

**SDS PODCAST
EPISODE 172
FIVE MINUTE
FRIDAY:
GO BEYOND THE
DATA**



Hey guys, welcome back to Super Data Science podcast, this is FiveMinuteFriday, and today, we're talking about how to go beyond data.

All right. So, first of all, before we get started, a huge shout out to all of our students from Egypt! I mean, we're done there right now, it's an amazing place, very exciting country. I've been the pyramids a couple of times. I'm going for a few scuba dives in the coming week, and very, very excited about Egypt.

One thing I do wanna say, it is so hot in your country! This is crazy, guys. How do you survive here? Yesterday, it was 42 degrees, and I was burning. Right now, it's so hot, I found this one shady place that I could record this podcast in. So, huge shout out if you are from Egypt. Thank you so much for being a student. I know we have at least a couple thousand students from Egypt. You guys have an amazing country, and it just generally rocks.

The other thing is, if you're watching the video, you can see these funky sunglasses that I'm wearing. I wanna do a huge shout out to Tree Tribe. Not affiliated with them in any way, just some cool guys and girls put this company together, and they're running this. It's about creating products that are sustainable and helpful for environment. These sunglasses are actually made out of bamboo, so that is environmentally friendly. Second thing is, they float if you drop them in the water, so they don't get lost. Third thing is, every time you buy a product from them, they plant 10 trees. Amazing. And fourth thing is, they look pretty cool.

All right. So, that is our intro part, and now, let's talk about the topic. I was flying here from ... It's quite a long flight. Several short flights, but you know, two hours each but overall, a long day. On my flight from I think Khartoum to Cairo, I was sitting next to a lady who works, guess where? At World Bank. How cool is that? Because, if you've done ... Of course, you know that we use data from World Bank quite a lot. We use data sets such



as different demographics of different countries, we use the overall geopolitical situation, or socio-economical statuses of different countries and regions. For instance, we might get data on access to internet, and then cross-tabulate that with data on population and see how those two are interlinked. So, we've done quite a few of those things in our courses.

So, this lady works at World Bank, and we got to talking about data sets and how she's involved with them, and what data sets she uses in her own. She mentioned quite an interesting phrase that I actually really liked. It was, "Sometimes you need to go beyond the data." At first, I thought, why would you go beyond the data if the data can tell you everything, right? Everything in our life is data. Right now, this gust of wind is blowing, and some leaves are moving. This is actually data, but we're not capturing it, per se. If we captured it, and stored it, and analyzed it, we could make use of it. But the fact that we don't capture it doesn't mean that it's not there. Basically the point is that everything in the world is actually data, so why would you ever go beyond data. Why would you bother with other things apart from data?

The thing is, if we had all the data in the world, like every single data piece, then yeah, then we could analyze and extract any insight. However, we don't. A lot of the time there are things that are still hidden that we cannot extract, but just look at data. A good example is, even working at World Bank, she was ... How does that look? That's good, yeah?

She was, sometimes she needs to analyze the data, especially forecasts for different countries because she works in the part where they invest into new private projects and helps in growing private companies in Madagascar and Comoros and other countries in the Third World, to grow. And so, sometimes she needs to analyze the forecast for different countries' developments, GDPs, and so on. She compares World Bank data to IMF data. IMF is the International Monetary Fund, and they do similar things



sometimes. They sometimes, also, do forecasts or gather data. Sometimes the date matches up, sometimes it doesn't match up.

So, what do you do when it doesn't match up? I'm sure you've had those situations in your career, or you will have this situations in your career where you have data from two sources, and they're both valid and relevant, but they don't match up for some reason. What do you do?

Well, that's when you need to go beyond the data. You need to go and either talk to the people that came up with these data sets, or that collected them, or that gave them to you, or that made those forecasts, made those predictions. Or, you need to see, what are the different motivators behind them? What are the different drivers? Maybe the analytics that went into those data sets or methodologies are slightly different, and that makes a difference. Maybe the intentions, motivators, are different, and therefore these people wanted different results.

So, you need to understand, lots of different factors could have influenced the data. That's why sometimes, and actually, often, you need to go beyond the data. Even what that's not happening, that might be just because you're analyzing one single data set, and you're not comparing it to anything or you don't have anything to compare it to. If you did, you would see that there was a discrepancy and that you do need to investigate it further. So, it's always a good idea to go beyond the data, and understand how did it originate? How did the people collect it? What thought went into it? What was the purpose of collecting this data? Is that the exact purpose that you're collecting for, or do you need to make adjustments and account for the differences there?

So, think about that, that going beyond data is a skill ... It's like a next-level skill for a data scientist. Building a machine, learning algorithms, crunching numbers, creating insights, that's all great. That's the essential foundation that you need in this profession, but going beyond data is an extra consideration. Like, you know, what's the difference between a three-

Show Notes: <http://www.superdatascience.com/172>



start hotel and a five-star hotel? Or, a four ... What takes a four-star hotel to a five-star hotel? Where they do that extra care, they give you extra care, and they look after you a bit more, and they make sure you're happy about everything. They're very thoughtful, right?

And so, going beyond data sometimes is absolutely necessary when you're just put into those circumstances, but sometimes it's a good habit because you want to be thoughtful. You want to take extra care, and do your job extra well to make sure that everything matches up, and everything adds up, and that you are driving the right insights.

So, there we go. That's going beyond data. I guess, this weekend, what I would encourage you to do is look back on your past week, look back on your past month. See where you didn't go beyond the data, where you just used the data plainly, but you could have had. You could have gone beyond the data, and seen what else there is, what other motivators and drivers behind how this data was created.

On that note, hope you have a fantastic weekend, and these are from treetribe.com. Look forward to seeing you back here next time. Until then, happy analyzing.