**RedZone Podcast Episode #77: How to Get Your Team’s Big Data Design & Visualization Skills to a World Class Level? Tactics+Tools to Get You There – with Seth Familian**

Bill: ... started. Seth, I want to welcome you to the show today.

Seth: Thanks, Bill, and thanks for having me.

Bill:

[00:00:30] I'm very, very excited about talking with you because we have two things that we want to talk about today and educate the enterprise, leaders, and entrepreneurs. They're on, they're listening. One of these topics is visualizing data and big data and then how you storytell around that, and I think that's super interesting. The other is how we deal with cloud automation, but before we do that, let's talk about you for a second on how you got to where you are. You have an interesting background. Let's just give the audience an idea of how did you get to this point where you're considered world-class at these topics. Can you give us a short story on that?

[00:01:00]
Seth:
Sure, it's a pretty bizarre background. I actually studied liberal arts in history and literature in college, and then afterwards, I served as a speech writer to the New York City Fire Commissioner after 9/11. Before, I worked for a marketing strategy consultancy and started learning more about storytelling from a quant side instead of just the qual, delivering speeches to a larger audience.

[00:01:30] That, in combination, inspired me to head to Berkeley for business school, and in 2008, after the economy tanked and after I'd been working in corporate strategy for a couple larger companies, I saw that companies really wanted to hire me to outsource their strategy. I effectively become a small strategy group of one that they could parachute in for particular problems.

[00:02:00] When I worked with companies on strategy, it really evolved into working with them on data automation and analytics and data visualization and visual storytelling. What I really found I loved doing and where my talents intersected with my passions was that I loved helping my clients build and improve on their data automation capabilities, so to enable them to develop a stream of data on a software application or a website that had so much potential to be so rich with content to analyze.

[00:02:30] Once we had the right stream of data in place, I could, then, analyze the data, derive insights out of it, build more automations on top of it, and then CCMOs and their marketing teams take what was previously an overwhelming amount of information and turn it into a highly compelling and actionable story with meaningful outcomes.

Bill:

[00:03:00] Let's break apart some of the words because I'd love to know some of the deeper meaning around it. When you say a rich stream of data and data automation, just so we know, what do you mean by data automation and the stream of data, and you could take an example, you referred to like a web property that a business wants to exploit. Is that what you're referring to? That's my one question, and then what do you mean by the automation part of it?

Seth:

[00:03:30] Sure. When I'm talking about a stream of data, I'm referring to a digital property. That could be a website. It could be an application. It could be an eCommerce platform. It could be anything where there's user interaction and user behavior that creates a stream of events that's taking place on the platform. Those events could be anything from visits to a webpage and time spent on a page to transactions and purchases and repeat transactions, or it could be things like clicking on buttons and replying or acting on certain prompts inside of a web application or a mobile application.

[00:04:00]

[00:04:30] In any of those contexts, what a lot of companies typically don't think about out the gate is how were we measuring exactly what is happening on those applications, web services, or web pages, and typically, they'll just put Google Analytics on the page, and they'll analyze their page views and their visits, and they'll maybe look at return visitors, or they'll put an eCommerce tracking tool in place and look at their revenues, but it sort of stops there. What they're not doing is they're not creating a robust stream of information on the events that are happening on the property and information, contextual information on each user that's performing those events, so they can then segment the events by user and cluster those events together by user type in order to have a really robust set of data.

That's the data stream piece of it. Does that sufficiently answer your question on that part?

Bill:
[00:05:00] Yeah, definitely. I'm quite familiar with Google Analytics, and I have not been exploring it recently so my information might be a little bit, my knowledge may be a year or two dated, but are you ... I was going to ask you the question the common mistakes people are making, and certainly you just answered one of them, which is potentially just using Google Analytics at its basic levels, sort of the macro higher level. What you're suggesting is going down and you could actually see context around users. Is that correct?

Seth:
[00:05:30] That's right. The basic use of Google analytics or a number of other platforms, Kissmetrics is another one to look at transactions, is you just put a snippet of header code on to a website or you instantiate the analytics tool at the beginning of your applications launch, and then you let it automatically track everything that's happening. Every page that's loaded, it looks at the metadata. It measures the time on page, things like that.

[00:06:00] What those applications aren't doing out of the box is they aren't identifying the users that are coming into the app, so it's not saying, "Hey, this is actually Bill because Bill signed in with his email address and Bill performed each of these actions on the page." That's the first thing they're not doing out of the box. They need to be customized for that.

[00:06:30] The second thing that those apps aren't doing without additional, what I call instrumentation, is every time you visit a page or you do an event, there are number of additional contextual properties that are associated with that page or event that aren't intuitive to the analytics tool when it's just looking at the header code of the page.

[00:07:00] One example, I have client that's a content provider, and they publish five different types of stories from short form to long form and podcast and videos, and there are a bunch of keywords that are associated with the stories, and maybe there's a headline style that's associated with each of the stories. Well, when they load Google Analytics in their header code, all Google Analytics is able to see is what was the page title, what are the keywords in Google organic search that led people to land on a page, how much time did people spend on the page, what page did they come from previously, what page did they go to afterward.

Bill: Right, right.

