anyway, although what's been interesting and I know there was quite a lot of glee in the traditional markets, the traditional world in the months that followed in 2016 where we saw a bit of a reversal of that.

I think your understanding and you are the guy who understands the data, is that we need to understand that reversal as well.

Data Guy: Exactly. I'm going to issue spoilers for an upcoming report, but in the last couple days I did another large-scale Amazon scrape, a couple hundred thousand titles. Things look pretty similar to how they did in October, but the indie share is starting to creep back up again.

There was one rather dramatic change happening within the realm of traditionally published eBooks, and that is for the first time since we started tracking this, small and medium publishers have overtaken the big five in both dollar and units. That's kind of an eye opener. This has not been a good three months for big five eBooks.

James Blatch: Right. I mean, I guess that's inevitable and you can just look at the way other industries have been affected by digitization is that the larger the organization, the larger the corporation, and we know this from every industry, the more difficult they find to adapt.

It doesn't surprise me that the medium and smaller enterprises are the ones that are showing some agility to take advantage of the changing market.

Data Guy: Some of it is agility and capability of the organization's benefit. Some of it is a very deliberate policy of which agency was a key tool in the tool kit, which is to slow down the adoption of digital and to keep people buying print.

If you were a big five publisher that would make tremendous sense to you. You would look at it and say this is a realm where I have huge advantages, distribution advantages, preexisting agreements, a lock over a lot of the shelf space and the ability to command what is included and what is not included.

In the digital realm I have none of those advantages and my organizational scale imposes an overhead that puts me at a tremendous disadvantage. It is a very logical thing for a big five publisher to want to keep the customers where they have the advantage. They don't want them walking over to the other table.

James Blatch: Is there any longevity to that policy?

Data Guy: It depends on your goals. I mean, if your goal is to maximize shareholder value in the short term maybe that makes sense, but long term it's a fool's errand really.

James Blatch: Yeah. Ultimately we are seeing people moving over to digital consumption as they did in film, as they did in music.

I was working in the film industry with Mark Dawson. I started with the BBOC in '06. When digital started happening, which I guess was about '08, it took about nine months. I mean, it was unbelievable how quickly we stopped watching films. The DVD submissions just dropped away and the bigger the organizations, in those days it was Fox and Paramount and so on who didn't quite know what to do.

The ones who moved last to digital are the ones who had the smallest share, and that's happening in books now. It's consumer led more than anything else, people deciding that, "You know what? It's actually quite convenient to have a Kindle in your handbag rather than a big book," or two books, or how many you want to take on holiday.

I always feel that if a big company is doing something to try and buck that underlying trend, ultimately as you say, it's not a long-term solution to their problem.

Data Guy: No, your insight is absolutely golden there, which is if you decide to fight your customer and their preferences, you are basically going to find yourself with far fewer customers. That's a no-brainer.

James Blatch: I just want to go back technically a little bit, because I think I said this to you in the email exchange we had before the interview. I am fascinated by the way this actually works.

If people are watching this on YouTube they'll see your spider logo sort of represents you, Data Guy as a proper metaphor. You talk about the scraping. Now, just bear in mind that most writers learn their little bits of technology like WordPress and all the rest of it, but aren't necessarily technical people.

When you're talking about scraping, the way you get this data and the little programs you have that crawl across the internet.

Can you just give us kind of layman's explanation of how that works?

Data Guy: Absolutely. When you go to Google and you type in a search request you're going to get a long list of search results. Every page on the internet that in some way ranks for a term or set of terms that you entered in the search bar shows up.

How does Google poll that data? They basically send out programs to act or pretend as if they are web browsers on a massive scale and explore the internet. They basically suck down the pages just as if they were web browsers and then take them apart and extract the data that they need. Scraping is a very, very standard piece of technology. Online retailers use it to do price checking and price matching with each other. Amazon does it. Barnes and Noble does it to some extent. eBay, all sorts of places use this. It's basically a core internet technology.

