Tutorials and worked examples for simulation, curve fitting, statistical analysis, and plotting.
https://simfit.uk
https://simfit.org.uk
https://simfit.silverfrost.com

The format required to input data into SimFIT is extremely simple: all that is required is to read in a rectangular table of numbers with no missing values, either as a file or from the clipboard.

Data tables

However there are four alternative ways to indicate separation of data values into columns that will be accepted by SimFIT as now discussed.

1. **Space separated variables**

   Here the columns are separated by spaces as in this example.

   ```
   11 12 13
   21 22 23
   31 32 33
   41 42 43
   ```

   This is how results files are formatted by SimFIT, but this is not acceptable for table format using some word processors which generally require columns to be separated by tabs.

2. **Tab separated variables**

   Here the columns are separated by tabs (indicated by –>) as in this example.

   ```
   11->12->13
   21->22->23
   31->32->33
   41->42->43
   ```

   This is acceptable for tables using word processors and SimFIT results tables can easily be transformed into this format, e.g. reading into Excel as a text file then exporting as a table for use by Word.

3. **Comma separated variables**

   Here the columns are separated by commas as in this example.

   ```
   11, 12, 13
   21, 22, 23
   31, 32, 33
   41, 42, 43
   ```

   This is the most common format used for data archiving, for instance *.csv files exported by Excel, and such tables can easily be transformed into tab format, e.g. reading into Excel as a *.csv file then exporting as a table for use by Word.

4. **Semicolon separated variables**

   This is when the continental practise of using commas for decimal points is used so that *.csv files are then exported from Excel in the following format

```
1,1; 1,2; 1,3
2,1; 2,2; 2,3
3,1; 3,2; 3,3
4,1; 4,2; 4,3
```

instead of this format which is universally used in scientific work.

```
1.1 1.2 1.3
2.1 2.2 2.3
3.1 3.2 3.3
4.1 4.2 4.3
```

Input and output formats

SimFIT does not distinguish between integers and floating point numbers in data sets supplied for analysis, but does distinguish three data types of numbers in the files that are automatically created to archive results from analysis.

1. Integers such as 1, 1000, 1000000

2. Probability estimates such as 0.0004, 0.000039

3. Floating point numbers such as 1.2345 or, in exponential format, 1.2345E+00.

Of course all numbers used by SimFIT are represented internally either as integers to full significance or floating point numbers in 64-bit precision, but there are reasons for the way that SimFIT outputs floating point numbers.

Probabilities calculated for significance tests are at best only approximations and, as the numbers must lie between 0 and 1, four or sometimes six figures are output after the decimal point. More than this are for the blind leading the blind.

Calculations done with 64-bit precision have up to fifteen significant figures and it could be argued that results should be output with a similar number of significant figures. However experimental data rarely have more than four or five significant figures, and most calculations involve iterative procedures or approximations in any case so probably up to six significant figures should be sufficient.

Scientific notation

From version 8 users can configure SimFIT to use a selected number of significant figures in standard format or to use exponential notation with exactly six significant figures as follows.

Exponential	Standard
1.23456E+00	1.23456
1.23456E−01	0.123456
1.23456E−02	0.0123456
1.23456E−03	0.00123456
1.23456E−04	0.000123456
1.23456E+01	12.3456
1.23456E+02	123.456
1.23456E+03	1234.56
1.23456E+04	12345.6

However, to avoid monstrosities like 0.0000000012345 in output tables, SimFIT always outputs very small or very large numbers in exponential notation The advantage of exponential notation is that all output tables will

be neatly formatted irrespective of the size of the numbers, and orders of magnitude can be seen at a glance. The disadvantage is that users not conversant with exponential notation may find confusing a notation that even programs like Excel have to resort to for very large or very small numbers. Actually it is very easy to replace exponential format by floating point format by importing a table formatted by SimFiT into Excel as a text file and then changing inside Excel from [scientific] format to [number] format using the [cells format] option. Such tables can then be imported into Word.

Row and Column labels

It has been explained that any rectangular data table with the same number of columns (i.e. variables) for each row (i.e. cases) can be imported into SimFiT by clipboard or file. However sometimes, by design in the case of multivariate data, or by accident otherwise, a full data matrix including row and column labels is submitted to SimFiT for analysis.

There are two ways when this is acceptable.

1. **All labels consist of a single word**

   In this example the interword spaces have been replaced by minus signs but any non-blank character such as an underscore can be used. Note the use of * to indicate that this is neither a row or a column label and so will be ignored.

   ```
        *     column-1  column-2  column-3
      row-1      11        12        13
      row-2      21        22        23
      row-3      31        32        33
      row-4      41        42        43
   ```

2. **All labels are included within double quotes**

   This is done in the next example.

   ```
      "*"     "column 1"  "column 2"  "column 3"
    "row 1"      11          12          13
    "row 2"      21          22          23
    "row 3"      31          32          33
    "row 4"      41          42          43
   ```

Conclusion

Any matrix with $n \geq 1$ rows and $m \geq 1$ columns can be submitted to SimFiT for analysis from file or clipboard as long as two criteria are satisfied

1. All n rows have precisely m columns and there are no non-numerical or missing values.

2. If it is intended to use a data matrix with non-numerical data for row and column labels these must be as single words or be quoted.

The columns can be space, tab, comma, or semicolon separated and labels are not limited to one or two words as long as they are filled out with non-blank characters as in Time_of_Day, or surrounded by double quotes as in "Score on a scale of 1 to 10". For Excel users the macro simfit6.xls in the C:\Program Files\Simfit\doc folder can be used to transform arbitrary matrices into SimFiT format.