



How can you determine which programming languages and technologies are most widely used? Which languages are gaining or losing popularity, helping you decide where to focus your efforts?

One excellent data source is Stack Overflow, a programming question-and-answer site with more than 16 million questions on programming topics. Each Stack Overflow question is tagged with a label identifying its topic or technology. By counting the number of questions related to each technology, you can estimate the popularity of different programming languages.

In this project, you will use data from the Stack Exchange Data Explorer to examine the relative popularity of R compared to other programming languages.

You'll work with a dataset containing one observation per tag per year, including the number of questions for that tag and the total number of questions that year.

stack\_overflow\_data.csv

Column	Description
year	The year the question was asked (2008-2020)
tag	A word or phrase that describes the topic of the question, such as the programming language
num_questions	The number of questions with a certain tag in that year
year_total	The total number of questions asked in that year

index	...	↑↓	year	...	↑↓	tag	...	↑↓	num_questions	...	↑↓	year_total
1			2008			treeview			69			
2			2008			scheduled-tasks			30			
3			2008			specifications			21			
4			2008			rendering			35			
5			2008			http-post			6			
6			2008			static-assert			1			
Rows: 6												

index	...	↑↓	year	...	↑↓	tag	...	↑↓	num_questions	...	↑↓	year_total	...	↑↓	percentage	...
1					2020	acl			165			5452545				
2					2020	waitforsingleobject			7			5452545				
3					2020	escaping			423			5452545				
4					2020	fillna			81			5452545				
5					2020	eslint-plugin-react			3			5452545				
6					2020	negative-number			34			5452545				
7					2020	aws-ebs			35			5452545				
8					2020	macos-catalina			934			5452545				
9					2020	scala-2.13			9			5452545				
10					2020	comonad			2			5452545				
11					2020	bootstrap-datepicker			100			5452545				
12					2020	rtp			122			5452545				
13					2020	code-metrics			8			5452545				
14					2020	java-ee-7			115			5452545				
15					2020	sslssocketfactory			13			5452545				
16					2020	compiler-construction			365			5452545				
17					2020	onepoint3			20			5452545				

Rows: 20,000 ⚠ truncated from 44,335 rows

index	...	↑↓	tag	...	↑↓	total_questions
1			javascript			
2			python			
3			java			
4			android			
5			c#			

Rows: 5

How likely are you to recommend DataLab to a friend or co-worker?

Not at all likely 0 1 2 3 4 5 6 7 8 9 10 Extremely likely

powered by InMoment