The advantage that I have in scraping only one site and only particular pieces of information is that I know what those pages look like. There's a surprising number of variations of an Amazon product page, but they all contain the vital info that you need to build this picture.

They contain the name of the book, the publisher who it's sold by, the digital list price and the sale price, and above all the sales rank of all books

sold on Amazon and all eBooks sold on Amazon. There's two different rankings there.

James Blatch: Yeah.

Data Guy: You can use that ranking. There's a very consistent relationship over a period of time. It's not a one day period, it's basically a recency weighted average of what you sold today plus half of what you sold yesterday, a quarter of what you sold the day before, an eighth of what you sold three days ago, a 16th, 32nd, et cetera going back in time.

There's a very key mathematical reason they use that particular type of formulation, and it's because it collapses into a very neat, closed form, which allows you to calculate your ranking based on not looking at all of history, but just number which you maintain plus today's sales. You combine those two and you have a ranking.

The reason Amazon needs that efficiency is because they're not just doing it for the 35 million books or the four million eBooks that they sell. They're doing it across over 200 million or 300 million products that they sell, I believe, every day. In fact, several times a day.

The mathematical efficiency of using that type of formulation just makes sense, and it's crystal clear when we put it into a large-scale mathematical statistical program that we use and then apply that formula. Things collapse into a very neat line.

It's actually log quadratic, but the minute you propose that from everything collapses into this lovely relationship. Then you look at daily sales and you go, "Yup. My sales and my ranking fall exactly on that curve."

James Blatch: Math can be beautiful with its lovely curves.

Do you only use the scrape data and extrapolate your results from that, or do you also combine it with maybe some paid for market analysis to bring some context to it?

Data Guy: From day one we were able to say very definitive and authoritative things about the relative number of sales between different writers, different publisher types. But there was a high amount of let's say, skepticism or dubiousness that we had the overall size of the pie accurate. So in 2016, late 2015 early 2016, we refined the methodology. A panel of about a dozen plus authors, it's actually getting close to two dozen now, who for obvious reasons are going to stay anonymous, but who between them have almost 1,000 books and from a good mix of genres have been sharing daily sales data with me absolutely unedited.

Basically what they see in their KDP dashboard is what my spiders extract. At the same time I'm tracking the rankings on their books, so over the course of a quarter we get a large number of data points, about a million, which we then use to refine that curve. That's the only non-public data we're using in this process, but it's not paid for market research. It's like most things in the self-publishing realm, authors volunteering to help other authors.

James Blatch: Just effectively it's validating your findings.

Data Guy: More than validating. It's being used to refine the curve that we use so that we can now say with definitive certainty not only are our pie charts correct that show the relative measures, but when we say Amazon sells one million and 64 thousand eBooks a day that's within 2% or 3%.

James Blatch: Yeah.

Data Guy: It gives us an absolute benchmark as well as being able to look at relative market scale.

James Blatch: It's always heartening to hear that type of figure for authors of how many eBooks get purchased a day from Amazon.

Data Guy: I will share one other thing from a calibration standpoint. Behind the scenes I've been working with some of the top traditional publishing analysts to calibrate our data. It's gone from, "Oh, this is some toxic garbage that can't possibly be true," from their perspective to, "Wow. Those predictions are actually coming true."

They're able to predict months in advance of what traditional publishing can see what the overall industry trends are. The reason for that is not that we have a crystal ball and can see the future, it's just that our data collection is immediate, whereas the traditional publishing numbers that you see from the AAP or book scan are much delayed.

With the AAP data it's about four months, five months, sometimes even six months delayed. So there was a lot of interest in seeing what traditional publishers and traditional book organizations could glean from this data about a side of the market that is basically invisible to them.

They see that a large percentage of the books on the best seller list that Amazon has daily are not books that they have any data on, but they don't know what that means in terms of market share. So that's where the interest comes from.

