# Write-up

### Hypothesis

Merge sort will outperform insertion sort for all but very small lists, with the difference in speed becoming particularly apparent for very large lists. As for just how small the list will have to get for insertion sort to outpace merge sort, I'd guess low double digits.

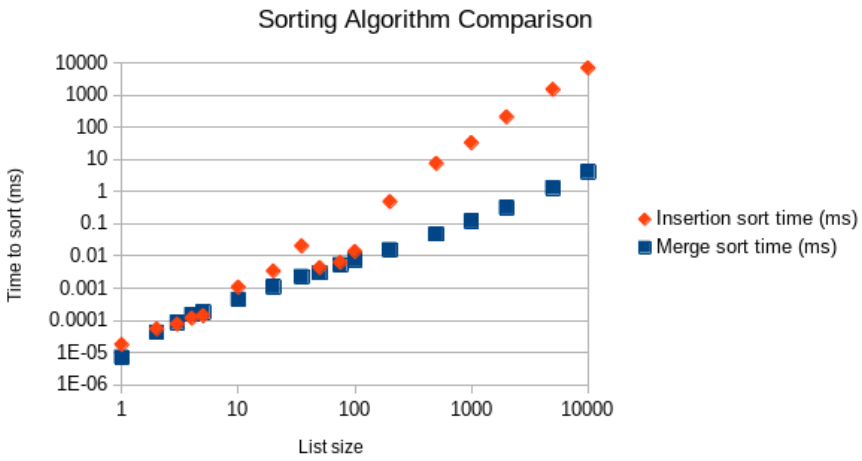
### Methods

The code I used to test these algorithms can be found at https://github.com/khuldraeseth/cse431hw4. Clone the repository, move into `p1`, and `stack run -- --output wherever-you-want-the-results.html` to run these tests yourself. Compiles with `-threaded`.

These are my own implementations of insertion sort and merge sort, written in Haskell. Merge sort does not split the list at the middle—rather, it uses the inverse of an interleave (which I have named `unterleave`) that takes `[a,b,c,d,e] ↦ ([b,d], [a,c,e])` as an example for what I expect may be a moderate speedup. Since merge sort is not stable anyway, nothing is lost here.

### Results

See below for Criterion output. The results are summarized in this plot:




### Discussion

Not quite as I had expected. For starters, merge sort is faster in these tests than insertion sort for very small lists as well as sufficiently large ones! This makes sense in retrospect—I have special cases in my merge sort code for empty and singleton cases, so these are just one pattern match from being sorted, rather than a call to `foldr`. Then insertion sort has its time on top, outperforming merge sort for lists of length 3, 4, and 5. The two cross somewhere between 5 and 10, and repeated tests show no consistent particular value.

Interesting is that insertion sort, after being outpaced by an order of magnitude on lists of length 35, makes a resurgence and competes with merge sort again for lengths 50, 75, and 100 before falling off again at 200 and never having a third chance. I'll attribute this to some compile-time optimization.

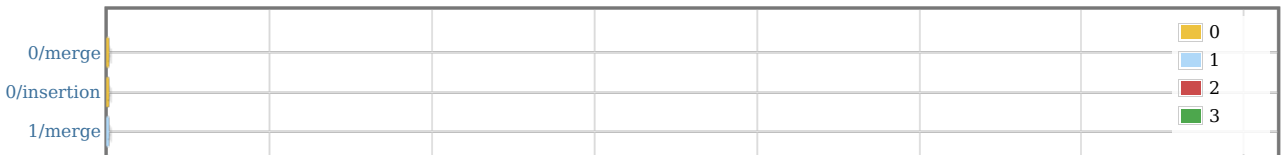
### Conclusion

Under the conditions tested, merge sort was faster than insertion sort on lists of all lengths except those between 3 and 10.

# criterion performance measurements

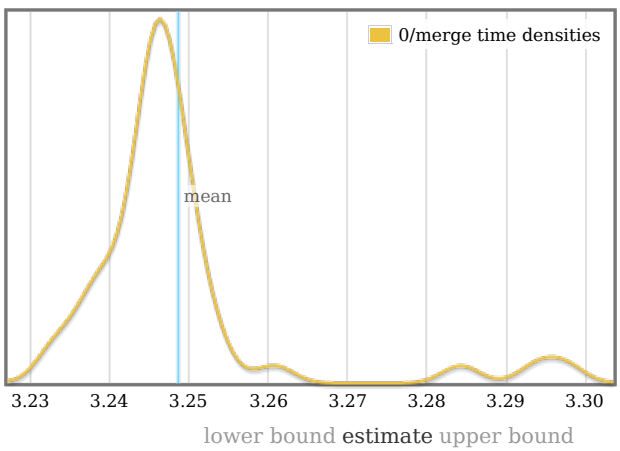
## overview

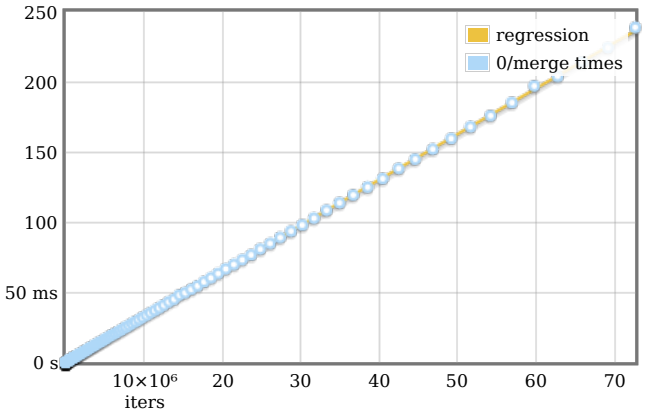
want to understand this report?




4
5
10
20
35
50
75
100
200
500
1000
2000
5000
10000

# 0/merge

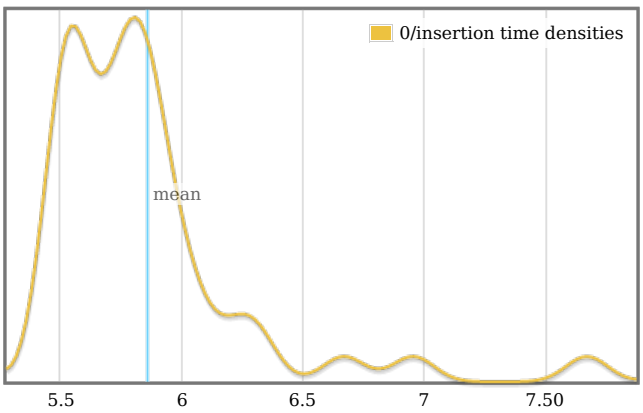

0/merge time densities
mean
3.23
3.24
3.25
3.26
3.27
3.28
3.29
3.30
lower bound estimate upper bound

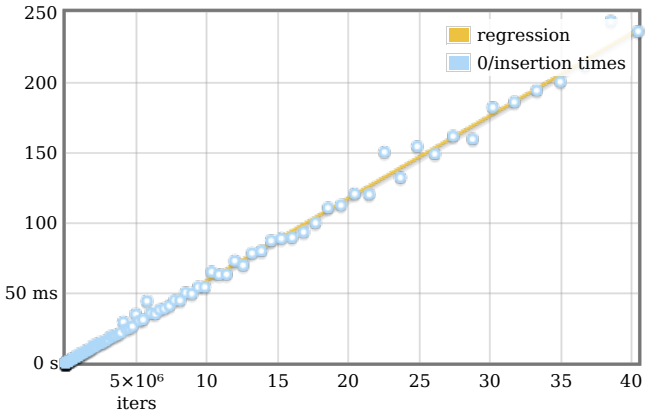

250
200
150
100
50 ms
0 s
regression
0/merge times
10×10⁶
iters
20
30
40
50
60
70

| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 3.24 ns | 3.25 ns | 3.26 ns |
| R² goodness-of-fit | 1.000 | 1.000 | 1.000 |
| Mean execution time | 3.25 ns | 3.25 ns | 3.25 ns |
| Standard deviation | 8.39 ps | 13.0 ps | 19.2 ps |

Outlying measurements have no (0.3%) effect on estimated standard deviation.

# 0/insertion

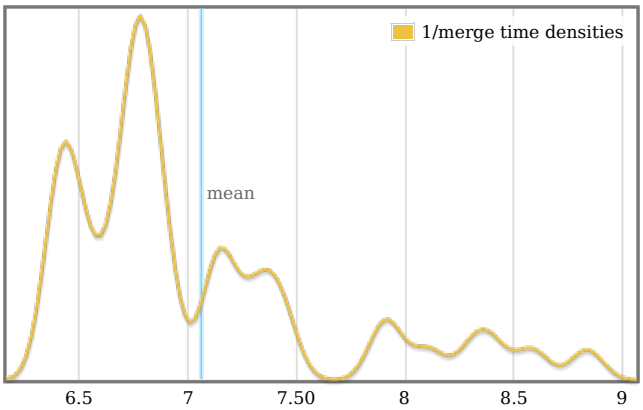


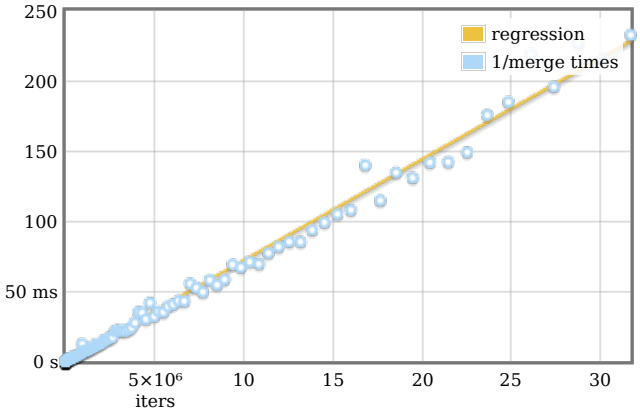




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 5.77 ns | 5.87 ns | 6.00 ns |
| R² goodness-of-fit | 0.996 | 0.998 | 0.999 |
| Mean execution time | 5.76 ns | 5.86 ns | 5.99 ns |
| Standard deviation | 270 ps | 417 ps | 639 ps |

Outlying measurements have severe (85.7%) effect on estimated standard deviation.

