simplestats

Simple statistics mini-tool. Work with copy-paste text boxes in which tab-delimited data are pasted (partly in specific formats still to be decided.

Mostly one input box for the data to be analysed, and one output box which provides the test results (always including some explanation of how to interpret the test results).

**1. Descriptive statistics**

- One box on the left where user pastes numbers (one value per line, line breaks separate the different values).

- Then one button "Calculate"

- And one box on the right where the following descriptive statistics are shown (one per line), always preceded like the text in the following so it is clear what is what:

Mean:

Median:

Minimum:

Maximum:

Standard deviation:

Standard error:

Sum:

Mean ± standard deviation (Minimum-Maximum):

[make sure the tool can process a large number of values, I would say, it should allow for at least 1000 or 5000 values]

**2. Test for normality**

- One box on the left where user pastes numbers (one value per line, line breaks separate the different values).

- Then one button "Calculate"

- One box on the right where the results of the normality test is shown [it is important that the user can copy-paste the results, so they must be accessible in a box for marking and copying]

One good example how this could look like is the Kolmogorov-Smirnov calculator here: https://www.socscistatistics.com/tests/kolmogorov/default.aspx

I like this site especially because the results are well explained, this can serve as a model.

But if possible several tests should be performed. Like on this website:

https://www.gigacalculator.com/calculators/normality-test-calculator.php

But here the Kolmogorov-Smirnov test is missing, that one should be added if possible.

To keep it simple, calculate all tests at once (don't let the user choose which will make the window unnecessarily complex.

**3. Binomial test**

Very simple. Follow model on this website:

https://www.graphpad.com/quickcalcs/binomial1/

This website is good because it gives nice explanations, I would suggest to use the same text, maybe with very small modifications

**4. Fisher's exact test**

This model here is good:

https://www.socscistatistics.com/tests/fisher/default2.aspx

Also gives nice explanations that we could also use, but rather than smoking, some example from taxonomy needs to be given (I can change that later).

On the website, the data needs to be entered stepwise, but I think in our tool it would be OK to enter all (column and row names and values at once, then press calculate, and all is calculated at once.

**5. Chi square test**

Very similar in its implementation to Fisher's exact test above. See model here:

https://www.socscistatistics.com/tests/chisquare2/default2.aspx

**6. Compare two independent samples (t-Test and U-test)**

Here we need two boxes to input data, one for the first set of data, one for the second set of data. Then press "Calculate" button and at the same time, results of a (Student) t-test and of a non-parametric (Mann Whitney / Wilcoxon) U test are outputted in a third box.

I really like the implementation here:

https://www.evanmiller.org/ab-testing/t-test.html

We don't need all the fancy graphics output, just the general output and explanations (although of course the graphics are cool!)

**7. Bonferroni correction**

I would suggest to implement here a very simple version of the Bonferroni correction, like here:

https://www.easycalculation.com/statistics/bonferroni-correction-calculator.php

But with some better explanation of the results, see this website and run the calculator:

https://www.quantitativeskills.com/sisa/calculations/bonfer.htm

There, the program explains well to the user what the results actually mean.

[this second website also includes the Sidak correction, and various Holm etc. corrections (if selected in checkbox), but coding all this would probably be overkill - if easily possible and formulas can be found, then it can also be includedin our tool, but not needed - a very simple version like on the first website would be sufficient]

**8. Boxplots**

I would say, 5 different boxes where data can be pasted. If data are pasted e.g. only in two boxes, then the plot will only use these.

Button calculate

The program calculates a boxplot with the data that have been provided. (and outputs it as PDF)

**9. Histogram**

Box to paste values

Calculate

An automated histogram of the values is drawn (and outputted as PDF).