As a result of that, I've been working behind the scenes with them to calibrate and compare and use the AAP data, for example, to fine-tune our math.

James Blatch: In terms of transformations the data business that you're in is on an equivalence to the digital transformations that we've seen in the other industries that we've talked about.

I've worked in traditional industries in the past where you would pay \$1,000 for a report that comes out quarterly or twice a year, and was already out of date. What you were paying for was access to company data that they had in turn paid for, whereas now you've got your clever little spiders that are crawling across the internet and getting, I mean, not just an equivalent amount of data but really an exponential more data.

Data Guy: Oh, absolutely.

James Blatch: That is industry transformation.

Data Guy: Yeah. What people fail to understand when they hear about our methodology, the most common thing they fail to understand is the scale at which it operates. I won't geek out on the number of servers. Okay, actually I will.

For the October report I ran over 2,000 servers for a very short period of time, every one of them scraping through multiple connections. Let's put it this way. If it was a smaller site I have to be careful not to knock them down. With Barnes and Noble, when we scraped Barnes and Noble we kept breaking their site because it's kind of flimsy. I had to deliberately scale back the rate at which we were drawing data from it because it was throwing errors and puking.

With Amazon, there's such a large, robust infrastructure that I can hit them with literally three million requests in under an hour, each of those requests drawing a product page, and they don't even care. There's absolutely no impact on their infrastructure with that kind of scale.

We're pulling close to two terabytes of data an hour when the thing is up and running at full speed.

James Blatch: That in itself tells a story.

Going back to what we were talking about when we were discussing the actual data and the results you draw, the inferences you draw from it. This is an alarm bell to the traditional industry about having to modify how they work.

I was really struck by something that came out in a presentation you recently gave that's on your site where you could show that even the print book consumers, the people who they're desperately trying to target their pricing policy towards, are moving away from places where they can't easily buy eBooks and indie author material.

Really, it's wake up and smell the data now, isn't it for them?

Data Guy: It is, although there's a fair amount of denial about this spill in the industry, because the message that print is back and digital is going down is so reassuring if your business is based on advantages you hold in the print distribution realm.

To hear someone say, "Well, actually print is not going to save you because print is moving online even faster than print readers are moving to digital." The thing is, the large publishers have enough scale that they see this. They know what percentage of their book sales come from online retailers. But they're not well incented to advertise that.

Nielsen BookScan, which does collect print data and the majority of the print sales statistics that you hear are based on BookScan numbers, also knows this. But because of the agreements that they've signed with the retailers, to be allowed to have access to that data they're not allowed to talk about how much of it came from who.

That's why when we see publishers weekly present stats and say, "Oh, print is up X%," they never say what share of it is online. Now, with my spiders running I can see that, because I can scrape Amazon's print sales as well. I was absolutely stunned when we first did that to see what a large percent of

Nielsen's reported 675 million print books a year is actually Amazon sales. It's 42% now.

James Blatch: That's incredible.

Data Guy: One puts that number out there.

James Blatch: Yeah.

Data Guy: Almost half of it is Amazon sales.

James Blatch: Guess what, guys? Even when you think you're winning you're not really winning.

Data Guy: I would argue that they are winning because selling books on Amazon is a lot more efficient cost-wise than selling them through brick and mortar. That's because your returns drop to very close to zero. The brick and mortar channel, about 28-32%, and it varies by format. For mass market it's over 50%. Mass market is a dying format. But 50% of mass ship to stores, they get the covers torn off and they get destroyed some months later.

James Blatch: Wow, really?

Data Guy: Yeah. For hard backs and for paperbacks, big paperbacks, the larger six inch by nine inch format, the numbers generally run in the 25-30% range. Surprisingly it's much higher for more popular authors because over-ordering is part of the strategy for creating a stack of James Patterson's latest or Stephen King's latest in the bookstore window, which itself prompts people to go, "Oh, hey. Maybe this is a really big book. I need to buy it."

