

Topic: “How” Should you Start with SAS ?

Case Studies in SAS Integration



Welcome to the Case Study!

Today's Lecture

- If you are new to learning SAS – especially in college courses – you may wonder: What is the future of this tool?
- Other reasonable questions: How does SAS work with open source tools? Is it a competitor – or do they integrate?
- SAS has been used for decades for different things. In 5 or 10 years, what will it be used for?



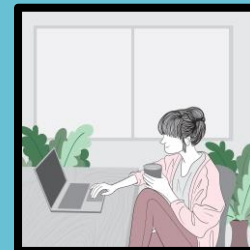
Try all of our career-building data science resources!



Join our group online data science mentoring program:
<https://buff.ly/3Sw4Fsa>



Take our exclusive online data science courses aimed healthcare analysts:
<https://monika-s-school-e295.thinkific.com>



Watch our tutorials on YouTube:
<https://buff.ly/2W6ijuc>

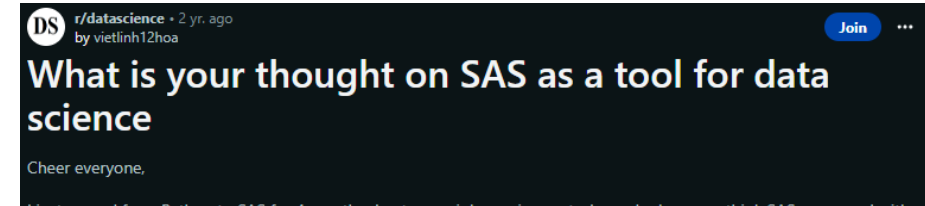
Take our LinkedIn Learning courses: <https://buff.ly/2lhd4Rq>



What if You are “New” to SAS?

- Should you throw yourself into mastering a challenging data science programming language that has been around since the 1970s?
- I learned SAS in the 1990s-early 2000s. At that time, I asked myself the same question.
- It would not have mattered whether I said “yes” or “no” at that time, because SAS was the only game in town.
- But what about now?

Today's Resource



2 years ago:

I just moved from Python to SAS for 4 months due to new job requirements. I wonder how you think SAS compared with other languages, any future.

<https://buff.ly/3rCp3je>



Open-Source Solutions Work Better and Are Cheaper

2 yr. ago

Our department got rid of SAS due to infrastructure issues and high costs. We converted all the SAS workflows into Python and saved a lot of money.

133 Reply Share ...

2 yr. ago

my company is doing the same thing, they are porting everything over to python and some stuff to KNIME

31 Reply Share ...

+ 1 more reply

- Easier to get an ROI with an open source solution
- Even people in SAS shops try open source solutions first (before looking at adding SAS components)



SAS is “Spoiled” – Tries to Charge Customers High Fees for Mediocre Software/Service

Person 1

2 yr. ago

I've been gone from the SAS world for almost 8 years now, and I really do not miss it. Their answer to *everything* was another expensive add-on module for stuff that you can trivially do for free with high-quality and well-supported open-source libraries in R and Python (specifically, I'm thinking of the time when they tried to sell us bare-bones natural language processing support for something like \$60K a year, when NLTK is *right there*).

Person 2

Among other factors (like the license fees), I am sure that the emergence of more remote work isn't going to help SAS' market share, since companies can now more easily hire for or contract out simple tasks at a fraction of the cost that SAS would charge.

Person 3

Speaking from my experience, the main factor for the company I worked for at the time was less "we can't find someone with the skills (locally or otherwise) to do the work" and more "we've already built everything on top of SAS and integrating something from outside of that ecosystem would be *scary*." Even then, we didn't go forward with the text analytics add-ons they offered us; it was too transparently a bad deal for the limited functionality they offered. Seriously, it was... sentiment analysis, maybe basic stemming and topic modeling, if that.

- SAS is always trying to sell us some expensive add-on that doesn't work as well as Python or R solutions.
- Hard to do remote work with SAS – easier to just hire open source consultants on-site anywhere.
- Person 3 summary: “We can’t find someone with the necessary SAS skills” but “we already built everything in SAS” so we were stuck.
- But then, we could not afford a simple sentiment analysis add-on from SAS...
- ...something we do all the time with open source tools efficiently and for free.



“SAS being intertwined with education was a big deal.”

Person 1

I think the thing that will drain the SAS market share more than anything else is that it's competing against tools that are free for anybody to learn and experiment with and extend. They can (and frequently do, I think) make partnerships with schools to get their software taught to students, but all the cutting edge stuff that's happening in both academia and industry tends to be open source and decidedly not SAS-based. Most of the less-traditional educational initiatives out there for data science and analytics lean heavily on open source tools because you can point anybody at them and get set up for free in no time, no matter where that person might be or what platform they're using.

⊖ ↑ 14 ↓ 💬 Reply ↗ Share ...

Person 2

2 yr. ago

SAS being intertwined with education was a big deal. When I went to grad school it seemed every analyst had been trained on SAS in school.

- **SAS has an effective monopoly in college education in some fields (e.g., clinical research, healthcare)**
- **SAS is hard to learn – so SAS training “supplants” other data science, informatics, management, or biostats training**



Banks are Stuck in SAS

2 yr. ago

This is my role at a big bank. Theyre sooooo behind.

6 ↑ ↓ Reply Share ...

