

SDS PODCAST
EPISODE 553:
THE STATISTICS
AND MACHINE
LEARNING QUESTS
OF DR. JOSH
STARMER



- Jon Krohn: 00:00:00 This is episode number 553 with Dr. Josh Starmer of StatQuest.
- Jon Krohn: 00:00:12 Welcome to the SuperDataScience Podcast, the most listened to podcast in the data science industry. Each week, we bring you inspiring people and ideas to help you build a successful career in data science. I'm your host, Jon Krohn. Thanks for joining me today. And now, let's make the complex simple.
- Jon Krohn: 00:00:43 Welcome back to the SuperDataScience Podcast. I cannot believe that I'm joined today by Dr. Josh Starmer, the creative musical genius behind the wildly popular YouTube channel StatQuest, which provides uniquely clear statistics in machine learning education, earning the channel over 650,000 subscribers so far. In addition to his brilliant StatQuest channel, Josh is Lead AI Educator at Grid, a company founded by the creators of PyTorch Lightning that enables you to take an ML model you have on your laptop and train it seamlessly on the cloud. Previously, he was a researcher at the University of North Carolina Chapel Hill for 13 years, first as a postdoc, and then as an Assistant Professor applying statistics to genetic data. He holds a PhD in biomathematics and computational biology, and he holds two bachelor degrees, one in computer science and another in music.
- Jon Krohn: 00:01:35 In this episode filled with silliness and laughs from start to finish, Josh fills us in on his learning and communication secrets, the single tool he uses to create YouTube videos with over a million views, the software languages he uses daily as a data scientist, his forthcoming book, The StatQuest Illustrated Guide to Machine Learning, and why he left his academic career, and a question you might want to ask yourself to check in on whether you're following the right life path yourself. Today's epic episode is largely high-level and so will appeal to anyone who likes to giggle while hearing from

one of the most intelligent and creative minds in education, on data science, machine learning, music, genetics, and the intersection of all of the above. All right, you ready for this incredible episode? Let's go.

- Jon Krohn: 00:02:29 Josh Starmer, welcome to the SuperDataScience Podcast. I can't believe you're here and that I get to speak to you in an interview. I've been excited about this for so long. Where in the world are you calling in from, Josh?
- Josh Starmer: 00:02:44 I'm in Chapel Hill, North Carolina.
- Jon Krohn: 00:02:47 Nice. We're filming in the winter. I suspect that winter in North Carolina is pretty idyllic.
- Josh Starmer: 00:02:54 It's relatively mild, but for some reason we've been getting snowstorms for the past two weeks and I think we're going to get another one. So it's a little bit of a bust this year, but usually it's nice. It's like highs in the fifties and lows in the thirties, but relatively good sunshine. So I like it.
- Jon Krohn: 00:03:12 Yeah. I guess it all just melts away right away anyway. Or, yeah.
- Josh Starmer: 00:03:16 Yeah.
- Jon Krohn: 00:03:17 Cool. So we know each other through Matt Dancho. So Matt was a guest on episode number 463 of the program. He runs a educational platform called Business Science, creates lots of great educational videos on particularly R as a tool for slicing and dicing data, for running statistical analyses. And it sounded from the email intro that he made that you might have collaborated with him in the past.
- Josh Starmer: 00:03:45 Yeah, we've done a couple of videos together. I'll be honest. Matt has been a huge help for me. When I first

left my job to do YouTube full-time, he was like holding my hand those first couple of months. He would check in like, "How are you doing? Let's do something together. Here's some tips on how to get going." And so he was super helpful and very encouraging and a big part of my ability to keep a positive mindset and think, oh, I can make this work because Matt's helping me out. It was just a huge help.

- Jon Krohn: 00:04:27 And it has clearly made a big difference. So I could imagine six years ago when you were starting out creating the StatQuest YouTube channel that you don't know how things are going to be. And I understand that. I mean, I recently started making videos and you're like, I think I'm making good content, but no one's really watching it. And now you're in a different scenario because now you have 650,000. Yesterday, I looked at your YouTube channel because I was like, okay, 650,000. Today I looked, 651,000. That is probably about your growth rate. Thousands of people per week subscribing to your channel. You've created tons of amazing videos. As the name implies, there's a lot of statistics in the channel, including you'll do a whole video series on specific aspects of stats like regression models, for example, but you also have lots of videos on machine learning. And machine learning in general, like what does it mean, training sets versus test sets, cross validation? So these general machine learning principles, but then also you've done video series on specific techniques. So neural networks like deep learning neural networks.
- Jon Krohn: 00:05:41 It is unsurprising to me that you have so many followers, because you have amazing visual explanations. You clearly put a lot of time into not only figuring out how to convey the information effectively, but then how to visualize it and create a really crisp video. So without dumbing anything down, you have these visual explanations with clear steps. You tell the viewer, this is

exactly what other videos you're going to need to watch before you tackle this one. It is no surprise to me that people love the channel so much. You have several videos with over a million views, including ones on principle component analysis on logistic regression. In case it isn't obvious to the listener already, if there's any topic that you're looking to shore up in statistics or machine learning, the first place to look is Josh's channel to see if he's made the video on that yet, because that is the way to learn it.

- Josh Starmer: 00:06:44 Oh, thank you very much. I really appreciate that.
- Jon Krohn: 00:06:47 Yeah. And it's not just stats and machine learning. You also have some original songs on the channel. So you are a musician. I know for sure you play a cello, a ukulele. You're a vocalist. What other instruments do you have in the bag so to say? Oh, you got a piano behind you there.
- Josh Starmer: 00:07:09 I've got those tabla drums. I haven't played them in a long time. But I've got a guitar, I've got a mandolin, I've got a banjo. I've been playing the bass a little bit lately.
- Jon Krohn: 00:07:23 Yeah. The guitarist loves to make fun of the bassist. Obviously you can joke that if you can already play guitar, of course you can play a bass. Now a really good bassist is something else, but...
- Josh Starmer: 00:07:38 Well, I have a secret actually that people not... I'm only going to tell you, nobody else will know. I don't actually know how to play a real guitar. So the bass is a real stretch for me. It's very different. I play a tenor guitar, which it only has four strings.
- Jon Krohn: 00:08:00 Wow.

Josh Starmer:	00:08:01	And it's tuned in fifth. The lowest note is C, G, D and A. And it's only four strings and it's tuned just like a cello. So, if you grew up playing the cello-
Jon Krohn:	00:08:17	Wow.
Josh Starmer:	00:08:18	... you can play the tenor guitar, just like out of the box. You don't have to learn anything. But if you give me a six string guitar, I'll just look at it and be very scared, because it's got six strings and that scares me.
Jon Krohn:	00:08:30	You mentioned mandolin, ukulele, those are all four string instruments. Do they all have the same tuning or did you have to learn at least some other tuning? A ukulele surely is different from tenor guitar.
Josh Starmer:	00:08:38	I mean, it's like the top four strings on a guitar. The mandolin is also tuned in fifths and I'm natural in fifths, so mandolin was like, dang, I can do that. It all started because I played the cello and the cello's a great instrument. I love the cello, but it's huge and it's terrible to transport.
Jon Krohn:	00:08:59	Right.
Josh Starmer:	00:08:59	And if you want to go like backpacking, you can't take a cello backpacking. So the first instrument I got was this thing called, I got a mandolin and that was super fun and I loved it. But then I found these tenor guitars and they're just magic. That's all I pretty much play these days.
Jon Krohn:	00:09:18	Cool. I don't suppose, I mean, the listener now knows that you've got a tenor guitar right there. I mean, could you just play a little melody for us on it? I mean...
Josh Starmer:	00:09:32	I mean, of course. I mean, I can play... So there's a... Just to give you a sense of the crossover between the tenor guitar and the cello is... One of the most popular cello

pieces is the Bach G Major Cello Suite and it kind of goes...

Jon Krohn: 00:09:46 Oh yeah.

Josh Starmer: 00:09:57 And so it's really easy to anything I can do on the cello I can do on this guitar, but it's also, because the chords are super straightforward, there's really only like two shapes in your bar chords. You only have to cross four strings instead of six. So it doesn't take much muscle at all.

Jon Krohn: 00:10:18 Right.

Josh Starmer: 00:10:19 I highly recommend it. It doesn't sound as rich as a full six string guitar because, not to get all geeky about guitars and tunings, but the distance in pitch between the strings on the tenor is wider than on a guitar. And that closeness in the harmonies that you get on a guitar contributes to a sound of warmth and richness and you don't get this... This is a slightly colder, icier sound. It's not as warm and rich.

Jon Krohn: 00:10:53 So to me, a cello does sound really rich, however. Is that because I'm used to hearing it with other instruments as well, perhaps?

Josh Starmer: 00:11:02 Well, it's just because the cello is a big, huge monster instrument with so many overtones coming at you. It's these thick strings, just getting vibrated like crazy. And as a result, it's just this dark, woody instrument and it's going to sound super rich just by the way it's been built. I guess what I'm talking about is, when you're strumming chords-

Jon Krohn: 00:11:29 Right.

- Josh Starmer: 00:11:30 ... you get a much more open sound on the tenor and you get a much closer, sort of richer thing on a guitar, but a cello, almost just always... Very rarely do you play chords on a cello. So that was kind of a new thing for me.
- Jon Krohn: 00:11:43 Great.
- Josh Starmer: 00:11:43 A cello is almost all melody.
- Jon Krohn: 00:11:46 Actually, that was going to be my next question is, do you ever play chords on a cello? That does happen sometimes?
- Josh Starmer: 00:11:50 It does. Yeah, it does. In fact, those Bach cello suites that are pretty popular have a lot of chord stuff in it, but it's rare. And one of the big differences in sort of learning how to play music on the cello versus my tenor guitar, which I'll just call a guitar for now because the concepts are the same, on the cello, you think of everything in terms of melody and on the guitar you think in terms of chords. Someone just says, oh, you just play D, G, A, C and on the cello, I'd be like...
- Jon Krohn: 00:12:25 Or just play a G for four beats, just a single note.
- Josh Starmer: 00:12:32 Yeah. Exactly. So that was a tough transition to just wrap my brain around thinking about songs in terms of chords, as opposed to songs in terms of melodies.
- Jon Krohn: 00:12:44 Right.
- Josh Starmer: 00:12:46 I think that melody... I do write my own music, not just... Obviously to clarify, a lot of people know me as the guy who sings a silly song at the start of all my videos, but I actually do like serious music. And when I do my serious music, I think of it as layers of melodies rather than chords.