They know full well that 40, 50% of those are going to be considered returned, although for hardbacks and paperbacks they generally don't actually return them. They just mark them down so no royalty there. James Blatch: No royalty and a lot of paper being printed and all the energy that goes into that.

One of the other areas I'm really interested in, I don't know how much you can shed light on this, because it kind of requires a good deal of knowledge of buying habits before the digital market really started to develop.

That's the fact that the gate keepers, the publishing agents and editors who are no longer controlling a good chunk of the market, half the market, is whether that has had an impact on readers' habits in terms of genre and type of books they're reading.

I wonder how much that has changed since the digital revolution? Is this something you can look at, is in your data?

Data Guy: It's not something I can look at in the data other than at a very high level where what I can say is this. I've seen this massive growth of indie and other non-traditional, meaning Amazon imprint, sales.

At the same time we haven't seen traditionally published sales collapse in half. They're down maybe 10%, 2 or 3% a year, and so the only way those two things can simultaneously have occurred is if book purchasing has gone up.

It's always easy to come up with subjective justifications to explain something, so I'm a little reluctant to do it but I will say this. Two things prevent the most avid segment of readers from buying more books. That's cost and availability of content in their preferred reading niche.

When supply was limited to what was on the bookshelf, and I can speak personally about this because I was the guy who used to walk through a Crown Books or a Walden Books running my finger along the shelf of thrillers and science fiction books and going, "Well, I don't see anything new here. Everything that I like I've read." There's maybe one or two new books a week that might interest me.

Now there's an almost unlimited supply, and so for readers like me we've gone from 30 books a year to north of 200 books. Now, that only affects the most avid segment of readers, your two book a year, "I picked it up on an international flight at the airport bookstore," or, "Hey, there's this book everyone's talking about and I'd be embarrassed that it's not on my coffee table." Those readers haven't really changed their habits.

James Blatch: Yeah. Well, that includes all of us, doesn't it? The one book that you see everyone reading on the tube, you kind of grab.

I'll tell you what happened to me, also a little bit of anecdotal evidence, is that I got married about 18 years ago, something like that. For most of that time it curtailed an important part of my reading time, which is very late at night. My wife goes to bed before me. I usually go to bed afterwards, and I used to read for half an hour. But you can't go to bed with somebody else in the bed and turn the lights on, because that's really antisocial.

That all changed with the Kindle. It completely changed with the Kindle. I've probably at least doubled, probably more than that, how many books I get through a year, because I can read from my Kindle in a convenient position. That's a specific example of being in bed late at night, but the Kindle itself, I think has probably been responsible for just giving us access to books, supply, and reading, which goes a long way to explain what you just said.

Data Guy: You're absolutely correct. There's the time that one can now spend reading in bed without disturbing one's spouse or partner. There's

the five minutes, 10 minutes, 20 minutes, half an hour that ordinarily you wouldn't happen to have a book with you where you're taking a train or you're dropping a kid off at soccer and waiting before it's time to pick up the next one at ballet. Those micro-windows of time that become available during the day, I think a lot of people are starting to fill those with digital reading.

James Blatch: Yeah. I mean, you started off with a very optimistic note. I rely completely on your judgment on this because you are a man with your head buried deep in the data that drives our industry. So I'll finish with another look forward to the future, optimistically hopefully for our listeners, most of whom are like me trying to get going in the indie sphere. Some are very successful. Others are just starting out.

Do you have as a result of looking at the data and the trends, particular tips for people beyond simply saying, "Don't give up. Keep going because there's a growing market there."

Is there a particular area do you think you could point people towards they should be looking at?

Data Guy: Absolutely, and that is marketing and advertising.

What is conventionally understood by traditionally published authors to be important absolutely isn't. Newspaper and radio ads, book signings at the occasional bookstore, they're fun. They are enjoyable. I've done them. I've really enjoyed as an indie author signing at Barnes and Noble day.