# 1/merge

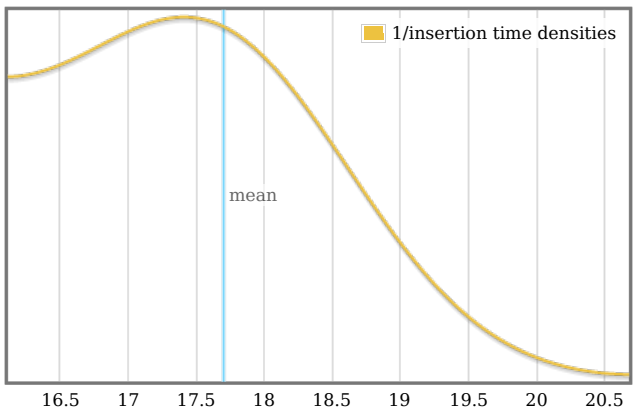


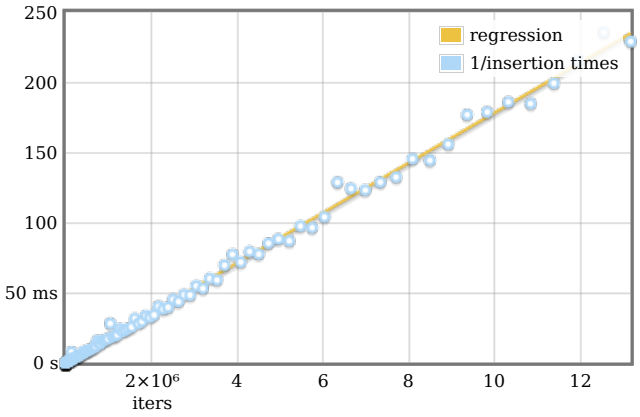




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 6.94 ns | 7.20 ns | 7.44 ns |
| R² goodness-of-fit | 0.990 | 0.993 | 0.997 |
| Mean execution time | 6.92 ns | 7.06 ns | 7.29 ns |
| Standard deviation | 482 ps | 637 ps | 822 ps |

Outlying measurements have severe (90.5%) effect on estimated standard deviation.

# 1/insertion

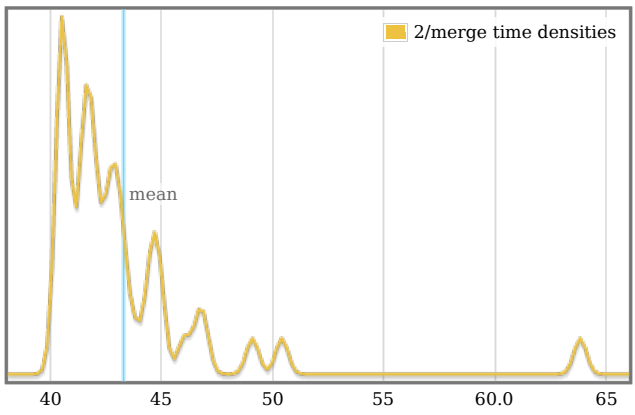


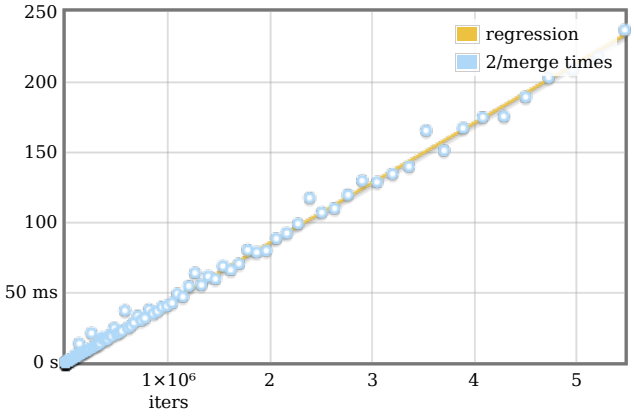




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 17.5 ns | 17.8 ns | 18.1 ns |
| R² goodness-of-fit | 0.996 | 0.998 | 0.999 |
| Mean execution time | 17.4 ns | 17.7 ns | 17.9 ns |
| Standard deviation | 734 ps | 924 ps | 1.19 ns |

Outlying measurements have severe (75.1%) effect on estimated standard deviation.

# 2/merge

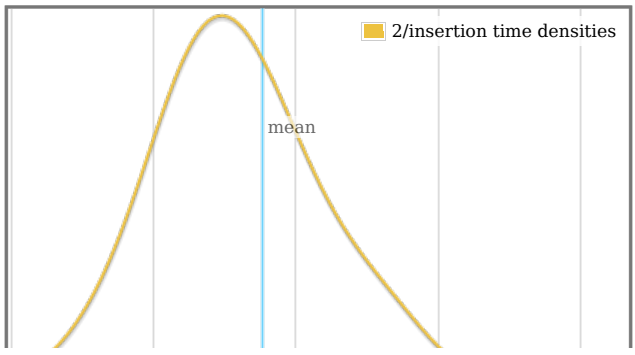


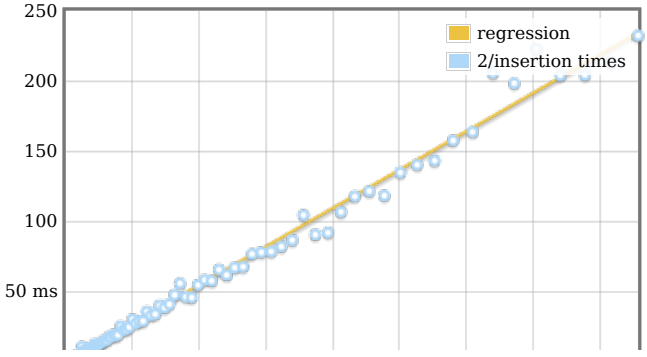




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 42.1 ns | 42.6 ns | 43.3 ns |
| R² goodness-of-fit | 0.996 | 0.998 | 0.999 |
| Mean execution time | 42.5 ns | 43.3 ns | 45.6 ns |
| Standard deviation | 2.16 ns | 3.93 ns | 6.72 ns |

Outlying measurements have severe (89.7%) effect on estimated standard deviation.

# 2/insertion

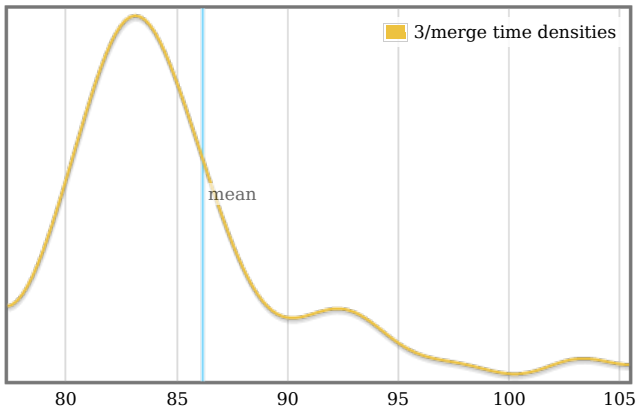


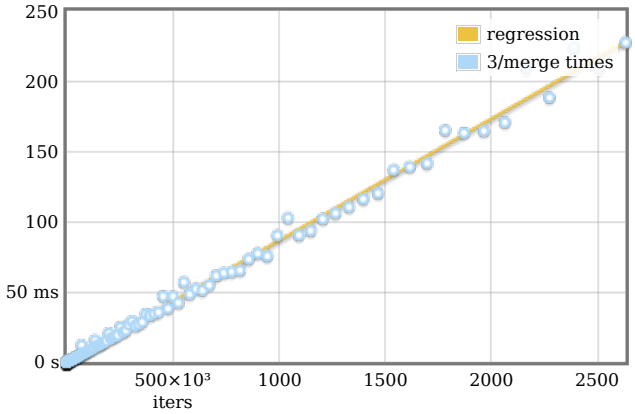



| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 52.9 ns | 54.6 ns | 56.4 ns |
| R² goodness-of-fit | 0.990 | 0.993 | 0.998 |
| Mean execution time | 52.8 ns | 53.8 ns | 55.2 ns |
| Standard deviation | 3.22 ns | 4.05 ns | 5.39 ns |

Outlying measurements have severe (85.2%) effect on estimated standard deviation.

# 3/merge

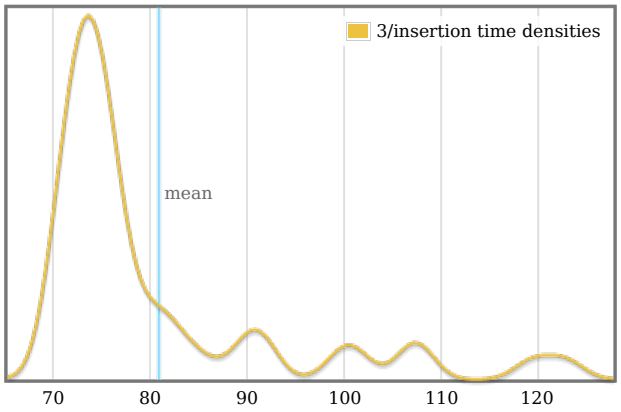


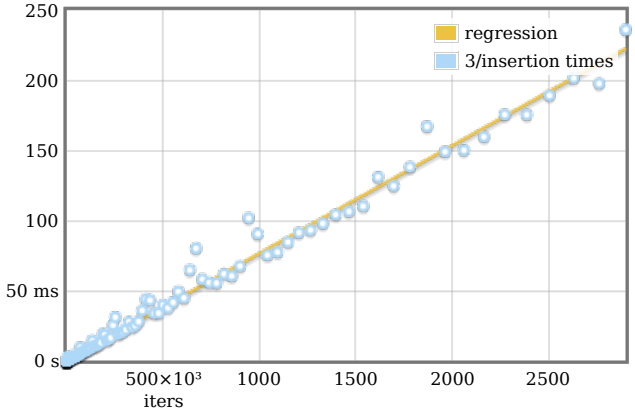




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 84.5 ns | 86.5 ns | 88.8 ns |
| R² goodness-of-fit | 0.994 | 0.996 | 0.998 |
| Mean execution time | 84.3 ns | 86.2 ns | 88.0 ns |
| Standard deviation | 4.49 ns | 5.99 ns | 8.07 ns |

Outlying measurements have severe (82.7%) effect on estimated standard deviation.