Seth:

[00:07:30] But when my clients ask questions like, "Which of our five story types are trending in terms of having the most sticky users?" or, "Hey, Bill signed up for our email list. What are the last five stories that he looked at?" those questions can't be answered unless you do much more customized instrumentation. That's where you can either send custom properties just to Google Analytics, but what I found was there's a really interesting problem, which is that as soon as you get into the world of sending custom events into one platform, what you start to realize is that each of the analytics platforms that out there whether it's Google Analytics or Mixpanel or Amplitude or Intercom, they're a gazillion of platforms out there, they all have their strength.

[00:08:00] Some of them try and do a lot of things. I call those Swiss Army knife platforms, so GA tries to do a lot ton of things, but what it's really good at is measuring traffic sources, understanding where your traffic came from, and what your time on site was like for a given user who performed over a given number of sessions.

[00:08:30] As soon as you start asking yourself questions around what flow of events led a user to convert to a certain action or what is the lifetime value of my customers or if a user or a writing and marketing automation rule that says once a user views these three articles, I want to send them an email that inspires them to view these other three articles or promote one of the three articles that they read. None of those things are actually possible in GA, so what you end up doing as a result, there are two different paths you can take, and this is where the world of automation comes in.

[00:09:00] One option is brute force, which is you start incrementally adding more and more tools to your header code and to the body of each of your pages or each of your app loads. What that means is, okay, let's say I'm tracking every time someone clicks a button on a page with Google Analytics.

[00:09:30]

[00:10:00] Well, now I realize I want to track that button with Mixpanel as well, so then after I have my developers implement Google Analytics, now I call my developers up again and say, "Hey, I know we just Google Analytics in. Now I actually want you to do the exact same thing with Mixpanel. It's really great for analytics, and it does things differently than GA, so I want you to put Mixpanel's header code in there, and then every time that you fire that event that says button clicked for GA, I want you to copy that code and paste underneath it the same event-tracking code for Mixpanel. By the way, I also want you to identify who my users are once they've signed in, and I want you to send that information anonymously to GA and non-anonymously to MixPanel because the two have different privacy settings, and they handle data differently."

[00:10:30] Your developers begrudgingly says, "Okay, fine. I'll do it," and then a couple months goes by, and you go, "Wow. Now we have actionable data. Now I have a set of rules that I want to send emails around or I want to prompt users with in at messages around," but I can't really do that well in Mixpanel across both mobile and desktop, so instead, I want to use this thing called Intercom. Intercom is an amazing application that lets you effectively compile a user database of all your users, send events into each of your user profiles to understand what they've done in your application, and then write automated campaigns based on triggers associated with the events they've done.

Once a user looks at a page title this or has done this event this many times, then send them this email campaign or pop this in at message.

[00:11:00] Well, when I'm ready ... I, as a marketer and a product person, get super excited about something like Intercom, so then I call up my development team and again, and say, "Okay, dev team. Now I want you to put Intercom's header code into our header, and I want you to copy the stuff that you did for GA and Mixpanel and do it for Intercom as well." Now, my developers are getting frustrated because there's nothing a developer hates more than having to repeat the same task over and over and over again.

Bill: Yeah, right, right.

Seth:
[00:11:30] It's just not efficient. Then they become reluctant to make any changes or improvements to the tracking as things change on the website, and every time I ask them to integrate another tool or remove a tool that isn't working very well, the frustration increases.

[00:12:00]

[00:12:30] In comes automation. There are platforms like Segment, and another one is called Google Tag Manager. What Segment does, actually, better than, I think, anyone else is they provide a way to write the tracking code once. You put Segment's header in your website, and you track all of the events and the identify calls once with Segment, and then once that stuff is tracked a single time, I, as the marketing manager or product manager, can go into the Segment interface and turn on and off different tools, like a Mixpanel and Intercom by piecing my API key in and things like that and instantaneously deploy very, very powerful cloud-based applications that are custom-tailor to deliver really powerful functionality.

Bill:

[00:13:00] The, essentially the Segment tool, and the other one is Google Tag Manager, but you're saying the Segment would be like a strong, almost like a glueware. It's sort of gluing ... Instead of stacking these tools in a, like Mixpanel and, that manage the flow of it, since that is tagging, stacking, you can actually have more of a framework like the Segment, and then you can glue, not glue, but basically it has extensions into these other tools, so it's a little bit more efficient?

Seth: That's exactly right. Segment and Google Tag Manager are the same in that they both enable you, instead of just shoving a bunch of things into the site, they create a door, and then they let you pass things through that door into the site very efficiently and easily without developer intervention. In that way, they're identical.

[00:13:30]

[00:14:00] Where Segment is even more powerful is that not only does it create the door, it actually creates a full framework for tracking all of the events that are happening on the site, and then as you turn on different integrations with tools that it is compatible with, Segment does the mapping in real time from the way it is tracking data and the way those tools expect to receive the data. You never need to worry about, let's say Mixpanel changes its events API. Segment is on top of that, and it changes the way it writes data over to Mixpanel so your data flow is seamless, and you receive the data immediately in Mixpanel.

[00:14:30] One of the stories I have around this kind of efficiency was I was working for a client and they needed an analytics infrastructure, and so we put in place a strong instrumentation around all of the events that were happening, identifying their users, and we passed all the data to Google Analytics and Mixpanel, and then we also kept a regular Google Analytics instance just sitting hard-coded on their website so we could compare the two and triangulate the data.

