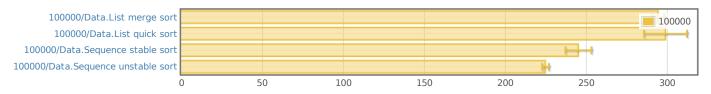
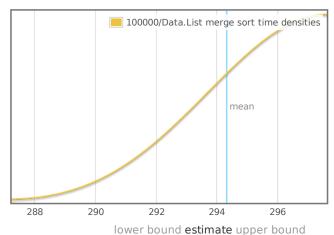
# criterion performance measurements

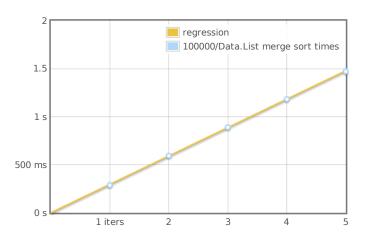
#### overview

want to understand this report?



## 100000/Data.List merge sort

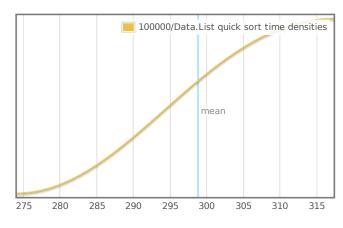


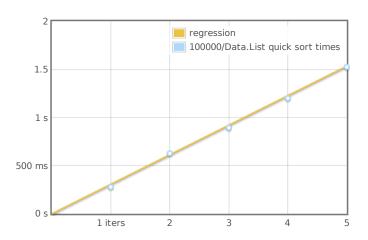


OLS regression 291 ms 296 ms 301 ms  $R^2$  goodness-of-fit 1.000 1.000  $R^2$  ms 294 ms 296 ms Standard deviation 657  $\mu$ s 3.61 ms 4.73 ms

Outlying measurements have moderate (16.0%) effect on estimated standard deviation.

## 100000/Data.List quick sort





lower bound estimate upper bound

 OLS regression
 286 ms
 307 ms
 318 ms

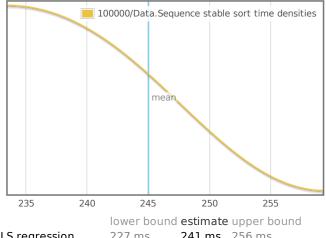
 R² goodness-of-fit
 0.993
 0.999
 1.000

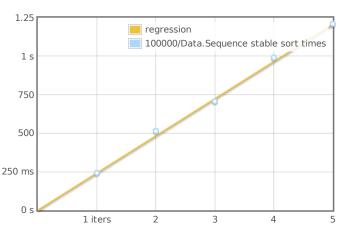
 Mean execution time
 287 ms
 299 ms
 306 ms

 Standard deviation
 6.20 ms
 13.3 ms
 19.7 ms

Outlying measurements have moderate (16.0%) effect on estimated standard deviation.

#### 100000/Data. Sequence stable sort





 OLS regression
 227 ms
 241 ms
 256 ms

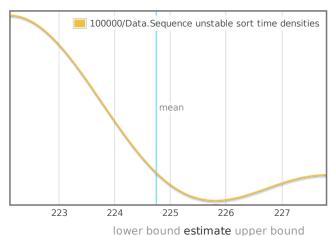
 R² goodness-of-fit
 0.988
 0.997
 1.000

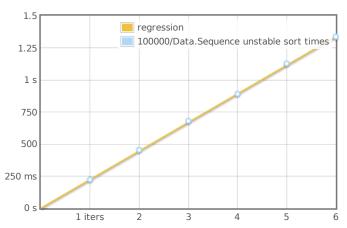
 Mean execution time
 239 ms
 245 ms
 252 ms

 Standard deviation
 4.31 ms
 8.00 ms
 11.7 ms

Outlying measurements have moderate (16.0%) effect on estimated standard deviation.

#### 100000/Data.Sequence unstable sort





 OLS regression
 219 ms
 223 ms
 226 ms

 R² goodness-of-fit
 0.999
 1.000
 1.000

 Mean execution time
 223 ms
 225 ms
 226 ms

 Standard deviation
 1.55 ms
 2.09 ms
 2.53 ms

Outlying measurements have moderate (13.9%) 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*<sup>2</sup> *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<sup>2</sup> 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 <u>criterion</u> benchmark execution and performance analysis tool.

Criterion is developed and maintained by <u>Bryan O'Sullivan</u>.