

**SDS PODCAST  
EPISODE 520  
FIVE MINUTE  
FRIDAY:  
THE HIGHEST-  
PAYING  
PROGRAMMING  
LANGUAGES FOR  
DATA SCIENTISTS**



(00:05):

This is Five-Minute Friday on the Highest-Paying Programming Languages for Data Scientists.

(00:19):

The beloved technical publisher O'Reilly recently released their 2021 Data/AI Salary Survey. It covers responses from 3000 survey respondents in the US combined with another 300 respondents from the UK. All of the respondents are subscribers to O'Reilly's Data & AI Email Newsletter. So, that's where the biases in the results come from.

(00:44):

There's quite a lot of detail in the report on how the salaries of data professionals vary, including by gender, level of education, career stage, industry, and geographic location. For today's episode, I'm focusing on how salaries vary by programming language since that's an attribute that you can easily change about yourself, simply by learning something new.

(01:08):

Unsurprisingly, the most popular programming languages are staple languages for data scientists that we mention all the time on this show: Python is used by 61% of respondents, making it the most popular language overall, SQL is next used by 54%, R is used by a markedly smaller 20%, but that was nevertheless enough for it to land in sixth place for popularity, JavaScript and HTML for designing websites, the Bash command-line language, and the classic software development language Java all fell ahead of R but behind Python and SQL — this could be partly because, again, the bias, the O'Reilly audience consists heavily of software developers.



(01:57):

In terms of salary, Python, SQL and R were all pretty close to the average annual salary across all respondents, which was \$146k. Python was slightly above the average at \$150k while SQL and R were just below it at \$144k. Since all three of these languages are so widespread, they've become the norm for data scientists and so it's unsurprising that they fall in the middle.

(02:25):

In contrast, Julia, another programming language designed specifically for processing data efficiently such as with statistical or machine learning models, is associated with a much higher average salary — about \$170k — relative to the much more popular Python, SQL, and R languages. Since Julia is less widely used, the data scientists who know it may have demand on their side and so they can command a higher salary. Julia is also relatively new, so knowing it may be a signal to prospective employers that you're still learning new tricks and you're staying up to date with recent data science languages.

(03:03):

Only a handful of programming languages were associated with salaries higher than Julia and they were mostly newer functional-programming languages like Rust, Go, Erlang, and Scala. Most of these languages are used by only a few percent of data scientists — they are more common in software developers though even amongst software developers, they're relatively rare. If you're thinking, however, about becoming a software developer or topping up your engineering toolkit, then perhaps consider taking up one of these functional programming languages. Rust in particular commanded the highest salary of all at over \$180k while Go and Scala were just behind at \$179k and 178k, respectively. For data science, Scala might prove particularly helpful to you since its syntax is used in the



Spark distributed computing platform that is popular for working with massive quantities of data.

(03:55):

On the flipside, from a salary perspective at least, a data science language you can probably ignore for now is Matlab as well as its open-source cousin Octave, as these languages came in below the middle of the salary pack. Visual Basic for automating aspects of Microsoft Excel came at \$135k, even lower, right near the very bottom.

(04:18):

Likewise, if you're interested in web applications, JavaScript is where you may want to focus your energy as it was middle of the pack — right around the \$146k average — while other web development languages like PHP, HTML, and CSS all collected near the \$135k bottom.

(04:35):

All right, I hope that was interesting for you and maybe gets your mind tingling about some new languages that you could be diving into next. In next week's episode, we'll go beyond programming languages alone to dig into the popularity and salaries associated with other tools and platforms like PyTorch, TensorFlow, scikit-learn, Spark, and Excel. If you'd like to check out the full report in the meantime, we've included a link in the show notes.

(05:03):

All right, that's it for today's episode. Keep on rockin' it out there folks and catch you on another round of SuperDataScience very soon.