[00:15:00] All was good and well, and we saw that the numbers were looking really compatible and the Mixpanel data was actually providing some really nice funnel analysis to understand how users were stepping through certain flows like there is a five-step welcome flow where first you had to provide your username, then you had to provide a bio, then you had to upload a photo, then you had to invite new users, and you could drop out of that flow at any given time.

What we could see in something like Mixpanel that is actually much harder to do in Google Analytics is how many people entered step one, how many completed step one and went on to step two, et cetera, to three, four, and five, how much time elapsed between each of those completions, and what was the total completion rate for our funnel, and how did the funnel completion rate trend over time?

[00:15:30]

[00:16:00] Once I got that in place, the company I was working for was totally floored, and then I said, "Well, wait guys. We're missing an opportunity," and they said, "What?" and I said, "Well, what about all the people who abandon the funnel? Don't we want to engage them and actually regain them as users because those are lost, your users?" They said, "Great, but how do we do that?" and I said, "Well, we need to turn on Intercom." They said, "Well, okay, but we don't have any engineering resources to turn Intercom on," and I said, "All I need is a credit card and authorization, and then we can turn on Intercom."

I'll never forget this meeting. It was in this big board room with 20 people, and I literally had, I just got the main decision maker to give me his credit card after the meeting-

Bill: Awesome.

Seth: ... and when I entered the credit card number, by the end of the day, we had 50,000 users that were inside of Intercom who we could automatically message both in the app and by email because all the heavy lifting had been done.

[00:16:30]
Bill:
Right, right, right. Okay. That is fascinating, so our most, so we're, we talked about one of the big ... You probably run into situations where people are misusing tools or maybe are not, or expecting more out of their, when you say GA, it's Google Analytics, so they're misusing Google Analytics, and they're being too simple or they've over engineered and they're doing too much.

Seth: Yes.

Bill: Okay. Is it-

Seth: So-

[00:17:00]
Bill:
Does it take much talent? Is it a talent issue or is it just too much complexity that people are just, aren't understanding how to rein in the complexity and keep it simple yet still get the outcome they're looking for?

Seth:

[00:17:30] It's ... That's a great question. I'm really glad you asked it because you know, it's funny. I have two types of clients. I have clients who have never measured anything. They're typically startups, and they need an instrumentation infrastructure that we write from scratch together, and then I have clients who have already done instrumentation through Google Analytics or through Mixpanel or, usually it's just through GA, and they sent an array of events, a mind-boggling number of events to Mixpanel and Google Analytics because those are, typically, the two that are used the most.

[00:18:00] For the clients that are just starting up, it turns out those are much easier projects because all I need to do is I work with a client to understand their business and what their core metrics are and what their KPIs are, and then we collaborate on a system of events that we're going to measure and properties that are associated with the events and users that we want to identify and traits that are associated with those users, and we can create the infrastructure quite efficiently.

[00:18:30] When I encounter clients who already have an infrastructure, there are two other major steps that are required, so not only do I have to understand their business and identify the best future state for what their analytics should look like, I also need to deconstruct what they've already built and reconcile what they've already built with what we need to have in order for the analytics to be really effective.

[00:19:00] It's typically not, it's never oversimplified. I've never seen an oversimplified analytics infrastructure. It's oftentimes the opposite where, through no fault of their own, my clients are sending way too many events with too few properties, and they're also not selecting the best events and the most effective events to send that are going to give them the strongest understanding of who is a high-quality user and who is a low-quality user.

[00:19:30] Typically, companies are focused on vanity metrics. They're focused on metrics that make them feel really good about the application that they've brought into the world, and it's very easy to produce a vanity metric. I worked for a company with ... This is actually my first project where I worked significantly with Segment and Mixpanel and events optimization, and it was so funny. They were spending an extraordinary amount of money on their Mixpanel bill every month. They kept breaking their quota in terms of the number of events that were getting fired, and they didn't have that many users relatively speaking.

[00:20:00] I went, I looked at the backend, and I analyzed all the data, and it, they were a, the easiest way to describe what they did is they were a Google Docs for code, so they enabled engineers to collaborate on code writing together, and they were trying to understand what an active user looked like, but the way they were measuring what an active user looked like is they were measuring every time a piece of data was saved to one of their servers, they recorded that as an event.

Bill: Oh, okay. Well, since you say-

Seth: So ... Yeah. So-

Bill:
[00:20:30] What do you find, do you find ... Are the companies you work with, what if they don't have a CMO that's very smart or intelligent, but their CIO or their digital leader has the native talent to basically understand data and understand the analytics?

[00:21:00]

[00:21:30] I guess what I'm trying to say is, I see a lot of kinetic physical traditional ... Like, I have a 400 billion company that sells Caterpillar, Earth-moving equipment. They're the largest in the area. They'd lose a lot of opportunity because they are getting their clock cleaned at the lower margin arena with some smaller type of tractors and bobcats, the smaller little landscaping deals, and those are going to case, and John Deere and such, but they could capture that if they had an infrastructure that could capture people looking doing quick searches for Earth-moving, landscaping gear, for example, and they needed the ability to marry online with the ability to take online data and convert to original sales force material where they can reach out to people with a phone call. How do you reconcile that in today's economy with people?

Seth: How do I reconcile the need to marry online and offline data?

Bill:
[00:22:00] Yeah, like how would you do that so that it produces action? I guess my biggest concern is that the data from a marketing angle sits and tells a story, but how does it, how do you convert people to actually making money from that information?