But 70 books in a day in print, where you basically earn very little with your POD books, is not comparable to selling 1,100 or 2,000 books in a day, which is what you can do with an online promotion without too much difficulty if you plan it right.

Focus the energy on what works today, and that includes Facebook ads.

Here's a little piece of anecdotal support for the fantastic Facebook course that you guys have put together and that Mark pioneered. I was able to take two books that are more than three years old and keep them in Amazon's top 2-3,000 overall at full price for months and make money doing it. That was with the power of Facebook advertising.

People who are not staying on top of what works, whether it's Facebook advertising, AMS, writing in series, whatever the knowledge is, and it's constantly evolving, that authors share amongst themselves about what works, we must all be students of. Because the idea that your book, if it's "good enough" is going to get discovered and spread via word of mouth. That's a rarity.

It requires a push to even reach people's attention and become visible. I think we must all be students of what's working now and understand also that what works now may require tweaks or may require a change in strategy six months down the road or a year down the road.

James Blatch: That's very much what our community is all about, sharing, and it was great to hear your example there. Actually, I know you're in Palo Alto, the home of Silicon Valley, but despite that your broadband did just drop out as you said the price of your book that you kept in the rankings.

What full price was that?

Data Guy: \$3.99

James Blatch: Oh, \$3.99. Okay. Let's just very quickly deal with price then because we're coming out to our kind of 40 minute mark where we try and pitch our interviews out.

That's one of the things that affected the other industries, music and film in particular, is the price per unit did start really tumbling when people didn't

associate the same value to a set of ones and zeros that they did to a hard product.

In terms of price, how has that held up in the indie market?

Data Guy: We've actually not seen that happen, perhaps thankfully, I should say. I'm looking at data for 2016 right now and indies are earning the most money in aggregate at \$3.99 and \$2.99 and to a slightly lesser extent \$4.99. There's still a lot of 99 cent sales, but I suspect those are mostly promotional spikes.

James Blatch: Yeah.

Data Guy: When I looked at 2016 versus 2015, and this came up in the comments on the article which you're referring to, the Digital Book World presentation that I did, it actually looks like the average price paid by a consumer for an indie book inched up about 10 cents or so between 2015 and 2016.

Now, the one thing that we don't have a lot of visibility into is how much the "supply" has increased. The prices hold steady and sales go up a little bit, but are there far more authors today than there were a year or two ago? I don't have visibility into that.

James Blatch: Yeah. Well, Data Guy, it's heartening to hear your extrapolations, so I want to say-

Data Guy: It's an absolute pleasure and somewhat of an obsession to talk about this with an audience that appreciates it. When I'm standing up in front of a room full of traditional publishers I have to self-edit somewhat.

James Blatch: Yeah. Be polite in places. Well, you don't have to self-edit here. We're lapping it up. Thank you so much indeed for coming on to the podcast.

We should tell people that the course they go to, authorearnings.com, you can get all this data.

You do your reports quarterly, I think, Data Guy?

Data Guy: Quarterly is what we tend to do, yes.

James Blatch: Yeah, and they're great reading for somebody who wants to learn a little bit more about the industry and the way it's going. Keep your spiders going. We'll catch up with you hopefully in the future. Good luck with your own career, Data Guy.

I know you're anonymous in the public zone, but you are an author yourself. Right?

Data Guy: Absolutely. In fact, I would recommend if you're going to run spiders against retailers at the scale that I do and you're also an author, that you keep those two endeavors separate.

James Blatch: I can understand completely why. There's definitely, by the way, a book there about the spiders and the scraping coming together in some kind of a life form in the future. But we'll part that one for the moment and think about it. Data Guy, thank you for joining us.

Data Guy: You bet. Take care, James.

James Blatch: There you go, fascinating. I mean, I was really interested in how this guy set this stuff up. He's not the only person doing this obviously, but he's obviously a bit of a master at this scraping.