- Jon Krohn: 00:13:11 Right. Right. And that shows how much more sophisticated you are with music than me, because for me... So I'm a rhythm guitarist. I sing and I play rhythm guitar. In rock bands, that's really all I ever did, except that I could trivially play the bass if I had to and have done it because then it's the same, because I see the chord charts and I'm like, well, I will play G just for this whole bar. Yeah. So for me, yeah, it's the dream. I have this amazing...
- Jon Krohn: 00:13:45 So I love the Beatles, I think pretty much everyone loves the Beatles. It's a cliché, but I love the Beatles too, just like everyone else. And I have this Beatles chord songbook that I bought almost two decades ago. I grew up in Canada. So I was still living in Canada when I bought this. And then I went to Oxford and did my PhD there. So this chord songbook of all the Beatles songs, it's comprehensive. And the person who did it, did an impeccable job. This guy's name is Rikky Rooksby. And so Rikky Rooksby he... Because to do it very well, especially when you're converting melodies, like somebody playing a melody on a guitar, as opposed to just strumming chords, to convert that into a sensible chord and have it work, that requires a lot of ingenuity. And so Rikky Rooksby did it immaculately in this book.
- Jon Krohn: 00:14:42 When I started at Oxford in... So you might not know this about me, but like you, I used to be in genomics. So my PhD was in neuroscience, but I was specialized in creating machine learning algorithms or statistical algorithms, much like you, to analyze genomic data and brain imaging data. Anyway, the Genomics Center in Oxford is a bit away from the center of town. And it's in this area called Headington outside of Oxford, Oxford. And we get this email sent around the Genomics Center saying "guitarist teacher seeking students". Rikky Rooksby lived a few minutes' walk from the Genomics

Center and I had guitar lessons from him and he signed my book and we're still in touch.

Josh Starmer: 00:15:31 Wow. That's so cool.

Jon Krohn: 00:15:33 Yeah.

Josh Starmer: 00:15:34 Wow.

Jon Krohn: 00:15:36 When I talk to leaders in data science, I notice they all make time for learning and encourage the same of their teams, but with your actual everyday work to do, all day trainings aren't possible for most of us. That's why an on-demand learning platform like Udemy Business makes sense. With Udemy Business, you can access over 500 cutting-edge data science courses taught by real-world subject matter experts and validated by other learners' realtime reviews. Amongst these 500 courses, you'll find my own Mathematical Foundations of Machine Learning course, as well as dozens of mega popular courses from other SuperDataScience instructors. To hear the latest on the state of data science in the workplace and discover how you can democratize data science learning in your teams through Udemy Business, check out the new video series called Insights on Demand: Diving Into Data Science. To watch this series and learn more, visit business.udemy.com/sds. That's business. udemy.com/sds.

Jon Krohn: 00:16:34 Yeah, we have digressed a lot. People came here for stats and all they're getting is tenor guitars. Oh well. Yeah. So the reason why we got here is because your StatQuest channel does also have some original songs and it also has some covers. So I particularly love this Psycho Killer, [Keskes 00:16:56].

Jon Krohn: 00:16:56 (singing) Yeah, got into that. And yeah, as you mentioned, even your videos, you break up. I mean you break up

technical content with levity in a lot of ways. So your tone of voice, it creates levity and enjoyment. You have musical jingles that you'll interlude with or start or end videos with. And you also have, you have a catchphrase that is brilliant on so many levels, Josh. So for marketing purposes, for ease of marketing, for memorability, for listeners, just for creating that levity in videos and also then for creating social media engagement, it's just a single syllable.

- Josh Starmer: 00:17:45 BAM!!!
- Jon Krohn: 00:17:49 What is that? BAM!!! Exactly.
- Josh Starmer: 00:17:50 BAM!!!
- Jon Krohn: 00:17:50 And then I love... Yeah. I even saw it when I asked listeners if they had questions for you. And so I wrote this post on LinkedIn and it's ended up being... We're going to get to audience questions later at the end of the episode, but it's the most viewed post I've ever made. And part of what's driven the social media engagement is people writing "BAM!!!" as comments and then people replying that and saying "Double BAM!!!" which is another. I don't want to work in a triple BAM!!! too early into this episode, but maybe we'll say something of such significance that will merit it.
- Josh Starmer: 00:18:33 Yes, let's say that.
- Jon Krohn: 00:18:38 Yeah. So again, clearly your channel is one that I highly recommend people checking out. So tell me, what is your inspiration behind your topic choices? I mean, it was probably different in the beginning than it is now, but I'd love to have some insight into that.
- Josh Starmer: 00:18:53 Well, let's go all the way back to the beginning. So at the very, very beginning, my inspiration was my coworkers. I

used to work in a genetics lab doing statistics for my coworkers. They were all what we call wet bench researchers, meaning they do experiments that involve wet things as opposed to dry things. A computer is a dry thing, but they're like measuring small volumes of liquids from one tube and squeezing them into other tubes. So they did all these experiments and they had me doing the statistics and it's an academic lab and that means there's new people coming and going basically every six months, sometimes even shorter periods than that. I would do the statistics, but I didn't want them to think what I was doing was magic.

Josh Starmer: 00:19:49 And so what I did is I started these little like, it wasn't really a seminar series, but I started... We had a lab meeting every Friday and part of that lab meeting, I would do a little stat chat, like stats corner or something like that. And I'd try to teach my coworkers, the statistics that I was using on their research. So it was real obvious what to talk about, because it was like, they'd say, "Hey, Starmer, can you do a T test?" I'm like, "Yes, I can. And I can teach you about it." And so I'd do those little things, but like I said, it's an academic lab and every six months there's new people. And I was like, I could repeat myself every six months for the rest of my life, or I could just record a YouTube video and put it online. And then, when someone in the lab needs to learn about R-squared, they just go to the link and they learn about R-squared. And it's the same presentation I would've given anyways.

Josh Starmer: 00:20:47 And so if you watch the early videos, they're all about mice, because it was a mouse genetics lab. And so all I'd do is talk about mice and things you can do with mice. But to be honest, when I was starting out, I mean the channel, I thought was going to be relatively... I wasn't intending for it to be private, but I just figured it was going to be private because let's be honest, back then I was like, who wants to watch a video about R-squared

other than someone in Dire Straits? I just assumed no one would watch that kind of stuff. And I remember at the time-

- Jon Krohn: 00:21:23 You think members of the band Dire Straits are really keen to know the proportion of variance?
- Josh Starmer: 00:21:30 Mark Knopfler's dying to learn about R-squared.
- Jon Krohn: 00:21:35 Well, finally R-squared with mice, this is what we've been waiting for.
- Josh Starmer: 00:21:39 Yeah, exactly. Yeah. He text me like five times a day. He's like, "When's it coming out?" It drives me crazy. I'm like, "Come on, you got to wait for it." But anyways, and I'll be honest in the first year, the only people that watched my stuff were people in my lab and that was fine because that's who it was for. But I'll be honest, I'm not like some Zen guru who's just like, I just want to do it for the... I had some fantasies about people watching my videos and I thought, wouldn't that be cool if people watched them? But I had a friend of mine, he was starting a guitar channel and it was like blowing up. And I was like, dang, if only I was teaching guitar stuff-
- Jon Krohn: 00:22:23 If only I could play the guitar. All I've got is this tenor guitar and no one likes listening to it.
- Josh Starmer: 00:22:29 No one wants to learn how to play the tenor guitar. So I was like, man, that's the way to do it. He's got the channel. And I'll always just have this little tiny channel and it's fine because it spared me the agony of having to present the same material every six months. And so it served its purpose, but in the end it kind of blew up despite my best efforts to make it sound obscure, going, well, let's weigh another mouse. And I've got all these business people watching and now they're like, "Could

you use some business examples instead of mice?" And I'm like, "Well..."

- Jon Krohn: 00:23:03 You've got a mouse with a suit.
- Josh Starmer: 00:23:06 Exactly.
- Jon Krohn: 00:23:07 That goes to the office.
- Josh Starmer: 00:23:08 A mouse goes to Wall Street.
- Jon Krohn: 00:23:09 And weighs himself.
- Josh Starmer: 00:23:12 Yeah. So, that's how it started out. Anyways, to go back to the original question, what is my inspiration, the inspiration for that was my coworkers. It started out that way. And then I was at a conference once and this guy who was also at the conference and he was a chemist and he goes, "You're in statistics, right?" And I go, "Yeah." And he goes, "Do you know anything about random forests?" And I said, "No." And he said, "Yeah, because I just read a paper and it had a random forest. And I have no idea what it means." And I said, "I don't either." I was like, "But I'm kind of curious." And so when I came home from that conference, I researched random forests and random forests are made from decision trees. And so I had to research decision trees and I was like, I'll tell you what, I'm going to make a video on decision trees.
- Josh Starmer: 00:24:02 And that was the start of StatQuest's transformation from being all stats all the time to being some stats with a lot of machine... I think I've got more machine learning videos than I have stats now. And sometimes I worry that maybe the name of the channel is holding me back because I think a lot of people just say, oh, Josh Starmer's great for stats. And you got to go to this other channel for machine learning. And maybe that's true and maybe that is the best way to learn. But I sometimes

wonder if it'd been more like DataScienceQuest or something like that. But anyways, it is StatQuest. It's what it is.

- Jon Krohn: 00:24:37 I've got a great name for you. It's StatSymbolQuest.
- Josh Starmer: 00:24:38 StatSymbolQuest. Yeah. I like that.
- Jon Krohn: 00:24:45 When did it change from... You jokingly called it a Stats Corner on when it was like the Friday. When did it become StatsQuest? StatsQuest.
- Josh Starmer: 00:24:54 StatQuest. Yeah. And now you can't say it anymore. You're going to always say StatSymbolQuest.
- Jon Krohn: 00:25:00 StatSymbolQuest, as it'd been known for minutes.
- Josh Starmer: 00:25:01 Exactly. The channel. I realized that there might have already been a channel called StatChat or a website called StatChat or something like that. And so I had a vote and I had the lab vote on what it should be called. And we went with StatQuest. And I love it. It's funny because in hindsight my whole life has been nothing but quests. I'm preparing a talk for McGill that I'm going to give on Friday and they want to know about my path. And so I was going way back to high school. Like what did I do in high school? Did anything I do in high school, give me an indication of where I would be later in life?
- Josh Starmer: 00:25:49 It turns out that I taught myself how to program to build a game. And the game was called Jimmy Quest. And really, all I did, I mean, I had a few... It was like an adventure game. And like, you go through a door and you find some stuff, but there really wasn't a whole lot to the game. The point of the game was to have the theme song. And there was a theme song for Jimmy Quest. And it went, Jimmy, Jimmy, Jimmy, Jimmy, Jimmy, Jimmy, Jimmy, Jimmy, something like that. And you had a yellow

bouncing ball going over all the words of Jimmy. So it just never... Like why even follow... It was a joke that my friend of mine came up with. Anyways. I was like, wow. It goes all the way back there. I've been on a quest forever.