Seth:

[00:22:30] Yeah, it's a great question. In my mind, the most important thing is what's the shared key that links the online and the offline data. If you have ... I was actually reading about this recently where one of the simplest hacks you could do on your website, and this is a very rudimentary example, is if you provide a phone number on your website, then you track an event associated with the click on that phone number, and you track the campaign as phone-

Bill: Oh, interesting. Yeah, yeah.

Seth:

[00:23:00] ... and then you can look at all the phone conversions. That's one way to do it. I am on the board of a small business, a small self-storage business in Idaho, and I mean, it's reasonable revenues on an annual basis, but we, oftentimes, have the conversation about how effective is our offline print advertising because we have a nontrivial advertising budget.

[00:23:30] I have, now, mentioned a couple of times to the lead manager, and she's now adopting this methodology is, we're putting in offer codes in every single magazine, very simple offer codes, but each of the offer codes suggest what that ad was. If you say, basically you give a coupon code that refers to one magazine versus another, and then you know which channel was most effective.

[00:24:00] There are interesting tricks that you can use for bridging online and offline. When it comes to marrying a sales force database that something where it's a phone call away, yeah, it actually is quite tricky unless you have some sort of common identifier. Email addresses tend to be very powerful if you get someone to sign up for an email newsletter, and you know their email and you've contacted them by a phone, then you can join the two records on email. Right? It-

Bill: Yeah, absolutely.

Seth: But without that common identifier, it becomes kind of tricky, and so that becomes the big challenge.

But I want to go back to your other question, which is where, I think it was a bigger question around, well, what if you have an amazing CIO and an incredibly technically talented team but the marketing side isn't as savvy or sophisticated.

Bill: Yeah.

[00:24:30]
Seth:
It's a really interesting question because I feel like over time, the worlds of marketing and IT are merging closer and closer together.

Bill: 100%. 1,000,000%. I think if IT guy doesn't get it, the job's going to be on the line. I mean, maybe not tomorrow, but-

Seth: Yeah, it's true.

Bill: ... [crosstalk 00:24:44] very quickly.

Seth: And then-

Bill: I think the smart CMOs that understand what you're talking about are going to win, and-

Seth: Yup.

Bill:
[00:25:00] ... I believe that a lot of companies don't have strength at both positions. They just don't. I mean, people just don't get lucky. You don't have an A or B-plus player, you might have a C or B, and then what do you do if your CIO is actually a rockstar, and he could actually run data analytics and has people ... Well, you would know more about this. What do you see when you're in the market?

Seth: Well, it's funny because first of all, I think, in general, whomever is stronger needs to be able to step in and establish between both parties that they are not at odds, and instead, that collaboration is really where the win-win happens.

[00:25:30]

[00:26:00] If the CIO is stronger and has a rockstar analytics team, it's actually the, I think it's CIO's responsibility to say, "Hey, marketing team, we have incredible analytical horsepower. Let us help you do targeted marketing better." When it's the marketing team that has stronger prowess than the IT team, then it's the marketing team's responsibility to go to IT and say, "Hey, IT, believe it or not, we've developed a data infrastructure that you guys need to understand how to leverage throughout the whole organization."

There's a shared responsibility for lowering the stakes around, "Hey, one ... My position is going to overtake yours," and instead, understand it as a true collaboration.

[00:26:30] Where I've experienced that most acutely is one very large client I had that was in the telecommunication space. They brought me in to do automation analytics and marketing strategy for a software company that they had acquired that was like a Slack competitor, so it was like a chat application. The telecommunications company was playing across all Telco, so digital voice and video conference and across the board, but they wanted a chat player.

[00:27:00] The acquired company's strategist brought me on board to clean up their data, because typically, whenever they have messy analytics, the other thing that they, often time, say to me is, "I don't trust my own numbers. Something feels weird. I know we're growing, and I don't see us growing," or, "These numbers look funny," and so the first exercise is let's do a triangulation and build trust around the numbers.

[00:27:30] I did that. I understood exactly what was happening, then we implemented a segment-based automation where we're sending data in a really powerful way, and about six months in, all of a sudden, we had this crazy, power infrastructure, and all of this insight into not just users of this small software company's platform, but because of the data we were passing in on the users of the platform, half the users of the platform were customers of the larger company.

[00:28:00] Now, all of a sudden, we had insight into how customers of the larger Telco were using the platform or weren't using the platform. We could also identify, we had a sales pipeline where we could also identify anyone who is using the platform who wasn't a customer of the telecommunications company, and we could, potentially, if we did it within the right way where we're playing nice by regulations, we could upsell or cross sell the Telco companies other services true users of the software that weren't using other services for the Telco.

[00:28:30]

[00:29:00] It became a huge IT win, and my experience was we started from marketing, and then my last project before we concluded my entire engagement was to work with the IT team as a whole and unify the databases, the database that we had created just for this one software company and link it deeply with the databases for the entire company. It was a really interesting project. It was a really interesting study in ... At first, I think a lot of people were threatened by what we were building for this one software company because the business intelligence team, for instance, thought their jobs were on the line.

Bill: Yeah, yeah.

Seth:

[00:29:30] [crosstalk 00:29:10] I was doing is I was automating them out of a job, and so then my responsibility became to train the BI team and how to more effectively use the tools I had created so they could actually leverage them and answer data questions that they couldn't actually answer as effectively with their other tools, and then trust built, and then we sort of moved on from there.

