1. Data entry

For a certain dataset (with n elements) the time to sort the results is measured and recorded in an excel worksheet. The table shows the number of elements (in column B) and the time (in seconds) it took to sort them (in column C). How do these numbers relate to each other?

Copy and paste this data into a new sheet..

B	C
n	secs
280	0,015
316	0,016
494	0,016
1347	0,031
1463	0,031
2872	0,063
3302	0,094
3717	0,094
4711	0,125
5408	0,14
6410	0,156
6417	0,156
6656	0,172
7251	0,187
7294	0,188
7879	0,204
7883	0,203
8097	0,187
9684	0,25
9901	0,25

2. Data analysis test 1, try log-function.

Let's see if this data correlates with $f(x) = \log_{10} x$…

Add column D: give it a header: "log n" and add the formula in D2: =LOG10(B2). Extend this value for the full list…

B	C	D
n	secs	log(n)
280	0,015	2,447158031
316	0,016	2,499687083
494	0,016	..

3. Data analysis test 2, try square.

Let's see if this data correlates with $f(x) = x^2$…

*Add column E: give it a header: "n^2" and add the formula in E2: =B2*B2. Extend this value for the full list…*

B	C	D	F
n	secs	log(n)	n^2
280	0,015	2,447158031	78400
316	0,016	2,499687083	99856
494	0,016	..	…

4. Data analysis test 3, try n*log-function.

Let's see if this data correlates with $f(x) = x \cdot \log_{10} x$…

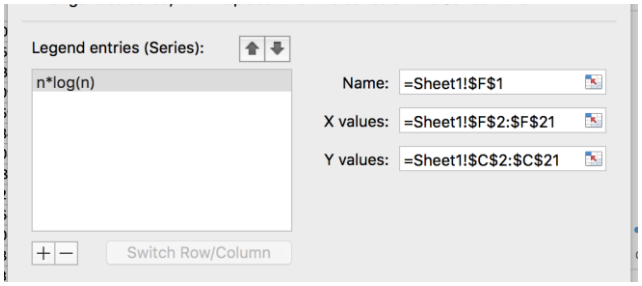
*Add column E: give it a header: "n*log" and add the formula in E2: =B2*D2. Extend this value for the full list…*

B	C	D	F	E
n	secs	log(n)	n^2	n*log(n)
280	0,015	2,447158031	78400	685,2042488
316	0,016	2,499687083	99856	789,9011181
494	0,016	…	…	…

5. Visual analysis, preparation

Create graphs for each combination (D, C), (E, C) and (F, C). Take care of the fact that you assign the values of column D, E, F as x-value.

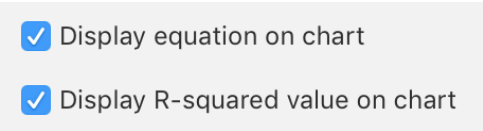
Add "Scatter charts" to represent the data, see the example below…



6. EXCEL analysis, add trendline

Complete the chart and add a trendline. To do so, right-click on any data point in the chart and choose "Add Trendline…".

Add trendlines on all charts. Check the options "Display equation on chart" and "Display R-squared value on chart" to show these in the chart.



6. EXCEL reading, determine

The R-square (R^2) value indicates the correlation. Find the chart where R^2 is the largest. The closer it is to 1, the more it resembles the original function.
In this example the last chart has the best option The equation that goes with this correlation is y=6E-06x + 0.0098.

This means the same as $y = 6 \cdot 10^{-6}x + 0{,}0098$. Because y represents the number of seconds and x equals $n \cdot \log_{10} n$ this formula stands for

$$sec = 6 \cdot 10^{-6} \cdot n \cdot \log_{10} n + 0{,}0098$$

which states that the algorithm behaves as *O(nlogn)*.