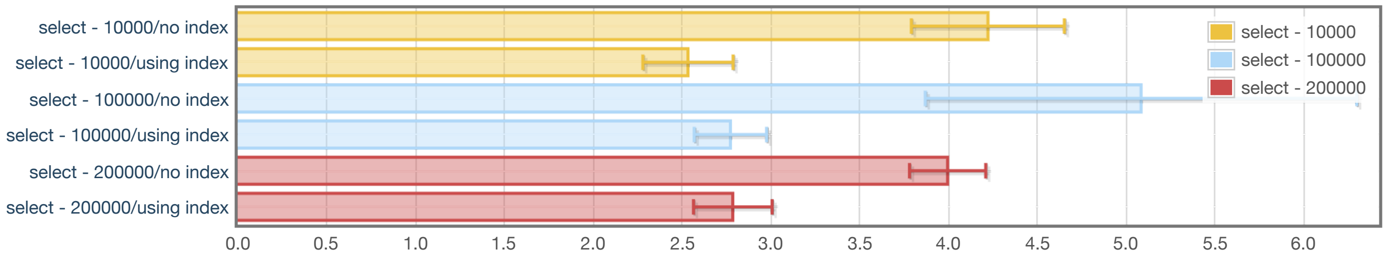
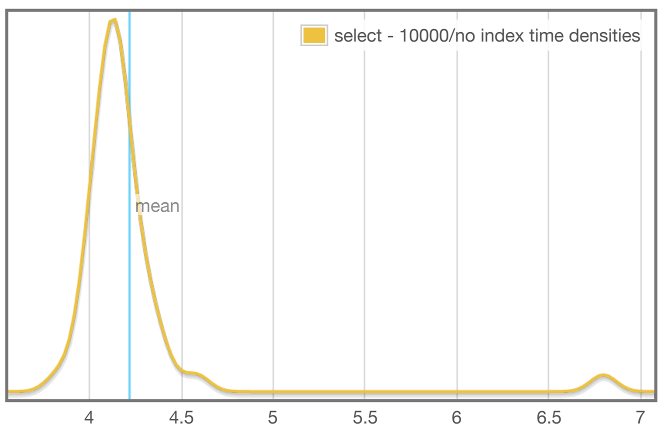
criterion performance measurements

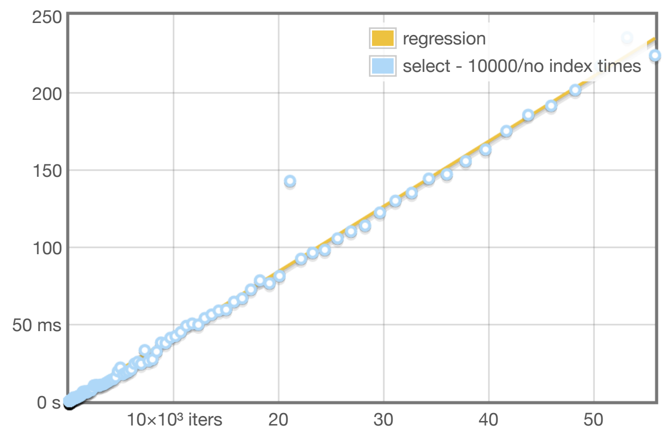
overview

want to understand this report?