Does that answer your question?

Bill: It does. It does, and what's really striking for me is as you're talking is that I almost am really starting to dislike the word big data, and I like the word little data better, because I ... I'm speaking primarily from the security world right now because it's context is king, context around ... I mean, a perfect example is when you are authenticating into the Microsoft platform.

[00:30:00]

[00:30:30] I was listening to the machine learning expert Russinovich talk to, this is two RSA conferences ago. He talked the complexity that's involved with a person that wants to access 365 email from, and they're based out of a company in California, and yet they're on travel to China, and their system across millions and millions of users has to reconcile the fact that where having a user that primarily is geographically located in California now having to appear in China to access the same systems. How does the 365 know that that, how can it reconcile that geographic issue for allowing someone access the system.

[00:31:00] That's a problem for machine learning to solve because you have to do it at scale, but that's context. That's a very granular context around a user, not just that that user has rights to the system but that user in a geography context has rights as well. It's amazing to me, and that's the challenge with security is context is king. Well, what you're talking about is deeply respecting context and not just at a macro that are going to launch your information, but really driving down into the context around it. That seems to be a huge angle that you're talking about.

Seth:
[00:31:30] Well, yeah, and I think this is a great transition from automation to data visualization as well and storytelling because I love the fact that you said, "This is not really ... I'm starting to hate the word big data because it's really about small data," and I also feel strongly that we live, Bill, we live in a world of big data. I call it the rise of data obesity or the start of databetes, and we need to solve the disease that is databetes-

Bill: I love that, databetes.

Seth:
[00:32:00]

[00:32:30] It's just too much, and it's too, we are literally over consuming. We're swimming in this vast ocean. I feel like there is a strong temptation, even among data scientist to show all of the data at the smallest possible granularity because you can, so I'm going to show a chart of second by second what's happening, but it turns out that's actually really not valuable. Having big data on its own is overwhelming and nausea-inducing and blood-sugar spiking. Instead, really, the art of great data storytelling is to aggregate and summarize that giant volume of data in succinct metrics that are easy to understand and are digestible and that you can actually trend in a meaningful way because at the end of the day, you're trying to show these numbers and provide this analysis, so there can be a strategic understanding of what's happening within the organization, and there can be strategic decision making on what to do next.

[00:33:00] Without zooming out and rolling up, I feel like there's a huge problem in the world of storytelling because there's an immense pride around the sheer volume of what's been captured.

Bill: This is ... I'm sure ... Are you familiar with Edward Tufte? I'm-

Seth: Mm-hmm (affirmative).

Bill: I went to one of his seminars like a year and half ago and-

Seth: Now, that's a business model, by the way, his seminars.

[00:33:30]
Bill:
What'd you say about the business model?

Seth: I said that's a business model. He's super-

Bill: He's a machine.

Seth: ... impressive. Yeah-

Bill: Because he's a total machine.

Seth: ... he's a machine, but I went to one of his seminars [inaudible 00:33:42] incredible. He's incredible.

Bill:

[00:34:00] A lot of things that you have, and I'm going to link out to this because you have an amazing, you're powerful using SlideShare, which I love SlideShare. I find you can find great expertise on SlideShare and the people that really deeply respect the visual, and there's data on this now. Our visual cortex is got so much more bandwidth.

[00:34:30] I think McCandless talks about this, and others, that we deeply, if we respect this channel of taking data in through our eyes, we can really use this big data and little data to our advantage, so that being said, how do you, in an engagement, where are you seeing the mistakes being made at the CMO, CIO data analyst side, and how do you quickly try to counsel or coach people into looking at a different paradigm of looking at the data?

Seth:

[00:35:00] That's a great question. I completely agree that we're visual beings, first and foremost, and it's funny, we gravitate towards text. I mean, look, I'm a writer, and I was a speech writer, but then I went into the world of writing or designing sides, and I have a background in graphic design as well, and I started realizing, wow, actually, we're much more about seeing than we are about reading. That's really what data visualization, effective data storytelling does, is that it acknowledges that we are animals that see, first and foremost, and we can react and respond instantaneously to something that we viscerally take in rather than something that we something read from right to left.

[00:35:30] Having said that, there's also ways that we digest visual prompts like a chart, and we actually do read it like a sentence. To answer your question, there are couple of big mistakes that I see my clients and my students making, and it's really what I really reach around.

[00:36:00] The first speaks to that notion of too much input, so there's literally too much content on the page. There's ... We're looking at too granular of a time scale. We're looking at numbers that are all the way out to, we're looking at millions of dollars, and we haven't rounded at all. Now, we're looking, if we're looking at data table, for each cell, there's seven or eight figures times however many cells there are on the page.

Bill: Oh, yeah, yeah, okay.

Seth:

[00:36:30] It creates literally a wall of text. It's this completely dizzying thing that there's no cognition around because if we are people who see, first and foremost, instead of read, how am I supposed to see the trend? I can see if things move from thousands to hundred thousands to millions, but beyond those giant spikes, how am I actually inferring quickly what's happening?

[00:37:00] The first thing is rolling up and simplifying in terms of the actual data that you have. Beyond that, there's also simplification of format. I can't tell you how many spreadsheets I've seen that are formatting disasters, and it's as if we feel like a spreadsheet is a piece of scratch paper that it's just a working document that we can use, and yet, if you present a spreadsheet or a model to a third party, that is their first interpretation of the data you're presenting is how it is displayed and portrayed.