# 3/insertion

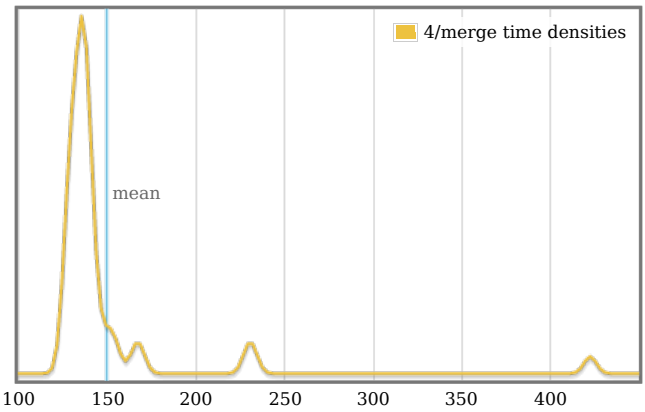


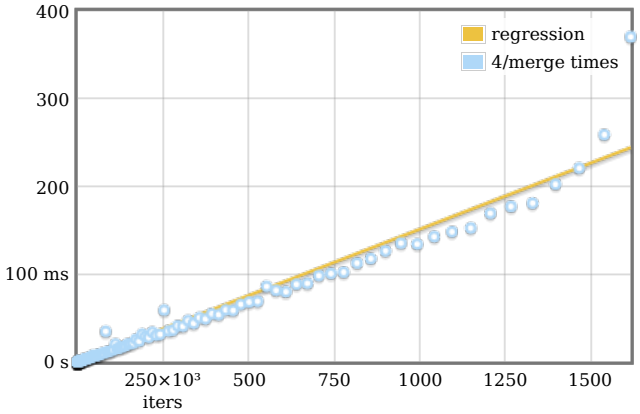




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 74.8 ns | 76.6 ns | 78.8 ns |
| R² goodness-of-fit | 0.984 | 0.991 | 0.996 |
| Mean execution time | 77.8 ns | 80.9 ns | 85.2 ns |
| Standard deviation | 9.52 ns | 13.1 ns | 17.8 ns |

Outlying measurements have severe (96.4%) effect on estimated standard deviation.

# 4/merge

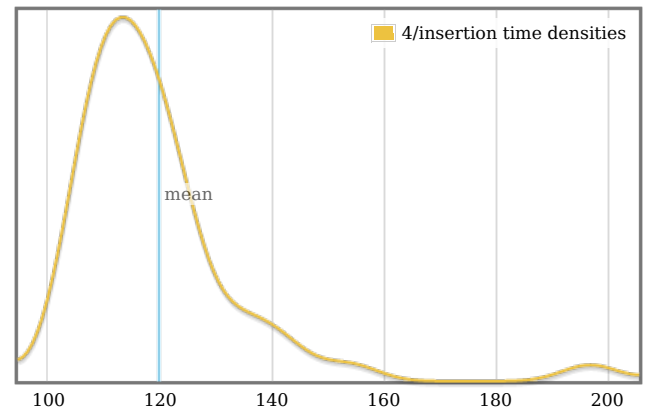


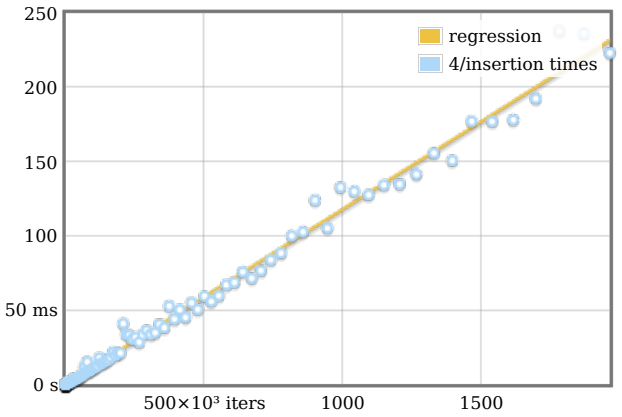




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 138 ns | 151 ns | 169 ns |
| R² goodness-of-fit | 0.936 | 0.961 | 0.997 |
| Mean execution time | 142 ns | 149 ns | 178 ns |
| Standard deviation | 20.8 ns | 47.0 ns | 91.4 ns |

Outlying measurements have severe (99.1%) effect on estimated standard deviation.

# 4/insertion

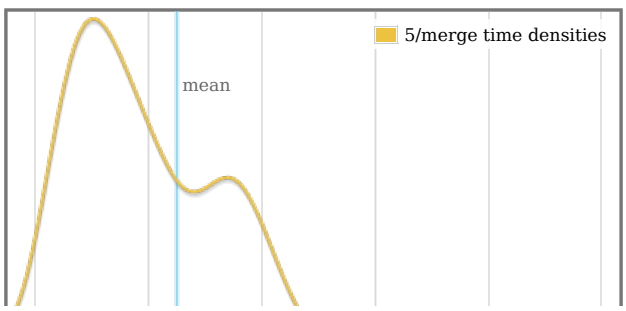


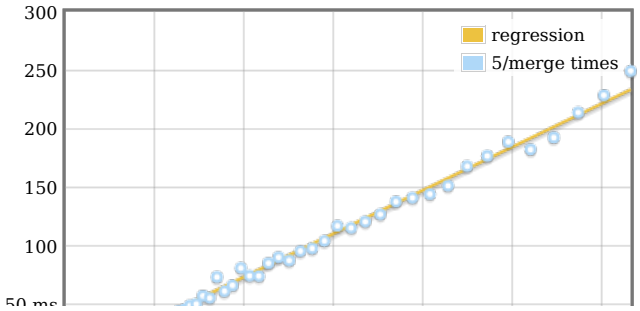




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 114 ns | 118 ns | 122 ns |
| R² goodness-of-fit | 0.991 | 0.993 | 0.997 |
| Mean execution time | 117 ns | 120 ns | 127 ns |
| Standard deviation | 9.62 ns | 15.7 ns | 28.0 ns |

Outlying measurements have severe (94.6%) effect on estimated standard deviation.

# 5/merge

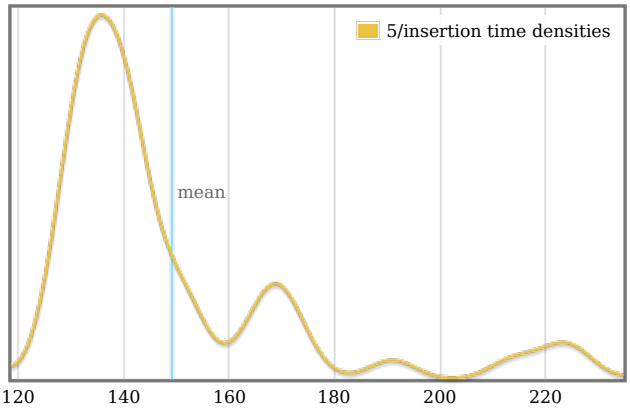


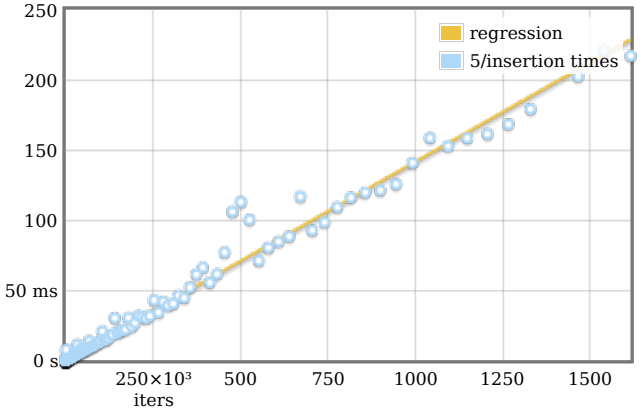



| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 181 ns | 184 ns | 188 ns |
| R² goodness-of-fit | 0.997 | 0.998 | 0.999 |
| Mean execution time | 180 ns | 182 ns | 187 ns |
| Standard deviation | 7.58 ns | 10.2 ns | 13.9 ns |

Outlying measurements have severe (74.1%) effect on estimated standard deviation.