select - 10000/no index

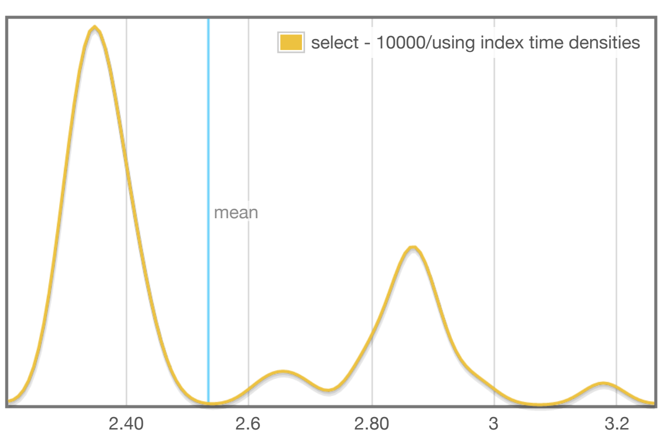


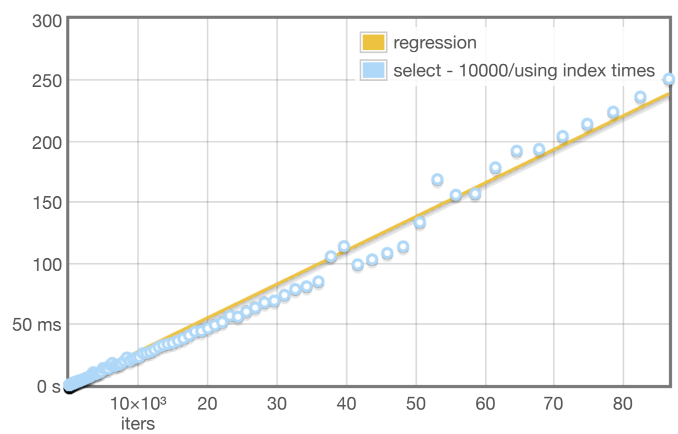


	lower bound	**estimate**	upper bound
OLS regression	4.13 μs	4.21 μs	4.30 μs
R^2 goodness-of-fit	0.980	0.992	0.999
Mean execution time	4.14 μs	4.22 μs	4.49 μs
Standard deviation	115 ns	431 ns	888 ns

Outlying measurements have severe (87.8%) effect on estimated standard deviation.

select - 10000/using index



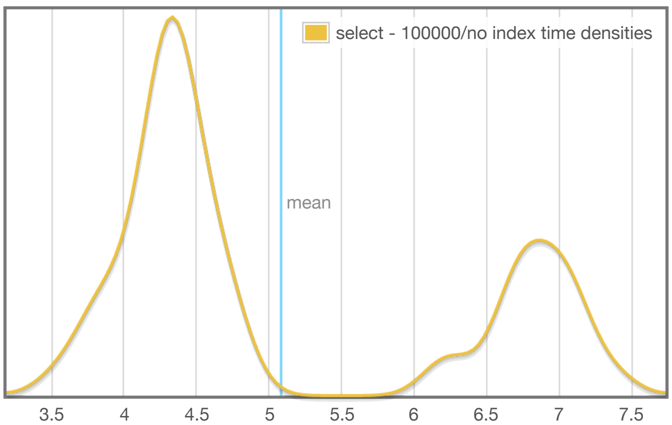


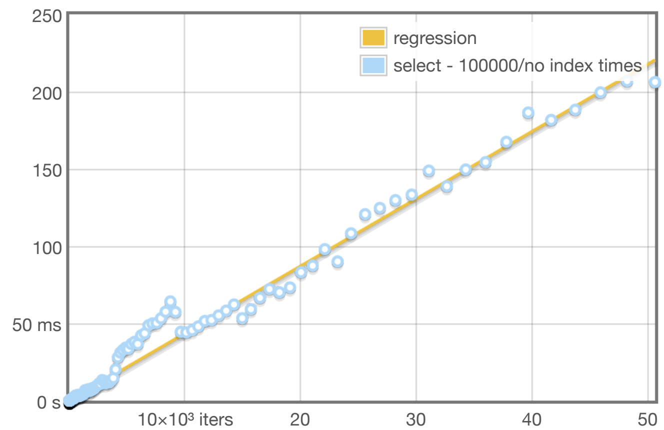
	lower bound	**estimate**	upper bound
OLS regression	2.66 μs	2.76 μs	2.82 μs
R^2 goodness-of-fit	0.988	0.991	0.994

	lower bound	estimate	upper bound
Mean execution time	2.46 μs	2.53 μs	2.61 μs
Standard deviation	207 ns	254 ns	299 ns

Outlying measurements have severe (87.8%) effect on estimated standard deviation.