[00:37:30] Whether or not, whether you like it or not, there's a very real process that their brain goes through to consume the information that's shown on that spreadsheet, so when columns aren't the same width or rows aren't the same height or fonts aren't the same and consistent or there's no information hierarchy where there's large text at the top and smaller text at the bottom and bolded text that signifies a summary, if design principles aren't put in place, it actually, you've made someone work much harder to infer meeting and it immediately frustrates the conversation because now we can't talk about what I'm showing you. We first have to talk about what it actually means or what am I actually saying before we can talk about what it means.

[00:38:00]
Bill:

[00:38:30]
This is really interesting. I have two questions. I'll start with the simpler one first. If I am getting what we're saying, if I'm a CMO or CIO or someone in charge of presenting the data to decision makers or to peers that are decision makers, what is a good ... In the mist, I'm going to share a story first, and then maybe you can elaborate and say yes, that would work, all right, here's an amplification to it.

[00:39:00] I had, one guy that I knew, he had a great team of database folks, and they started building dashboards and such to show the data. He would actually print out a sample dashboard, bring it to the business manager, and he gave the business manager literally 20 seconds to try to understand what the image was conveying. If he got a red light or a yellow light, he just went to the drawing board with his team. He's like, "If I can't get it within seconds, I'm done." That was his approach. What are you finding works for people to validate their intellectualness with visualization?

Seth:

[00:39:30] I think there are a couple of different approaches you can use. I think the one that you just described is a great one, which is constant validation and constant iteration and looking for instant insight. Is someone getting it with 20 seconds when I show them a piece of paper? But what if you don't have the opportunity to show them a piece of paper? Then it's up to you to become your own critic on the slide or the asset that you're building and asking ... What I typically do is I will ask myself questions. What can I remove without removing any meaning, and actually, what can I remove that will then emphasize the meaning on this chart or this table?

Bill: What a great, that's a great question.

Seth:

[00:40:00] Do I really need four lines? Can I gray out three of them and highlight one of them? Do I need the grid lines to show all of the contextual measures of the horizontal and vertical axis, or can I just point out the value that matters most? Can I literally put a box around the thing that matters most on the chart and let people's eyes draw to that.

[00:40:30] There's a technique I use called insight screenshot, and I find that to be extremely effective in terms of email communication because typically, I mean, you have to think about what's the format of the communication. The reason why your friend was checking in with his manager on every iteration of the dashboard and looking for 20 seconds of cognition was because the end product was a dashboard that was going to be delivered electronically where your friend would have no control over how that dashboard was received by the end user other than the design of the dashboard itself.

If they design a terrible dashboard, and the manager gets frustrated with it and can't understand it, then all of a sudden, the big conversation around what does this mean erupts. Your friend is heading that off with a pass, which I think is very, very smart.

[00:41:00] What happens when the medium of delivery is a presentation in the room? In that case, I show a chart and check for understanding in the room, and I engage the room in a conversation. If I don't have that luxury, and say I'm writing an email, then I'm going to rely on techniques that I know are really tried and true, and I find this one called insight screen shot to be exceptionally powerful. It's really easy.

[00:41:30] The concept is, it's just like a good essay. You tell someone what you're going to tell them. You hand-hold them, and you say, "Here's the insight I'm about to share with you in the form of a chart or a graph," and then you provide directly below that insight, which should be no more than one sentence, usually half a sentence long the chart or graph itself.

[00:42:00] The chart or graph should intuitively explain and show whatever point it is that you made in that sentence above. Then ideally, that, then, starts the foundation of an argument that you continue building with subsequent insights and screenshots, so you create a cadence in the email, and you say, "So, at first, this thing happens." Screenshot. "Then as time went on, this thing happens." Screenshot. "Now, this thing happened." Screenshot.

[00:42:30] If you do it effectively, and if you choose the right screenshot that is simple enough for someone to understand, it maybe has robust data that shows a complex trend, but you've provided the right context in your insights, all of a sudden, you've presented an incredibly power analysis in three sentences and three screenshots, and you've guaranteed that your end user is going to consider, at least consider doing an analysis because it is predominately visual rather than written. It's not a report. It's just a bunch of screenshots and explanations.

Bill: It's in the body, it's not something they have to, like an attachment they have to open up or anything. They can see the visual, the story emerging as the author is sending it to the person.

[00:43:00]
Seth:
That's right, and then there's one additional thing that I'll do for especially complex analyses or charts or emails that I need to write is I will use an annotation program like Skitch. I don't know if you're familiar with it. Skitch was bought by Evernote a couple of years ago. It was a brilliant acquisition.

[00:43:30] It's a really simple tool. It's a piece of desktop software. You download it, you put it on your computer, and then you take any screenshot that you'd like, and you copy it to your clipboard, and then you can paste it into Skitch, and Skitch gives you the ability to annotate the screenshot in vivid, colorful [crosstalk 00:43:36].

Bill: Oh.

Seth: So-

Bill: How do you spell that? Skitch is, S-K-I?

Seth:

[00:44:00] T-C-H. S-K-I-T-C-H. The default font color is bright magenta with a white boarder around it with a drop shadow, so anything you type, it's, instantly, there's contrast, and you can see exactly what the words are. It's not like it blends in with the image. You can drag arrows and point to anything, you can put boxes around things. You can also, which I really love, you can very easily blur things out by drawing, by clicking a little blur box and then dragging a box around sensitive data, and then blur it out if you need to share something [crosstalk 00:44:17].