Jon Krohn: 00:26:30 Oh yeah. I mean, hopefully most people could... Now that you've mention it, especially I expect people that are listening to a data science podcast, hopefully there's been some questing in your life. Yeah. I love that you've named them clearly the Jimmy Quest, now the StatQuest. Super cool. So then today, with now there being so much machine learning on there, so we had the beginning with T tests and specific techniques that would happen a lot in your lab, then the random forests conversation led to the random forest content. So now with all the recently machine learning, neural networks, I guess they might similarly have been things that you were like, I would love to understand that a lot better and maybe other people would too.

Josh Starmer: 00:27:27 Yeah, exactly. So not everyone knows this about StatQuest or me personally. A lot of people think that I know everything there is to know about both statistics and machine learning. And I'll be honest, I'm like one StatQuest video ahead of you. That's the extent of my knowledge. I mean, I've done this for a long time. So yeah, I've actually got some good depth at this point. But I did a series on neural networks because I was like, I got to figure out how these things work. And I've watched a lot of videos and I've read a lot of blog posts. And I was like, "None of this makes any sense to me." And so I had to do it my own way and I had to just make a simple neural network and plug numbers in and just see what happened. And once I saw what happened, I was like, "Oh, I get it. You just have to have a real simple input and that gives you a real simple output, and you can do the math and you can see what's happening." But before I made that video, I'll be honest, despite my best efforts, I

only had a vague concept of what a neural network was, other than a weird, spiderweb of messy things.

- Jon Krohn: 00:28:36 Yeah, yeah. I wonder how many people would be blown away. So if people have been working with regression models for decades and they see all the incredible things neural networks are doing, it's staggering that it's just a bunch of regression models.
- Josh Starmer: 00:28:55 Yeah, exactly. It's just fancy... It's a, what is it? What do they call it? A weighted regression model, or something like that, or something, I don't know. I remember when I was in graduate school, there were these two guys in the Stats Department and this one guy goes, "Have you heard about neural networks?" And the other guy goes, "Yeah, it's just a bunch of weighted regression." Or I think he might have said, "Weighted least squares." And I was like, "I don't even know what any of that means, ah." So I've been thinking about that for the last 13 years.
- Jon Krohn: 00:29:32 [crosstalk 00:29:32].
- Josh Starmer: 00:29:32 I was like, "How do I not feel intimidated by those guys?" Right? Because they clearly knew so much more than I would ever know in my entire life, just in that off-hand conversation that they had. So I've been working on it and I finally figured it out and I was like, "Oh yeah, sure. You could call it that if you wanted to."
- Jon Krohn: 00:29:49 Yeah, yeah, yeah.
- Josh Starmer: 00:29:50 But it's a little different in that, it's got a lot more flexibility in that, it's a lot easier to add more nodes and more hidden layers and things like that.
- Jon Krohn: 00:30:02 Yeah, yeah.



- Josh Starmer: 00:30:03 But with that greater flexibility, comes a lot... It's not as straightforward to put one together and have it work correctly. Whereas, regression is such a science, right? That it's so obvious as to how to assemble it. And I think that's a real advantage that regression can have over neural networks sometimes.
- Jon Krohn: 00:30:26 The alternative data industry is expected to grow tenfold in the next five years. So how will your skills and training stay competitive in this brave new world? Well, fortunately for us, the Alternative Data Academy has just released the world's first alternative data courses aimed at working professionals, through its free and on-demand training platform. The courses are taught by some of the biggest names in the alternative data space, including the prominent New York University professor, Gene Ekster, and renowned Oxford professor, Alexander Denev. Sharpen your alternative data skills for free today with the Alternative Data Academy. All the details are online at altdata.academy, that's altdata.academy.
- Jon Krohn: 00:31:06 It was in the past decade, neural networks have suddenly become really relevant. Not only because computers become a lot cheaper and storage has a lot cheaper, so we can have bigger data sets, which are two really big things that we need for neural networks, particularly deeply layered neural networks, deep learning approaches to work effectively. But the third piece alongside the cheaper compute and the bigger data sets, is that in the last 10, 15 years, people have come up with ways of more reliably training neural networks. And that's been a big innovation.
- Josh Starmer: 00:31:43 Yeah, they got rid of the sigmoid.
- Jon Krohn: 00:31:44 Yeah.



- Josh Starmer: 00:31:47 That was the big thing, that was the big turning point. Get rid of the sigmoid and use the ReLU activation function from now on. I mean, there's variations on that theme, but I think the sigmoid was a major inhibitor of neural network progress.
- Jon Krohn: 00:32:03 Yeah, yeah. And that for people who aren't aware of it, so these words, sigmoid, ReLU, they are functions that we apply to basically, the result of a regression at each of these little neural network nodes within a broader system of these neurons. And that's another really powerful thing that does differentiate the approach significantly from regression or linear regression, is that it then allows you to have nonlinear relationships. So you can have regression models, I think almost always at their core, there's going to be a linear relationship. You can have a polynomial input but... So neural networks can automatically handle non-linear relationships between inputs and outputs and interaction terms and so it means that... And there's a really great chapter four of Michael Nielsen's free online book called, Neural Networks and Deep Learning, has these little Java script applets that you can play around with in your browser, that proves to you that by having these non-linearities allowed by a sigmoid or a ReLU function, it allows you to approximate any continuous relationship between an input and output. So it's part of why neural networks are so powerful.
- Josh Starmer: 00:33:19 Yes. The other thing that's cool about neural networks, is they don't need a whole lot of data. For general linear regression, right, you need more data than you have parameters. Neural networks, you can actually have less data than you have parameters and you can still fit it, it's crazy right?
- Jon Krohn: 00:33:38 Yeah.



- Josh Starmer: 00:33:38 How's that possible? But it works.
- Jon Krohn: 00:33:41 Yeah. It's true, it does. It boggles the traditional statistician's mind to think that you're like, "Let's have a model with a billion parameters and 10,000 training data points." You're like, "That's never going to work. That's the stupidest thing I've ever heard." And yet somehow, it's a super, super powerful, cool technique. I mean, I don't know if you know this about me, but it's deep learning and teaching deep learning that got me started in communicating data science to people. So I could go on about this forever, but this episode is about you. So what's your process? Do you have a particular process for learning content?
- Jon Krohn: 00:34:29 I know from videos of yours that I've seen, that generally, you're trying to identify general principles as opposed to memorizing things, memorizing equations, so you actually... I can't remember what all of the examples were, but in a video that I watched recently, you showed several different equations where the concept was the same, because in the numerator we had explained variance and then the denominator was the total variance. And you had three very different looking equations used in different statistical approaches, that all really meant the same thing. So is that a key part of your process to learning? What other tricks do you have up your sleeve?
- Josh Starmer: 00:35:18 Yeah, I do try to... My dad, ever since I was a little kid, he was always like, "You got to focus on the main ideas." And so just my whole life doesn't matter what I'm doing, I've always thought, "I better focus on the main ideas." And yeah, I just hear that voice in my head and that's what I try to do. As opposed to getting lost in, what do they call? The bells and whistles. It's easy to get excited about bells and whistles and get overwhelmed by them. I sometimes use a car as an example of lots of bells and whistles.

When you watch a television commercial about cars, they're going to talk about bells and whistles as if they're the most important things ever. And to be honest, they'll be like, "This has got something cool feature. It's got a rear view window. It's got a..." Or what else they're going to say. "AM, FM cassette." Or something.

Jon Krohn:	00:36:24	Yeah.
Josh Starmer:	00:36:24	They're going to try to impress.
Jon Krohn:	00:36:26	Heated pad behind your head.
Josh Starmer:	00:36:27	Heated headrest, yeah.
Jon Krohn:	00:36:28	[crosstalk 00:36:28]. Heated headrest, yeah.
Josh Starmer:	00:36:29	He's going to say, "This has got a manual transmission." All these things or whatever. They're going to try to blow your mind with all these fancy, fancy [crosstalk 00:36:38].
Jon Krohn:	00:36:37	Extra manual...
Josh Starmer:	00:36:38	Yeah, exactly. "It's got a parking brake." And you're just like, "Wow. All these bells and whistles are so cool. I'm going to go buy this car." Well, fundamentally, if you're a smart car buyer, you're going to think about is the engine reliable? What's the gas mileage? Does it-
Jon Krohn:	00:36:59	How expensive are repairs?
Josh Starmer:	00:37:00	How expensive are repairs? Those are the main ideas, those are the main concepts of the car. And the bells and whistles are just marketing ploys.
Jon Krohn:	00:37:10	Yeah.
Josh Starmer:	00:37:10	And so I try to look past all those. And I'll be honest, it's easy, because everyone talks about the bells and whistles.

Especially with machine learning, you hear all in that hallway conversation, where like, "Did you hear about neural networks?" I don't know what they were saying, but it didn't make any sense to me at all, but it wasn't also... I mean, everyone's talking about the latest, greatest and then machine learning especially. There's always the latest, greatest, there's a new method coming out for neural networks every other week. How many of those have long term lasting value? I'd say only a small percentage of them. And so it's really important to try to sort the signal from the noise and the noise being all the latest trends in data science. And not to say that the trends are not important, the trends are important, but fundamentally, we need to... At least from my perspective, I like to focus on the main ideas of how things work, so that I can either see how they're related to other things that I already know, like seeing the relationship between neural networks and regression. Or seeing how things are different and going, "Oh, this is actually a fundamentally different process." And that's actually very useful to know, because when one method works, I don't necessarily want to use a method that's highly correlated with that first method. I want to use something that's orthogonal and it's going to be completely different and try that as a way to... Maybe that's got a better chance of working or just validating that the results are from this are consistently reproducible. And so it's nice to know what's similar and what's different. And I find the answers to those questions are given to me from the main ideas.

- | | | |
|---------------|----------|---|
| Jon Krohn: | 00:39:11 | Cool. I love it. So as a specific example, when somebody says random forest to you or neural network to you, and you're like, "Oh, I don't know anything about that." |
| Josh Starmer: | 00:39:21 | Yeah. |
| Jon Krohn: | 00:39:22 | You just start with a Google search? |

Josh Starmer: 00:39:24 I do. Yeah, the step one, Google search.

Jon Krohn: 00:39:27 Yeah.