select - 100000/no index

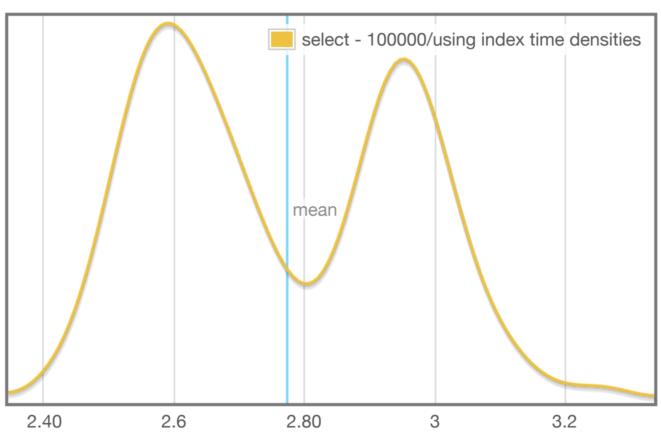


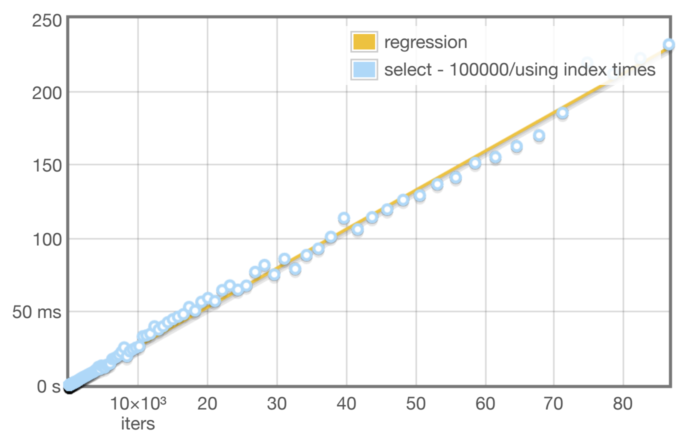


	lower bound	estimate	upper bound
OLS regression	4.27 μs	4.36 μs	4.45 μs
R^2 goodness-of-fit	0.977	0.986	0.992
Mean execution time	4.80 μs	5.08 μs	5.49 μs
Standard deviation	1.04 μs	1.21 μs	1.36 μs

Outlying measurements have severe (97.6%) effect on estimated standard deviation.

select - 100000/using index

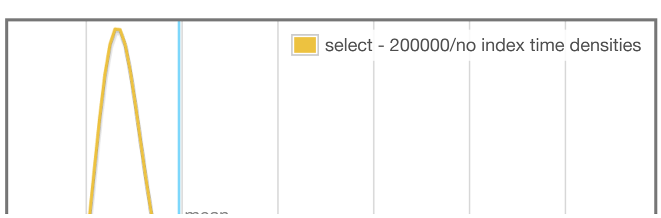


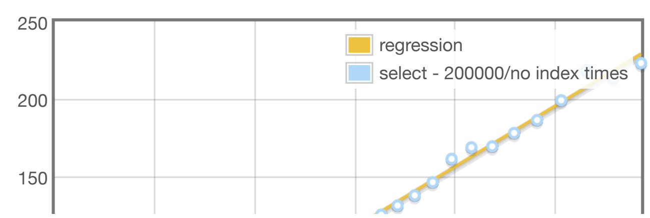


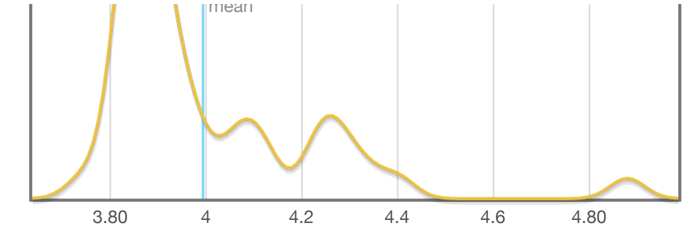
	lower bound	estimate	upper bound
OLS regression	2.61 μs	2.66 μs	2.72 μs
R^2 goodness-of-fit	0.995	0.996	0.998
Mean execution time	2.72 μs	2.77 μs	2.83 μs
Standard deviation	182 ns	204 ns	244 ns

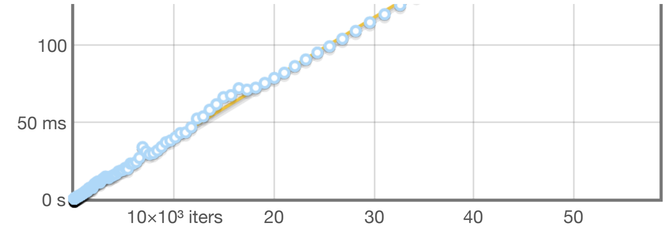
Outlying measurements have severe (79.6%) effect on estimated standard deviation.