Bill: Oh, yeah, yeah, yeah. That's interesting.

Seth:

[00:44:30]

[00:45:00] What's really powerful about Skitch, though, is that if you start with the asset of a chart, and you need to explain a trend that's happening. Let's say there's an upward movement for the first time period, and then a downward movement in the second time period, and then there's a spike in the third period. You can draw an upward arrow parallel with the trend line, and you can write with text that overlaps in the upward arrow. "This trend was upward because of blank," so you're using the arrow to drive your insight, and then you draw a downward arrow parallel with the downward trend, and you write an explanation as well. "It's going down because of A, B, C," and then if you point an arrow to a spike or you draw a box around the spike, you can, then, label the spike and explain, "This was because of this campaign that we ran."

You, then, copy that in Skitch, and that becomes your new base asset that you paste into your email. It's not editable. It's an image, but-

Bill: I love that.

Seth:

[00:45:30] But then what you do, and this is what becomes so exciting, in my mind, this is where I really nerd out on things is you've now created a visual sharable asset that ... My job, really, is to set my clients up for success and to make them look good, so I send this analysis to my client, and they, then, can send it on to their SVP where they can share it with everyone and they, basically, say, "Hey, here's how things are trending," boom. No forwarded email from me. They just take the asset, they pass it right along, done. So-

Bill:

[00:46:00] Yeah, so these are the tools that people need, though, because you're actually answering the question that I had because the biggest thing that we have to respect from context is our time-

Seth: That's right.

Bill: ... and the busier, the more higher up the food chain one is, they, I think what I love about our conversation is that, basically it's helping them skip time. "Okay, what are the tools? How can I do this fast? I have the data, but I know I gotta respect people's time around me, and I gotta put that, but how do I do it because PowerPoint, I know it's going to take me a lot of time. How can I, is there a way to shortcut this and hack this way to doing it," and then that's what you're explaining with Skitch.

[00:46:30]
Seth:

[00:47:00]
There's a couple of other time savers. It's funny, I teach a class called Digital Tools for Business, and the first second is all about personal productivity and how do we save time in email digest and there's a thing called Unroll.Me that lets you roll up all of your emails into a visual digest. You don't need to wade through junk mail. There's something called RescueTime that lets you understand how you're spending every single minute on your computer and assign applications and websites with productivity scores from negative two to positive two and then it gives you a grade at the end of every week. It's really insane. It's like Big Brother-y but in a wonderfully, positive way.

[00:47:30] But when it comes to data visualization, there are some amazing time sucks that, if avoided, can create incredible good will within a team. You describe them as hacks, and I think they're great hacks as well. One of them is the fallacy of building your own chart. I think a lot of people are convinced that Google Analytics hasn't done a good enough job of building the charts, so I'm going to download the GA data and build my own chart in excel or I'm going to make it an interactive chart in PowerPoint.

[00:48:00] That's the wrong answer. Why not just take a screenshot of the correct view in GA and put it into the presentation, or why not take a screenshot of the view in GA and put it to an email, highlight the screenshot, and link it back to the page in GA so you prompt your reader to click on the image and continue the analysis from there, to play it from there.

To me, one of the most important things is to sit on top of existing charting platforms that are already doing the visualization and leverage them whenever possible rather than reinvent the wheel.

Bill:

[00:48:30] Or like you said, with Skitch, people, you can take a GA, a Google Analytics chart you might not like that much, but it still makes the point, and you put it into Skitch, and then you can build the story around the image because you can just annotate right to that image, or right to-

Seth: That's exactly right, and it's precisely what I did with the GA dashboard very recently for one of my clients. They asked me, "How is one subset of content comparing to the overall content on our website in terms of performance around the key engagement metrics?" and there are about eight different engagement metrics.

[00:49:00] Well, I created a segment inside of GA, and that segment represented all of the pages that the client was interested in learning about, and I selected all traffic as one segment, and this custom segment I created is the second. I pulled up the page in GA, it showed all the comparative data. I did a screenshot of that data. I pasted it into Skitch, and then I started highlighting things that were different, like two times higher, time on page; five times higher, pages per visit.

[00:49:30] Then I was able to basically just tell the story and summarize it very quickly at the beginning of the email, but that became the sharable asset, and it saved so much time than building the custom presentation with all the charts. It would have been the exact same conclusion, but I did it in 20 minutes instead of two hours.

Bill: And in enough fraction of the slides, potentially just a handful of slides to tell the story.

Seth: Yeah, or one or two, max. I mean, it's kind of amazing.

Bill:

[00:50:00] Yeah, that's great. I just love this. Well, I can go on and on with questions, but I want to respect the time that we've spent today, and we can always have a round two at a future point. We can kind of, I mean, this can go on and on. These are two topics that are, I'm very, and I know you're passionate about. I am as well. Again, as we talked about, from a different angle, from the security angle, but they play out equally from the automation that we're talking about with the marketing and such.

[00:50:30] Is there anything that we left out that you want to make, before we go, any points that you're like, either I didn't ask you, or you're hoping I'd ask you that we could end with some value for our listeners?