Josh Starmer: 00:39:27 And then I just click on every single thing I see. Although at least I used to, I don't do that anymore, I'm a little more judicious. But I used to read everything, but now-

Jon Krohn: 00:39:40 You're just like, "Oh, I should go watch the StatQuest video."

Josh Starmer: 00:39:44 That's true. Yeah, that'd make it a lot easier.

Jon Krohn: 00:39:48 I'm digressing now, but does that happen that you made so many videos now over so many years that sometimes you come across something and you're like, "Oh, I made a video on that four years ago and I don't really remember anymore."? And then you actually watch your own video to remember. Does that ever happen?

Josh Starmer: 00:40:03 I do watch my own videos to remember all the time, because well, partly because people are always asking me questions about them and I can't remember the details.

Jon Krohn: 00:40:12 Right.

Josh Starmer: 00:40:12 They'll be like-

Jon Krohn: 00:40:13 Right, right.

Josh Starmer: 00:40:14 "Wait a minute. What's Lambda?" And I'm like, "What? I made that video three years ago."

Jon Krohn: 00:40:21 That is something, so if you're listening, listener and I guess, you are listening. So since you're listening, listener, when you ask a question of a YouTube video, please put a time stamp in, because I'm sure this happens to you at orders of magnitude more than me and it definitely happens to you at orders of magnitude more

than me, because there'd be this normal distribution of how much people are commenting on your videos. And then you take how many I get, and then you multiply it by a thousand. And that gives you how many comments you're getting on your videos. And so when you get these comments, so you're watching the video, there's something that doesn't make sense to you, or you have a question about, and so you just ask the question. But some of these videos are long, they're 20 minutes long and it might have been made years ago. And so I have no idea what you're talking about and I can't take the time to watch all of the video to try to guess where you were when you had that question. So if you put a timestamp, when you ask a question that would be super helpful.

- Josh Starmer: 00:41:21 That's a bam. You put the timestamp. I say, bam.
- Jon Krohn: 00:41:24 Bam. We've got one.
- Josh Starmer: 00:41:27 Yeah, it's a lifesaver. So yeah, people ask real specific questions about things I haven't thought of in a long time and that's okay.
- Jon Krohn: 00:41:38 Right. And so anyway, so I've taken you now several steps away from where we were on the intended course of conversation. So going back from answering questions. Yeah. I mean, just not even, how do you make the videos? Which actually is my next question, but how you learn? So you go into Google, you click on everything on the first page about random forest and then?
- Josh Starmer: 00:42:08 Then I do it on the second page and the third page and I have 30 tabs open and I read everything I can. And very little of it will make sense at all, I'm not very good with equations, I'm not really good with Greek mathematical symbols, a lot of them scare me. And so I glaze over when I see that stuff and I go, "Come on, someone's got to have something else." And I just go page to page and I'll pick

up a little bit of lingo while I'm doing that, I'll start seeing terms that pop up a lot. When I was researching neural networks, people started saying, weight, people started saying, bias, they start saying that a lot or activation function. And so it's like, I it's like my brain is building a histogram of terms that it frequently sees in association with this topic. And I'll go, "Okay, well, what's an activation function?" And so I'll dive into that. And so then it'll be another thing, okay. So lots and lots of reading and reading and reading and then I just go back. And once I've got to the bottom of that, or I can't go any further, because I'm going to go crazy, because you never hit the bottom. It's like that hole, you're just like ahhh, and you're falling forever. And that's the way the internet is sometimes. But anyways, so what I'll do is I'll go back to the top of that very first search and I'll go through everything again and I'll reread it and I'll go, "Hey, okay, now I know a little bit more about activation functions. This thing I read, now makes a lot more sense."

- Josh Starmer: 00:43:41 Or I'll say, "Oh, I've seen this equation, but I saw it, someone else broke it into smaller pieces." And so I understood this one piece. And so now when I see the whole thing, maybe it makes a little more sense. And it's an iterative process. And once I've got a very basic, very fundamental, rudimentary understanding of what's going on, my next goal is to make the simplest example, like the hello world of whatever it is. I make a very simple neural network. I make a very simple support vector machine, whatever it is, I try to make it as simple as possible. And then I just plug numbers in and I do the math by hand.
- Jon Krohn: 00:44:18 Right.
- Josh Starmer: 00:44:19 Those big scary equations with all the Greek, they're scary, but they're just like those things I had to do in high school. You just plug in a number and then you got to multiply it by E and raise it to some exponent and then

you've got to divide it by something else. I just do the math and as scary as it is, it's still just a little warm, fuzzy equation, that's not... As long as you plug in the numbers, you're going to get something as output. And that helps me a lot, it takes a lot of the fear and the anxiety that I usually associate with seeing a complicated mathematical formula and I go, "I just got to do it." And once I do it, I draw the output, I draw relative to the input and I start seeing what's happening. And that's a big, big help for me. And then I also, if there's an implementation, I'll drop in my little simple data set and I'll see if I get what I expect. And I never do, I always get something different. And that's where the learning really happens. I start, "Why didn't I get what I expected?" I have to answer those questions. And to me, it's very exciting. And it's scary, it starts off very scary. And it's funny, because you'd think after doing this for years, I'd be less scared and intimidated, but I'm always scared and intimidated, because those first, sometimes weeks I feel like I'm getting nothing done.

- Josh Starmer: 00:45:41 I'm like, "I'm never going to understand this." And I freak out, but I always do. And you think I'd learn that, "Don't worry Starmer. You're going to figure it out." But I always freak out. I'm like, "I'm never going to figure it out." But I do. But that's the process and it's very iterative and it's staged. So you start with reading everything and then you start with trying to do the math and then you do an implementation. And once I've done all three of those things and everything's lining up-
- Jon Krohn: 00:46:12 Yeah.
- Josh Starmer: 00:46:12 The implementation, my little doing the math, what I've read, everything is like, "Oh yeah, this is all consistent." Then it's time to start making the video.

- Jon Krohn: 00:46:23 Awesome. So you've gone through your weeks of trials and tribulations, your little mini quest, your random forest quest. "I'm in the random forest and I can't find the sword. Jimmy, where'd you go? You're lost, I know you must be around here somewhere." So you go through the reading stage, the math stage, the implementation stage. Now it's time to create the video. And I know from another video of yours that, so this was another thing that your dad imparted on you, was that there's always a better way. So that photo of the turtles and the one on the upside down turtle shell, rowing. So there's always is a better way to do things. And so do you end up... I mean, I guess you're about to tell me, but I'm guessing here, in addition to the iterative process of learning, there's probably an iterative process of creating. So from the dozens of places that you'd learn something, probably some of those stuck, you were like, "Oh yeah, that was a cool visual way of doing it. But actually if it had a bit from that other blog post, it'd be even a little better." And then all of a sudden, bam, you've got a new way of doing it.
- Josh Starmer: 00:47:43 Yeah. So the way I do it is when I'm researching, none of it makes any sense to me at all. But I ask the question, why? Why does it not make any sense? What's missing? What are they not saying? What parts are they have they omitted that prevent me from just instantly grasping this concept? So I ask myself those questions and that's what I try to... The answer to those questions of what's missing is what I try to provide with my video. And one thing I like about that strategy, is it means I'm not copying anybody. This is an original perspective. And what I like about that is...
- Josh Starmer: 00:48:22 I mean, the internet is in some ways a big echo chamber, someone comes up with something, someone says, "This is how support vector machines work." And then someone else put up an article on their blog or whatever, that's basically the exact same thing. And sometimes they even

use the same pictures and everything. And if blog number one did not help me, I'll guarantee you blog number two is not going to help me either, because they're saying the same thing. And that may reach some people, some people may read that and go, "I get it." Okay. But I didn't get it, when I read it, I didn't get it.

- | | | |
|---------------|----------|--|
| Jon Krohn: | 00:48:55 | Often when I come across that, I'm even not convinced that the author of the post gets what's going on here. |
| Josh Starmer: | 00:49:02 | That's true. Yeah, because they're just copying somebody else. |
| Jon Krohn: | 00:49:06 | Yeah. You're like, "What?" |
| Josh Starmer: | 00:49:07 | Yeah, anyways. |
| Jon Krohn: | 00:49:09 | Explain it in your own words. |
| Josh Starmer: | 00:49:09 | Yeah, exactly. Yeah, exactly. So anyways, so what I'm thinking when I see that I go, "Well, that might speak to somebody, but it doesn't speak to me." And so what I want my video to be, when people read that blog and they go, "That didn't make any sense." I want them to have a place they can go, that says it differently and fills in all the blanks that first person left out, so that they'll go, "Okay, well maybe it wasn't the first thing I looked, but the second thing that StatQuest video, that got the things, and I was able to understand it that way." So I try to focus on what they're not saying and the things that would've helped me understand it to begin with. So I guess, it's a lot of focus on me. And it's my quest, I'm trying to learn new and exciting things. And so I keep track of what didn't make sense and what they didn't say and what maybe if they'd said, maybe I would've gotten it. |
| Josh Starmer: | 00:50:05 | And to be honest, when I do read up on a topic that I think is like super crystal clear, it's super hard for me. |

And I don't even want to make a StatQuest on that. All I want to do is just say, "Hey, just check out this other website, because they do a much better job." And sometimes I'm wrong about that, sometimes I'm right, sometimes I'm wrong. But if I find something that's great to begin with, like this guy, Jay Alammur, he has some awesome visualizations. And it's like, "I don't even want to make a StatQuest on that, his stuff", bam, that's done. So there's just some great stuff out there and I don't want to overlap them with my own, I don't need to do that. And making a StatQuest takes an incredible amount of time.

- Jon Krohn: 00:50:53 I bet.
- Josh Starmer: 00:50:54 So I'd rather not have to do that. So when I find something that's good, I'm like, "Okay, that's done. I don't have to do it. I can move on to another topic."
- Jon Krohn: 00:51:01 Yeah. It's clear that you put a huge amount of time into your videos. So when you come up with the visualizations, the general flow of what you'd like to do, then it seems like you script it. I mean, you definitely do, because you put the text on the screen, so you script it. And so maybe you have some word processor file with the sketches of what you want to have and the text that will accompany it. And then you convert that into a slide format, I guess. Are you just using slides?
- Josh Starmer: 00:51:34 Yeah. I do everything in Keynote. I love Keynote. Yeah, I love it.
- Jon Krohn: 00:51:41 Yeah.
- Josh Starmer: 00:51:42 I wish Apple would give me a direct connection to some of their developers, I've got some ideas, if any of them are listening, I've got some great ideas. But I love Keynote, it's just... I'll be honest, I just use the default settings on everything, everything looks great.