select - 200000/no index





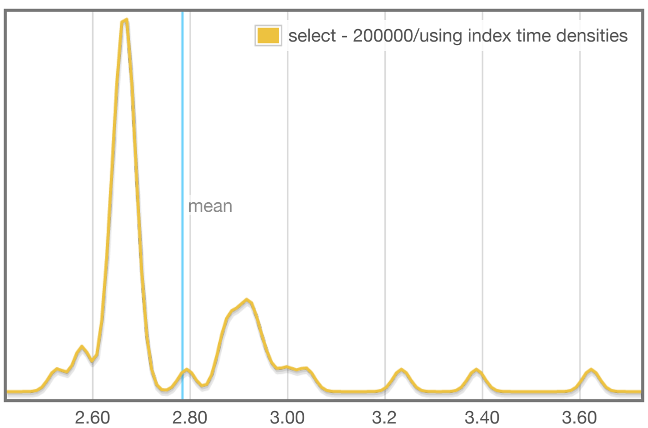


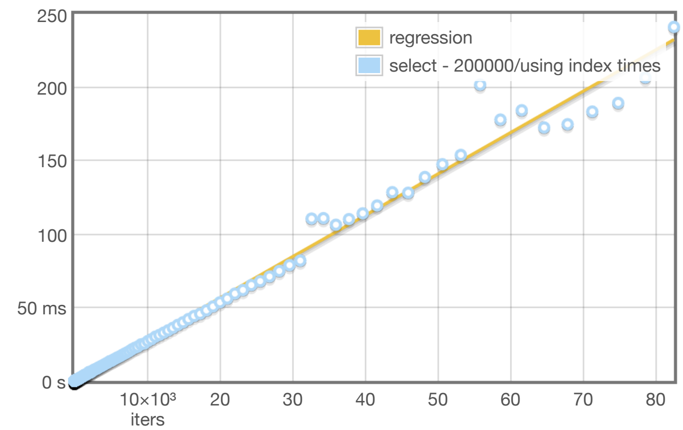


	lower bound	estimate	upper bound
OLS regression	3.87 μs	3.91 μs	3.98 μs
R^2 goodness-of-fit	0.998	0.999	0.999
Mean execution time	3.94 μs	3.99 μs	4.07 μs
Standard deviation	153 ns	214 ns	303 ns

Outlying measurements have severe (66.4%) effect on estimated standard deviation.

select - 200000/using index





	lower bound	estimate	upper bound
OLS regression	2.70 μs	2.82 μs	2.93 μs
R^2 goodness-of-fit	0.984	0.990	0.997
Mean execution time	2.74 μs	2.79 μs	2.87 μs
Standard deviation	160 ns	221 ns	328 ns

Outlying measurements have severe (81.7%) effect on estimated standard deviation.

understanding this report

In this report, each function benchmarked by criterion is assigned a section of its own. The charts in each section are active; if you hover your mouse over data points and annotations, you will see more details.

- The chart on the left is a kernel density estimate (also known as a KDE) of time measurements. This graphs the probability of any given time measurement occurring. A spike indicates that a measurement of a particular time occurred; its height indicates how often that measurement was repeated.
- The chart on the right is the raw data from which the kernel density estimate is built. The x axis indicates the number of loop iterations, while the y axis shows measured execution time for the given number of loop iterations. The line behind the values is the linear regression prediction of execution time for a given number of iterations. Ideally, all measurements will be on (or very near) this line.

Under the charts is a small table. The first two rows are the results of a linear regression run on the measurements displayed in the right-hand chart.

- *OLS regression* indicates the time estimated for a single loop iteration using an ordinary least-squares regression model. This number is more accurate than the *mean* estimate below it, as it more effectively eliminates measurement overhead and other constant factors.
- *R^2 goodness-of-fit* is a measure of how accurately the linear regression model fits the observed measurements. If the measurements are not too noisy, R^2 should lie between 0.99 and 1, indicating an excellent fit. If the number is below 0.99, something is confounding the accuracy of the linear model.
- *Mean execution time* and *standard deviation* are statistics calculated from execution time divided by number of iterations.

We use a statistical technique called the bootstrap to provide confidence intervals on our estimates. The bootstrap-derived upper and lower bounds on estimates let you see how accurate we believe those estimates to be. (Hover the mouse over the table headers to see the confidence levels.)

A noisy benchmarking environment can cause some or many measurements to fall far from the mean. These outlying measurements can have a significant inflationary effect on the estimate of the standard deviation. We calculate and display an estimate of the extent to which the standard deviation has been inflated by outliers.

colophon

This report was created using the criterion benchmark execution and performance analysis tool.

Criterion is developed and maintained by Bryan O'Sullivan.