# 5/insertion

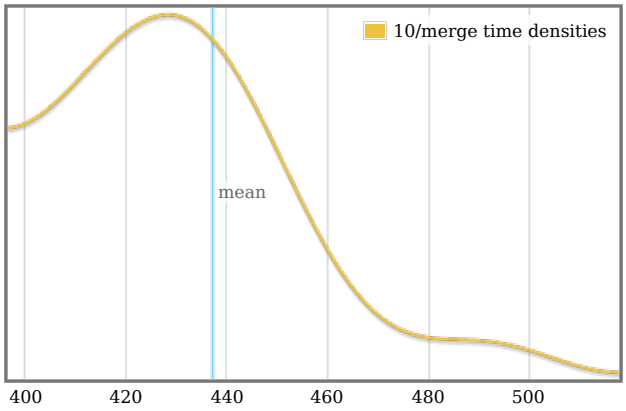


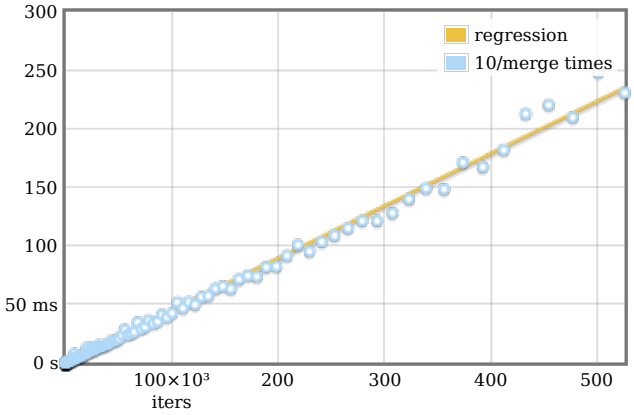




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 138 ns | 141 ns | 146 ns |
| R² goodness-of-fit | 0.975 | 0.987 | 0.996 |
| Mean execution time | 143 ns | 149 ns | 157 ns |
| Standard deviation | 15.6 ns | 24.0 ns | 32.5 ns |

Outlying measurements have severe (96.2%) effect on estimated standard deviation.

# 10/merge

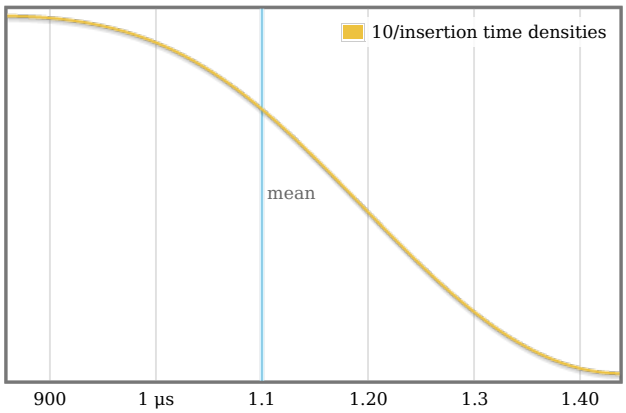


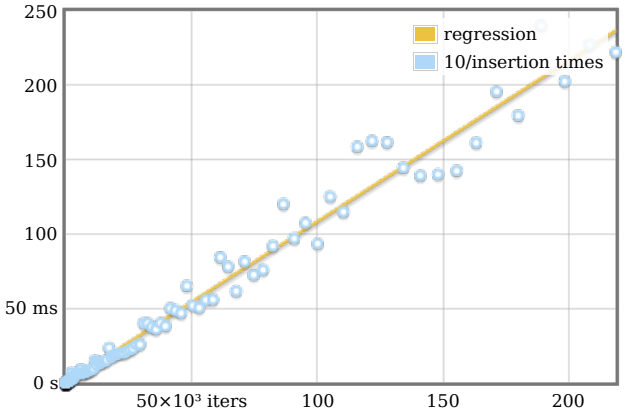




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 433 ns | 447 ns | 459 ns |
| R² goodness-of-fit | 0.994 | 0.996 | 0.999 |
| Mean execution time | 430 ns | 437 ns | 446 ns |
| Standard deviation | 19.9 ns | 26.1 ns | 32.7 ns |

Outlying measurements have severe (75.2%) effect on estimated standard deviation.

# 10/insertion

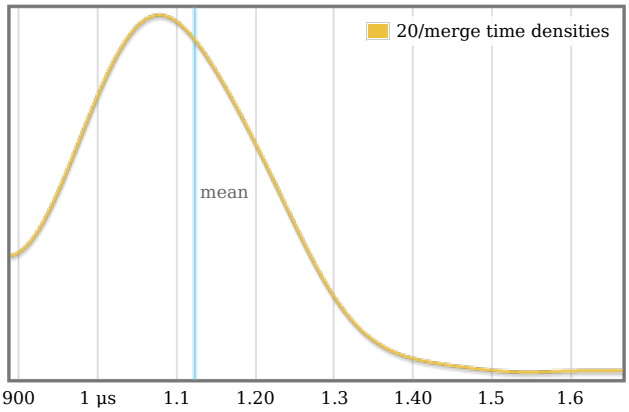


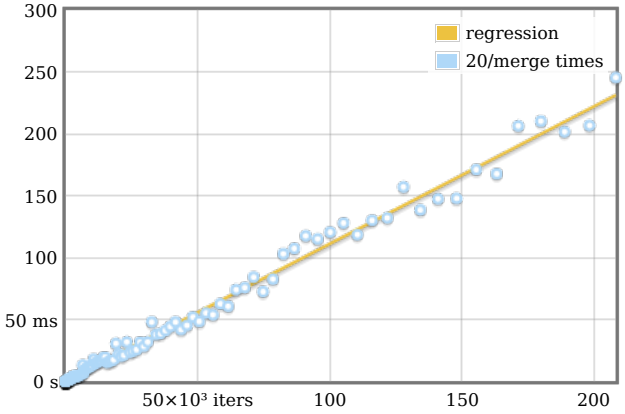




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 1.04 µs | 1.08 µs | 1.14 µs |
| R² goodness-of-fit | 0.976 | 0.983 | 0.990 |
| Mean execution time | 1.06 µs | 1.10 µs | 1.15 µs |
| Standard deviation | 120 ns | 141 ns | 165 ns |

Outlying measurements have severe (92.9%) effect on estimated standard deviation.

# 20/merge

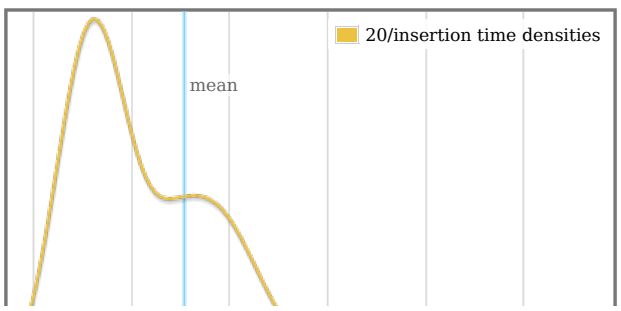


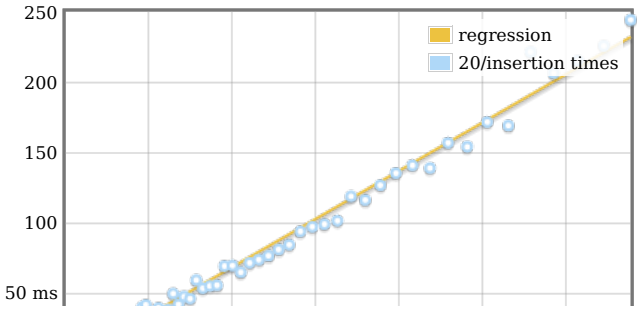




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 1.07 µs | 1.11 µs | 1.15 µs |
| R² goodness-of-fit | 0.990 | 0.992 | 0.995 |
| Mean execution time | 1.08 µs | 1.12 µs | 1.16 µs |
| Standard deviation | 98.7 ns | 133 ns | 187 ns |

Outlying measurements have severe (91.9%) effect on estimated standard deviation.

# 20/insertion

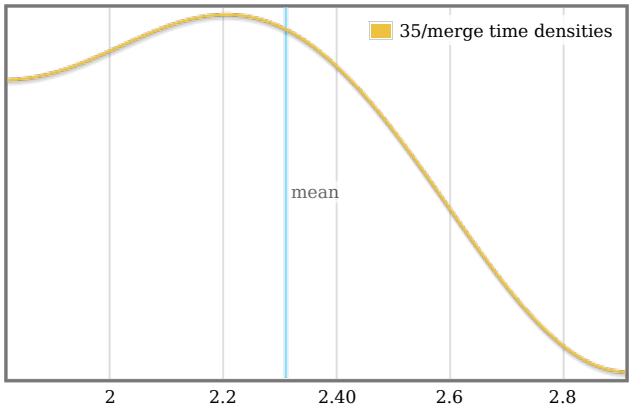


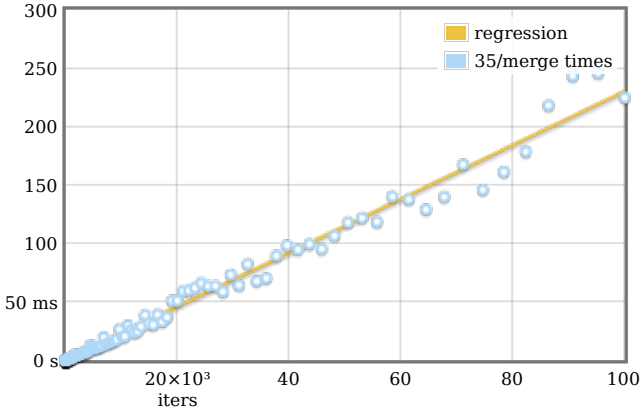



| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 3.32 µs | 3.42 µs | 3.52 µs |
| R² goodness-of-fit | 0.990 | 0.994 | 0.997 |
| Mean execution time | 3.30 µs | 3.38 µs | 3.48 µs |
| Standard deviation | 246 ns | 315 ns | 409 ns |

Outlying measurements have severe (85.8%) effect on estimated standard deviation.