2 yr. ago

Banks got comfortable, inertia did the rest. Most banking models don't need more than proc reg anyway due to regulations requiring interpretability. After a while it was just a little feature engineering and dumping the same 15 cars into a proc reg with the stepwise option.

I expect they will eventually come around; but they were stuck on mainframes and vsam files in the 2000s.... They will always be slow to change.

1 ↑ ↓ Reply Share ...

- “Slow to change” – but “how” to change?

- It takes a while to migrate from SAS to any other environment



...and Finally, the SAS Champions – the Realists!

Person 1: SAS works, and though R and Python are better at some things, SAS will still “be around for a while”.

I've used SAS for more than 25 years. It is a tool that can be useful. Many places still use it for data management and modeling on large datasets. Because SAS is a “read a record write a record” system you can process data that doesn't fit into memory easily.

I think some places are more comfortable with the stat procedures in SAS or Stata because they are well established and tested.

I have see R making inroads though. Even big SAS shop have R users now and it is being used quite a bit. Less so with python, but it will happen.

Many people talk shit about SAS, but i think it is still one of the largest privately held companies. They've quite a bit of adapting in the last 10 or 15 years. I don't see how they can fully adapt to R and Python, but they will be around for a while.

Person 2: SAS works, but only big companies can afford to keep paying for it and supporting it (e.g., with a dwindling labor pool).

SAS works but try to sell me on SAS over R or Python for any non-F1000 company and financially it makes no sense. Then logically back that out as to why at a F1000 firm you would incur this huge cost and a small labor pool for said software.

The documentation in SAS is one of the biggest upsides, but I know even utilities that were all SAS for decades have switched to python because the license fees are just so high.

- “Inertia” is the main reason why companies use SAS today (instead of migrating)

SAS is bad

1. I dont understand the difference between complexity and having to memorize a lot of syntax
2. SQL is good. Python and R also have packages to query their data with SQL syntax, if you want to.
3. Not just unique but expensive.
4. This is wishful thinking on the company's part. SAS has many problems, but when SAS has a problem they just pretend it is a new way to do analysis.

↑ 98 ↓ Reply Share ...

“...when SAS has a problem, they just pretend it is a new way to do analysis.

This is wishful thinking on the company's part. SAS has many problems, but when SAS has a problem they just pretend it is a new way to do analysis.

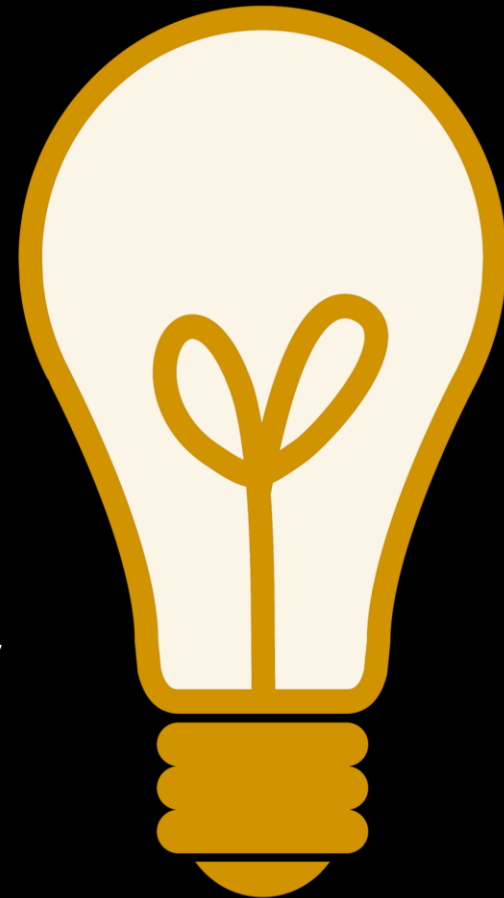
I have a lot of trouble believing that tools like SAS still being used at companies is a product of anything other than “this is the way we do things” or “[insert collection of middle-managers] used SAS, so they want the department to use SAS”.

- You can still get SAS tools to work for you today
- However, it's hard and expensive, and generally better in open source tools



SAS is the “COBOL” of Health Data Science

- My mom was a COBOL programmer. She reminds me regularly of how stuff is still in COBOL and that's why a lot of our public data infrastructure broke in the US during COVID-19.
- SAS is the COBOL of health data science, in that we will always have things in SAS, so someone will always have to know SAS.
- But SAS is probably not “the future” alone, *per se* (although analytics is still tip-top).
- What the future looks more like to me is trying to get data out of SAS servers and into the cloud





Questions?

Facilitator: Monika M. Wahi, MPH, CPH

Email:
dethwench@gmail.com

Blog:
<https://dethwench.com/blog/>

LinkedIn:
<https://www.linkedin.com/in/dethwench/>



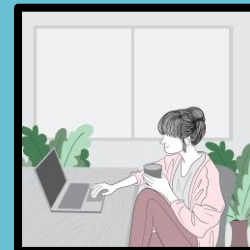
Try all of our career-building data science resources!



Join our group online data science mentoring program:
<https://buff.ly/3Sw4Fsa>



Take our exclusive online data science courses aimed healthcare analysts:
<https://monika-s-school-e295.thinkific.com>



Watch our tutorials on YouTube:
<https://buff.ly/2W6ijuc>

Take our LinkedIn Learning courses: <https://buff.ly/2lhd4Rq>