

**SDS PODCAST
EPISODE 764
FIVE-MINUTE
FRIDAY:
THE TOP 10
EPISODES
OF 2023**



(00:02):

This is Five-Minute Friday on the most popular episodes of 2023.

(00:19):

Welcome back to The Super Data Science Podcast. I'm your host, Jon Krohn. First of all, we've had quite a few reviews since last week. One of my favorites came from Sean Harrison, the Chief Analytics Officer at Acentra Health, who said, "I've been a devoted listener of the Super Data Science podcast, valuing the depth of knowledge and insight your episodes consistently provide. Your show has been an indispensable resource, keeping me updated on the latest trends and developments in AI and ML. Episode 754 with Jason Warner... was one of the best ever!" Thanks for that review Sean and yeah, that is a pretty mind -blowing conversation with Jason Warner. We've got actually lots of historical episodes to talk about today coming up soon.

(01:03):

But, before we get to that, another cool review recently came from Anna Foard, who was our guest way back in Episode #225. Anna, who's now a Data Solutions Architect at a firm called Analytic Vizion said, "I was on the SDS Podcast in early 2019. Since then I've relied on this podcast to learn, now leaning into LLMs as I lead an analytics team. Thank you for inspiring me!" Thank you for inspiring all of us with your great episode, and I'm so glad that you're finding the show useful for learning about LLMs, Anna. We certainly have had a lot of content on LLMs in the past year, so heavenly.

(01:41):

All right, so thanks to everyone for all the recent five-star ratings on Apple Podcasts, Spotify and all the other podcasting platforms out there, as well as for the likes and comments on our YouTube videos. I appreciate the



support and feedback, so much, and the feedback I catch, I'll read on air like this!

(01:58):

The feedback no doubt helps new folks find the show and helps it grow. In 2023, that led to a new record of 4 million combined podcast downloads and YouTube views of the Super Data Science podcast. That's up from 3.3 million a year earlier, so pretty darn cool; thank you for your support listening, rating, sharing, liking, commenting on episodes and so on, and getting us to that 4 million downloads and YouTube views mark. Super cool. Most episodes I'm recording alone in my home studio, which feels quite lonely and so it's surreal that the all-alone home-studio-recording experience impacts such a large community of humans out there.

(02:42):

Speaking of 2023, the time has come for our annual "best-of" episode, where I provide you with a data-backed set of outstanding episodes from the prior year that you might want to go back and check out if you're hankering for more content. This episode also ensures that you didn't miss any of the most popular episodes from last year that sound interesting to you.

(03:03):

One thing you might be wondering is why I'm airing a best-of-2023 episode in March instead of in, say, January. Well, there are two factors. First, internally at the Super Data Science Podcast we use the 30-day mark after an episode's release as our KPI of how an episode's been received by you, by our audience. For our purposes today, using the 30-day snapshot also ensures that episodes released at the beginning of last year don't have an unfair advantage over episodes released at the end of last year, which haven't had as much time live to garner listens. So all right, if it's just a 30-



day data lag, why wasn't this episode out in February? Well, I typically record episodes several weeks ahead of their release date to give our production team plenty of time to clean them up real nice for you and to leave a few episodes in the pipeline in case I get sick or something. I also had tons of exciting news stories that I just wanted to get out in the last few weeks, so those ended up bumping this episode back a bit.

(04:05):

But, you ready? Explanations out of the way, let's dig into — quantitatively-speaking — the ten top-performing episodes of 2023. The tenth-most popular episode was a “Machine Learning 101” episode led by, not one, but two guests. Those were Kirill Eremenko, the founder and one-time host of this very podcast and Hadelin de Ponteves, the super-popular data science instructor.

(04:32):

The ninth-most popular episode starred the renowned data engineering instructor Andreas Kretz detailing, of course, how best to learn the data engineering skill set.

(04:41):

In eighth place comes the Hugging Face ML engineer Dr. Lewis Tunstall with an introduction to NLP with Transformers, based on his eponymous, bestselling O'Reilly book.

(04:51):

In seventh place is an episode featuring Prof. Tom Davenport, who's the author of more than 20 bestselling books in the data space and who coined the concept of data science being the sexiest job of the 21st century. In his episode, he detailed how A.I. will augment rather than replace human workers and he filled us in on whether he thinks data science is still sexy.



(05:13):

In sixth is another professor and many-time bestselling author; this time, it's Allen Downey. My conversation with Prof. Downey is one of my favorites of all-time; it's a mind-blowing one featuring topics from his latest book, called *Probably Overthinking It*, in which he expounds upon many startling statistical paradoxes that have profound implications for how we view the world. I highly recommend checking out that episode with Prof. Allen Downey.

(05:39):

Moving on, in fifth place is Dr. Ben Goertzel, one of the world's best-known futurists. His episode is another mind-bending one, in which he provides his roadmap for making AGI a reality within a mere seven years.

(05:53):

In fourth place is the extremely popular Chip Huyen, on designing ML systems — the topic of the Stanford course she developed and her subsequent bestselling O'Reilly book.

(06:05):

Our bronze medallist for 2023 is the angel investor and data science consultant Josh Wills on how to blend social impact with financial success in your data science career.

(06:15):

Our silver medallist is Vincent Warmerdam on open-source tools for NLP.

(06:20):

And... drum roll please! Our gold medalist, with over 19,000 listens within just 30 days of publication — a record not only for 2023, but an all-time



record for our show — is the very first episode of 2023, which featured Sadie St. Lawrence's predictions of 2023's data science trends, all of which ended up being spot on, by the way.

(06:46):

All right, so there you go. Ten incredible episodes to consider checking out from last year if you're hankering for some more episodes to dig into. Of course, we've got links to all of the episodes mentioned today for you in the show notes and, while of course tons of our back catalog is incredible, if I were to highlight everything, I wouldn't be highlighting anything at all, so that's it for today.

(07:07):

If you enjoyed today's episode or know someone who might, consider sharing this episode with them, leave a review of the show on your favorite podcasting platform, tag me in a social media post with your thoughts, I'll respond, and something that you might want to tell me about in that kind of form is letting me know what your all-time favorite episode is. And of course, if you haven't already, be sure to subscribe to the show. Most importantly, however, just keep on listening. We've got tons of amazing episodes to come in the future as well, and no doubt the show will continue to grow, thanks to you listening, sharing, yeah, all these things. It's because of you that we have this program and that more and more people are listening to it.

(07:51):

All right, so thank you. Until next time, keep on rockin' it out there and I'm looking forward to enjoying another round of the Super Data Science podcast with you very soon.