- Jon Krohn: 00:51:57 Actually, you totally do. I see that now, yeah.
- Josh Starmer: 00:52:02 Defaults on everything Keynote and I do it all Keynote and I just love it and I highly recommend it. So if you want to make something that looks just like mine, it's super easy. Just do it in Keynote with the default settings and it'll look just like a StatQuest. But yeah, I do it, I outline it and iterative, that's a very iterative procedure as well, where usually what I do, is I take... Like I said, during the research phase, I do the math and if I think doing the math is going to be helpful for other people, if it helped me, then I'll include the math and that little... Remember that the simplest example I can possibly come up with, that's always key. So a lot of that research funnels straight into the video. And a lot of the research I do on within Keynote, I just open up a blank Keynote thing and I just start typing in random thoughts and I'll draw out what I think is a simple example and I'll test it out. And for me, Keynote is there from the very beginning, all the way to the very end.
- Jon Krohn: 00:53:06 Awesome. And then, so the illustrations themselves are also made in Keynote?
- Josh Starmer: 00:53:12 Yeah. Everything's done in Keynote.
- Jon Krohn: 00:53:13 Cool.
- Josh Starmer: 00:53:13 Very rarely, occasionally, I have to do three dimensional graphs. Like in my gradient descent video, I had to draw three dimensional surface or something like that. And I'm like, "Ah, I can't do that in Keynote." But almost 99% of the time I do everything... In fact, I feel like I'm going to steal a punchline to maybe later on in the thing, but I actually wrote a book in Keynote. I wrote an over 300 page book in Keynote.

Jon Krohn: 00:53:46 What? Somehow, even in researching for this episode, I didn't know you wrote a book. Has it not come out yet? Because one person on Twitter when I said I was interviewing you, wrote, "I'm waiting for his book."

Josh Starmer: 00:53:58 Yeah. It comes out this spring. It's called, The StatQuest Illustrated Guide to Machine Learning. It's coming out in a few months.

Jon Krohn: 00:54:06 Oh, cool.

Josh Starmer: 00:54:07 But it was written in Keynote with all default settings.

Jon Krohn: 00:54:11 Very cool. Wow. Awesome. Well, I can't wait to check that out and neither can that Twitter user and I'm sure countless others. So how did that person on Twitter know that you're making this, you've talked about it before?

Josh Starmer: 00:54:28 Yeah. I've tweeted about it, I mean, especially on Twitter.

Jon Krohn: 00:54:30 Yeah.

Josh Starmer: 00:54:31 I post about things like that on Twitter and on LinkedIn. It is something I'm pretty excited about it. People have been asking me to do a book for a long time and likewise, I always like, "Why would you want a book? We got the Introduction of Statistical Learning and oh I love that book." But over time, I realized that we did need a book. There was a lot of things not being said. And I can also, when I say I wrote a book in Keynote, it's actually, I drew a book, it's all pictures. I mean, there's a little bit of text, but I drew a book and it's over 300 pages of me drawing a picture for... It's like a comic book, it took forever.

Jon Krohn: 00:55:20 Wow. I get it.

Josh Starmer: 00:55:21 But I don't think there's anything else like it.

- Jon Krohn: 00:55:23 Yeah. That does sound unique. So my book is called, Deep Learning Illustrated.
- Josh Starmer: 00:55:27 Oh, wow.
- Jon Krohn: 00:55:29 Yeah. And so there are a lot of illustrations. So I worked with, I'm fortunate to have a friend who is an extremely talented artist. She's actually is a fine artist. So her making illustrations for a deep learning book was really taking a step down for her. And when she does a gallery opening, it doesn't mention my book ever. But still, it's not like the book you're describing, still, it's a textbook. I did it in LaTeX, so it has LaTeX for the type setting of the language, and then I insert an illustration and I insert code, and so it flows very much like a normal book. We focused on visual explanations, wherever we could. It's something like a 500 page book I think, and there's something like 200 illustrations. So there's a lot of illustrations, but I have had some super scathing Amazon reviews that are like, "This isn't illustrated." And I'm like, what? I think what they're looking for is what you've described, where it's really every page is it's an illustration per page and the text is explaining the illustration. It's illustration first.
- Josh Starmer: 00:56:51 Yeah, it is an illustrated guide. You could take the words out. You could take all the words out and you could probably still read the book, and it would probably make a little bit of sense.
- Jon Krohn: 00:57:10 Awesome. Yeah. I just remembered one very specifically, which was like, "Would've given this a five star review, but it's not illustrated. So giving two."
- Josh Starmer: 00:57:20 Ouch.

- Jon Krohn: 00:57:21 Come on. Anyway. All right, so the book's coming out. That's something that I didn't even know. That's super exciting. In addition to that, so in addition to the StatQuest channel, in addition to the forthcoming book, you are also the lead AI educator at a firm called Grid AI, or Grid.ai?
- Josh Starmer: 00:57:42 Yeah, Grid.ai, or just Grid. There's two, or there's two lead AI educators. We're two leads, but the other one is Sebastian.
- Jon Krohn: 00:57:57 Yeah, Raschka.
- Josh Starmer: 00:57:57 Yeah, exactly. So we were just hanging out. We just spent last week in Milan together making our plan of attack and how we can collaborate. We do different things, but there's going to be some intersection, which is pretty cool, because he is a wizard at the nuances of actually implementing a neural network and making it be awesome. I'm more of like, well, I read about our neural networks and so what else is there to do? He actually knows those nuts and bolts, and so we're thinking of coming up with some great collaborations between the two of us where we combine my big picture and make sure you understand the main ideas with his and let's actually get this to work attitude. I think we got some cool stuff coming up. I don't know what you want to ask about Grid, but so far it's been awesome. I love what they do.
- Jon Krohn: 00:58:54 What is it? Would be a good one.
- Josh Starmer: 00:58:55 Yeah. So, Grid's a couple of things.
- Jon Krohn: 00:58:58 Two dimensions, and we've got vertical lines and horizontal lines.
- Josh Starmer: 00:59:03 Yeah. It's hard, so let me explain. I don't know if you've ever tried to do anything in the cloud.

- Jon Krohn: 00:59:10 Sure.
- Josh Starmer: 00:59:11 I'm intimidated by the cloud in terms of layers of overhead and tedium and things that, of you got to log into a node, you got to configure the node, you got to get your stuff on the node. How many nodes do you need? You don't even know. It's like, I don't know, the cloud scares me with all these implementation details, and Grid, say you've written a program in Python and you're like, "I've ran on my laptop and it works great. Let's run on the cloud now," all you've got to do is add one word to your command line that proceeds Python, and boom it's on the cloud.
- Jon Krohn: 00:59:50 Wow.
- Josh Starmer: 00:59:51 Yeah. That's what I'm talking about. That's my kind of thing. I mean, some people love that kind of stuff. I'm not a big fan of all that stuff. I just want to run my program on the cloud and not have to worry about all those details, and Grid takes care of it. That's super awesome. Grid, they also support something called PyTorch Lightning, which is a layer on top of PyTorch that, again, what it does is all those technical details of I want to run this on this GPU and I run this somewhere else, blah, blah, blah, all these little technical details, PyTorch Lightning takes care of that. You don't have to worry about it. You just worry about making the best neural network you can without having to worry about any technical details.
- Josh Starmer: 01:00:42 I love that they're basically making everything something that I would want to do, as opposed to just look at it and go... Because I'll be honest, I don't have a whole lot of time. I mean, I spend so much time on my StatQuest. I don't have a whole lot of time to be dinking around with getting stuff running on the cloud, but I think it's important. It's something I want to do, because I'm kind

of curious about it. I hate to be that guy that's like, "Well, I can talk about it, but I can't do it." I don't want to be that guy. I want to be the guy that actually can do it, and so Grid is saving me by making it a lot simpler to do that. I mean, I got to be honest, I love that it's a company that values me for what I do. They're like, we want to make this product something that anyone can use, and that's why we're bringing on Josh Starmer.

- Jon Krohn: 01:01:40 Because if Josh can use it, anyone can use it.
- Josh Starmer: 01:01:43 Yeah, and he should be able to explain it to people. I think it's a great product, and what I love about it is it takes all the, what is it? Sort of the overhead of expertise for cloud computing and infrastructures and all that, and it makes it go away. And it means anyone anywhere, they could watch a StatQuest video and then go, "Oh, I want to try that too," and then without having to then spend the next four months reading about how to do it on the cloud, they can just start doing it on the cloud from anywhere. So even if they've got the crummiest laptop ever, it's a hand-me-down from their grandfather, and they're not going to be able to run anything locally, they can still do real industrial grade machine learning anywhere in the world with anything.
- Josh Starmer: 01:02:39 It doesn't matter. The hardware you bring to the party doesn't matter, I love that because a lot of my audience is all over the world. Some places are super rich countries, and sure, they probably have relatively easy access to fancy computation, computational tools and fancy computers. But other places, they may not have as easy access to the state of the art that other people do. And I love that Grid is making it so that you don't have to be super rich, you don't have to be living in a rich country. You can have access to all this stuff anywhere in the world, no matter what, and I love that too.

- Josh Starmer: 01:03:20 The other thing, I hate to keep talking about how much I like Grid, but the last thing is they're actually coming up with a way to, because on these new Apple laptops, those processors are so discreet and yet powerful at the same time, that they're trying to come up with a way that people can donate processing power to anyone. Presumably this is going to be people that need it to be donated to them, but you could donate computing power to anyone for free. And so then people could come in from any background, any situation and go, "Okay, I want to learn about machine learning. I want to do machine learning. I want to do it in the cloud, but I've got this crummy grandfather hand-me-down computer and I don't have much money for cloud compute time." Well, they're figuring out a way to give away for free cloud compute time, and I think it's amazing.
- Jon Krohn: 01:04:18 That is awesome.
- Josh Starmer: 01:04:20 I'm a big fan.
- Jon Krohn: 01:04:21 I hope they find a way to prevent people from mining Bitcoin with that free compute.
- Josh Starmer: 01:04:28 Yeah, that's true. I'm right there with you, that's kind of an abuse of all kinds of things. It's kind of frustrating.
- Jon Krohn: 01:04:35 Oh, super frustrating. Unbelievable. All right, that sounds super cool, and you are doing a brilliant job of explaining how cool Grid is. I love that. Also, if you, Josh, or anybody listening ever knows how I can get Will Falcon, the creator of PyTorch Lightning on this show, I'd love to have him on the program.
- Josh Starmer: 01:04:57 Oh, sure. I'll see what I can do. I will.
- Jon Krohn: 01:04:58 Wow.