Seth:

[00:51:00] Sure. I mean, one of the things that comes up with data visualization, I think when everyone thinks of data visualization, they rarely think of static charts and graphs. I think they often think of amazingly immersive ... Actually, I think when some people think of data visualization, they think of Minority Report and Tom Cruise and basically using his hands to move everything around, but beyond that, I think when people hear data vis, they're often thinking of an interactive data visualization, something with animation, something that moves around or is controllable by the end user.

I just want to touch on that really fast because it actually speaks really strongly to both what's possible and the powerful potential of interactive visualization, but also the time-suck nature of it as well, and there are couple of fallacies within the world of interactive database that I think is worthwhile for your listeners to just keep in mind.

[00:51:30] It turns out that just because you could potentially, and let's think about it in terms of slides. How often have you found a PowerPoint slide or a PowerPoint deck that overuses animations for transitions or for builts?

Bill: Oh, I mean, that is, yeah, frequently used, get attention, right?

Seth:

[00:52:00] Right, and it's a little, they can be a little cringe-worthy, usually. I mean, sometimes, they're amazing, but oftentimes, it's like, really? You had to use a spinner to animate the thing? Like really?

[00:52:30] That actually is a risk or a fallacy that happens inside of the data vis world as well where there's a temptation just to over animate transitions between charts because you can. I think it's really important for us to keep in mind that as we build really complex data visualizations, we have to fight the temptation to animate everything, and instead, to use animation as a tool that helps reinforce the storytelling aspect of the visualization.

[00:53:00] That's where this concept of scrollytelling comes in. It's a really interesting new trend in data visualization, and what it does is it doesn't rely on the user to click on buttons in an interface to transform one chart into another. Instead, it uses the action of a scroll through a webpage to trigger transformation's individualization.

[00:53:30] Let's say I start with a stack bar chart looking at the distribution of types of jobs over time, and then as I scroll down the page, after a paragraph of text scrolls by that says, "Notice that for the top five biggest volume jobs, there's really been a decline over the last five years," and then as I scroll past a certain point on the page, all of the other jobs in the stack bar chart fade away, and the five stacks that were the biggest expand into their own individual time series.

That's really powerful, because now what I've done is I've handheld someone through a data visualization argument through a data story, and it's sort of the like the next level beyond the insight screenshot thing that I was describing in an email. It's the interactive version of that.

[00:54:00] When used effectively is incredibly powerful, and there are some wonderful examples of scrollytelling. Actually, if you just do a search on scrollytelling, you'll find a couple of Google results, and there's one that The Guardian recently published on a subsection or a neighborhood inside of Chicago called Homan Square, and about arrests that happen in Homan Square. Excellent example of visual storytelling and scrollytelling, and from there, you can find a number of others as well.

[00:54:30]

[00:55:00] I find that it's funny. As we enter this brave new era of needing to be visual and succinct as possible with communicating ever larger sets of data that as data visualizers and as data storytellers, we have a responsibility to hold back and not overuse the animation and understand that that kind of interaction, that kind of animation is also a super important tool that, when used effectively, is very powerful, but when overused, can do as much damage as a giant wall of text.

Bill:

[00:55:30] I love that, and at heart, we're storytellers, I mean, going back to the caveman days, I mean, we have thousands of years of history of putting visual images to represent our stories, and the platform of the day's not a cave, but it's these, our mediums that we're using right now. What I love is the fact that we've gotten down into time is one of the biggest things we want to convey stories powerfully in respecting the time of the day because I think, you know, it's interesting with AI and some of the automation that's happening of the data, that's just automating stuff that probably should be automated anyway, in some respects, but we still are left with the responsibility to convey the story in the message.

[00:56:00]
Seth:
That's right. Time is that fleeing resource, and it's amazing that if you have a good story to tell in a concise way, then you can capture someone's time, and you really make the most of it.

Bill:

[00:56:30] Well, this is great, Seth. I really appreciate the time we have spent today. I'm going to, I think you and I just have to, I want to make this blog post powerful for people so they can link out to the resources to learn more about the tools because you have specific resources on tools, you have, really, the data visualization piece, which had half a million downloads, and then you also have the big data automation part as well, which we covered in the first part, but you have some really good assets that people can learn more about you, how to reach out to ... What is the best way for people to reach out to you? Which platform do you prefer?

Seth:

[00:57:00] My website familian1.com is the easiest way to find me, to get in touch with me, and see all of my presence on social media, as well as get examples of all the presentations I've done. There's also an interactive visualization of all the different tools that I love that you could explore.

LinkedIn and SlideShare are also a place that I have a very strong presence, please feel free to connect with me on either of those two platforms. You'll find Twitter on my website as well. Twitter is actuality a place that you can find the most compelling visualizations that I stumble across. I typically cross post to Twitter and LinkedIn.

Bill: Well, this is great, Seth. I appreciate you for your time. I'm sure at some point, we're going to have a round two.

[00:57:30]
Seth:
I would love a round two. In the meantime, for all you listening, if you're interested to learn more, I really welcome any sort of communication, questions, or any ways that I can help you. I really love helping people solve their automation, data vis, and analytics problems.

Bill: That's great. I'm going to put direct links for people to link out to your website and to connect with you on LinkedIn and Twitter. This is good. This is good.

Seth: [inaudible 00:57:54].

Bill: I appreciate your time. Thank you, Seth.

Seth: I appreciate all your great questions, Bill. Thanks for a wonderful conversation.