Where websites set themselves up, companies keep a lot of stuff commercial. I've worked in the commercial sector and you have as well in the past. You'll know that a lot of the turnover figures and stuff that have to

go out into the public domain are quite general. A lot of the detail is kept back and it's kept back for good competition reasons.

Actually in the modern environment, if you've got some ability to look at every single page and every update of every page, you can start to accurately put together what's happening to that company in really detail, and more detail than Amazon and the others would probably want you to have, but there's no way around it from their point of view. They have to have the ability to show the product and sell it so you can work these things out.

It's fascinating, seeing those figures and seeing the fact that the indie space is growing and is healthy and is only going in one direction?

Mark Dawson: Yeah, it is. It's really grand. I think the thing I take away from it is ... The macro level stuff is great in terms of trends and that kind of information.

Two things, really. If you're just fresh and you're about to publish something you might be thinking, "Do I go traditional or do I go independent?" Well, the Author Earnings report will give you very useful information as to how you might perform based on how people who've just come into the industry over the last couple years have performed, because they can tell that. They can work out will you make more money from a traditional deal or will you make more money being independent most likely? So that's very useful.

Then the other thing, and this is what Hugh Howey draws out quite a lot and I think it's very important, is that it's not helpful necessarily to focus on outliers. We had Holly Ward on recently. If you're just starting to write romance don't immediately think that it's all or nothing. You've got to do seven figures every year or you're a failure.

What they've enabled people to do is to look at numbers and work out ... There is a vast middle class of people who are making good money, and that good could be ... It's all subjective, of course, but it could be enough

to go and have dinner every month with your partner, or it could be enough to pay your mortgage off or pay some bills or leave your job. You don't have to be making the kind of stupendous amounts of money that people like Marie Force and Barbara Freethy, Holly Ward, they're making. It could be maybe \$1,000 bucks a month. What they have done is to demonstrate that that is possible and more than that. It's not just possible. There's an awful lot of people who have reached that kind of level, and that's a real service to the writing community as a whole. I think that's the most important takeaway from the work that they've done.

James Blatch: Yeah, absolutely. It's something that we say as well in terms of what a successful indie career can do for you. It depends on what you want it to do for you, and we're talking your ambitions. But you're absolutely right.

We've got plenty of listeners who've contacted us who are absolutely thrilled that they've cleared \$500. That's actual profit and they can go out and buy something. We always remember the interview in Orlando with Elicia Hyder who was so thrilled that she could buy a mailbox for her house from her writing. Now she is doing five figures a month.

Mark Dawson: She has a gold plated letter boxes now.

James Blatch: She has gold plated letter boxes now. But yeah, and gosh I would be absolutely thrilled if I could start paying some bills at some point. Hopefully that'll happen in the next 18 months for me. Okay, good. Data Guy.

Mark Dawson: So yes, just to reiterate. There's two giveaways. You can get the Data Guy giveaway at selfpublishingformula.com/download59

You can also get the Self Publishing vault, the vault of goodness or value, whatever we might end up calling it, but all of the transcripts from all of the previous 59 episodes will be in that eBook. You can get that at selfpublishingformula.com/vault.

James Blatch: V-A-U-L-T. Yeah, do go to [download59](#) and get that presentation. It's beautifully put together and really goes as a great accompaniment to what Data Guy was saying. If we didn't get into enough of the detail and enough of the trends in the interview, it's all there in that presentation. Thank you so much indeed for listening, for watching if you're on YouTube. We'd love to hear from you, podcast@selfpublishingformula.com, and we will be back next week.

Speaker: You've been listening to the Self Publishing Formula podcast. Visit us at selfpublishingformula.com for more information, show notes, and links on today's topics. You can also sign up for our free video series on using Facebook ads to grow your mailing list. If you've enjoyed the show please consider leaving us a review on iTunes. We'll see you next time.