- Josh Starmer: 01:04:59 One little pitch, Grid is hiring, so if anyone listening to this is like, "Hey, I want a job," go ahead and apply. That'd be great. They're trying to grow, it's a startup, I think there's like 50 people right now. Awesome people. Like I said, I just spent a week in Milan with these people that are just, it's all distributed, so we were in Milan as like a popup office for a brief amount of time, because there was just stuff we could do together. But the company itself is a hundred percent remote. So it doesn't matter where you are, there's people in Croatia, Switzerland, France, Mexico, obviously I'm in the United States. You name it. There's there's people in India, working for Grid anywhere in the world, people are working for Grid. There's about 50 of us. Now they're trying to expand as quickly as they can. So go ahead, if you think you've got some skills, please apply. But yeah, I'd love to get Will on the show.
- Jon Krohn: 01:05:54 Awesome.
- Josh Starmer: 01:05:55 It'd be awesome. He's a great speaker.
- Jon Krohn: 01:05:58 I know, he's top of my list. And so for those job openings, do you have any specific titles that you know of that are currently open, that people could be applying to?
- Josh Starmer: 01:06:10 I mean, I think they're looking for engineers, all kinds of engineers. They're looking for technical writers. So even if you're not super, super nerdy.
- Jon Krohn: 01:06:22 Ah, cool.
- Josh Starmer: 01:06:26 They don't just want super technical, technical writers. They want people like English majors and philosophy majors, people that know how to communicate clearly, know how to speak in active voice. That kind of thing. They're interested in a wide variety, so even if maybe you

don't think your background is in tech, and I don't know why you're listening to this podcast, but...

- Jon Krohn: 01:06:51 For the musical tips.
- Josh Starmer: 01:06:53 Yeah, exactly, because you're curious about tenor guitars, but maybe you know someone who's an English major and is out of work or something like that. You can say, "Hey, apply for a job at Grid." I mean, because they want the best at all kinds of things, and as someone who is an avid reader of documentation, I love that they're putting a premium on the quality of the documentation they want to create. They're willing to hire people outside of the box just to make it so that you go to the documentation page and you're like, I get it. Plus, there's going to be a StatQuest there to help you along too. So it's like a double support.
- Jon Krohn: 01:07:34 Nice. I want their documentation to rhyme. That would be great.
- Josh Starmer: 01:07:39 Yeah, exactly. We'll get some poets.
- Jon Krohn: 01:07:42 So awesome. Yeah, it sounds like an amazing place to work. It sounds like they're doing incredible things, and it would be great to be able to work alongside people like you and Sebastian. It sounds like an opportunity. What kinds of tools do you use, either at Grid or just... So you've kind of talked about when you are learning a concept, you will do things by hand, but then?
- Josh Starmer: 01:08:10 Yeah, so...
- Jon Krohn: 01:08:11 What programming languages?
- Josh Starmer: 01:08:13 When I was working in a research lab, I did everything in R for a long time.
- Jon Krohn: 01:08:17 I knew it.



- Josh Starmer: 01:08:18 Yeah. So I'm an R guy. I do a lot in Python too, don't get me wrong. Don't send me hate mail.
- Jon Krohn: 01:08:28 You better not side too hard on this one.
- Josh Starmer: 01:08:31 Yeah. So one thing I like about R is that it's super lightweight, when all you want to do is do a little bit of math and draw a graph. In Python, by the time you're done importing all your modules, I've already finished and I've got my graph. It's super fast for stuff like that. And so I do a lot of little prototyping in R, but I'm a big fan of Jupyter Notebooks. I like prototyping in Jupyter as well.
- Jon Krohn: 01:09:06 But you could do that, that could be R or Python, right?
- Josh Starmer: 01:09:09 Yeah, you can do it either way, but I lately I've been doing it in Python because I work for a company that creates Python stuff, but those are my two languages right now that I prototype in, and prototype is what I do. I'm hoping with Grid to get a little more serious, and with Sebastian to learn how to get a little more serious. But those are my two tools of choice that I go to. If it's super quick and dirty, I always go to R. That's sort of my native language, so it's just, by default that's what's there.
- Jon Krohn: 01:09:52 Yep. I completely understand that, and having also come from doing a PhD focused on genetics, applying statistics and machine learning. One of the chapters of my PhD and actually a spin out, a startup that came out of my research, we were doing everything at that time in MATLAB. I was mostly R and a bit of MATLAB, but yeah, more and more for making code that needs to go into production systems. I know that there are ways to do it in R, and actually we have a really great episode on that. It was episode number 491 with Veerle van Leemput, who has a Dutch name that I always mispronounce, and so apologies to Veerle, and again, I'm sure I just did it again.

But really great episode on deploying R into production, so it can be done, but relatively rare.

- Josh Starmer: 01:10:53 You can do cool stuff in R. There's just demand for both languages, and to be honest, I like being able to do both. It makes me feel cool.
- Jon Krohn: 01:11:04 Yep. I think there's a huge amount of value in knowing both as a data scientist today. I mean, why not? It's the same as the TensorFlow versus PyTorch argument. It's like once you know one, it's pretty easy to know the other, you might as well just spend a little bit of time on it. Cool. So now we know what tools you're using today. Are there any tools or maybe even statistical or machine learning approaches that you're excited about for the future? Maybe something that overcomes some limitation that we have today?
- Josh Starmer: 01:11:37 Oh, I don't know. The future.
- Jon Krohn: 01:11:41 That's a tough question.
- Josh Starmer: 01:11:41 That is a tough question. To be honest-
- Jon Krohn: 01:11:43 Predict the future. What's the future?
- Josh Starmer: 01:11:46 Well, something that makes it a little, especially hard for me, is I'm almost always looking at the past. Making my videos takes so much time that I kind of like to wait for things to kind of like... Stuff is invented and research is going on, and then something will come up and it'll settle down and it'll last longer than six months or a year. I usually wait for those little nuggets to percolate from the noise of research, and so oftentimes, I find myself looking backwards rather than forwards. And so in terms of what do you predict are they going to be the great trends? I'd be like maybe one day Python will get a good

implementation of a random forest, but right now it doesn't.

- Josh Starmer: 01:12:34 I dream about that. That's like one of the things that makes me go back to R, is R has an amazing implementation of random forest, and I actually like random forest for reasons that you cannot actually use in Python. Random forest can be used for clustering, and what's cool about using a random forest to cluster your data is you can throw any kind of data you want at it. It can be discrete, it can be categorical, it can be rankings, it can be continuous. You can take any type of crazy data, throw it into a random forest and use it as a clustering algorithm.
- Josh Starmer: 01:13:14 Not many people know that, partly probably because the Python implementation does not allow for that to be done, but in R you can do it. It's amazing. It's an amazing tool that I don't know many other clustering algorithms that are so agnostic to the type of data that they're given. It's one of the most powerful clustering algorithms out there actually, that nobody uses because there's no good Python implementation. If I could fantasize of one thing that could happen in the future, it'd be a good random forest implementation in Python would be nice.
- Jon Krohn: 01:13:51 Yeah. There's a cool open source project for somebody out there to implement. That sounds awesome. So just check out how they're doing it in R and then figure out how to write it in Python.
- Josh Starmer: 01:14:05 The guy who invented it, he's got the algorithm all on his website one step at a time. He's actually written a very good documentation of how it works and how it can be used, and you can watch the StatQuest, and then bam, do it.

- Jon Krohn: 01:14:19 Sounds amazing. All right, so here's a question for you. We have heard a lot about your StatQuest journey. We got a glimpse of how it started as StatChat back when you were an academic. So you were an academic for quite a while. You did, it looks like two bachelor's degrees, one in computer science, one in music, and then you did a PhD in computational biology or biomathematics, and then for 13 years after that, you were either a postdoc or then an assistant professor at an imminent research institution, University of North Carolina, Chapel Hill. And so why change? You're making StatQuest videos, you probably could have stayed an assistant professor and keep making StatQuest videos. But now in the last couple of years, you've gone off to do something else.
- Josh Starmer: 01:15:31 I had the very good, good, good fortune outside of graduate school to get a postdoc in an amazing lab with amazing people, with an amazing lab head. We call them Principal Investigator, PI. Terry Magnuson, he was just a...
- Jon Krohn: 01:15:47 Oh yeah.
- Josh Starmer: 01:15:48 ... phenomenal mentor for me, both as a science guy, because he did amazing stuff. But also it's just like on a personal level, he has been incredibly influential, and he is actually the reason why I left. I did my postdoc there and he's like, "Josh, if you want to stick around, we'd love to have you. In fact, blah, blah, blah, give you a raise, and blah blah," and all this nice stuff. So I stuck around and I loved working in the lab. What I really loved is helping people who needed help with statistics. I loved that some of the research was super basic research, because it's super fundamental. It's like how does the DNA even work? But some of the research was very what they call translational, so there was stuff being done in the clinic as well. I could actually see the results of the research leading to helping people actually, not just in the future of

like here's some interesting knowledge, but actually these people are living healthier, happier lives because of some of the research that I participated in.

- Josh Starmer: 01:16:57 That was super cool, and the people themselves were super cool to work with, but I'll be honest, my heart was not into the actual research myself. I loved helping people with their research, and I would, when a friend, like my friend Dom, when he's writing a paper and I'm on there doing the statistics and he submits it to a journal and the reviews come and they crush it, I will go into an arena and fight like a gladiator for that research. That is something, I mean, because I believe in this stuff. I hope that makes doesn't sound like I'm biasing myself, but I've seen the results. I've seen how it works. I've been there first person, and I know that if the reviewers don't like it, it's because we're not communicating clearly what's going on. Obviously, clear communication is very important to me, and so I'm not going to just say, "Hey, that reviewer is wrong." I'm going to take what they're saying as a failure on our behalf for clearly communicating things. But I will fight and fight and fight until we've got that. Until we've got what they need.
- Josh Starmer: 01:18:18 But when it comes to my own research, I came up with a clever way of when people do PCA, they just do it once and they've got no error measurements. And so I was like why don't we apply bootstrapping to PCA, to get an error measurement on how that's going to work, instead of just taking it at face value? I was doing stuff like that and I tried to publish it and they were like, "No, we're not interested in this." And I was like, "Okay. Well, if you're not interested, then I'll do something else." I don't know if it's imposter syndrome or what, but I just never really felt good if it was my own stuff. But when it was other people's stuff, I was like, "Oh, we need to work very hard on this."