# 35/merge

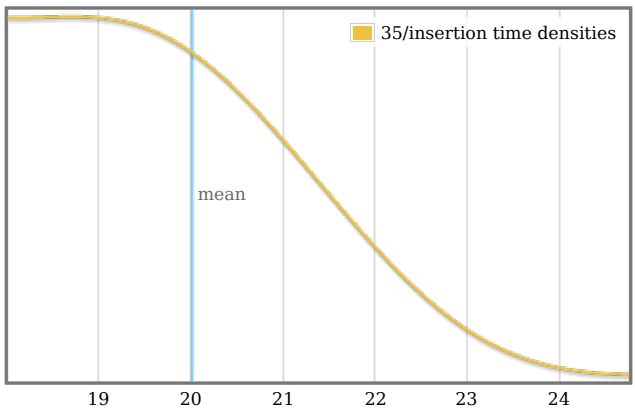


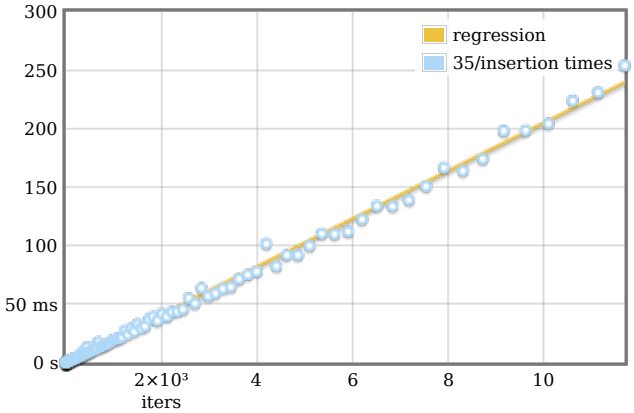




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 2.20 µs | 2.30 µs | 2.41 µs |
| R² goodness-of-fit | 0.983 | 0.987 | 0.993 |
| Mean execution time | 2.22 µs | 2.31 µs | 2.38 µs |
| Standard deviation | 225 ns | 261 ns | 298 ns |

Outlying measurements have severe (90.7%) effect on estimated standard deviation.

# 35/insertion

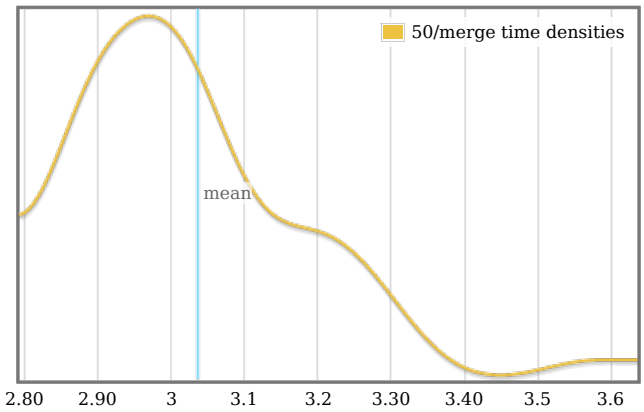


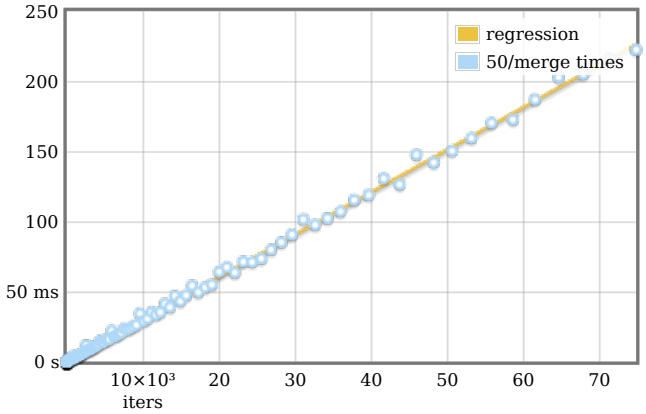




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 20.0 µs | 20.5 µs | 20.9 µs |
| R² goodness-of-fit | 0.995 | 0.997 | 0.998 |
| Mean execution time | 19.6 µs | 20.0 µs | 20.4 µs |
| Standard deviation | 1.04 µs | 1.25 µs | 1.88 µs |

Outlying measurements have severe (68.7%) effect on estimated standard deviation.

# 50/merge

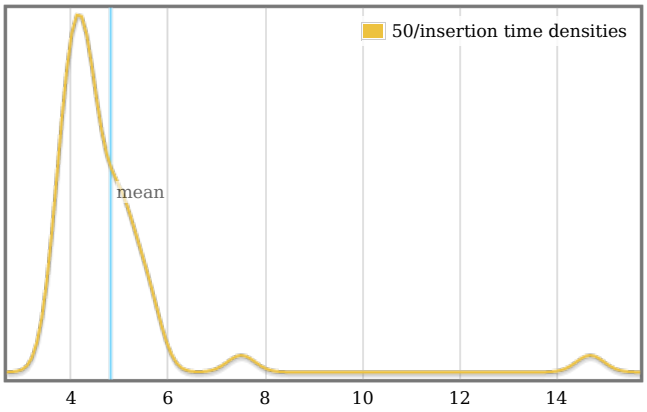


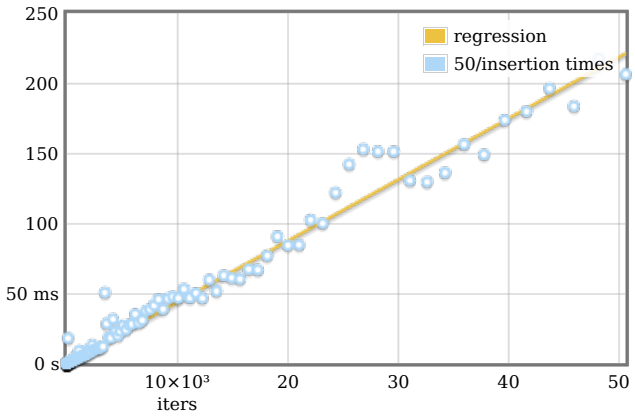




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 2.99 µs | 3.02 µs | 3.06 µs |
| R² goodness-of-fit | 0.998 | 0.999 | 0.999 |
| Mean execution time | 3.00 µs | 3.04 µs | 3.08 µs |
| Standard deviation | 120 ns | 151 ns | 204 ns |

Outlying measurements have severe (63.5%) effect on estimated standard deviation.

# 50/insertion

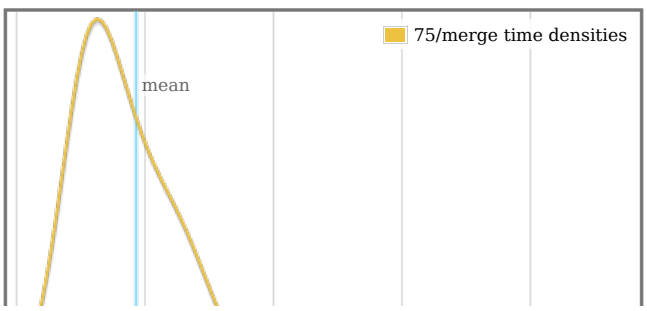


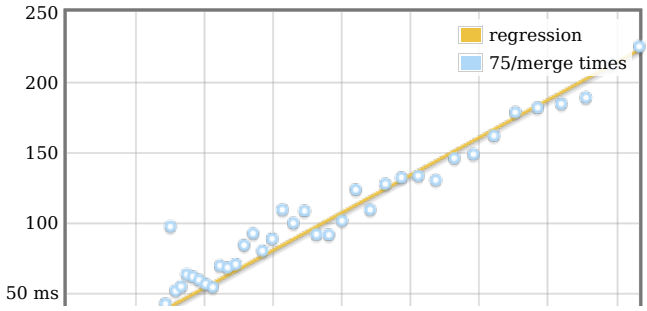




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 4.21 µs | 4.37 µs | 4.62 µs |
| R² goodness-of-fit | 0.972 | 0.981 | 0.990 |
| Mean execution time | 4.52 µs | 4.83 µs | 5.75 µs |
| Standard deviation | 577 ns | 1.64 µs | 3.25 µs |

Outlying measurements have severe (98.8%) effect on estimated standard deviation.

# 75/merge

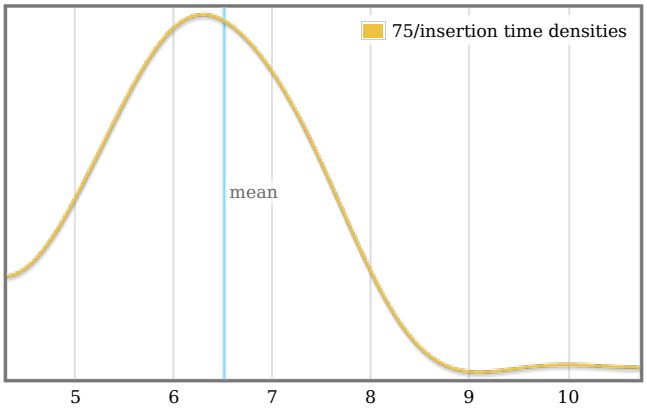


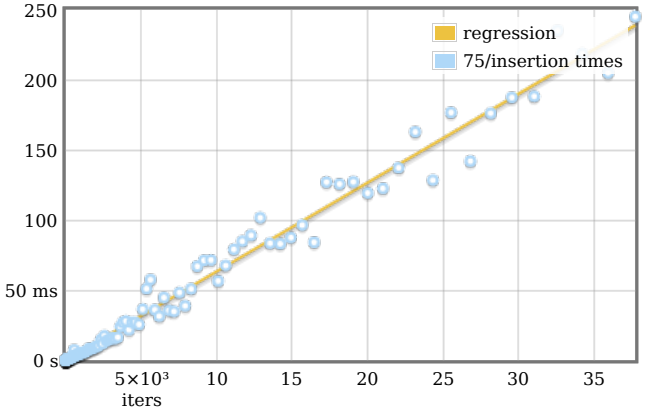



| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 5.25 µs | 5.35 µs | 5.52 µs |
| R² goodness-of-fit | 0.961 | 0.983 | 0.994 |
| Mean execution time | 5.59 µs | 5.85 µs | 6.50 µs |
| Standard deviation | 661 ns | 1.33 µs | 2.54 µs |