- Josh Starmer: 01:18:58 I just like helping people, and I'll be honest, the lab used to be very, very large and I could kind of get away, I could pay my salary, or not pay my salary, but earn my salary by helping everybody in the lab do what they need to do. That was a full time job, getting grad students through, getting postdocs through, getting research assistant professors doing what they needed to do. It was a big crew and I had a lot to do, but the lab did sort of downsize and it became more critical for me to start doing my own research. I told Terry, I said, "Terry, I'm going to be honest, my heart just isn't into my own research. My heart is in helping other people's research." And I said, "And it's also in StatQuest." And he goes, "Well..." I mean, he actually said a couple of things to me. One was he said his postdoc advisor gave him this advice, which is whatever you're doing, you have to believe it's the most thing in the world. And when you're doing research, things fail all the time, and the way you get through those failures and you keep getting out of bed is you firmly believe that what you're doing is the most important thing in the world. Well, I modified that and I say what I believe I'm doing is the most important thing for me to be doing. I'm maybe a little humbler, is the most important thing for me to be doing.
- Josh Starmer: 01:20:32 Me doing my own research is not that important. There's tons of researchers that are probably better at it than I am, but I'm actually pretty good at StatQuest videos, and it occurred to me that maybe that's most important thing that I could be doing, and anything that gets in the way of that is actually getting in the way of me doing what I do best. Terry eventually said, "I think it's time for you to go," and he was right. Two weeks later I left, and I cried, I totally cried.
- Jon Krohn: 01:21:10 That's all right.

- Josh Starmer: 01:21:11 It was such a big change. I'll be honest, I thought my whole life I was going to be a professor somewhere.
- Jon Krohn: 01:21:17 Right.
- Josh Starmer: 01:21:17 And I know everyone's like, "You're big on YouTube, that's a dream come true." But for me, it wasn't a dream come true, YouTube was always a side project. But once I got into the groove and now I know that this is the most important thing for me to be doing, I feel like I'm right where I need to be. I'm glad it all worked out the way it did, and I'm so happy that Terry, he let it happen. He let me develop StatQuest. He let me let it get big, and he told me to go when the time was right. He was like, "Now you can do this." He was supportive every step of the way, and he continues to be very supportive.
- Josh Starmer: 01:22:03 I do these drop in StatQuests every now and then for the lab, because he requests it. He's like, "Hey, why don't you come back and give us some StatQuest?" And so what I'll do is I'll prototype, I'll take a StatQuest that is 90% done and I'll run it by the lab, and those guys are great guys. They're not mathematicians, and they're not in a machine learning and they're not in the statistics, and so I can just watch. I can look at their faces and see how good a job I'm doing, and as soon as they go...
- Jon Krohn: 01:22:33 Nice.
- Josh Starmer: 01:22:34 They've got this look on their face. I go, "Okay, I need to work on that part and make it better." It's awesome. So that's a super helpful thing that they continue to provide for me. They've been so supportive all through the years, but that's what happened, is it was sort of a traumatic moment at the time when I realized that my lifelong dream was over, only to realize that the reality was better than the dream. Right?



- Jon Krohn: 01:23:01 Right.
- Josh Starmer: 01:23:02 I mean, I'd never even dreamed of being on YouTube, and yet it's better than what my dream was to begin with.
- Jon Krohn: 01:23:08 So cool. I love that story. I love that part of your quest. Yeah, and really well told. That's a really great message for anybody listening, probably even to me, that you should feel that whatever you're doing has to be the most important thing in the world. I mean, that doesn't necessarily mean that it is actually the number one most important thing that anybody on the planning could be doing, that would be tough. If you can figure out what that is, you should do that. But yeah, just for you, for your particular skillset and interests and the impact you can make, that's somewhere in the center of that three circle Venn diagram of your skills, your interests, and what you can get paid to do. That's a really nice spot to be. Super cool. So, we've got some audience questions.
- Jon Krohn: 01:24:03 I mean, I'm not out of questions because I still have... Listeners know that there's a couple of questions that I always ask at the end, but we do have some questions from the audience. So Serg Masis who was actually a recent guest on the show, he's an expert in interpretable machine learning and he wrote a book on it. He's at episode number 539 of the show, for people who want to listen to a great episode. So he wrote a question to you on LinkedIn. He said, human genetics are 99.9% similar. I think it's even much more than that. We're super, super similar to each other genetically. I think we're like 99.9% similar to a Chimp. He says I have no idea how similar music is, but in the realm of all sounds perceivable to the human ear, probably extremely similar. My question for Josh is given so many high profile copy write lawsuits, can music be identified by markers that demonstrate lineage or similarity much like it's done with genetics? So

cool question, because it ties together your genetics experience and your music experience.

Josh Starmer: 01:25:17 I mean to a certain degree, in pop music, you can listen to the beat, and the beat is almost always derivative of a whole tree of things, and the beat is... I don't know. How do you say it? You can say, oh this Justin Bieber song, is like this beat is super popular in the Caribbean. You know, it was origin in the Caribbean and they're the ones that sort of like pioneered it. And you can do that to every little thing. I used to play with this drummer. He was great. I mean, he was like a total like PhD drummer, because every little thing he did, he could cite. He could say, I've got the snare sound from Steve Gadd and on Asia, I've got the kick drum from John Bonham and Led Zeppelin like specific song. Every little thing and not just the sound he was getting from the instrument, but how he was playing it. He could cite everything. And music people can talk for hours and hours and hours about... What do they call it? It's not pedigree. Maybe pedigree is the right word. It's sort of like every little riff, every little jam, every little groove, every little beat has a pedigree of where it came from and where it's been on the way in between the original first time it was invented to when Justin Bieber uses it in his latest hit. Yeah. So yeah, so it all has a story, but yeah, it's all related as well.

Josh Starmer: 01:27:05 Music is an amazing thing, right? Because, just like DNA, you've only got four nucleotides and we get so much variety from it. Interestingly enough, one of the reasons why I got into computational biology was because I'm a musician and I learned about DNA. So, I studied classical music. I grew up playing the cello. And so, I learned about these things called fugue. And a fugue is a thing where each take a melody and you base the rest of the music on that melody with just minor changes to it. So, one thing you can do to that melody is you can take it and then you can overlay another copy of that melody above it, at a

different pitch. And that gives you harmony in the same melody. But another thing you can do is you can stretch it out. You can elongate it, you can make the melody take longer. You can make the melody take shorter periods of time, and you can also reverse it. And so the beat play the melody backwards or play it upside down.

- Jon Krohn: 01:28:15 Like Missy Elliott.
- Josh Starmer: 01:28:16 Yeah. And you can do all these things, but those are also the modifications you can do to DNA. They naturally occur. It's not like JS Bach is playing our chromosomes. But DNA, by accident, can flip, can reverse itself. You can have insertions that stretch it out. You can have deletions that shrink it down. All the modifications that could happen in a fugue, which I'd learned about. Ever since I was a kid, I used to love listening to fugues on the organ and whatnot. There's some real famous ones that you're probably familiar with. Even if you're not familiar with the term fugue, if you heard the melody, you'd be like, oh, I, know that one.
- Jon Krohn: 01:28:59 If you felt like you needed to demonstrate a fugue on say a tenor guitar-
- Josh Starmer: 01:29:07 I wish I could play one. That would be amazing. But it's interesting that like genetics, and you bring up genetics and like, and music because they both have a finite number of things that can happen. There's only so many beats in a bar, right? There's only so many notes in a scale, and yet we get so much variety and infinite amount of variety. And it's funny when you think about pop songs, right? Not only are they using the same structures, they're also using the same lyrics. Right? They're all about love. Or like you left me or I fell in love with you. It's like those two things encompass 99% of all music right there. And, it's amazing how much infinite, how just like every person is different, how we can continue to

make new songs. And we haven't run out of like... Someone is like, oh, this is the last song. We're done. We, got them all. No, it'll never be that way. There will always be new music. And that infinite for a variety is mind blowing. Yeah.

- Jon Krohn: 01:30:19 I sometimes think about that with film names, with movie names, I'm like, are there going to ever run out? Or like, we're just going to have to go back and call another movie, The Game.
- Josh Starmer: 01:30:32 That's true.
- Jon Krohn: 01:30:34 But yeah, it is really crazy. That is also reminds me of a stat and it's just math, right? How many letters in the alphabet? Is it 26, 24?
- Josh Starmer: 01:30:48 First to ask.
- Jon Krohn: 01:30:49 There's 20 to 30 letters in the alphabet. And so for every additional character in a string it's to the power of that number that I don't know, in terms of what the possibilities are. So when you have only one letter there's 26 possibilities, but then once you have two it's 26 times 26 or whatever the correct number is, that's not 26. And so you very, very quickly with just like five characters, all of a sudden it's like an absurd number of possibilities. And yeah, music is the same, DNA is the same. It's crazy.
- Josh Starmer: 01:31:34 I do have a funny, maybe digression about the number of letters and alphabet, but maybe it applies here. When I was right out of college, I interviewed at Microsoft, and they had me do something that involved the number of letters in the alphabet. And I couldn't remember what it was and I can't remember it now either. But I remember just like the interviewer just started laughing at me. Because I had all the math and it was off by like two digits because I couldn't remember or something. And I

was going A, B, C, D, E, counting on my fingers, trying to figure it out. And he was just like, this is ridiculous. How are you interviewing here?

- Jon Krohn: 01:32:12 And you're like, "I can only count to G, I'm a musician."
- Josh Starmer: 01:32:16 I'd give up.
- Jon Krohn: 01:32:19 I just looked it up. It's 26. But I bet it's super easy to think. 24. I don't know why, for me.
- Josh Starmer: 01:32:24 I mean, you got sometimes why, right? I mean, how do you count that?
- Jon Krohn: 01:32:29 Yeah, exactly. How many vowels are there? It's undefined. Awesome. What a great answer to that question. All right, so here's one from Nicolay Kurbatov, who is an AI product manager. And he often has great questions for guests. So here's another one. He's got three, but I'm going to pick the one of the three that's my favorite. And because this is the one that has frequently boggled my mind. And we might end up going off on some interesting philosophical quests here. So, why do all of the data on a big scale look Gosian almost all the time? I mean, is there a hidden law there or is this a kind of a mathematical constant like pie? I know that the question could look stupid, but I'm really wondering about this thing. Also I'm aware of the central limit theorem but it's not enough for me to understand this phenomenon.
- Josh Starmer: 01:33:35 But it's all real rooted in the central limit theorem. So the central limit theorem, the simplest version of it is that means that the mean of something is going to be normally distributed. But, also what that means the mean is just a summation divided by a constant, right? So if we ignore that constant, we can say sums are also normally distributed. So the sum, and that's why everything is normally distributed. Like the height of people. Because

we, I mean, in a metaphorical weird sense. You could say, oh, we're the sum of a lot of things. But like, anything that your cells are doing is the sum of a lot of processes. And the effects as a result mean that the output is going to be normally distributed, and that just cascades to all kinds of things. So, when we grow and you measure everybody in North America, you're going to see a nice normal curve, because everybody is basically the sum of tons and tons of things. And the outputs of all those summations are going to be normally distributed. And so that's where it all comes from. Is just the fact that every process is the product of... If I say the product of a sum, it's confusing, because that's math terminology. A product is multiplication. The output of everything is a sum that make sense.