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

# 75/insertion

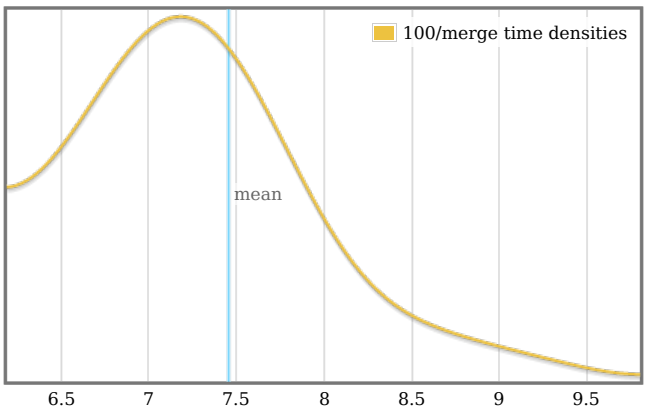


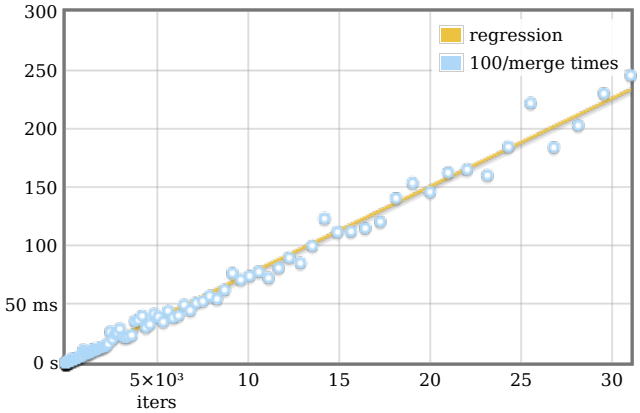




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 6.05 µs | 6.32 µs | 6.57 µs |
| R² goodness-of-fit | 0.977 | 0.985 | 0.991 |
| Mean execution time | 6.15 µs | 6.51 µs | 6.85 µs |
| Standard deviation | 821 ns | 1.11 µs | 1.52 µs |

Outlying measurements have severe (95.1%) effect on estimated standard deviation.

# 100/merge

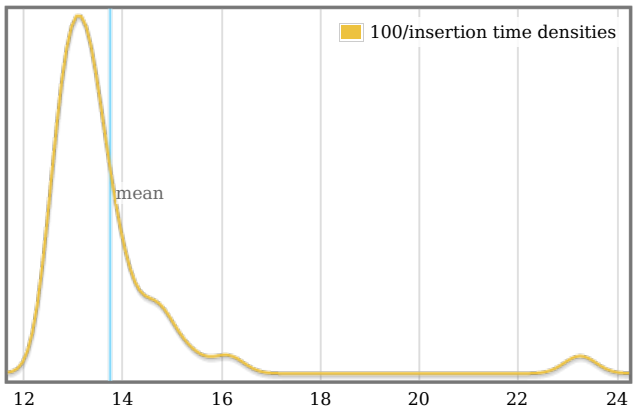


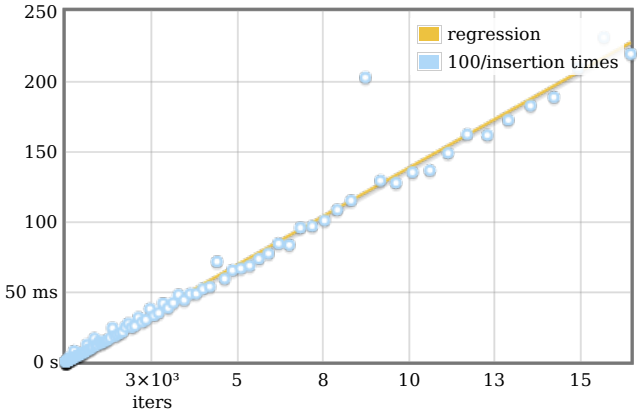




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 7.26 µs | 7.52 µs | 7.72 µs |
| R² goodness-of-fit | 0.989 | 0.992 | 0.996 |
| Mean execution time | 7.28 µs | 7.45 µs | 7.71 µs |
| Standard deviation | 627 ns | 764 ns | 936 ns |

Outlying measurements have severe (87.0%) effect on estimated standard deviation.

# 100/insertion

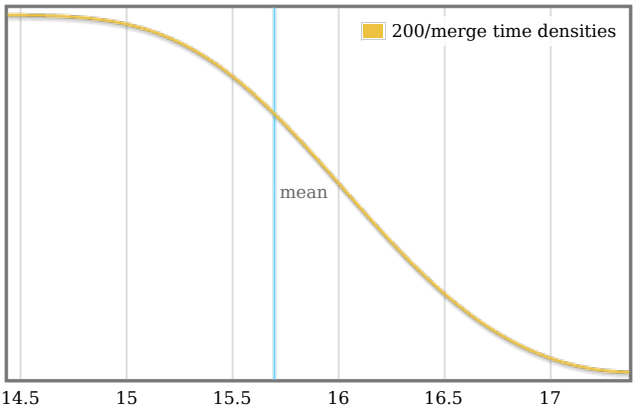


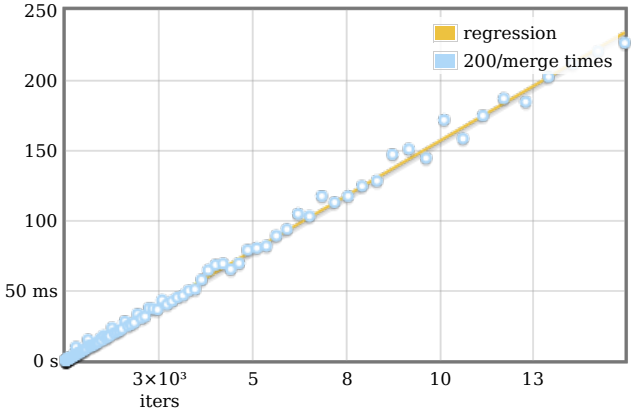




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 13.4 µs | 13.8 µs | 14.5 µs |
| R² goodness-of-fit | 0.951 | 0.981 | 0.999 |
| Mean execution time | 13.4 µs | 13.8 µs | 14.7 µs |
| Standard deviation | 737 ns | 1.70 µs | 3.34 µs |

Outlying measurements have severe (90.5%) effect on estimated standard deviation.

# 200/merge

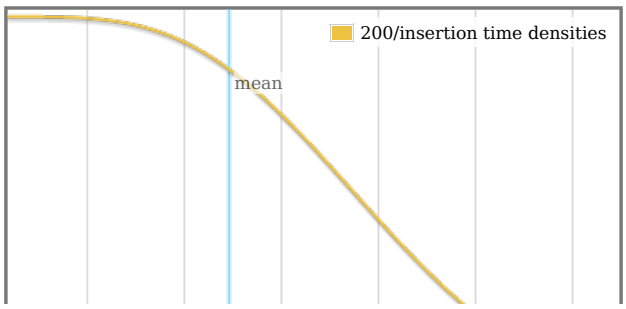


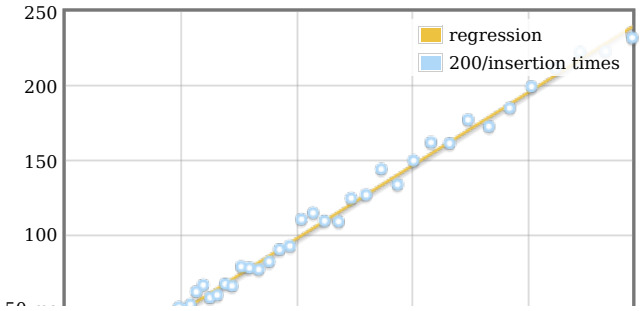




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 15.5 µs | 15.7 µs | 15.9 µs |
| R² goodness-of-fit | 0.997 | 0.998 | 0.999 |
| Mean execution time | 15.5 µs | 15.7 µs | 16.0 µs |
| Standard deviation | 672 ns | 777 ns | 897 ns |

Outlying measurements have severe (58.5%) effect on estimated standard deviation.

# 200/insertion

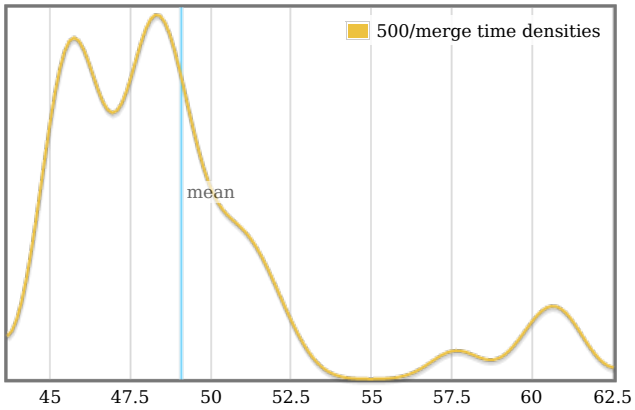


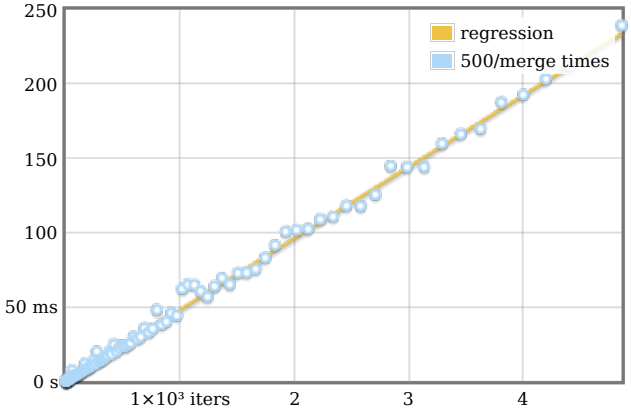



| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 481 µs | 489 µs | 498 µs |
| R² goodness-of-fit | 0.995 | 0.996 | 0.998 |
| Mean execution time | 482 µs | 489 µs | 500 µs |
| Standard deviation | 23.3 µs | 28.3 µs | 34.3 µs |

Outlying measurements have severe (51.4%) effect on estimated standard deviation.