- Jon Krohn: 01:35:08 Yeah, yeah, yeah, totally.
- Josh Starmer: 01:35:10 Everything to make something happen. Underneath it, everything cellular is a summation, basically.
- Jon Krohn: 01:35:17 Something that I think about a lot, related to this, which I think also is a way to show yourself that you can't possibly have free will. Which is that if you think about almost any thought that you have, for thoughts that you could put on a scale, how do you feel about this? So something like gay marriage, if you asked people a hundred years ago or 200 years ago should there be gay marriage? It's like an extremely small percentage would've said, yeah, I'm fine with that. Or definitely I highly support that. But then if you look at a chart of say people in the US over the last few decades, there's this dramatic swing. And we see it with anything. Like legalizing marijuana. It's like, we're a sum of all these influences. Right? So you have what you see on TV and what the conversations that you have at work and the conversations that you have at home. And so it's the sum of all these interactions that you're like, all right. I guess

people can smoke weed if they want to, or I guess anybody can get married if they want to.

- Josh Starmer: 01:36:42 That's regression too, I mean, isn't it.
- Jon Krohn: 01:36:45 Well, oh no. You're just being silly with language, because that's something completely different. Well, oh no, I see exactly what you're saying. So I often think about regression to the mean with like my own behavior, which is what I was thinking about there, with like if you do something out of character, you're likely to come back towards kind of mean, but yeah, you're right at a population level.
- Josh Starmer: 01:37:18 Yeah. That's what I was thinking. You start off with things on an extreme and then they just percolate to the median to where it's no longer thought of as extreme. It's more just run of the mill at that point. It's the average belief at some point.
- Jon Krohn: 01:37:39 Yeah. And the mean of that distribution is changing over time though. So like if you pull everyone in the US in the fifties and then the sixties and in the seventies, they're like their opinion on how they feel about gay marriage or legalized marijuana or whatever is shifting. And anyway, so yeah, I just find it interesting that it's like everything that I think and everything that I do is the product of some other interaction, fundamentally my genetics, and then every single interaction that the environment has had with my genetics, since. Which also actually people want to hear more about gene by environment interactions, episode 547, which came out about a month ago with Professor Jonathan Flint, who specializes in that kind of work at UCLA. It's pretty interesting. But anyway, so all right. Well, yeah. Great answer. And I got to say that a thing that I've been thinking about for a while with how we're just... No free will. We do a whole episode on that. All right. So, and then one last question here for

you. It was asked in a few different ways from different people like Jonas Christensen, who is an author, and Phil Fry, who is a Chief Data Officer. So they're all kind of asking the same question of like, what are the key stats fundamentals? So Jonas kind of asked it that way, what are the main fundamentals? But I actually really like the way Phil asks it, is what are the essentials of stats literacy that business people need to know?

Josh Starmer: 01:39:28 For me, the most fundamental concept in statistics that everyone needs to understand is variation. It's funny, because I remember for me it was like a light bulb moment. When I was a kid, I never thought about variation. I never thought that like things were different. My world was very small. I grew up in a small town, had my friends and I just never thought about variation. I never looked at flowers and thought all those flowers are different sizes and are all like slightly different shades of yellow. And I never thought about variation, but when I first started learning statistics, I realized that everything varies. Every single thing in the whole world varies. And statistics is a way to cope with that variation and to quantify it and to understand it and to take advantage of it. And I feel like it is the core. I mean, is the core of statistics in my opinion, is if you can understand that everything is always different. And that's kind of like a weird Zen thing, right? The same river never flows under the bridge twice. That kind of thing. No, but everything is always different. There's nothing consistent anywhere. And, especially in business, you've got manufacturing processes, you've got customer service stuff. It's all variable, right?

Josh Starmer: 01:40:57 And statistics gives us a way of understanding that, quantifying it, coping with it and taking advantage of it. And so I think that's a key thing. The other thing for me, conceptually, was visualizing residuals in many different contexts. Once I was able to understand how a T test can

be thought of in terms of residuals, how ANOVA can be thought of in terms of residuals, how regression can be thought of in terms of residuals. Once I understood that concept of like, it's all about the residuals, which are always different by the way. So there's a lot of variation there. But, once you can think about residuals, you can visualize all of statistics. You can just see it in your head. And so those are the two, main ideas that I think are the most valuable for anyone to know about statistics, and especially in business or whatever, a place where you don't really want to be a statistician. You just want to be someone who doesn't freak out when someone says statistics. If you can visualize residuals and understand what that means. And, you understand that there's always going to be variation and you need to quantify and understand what is that variation mean? How is it going to affect your business? Yeah, those are my two things.

- | | | |
|---------------|----------|--|
| Jon Krohn: | 01:42:18 | Love it. Great answer. And I agree. Those are awesome concepts for everyone to know. All right. So I think that wraps up the audience questions. So at the end of every episode, Josh, I ask for a book recommendation. I understand you have a great one. |
| Josh Starmer: | 01:42:33 | I have a fantastic book recommendation. I highly recommend the StatQuest Illustrated Guide to Machine Learning. |
| Jon Krohn: | 01:42:39 | Oh, come on. |
| Josh Starmer: | 01:42:40 | Available this spring. I love it. Anyways, obviously it's not out yet, so I can't really recommend it. But I will say, I'm a huge Neal Stephenson fan. I love reading Neal Stephenson. I don't know if you read Neal Stephenson. He's a science fiction. |
| Jon Krohn: | 01:43:00 | I've had it recommended to me many times. |

- Josh Starmer: 01:43:01 Yeah, exactly.
- Jon Krohn: 01:43:04 It's like, ah, geez. One of his big ones. It's like Quantopia or something.
- Josh Starmer: 01:43:09 Cryptonomicon
- Jon Krohn: 01:43:10 Cryptonomicon. Yeah.
- Josh Starmer: 01:43:12 Yeah. He's not for everyone. I will say that, he's not for everybody. I've talked to a lot of people that are like, yeah, I tried it. Didn't get into it. But I love it. I'll be honest. What I love about his books it's like, they're all huge. All his books are huge. I don't really like that. But what he does in one book is he writes three. He writes one book that has an exciting plot line. He writes one book where he's world building, he's creating exquisite detail about a world that we've never been to before. It's amazing. The things are so real in his books. You could touch them. You could probably stub your toe on it in the dark. And he's just thought about things in such a deep way. And he's describes it so clearly. So I love the world building, like he invents an entire universe just to write a book. And the last thing he does is, it's a philosophy book. He spends a lot of time sort of asking philosophical questions and trying to answer them. He doesn't shy away from stuff like that. And it's for those three reasons.
- Josh Starmer: 01:44:21 I mean, a lot of people complain when they're reading Neal Stephenson, that they love the action adventure part, but then they get bogged down during the philosophy part and they lose interest, or they get bogged down in the world building, and they're like, it's too many details or whatever. To me, I love all three of those things because I love sort of imagining fantasizing about being somewhere different. And what would it be like, especially if you're in the future. It's fun to really think about... I don't know. I've always been interested in the future and

the way he paints the picture of the future is so vivid that I love that. And, I love the philosophy because it makes me think. And I love thinking. So, he's got a new book that came out like a month ago or no, excuse me. It came out in November called Termination Shock. I'm reading that right now. Of the ones that I've read, that's one of my favorites. And I mispronounced the name of it. It's called an Anathem or Anathem. I don't know how it's pronounced, but I love it. I've read it twice. I've actually read everything he's written twice. So, I endorse Neal Stephenson.

- Jon Krohn: 01:45:28 That sounds like an epic recommendation. I have been looking for some fiction to dig into recently. And now this has got to go top of the list. Yeah. Thank you for that great explanation of why his books are so good. So yeah, that's it. This has been an epic epic episode. And so the only thing left to ask you is how people should follow you. So it seems like you're doing something on YouTube or something and they should maybe check that out.
- Josh Starmer: 01:46:05 I do. So, I do YouTube videos. And if that's your thing, you should definitely check it out. But if it's not your thing, and it's not everybody's thing, I'm on Twitter and LinkedIn. And I also do like, I'll put, I don't know, I call them flashcards. There's still value in following me, even if you hate videos and you hate silly songs, there's still some value in following me on these other platforms, because there's cool stuff coming out that is unique and not just rehashing that's on the video.
- Jon Krohn: 01:46:39 Nice. I love it. Well, we're going to have to have you back on the program sometime because we didn't earn a triple bam today.
- Josh Starmer: 01:46:48 Oh no, triple bam.



- Jon Krohn: 01:46:51 We can't just squeeze it in at the end like that. So we'll have to have you on the show again, sometime. Josh it's been so amazing having you on, so much fun. I've learned so much had so many laughs yeah. Can't wait to do it again.
- Josh Starmer: 01:47:06 Looking forward to it. Looking forward. Thank you very much for having me. It's been a real honor and it's been a great time. Had a great conversation.
- Jon Krohn: 01:47:17 Wow. Wow. Wow. I was excited to meet Josh, a personal icon of mine, and he did not disappoint. He's clearly deeply intelligent on a broad range of topics, unbelievably creative and a model communicator. In today's episode, Josh filled us in on his reading, to math, to simple implementation approach to learning something. How, when coming up with a way to teach a concept, he focuses on what would've been helpful to him when he first learned it, how he uses R for quickly prototyping while he might use Python for more involved work, how you can cluster any types of data using the R random forest package. And that in his opinion, the two stats concepts that everyone should know are variants and residuals.
- Jon Krohn: 01:47:58 As always, you can get all the show notes, including the transcript for this episode, the video recording, any materials mentioned on the show, the URLs for Josh's YouTube, LinkedIn, and Twitter accounts, as well as my own social media profiles at superdatascience.com/553, that's superdatascience.com/553. If you'd like to ask questions of future guests of the show like several audience members did of Dr. Starmer, during today's episode, then consider following me on LinkedIn or Twitter, is that's where I post who upcoming guests are and ask you for your thoughtful inquiries. All right. Thank you to Ivana, Mario, Jaime, JP and Kirill on the SuperDataScience team for managing and producing



another extraordinary episode for us today. Keep on rocking it out there, folks. And I'm looking forward to enjoying another round of the SuperDataScience podcast with you very soon.