# 500/merge

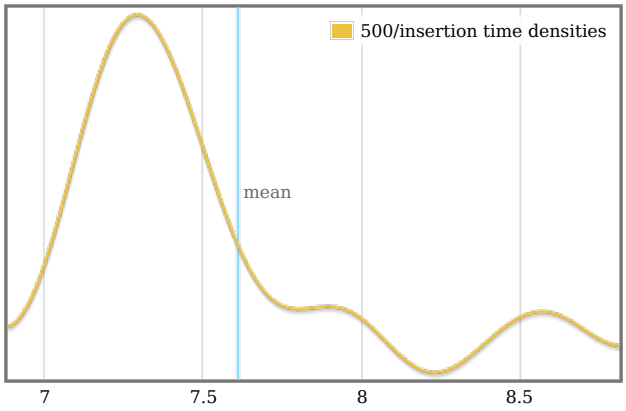


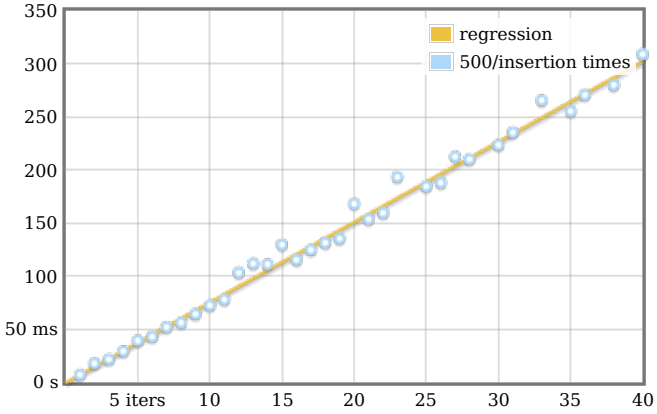




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 47.5 µs | 48.1 µs | 48.7 µs |
| R² goodness-of-fit | 0.995 | 0.997 | 0.998 |
| Mean execution time | 48.1 µs | 49.1 µs | 50.7 µs |
| Standard deviation | 2.72 µs | 4.06 µs | 5.60 µs |

Outlying measurements have severe (77.7%) effect on estimated standard deviation.

# 500/insertion

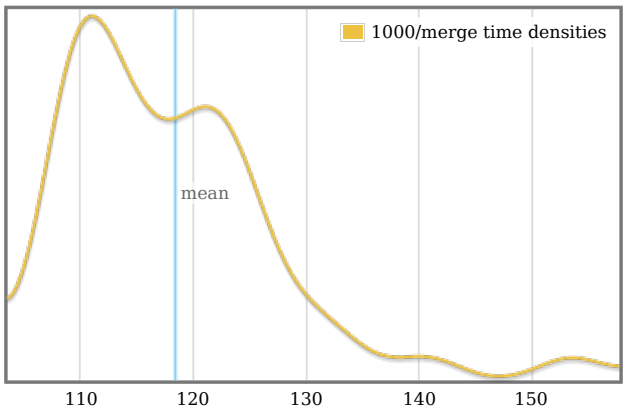


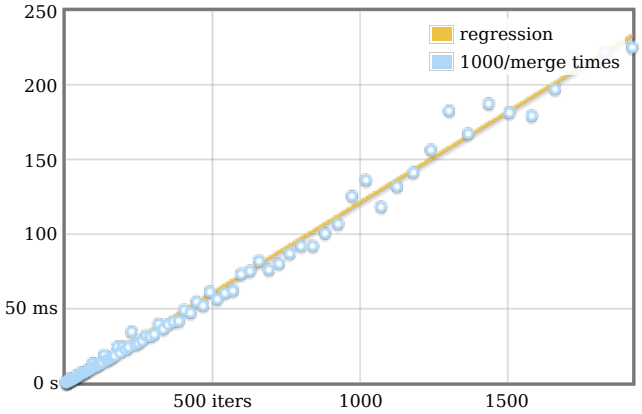




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 7.36 ms | 7.55 ms | 7.78 ms |
| R² goodness-of-fit | 0.984 | 0.991 | 0.996 |
| Mean execution time | 7.47 ms | 7.61 ms | 7.87 ms |
| Standard deviation | 382 µs | 496 µs | 617 µs |

Outlying measurements have moderate (36.7%) effect on estimated standard deviation.

# 1000/merge

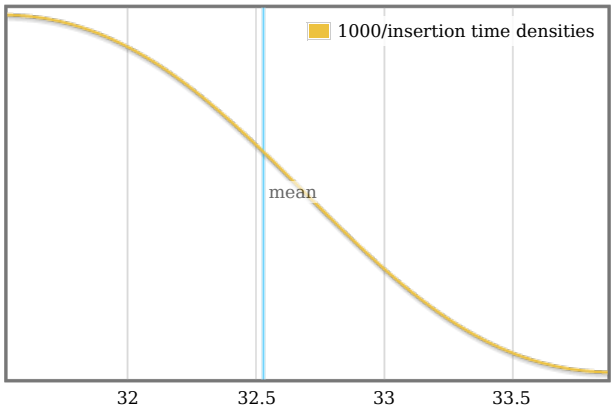


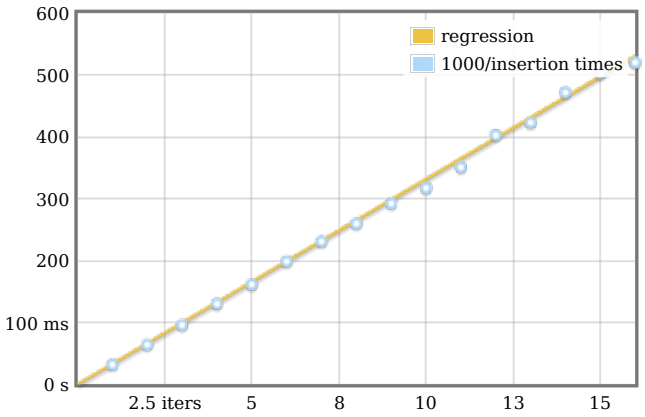




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 119 µs | 121 µs | 125 µs |
| R² goodness-of-fit | 0.991 | 0.994 | 0.997 |
| Mean execution time | 116 µs | 118 µs | 121 µs |
| Standard deviation | 7.16 µs | 9.43 µs | 13.4 µs |

Outlying measurements have severe (73.1%) effect on estimated standard deviation.

# 1000/insertion

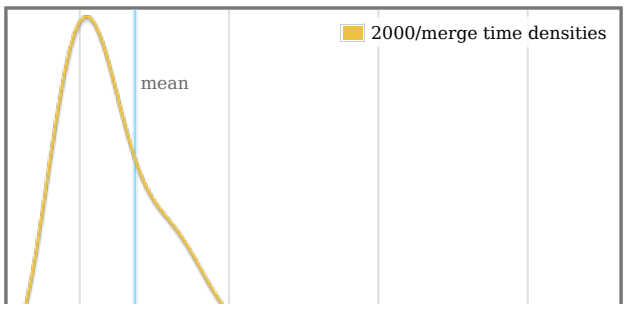


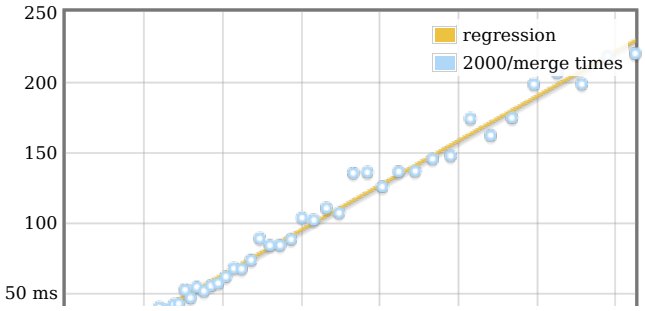




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 32.5 ms | 33.2 ms | 33.8 ms |
| R² goodness-of-fit | 0.997 | 0.998 | 0.999 |
| Mean execution time | 32.3 ms | 32.5 ms | 32.9 ms |
| Standard deviation | 498 µs | 661 µs | 866 µs |

Outlying measurements have slight (5.9%) effect on estimated standard deviation.

# 2000/merge

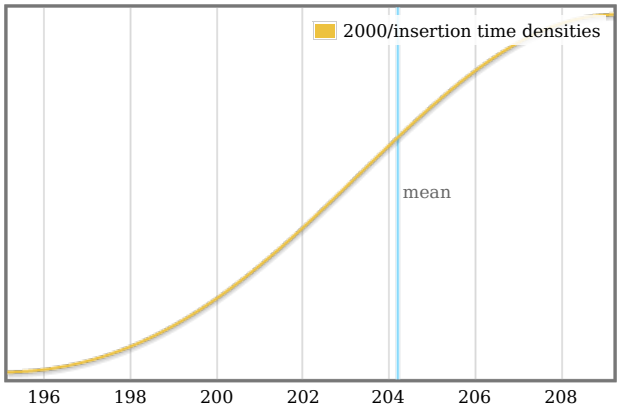


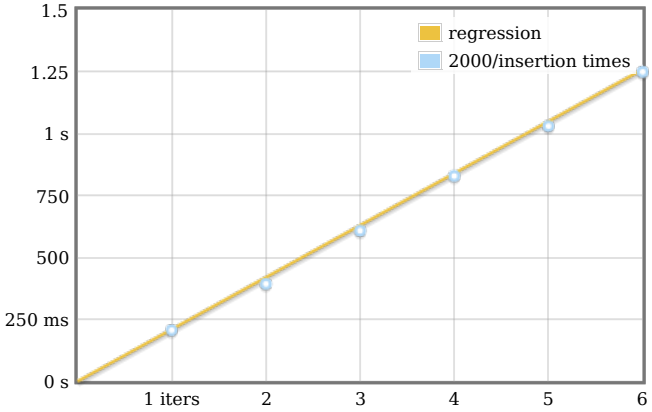



| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 309 µs | 316 µs | 325 µs |
| R² goodness-of-fit | 0.992 | 0.994 | 0.997 |
| Mean execution time | 311 µs | 318 µs | 331 µs |
| Standard deviation | 18.6 µs | 30.2 µs | 51.2 µs |

Outlying measurements have severe (76.3%) effect on estimated standard deviation.

# 2000/insertion

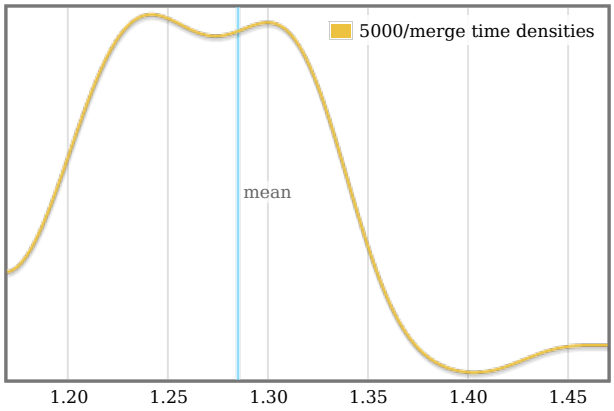


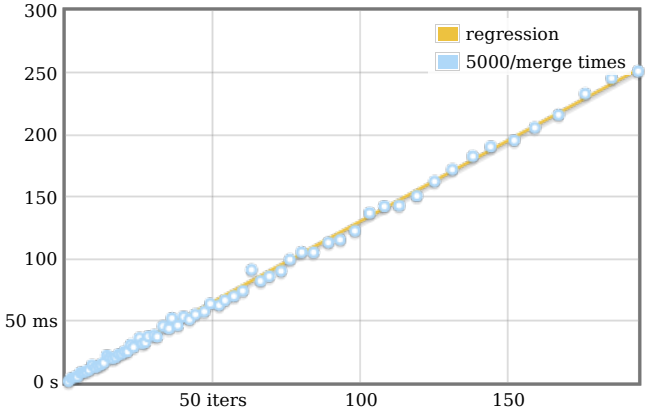




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 206 ms | 210 ms | 214 ms |
| R² goodness-of-fit | 0.998 | 1.000 | 1.000 |
| Mean execution time | 200 ms | 204 ms | 206 ms |
| Standard deviation | 1.73 ms | 4.30 ms | 5.86 ms |

Outlying measurements have moderate (13.9%) effect on estimated standard deviation.

# 5000/merge

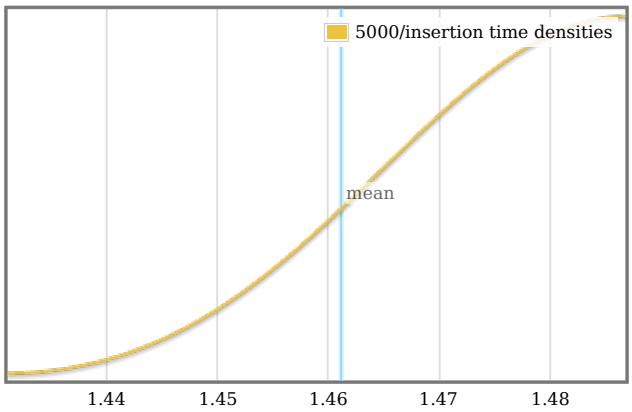


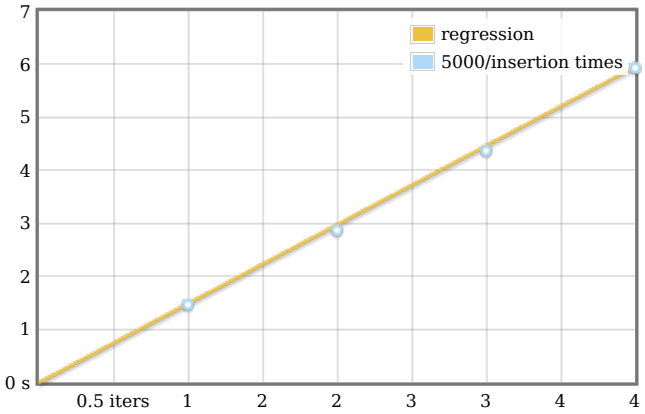




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 1.29 ms | 1.30 ms | 1.31 ms |
| R² goodness-of-fit | 0.997 | 0.998 | 0.999 |
| Mean execution time | 1.27 ms | 1.28 ms | 1.31 ms |
| Standard deviation | 50.5 µs | 64.4 µs | 81.9 µs |

Outlying measurements have moderate (37.9%) effect on estimated standard deviation.

# 5000/insertion

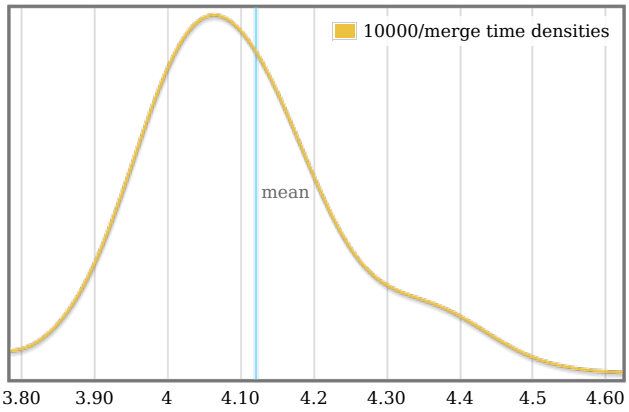


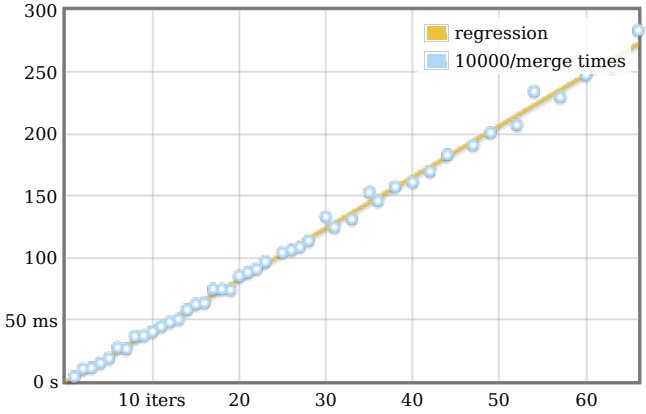




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 1.40 s | 1.49 s | 1.56 s |
| R² goodness-of-fit | 0.999 | 0.999 | 1.000 |
| Mean execution time | 1.44 s | 1.46 s | 1.47 s |
| Standard deviation | 6.88 ms | 19.9 ms | 26.8 ms |

Outlying measurements have moderate (18.8%) effect on estimated standard deviation.

# 10000/merge

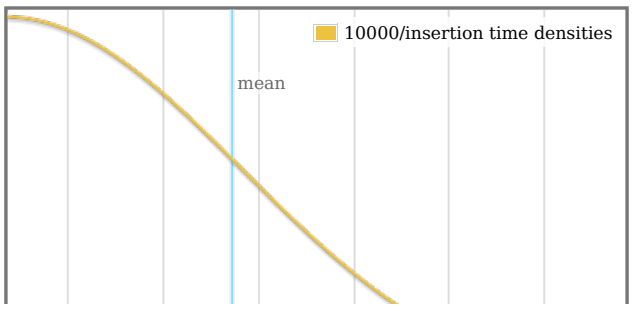


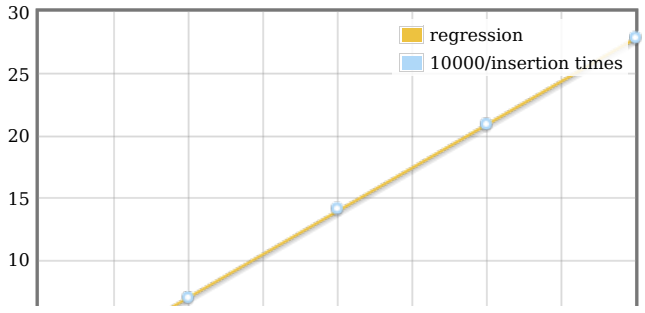




| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 4.05 ms | 4.14 ms | 4.23 ms |
| R² goodness-of-fit | 0.995 | 0.997 | 0.998 |
| Mean execution time | 4.08 ms | 4.12 ms | 4.18 ms |
| Standard deviation | 118 µs | 152 µs | 213 µs |

Outlying measurements have moderate (19.4%) effect on estimated standard deviation.

# 10000/insertion



| | lower bound | estimate | upper bound |
|---|---|---|---|
| OLS regression | 6.79 s | 6.95 s | 7.17 s |
| R² goodness-of-fit | 1.000 | 1.000 | 1.000 |
| Mean execution time | 6.98 s | 7.01 s | 7.06 s |
| Standard deviation | 4.54 ms | 49.3 ms | 62.5 ms |

Outlying measurements have moderate (18.8%) 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.
- *R² goodness-of-fit* is a measure of how accurately the linear regression model fits the observed measurements. If the measurements are not too noisy, R² 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.