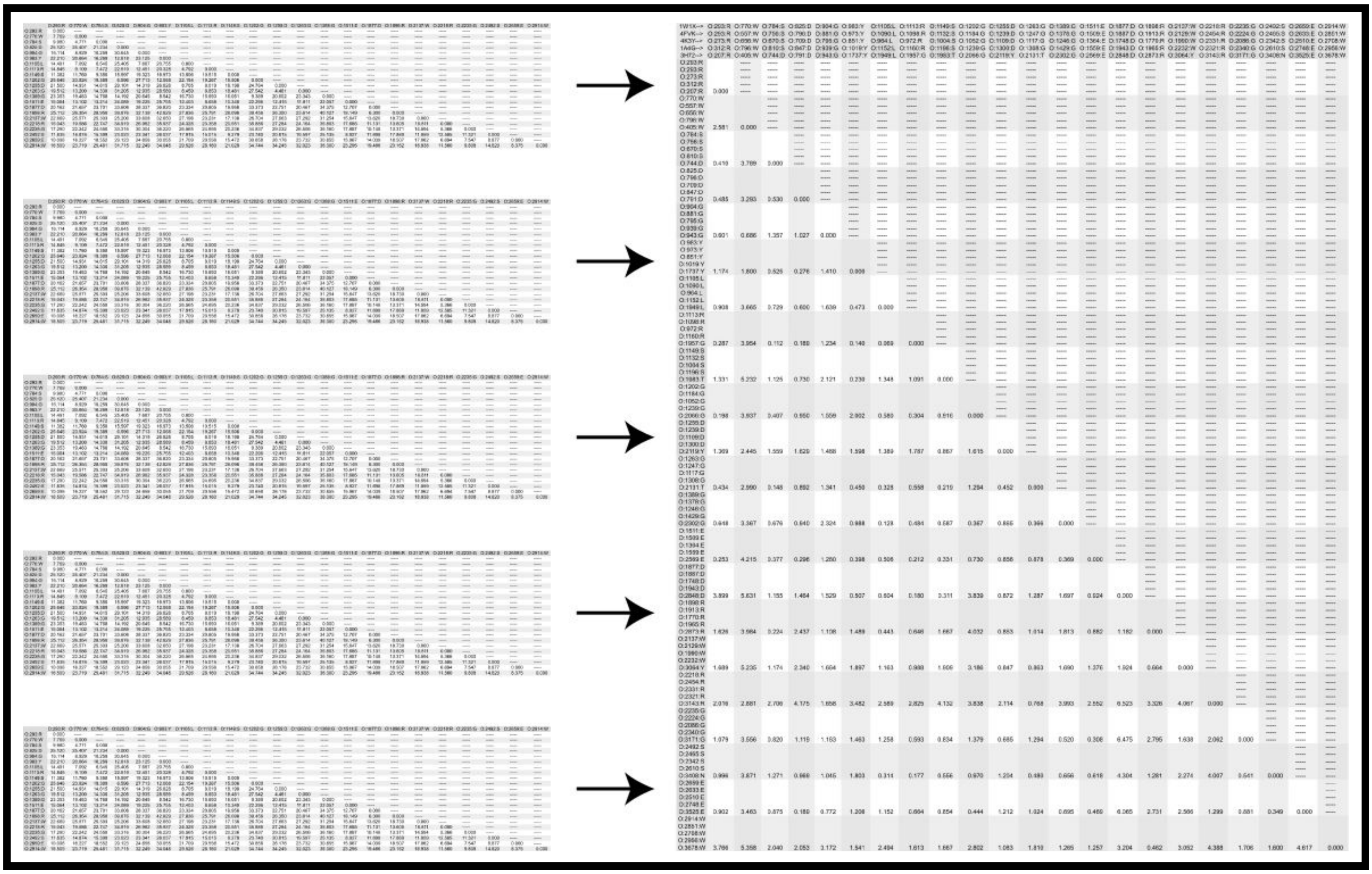

deviation Overview

DEVIATION

Overview

deviation Help Output

deviation Tutorial

deviation Overview

deviation reads in multiple files and writes out the population standard deviation of numerical values that match string token positions across all input files. Non-numerical values are passed directly to the output as sets of lines (one line from each file), with numerical values being substituted by the population standard deviation on the last line of each set of lines. (So if the input files are distance matrices, then the output file will be a population standard deviation matrix.) There is an option for producing condensed output (where only the last line of each set of lines is written to the output), and an option for substituting the mean of numerical values instead of the population standard deviation.

deviation Help Output (“deviation -h” output)

```
NAME

  deviation (version 1.0.1) -- population standard deviation of values from multiple files

SYNOPSIS

  deviation [-c] [-h] [-f #.#] [-o deviation_fn] <filename1> <filename2> [<filenameN>]*

  Command line values:
    filename1 .......... filename of first input file ......... (mandatory)
    filename2 .......... filename of second input file ........ (mandatory)
    [filenameN] ........ filenames of additional input files ... (optional)

  CHARACTER OPTION___KEYWORD OPTION_______DESCRIPTION_____________________________DEFAULT________
  -o <filename> .. --output=<filename> .. output standard deviation filename .... stdout
  -c ............. --condensed .......... produce condensed output ............. complete output
  -f #.# ......... --format=#.# ......... output format of standard deviation ... same as input
  -m ............. --mean ............... output mean (NOT standard deviation) .. std.deviation

  -h ............. --help ............... prints help (Enter 'deviation -h' for help.)
  <NO OPTIONS> .......................... shorter option synopsis (Enter 'deviation'.)
                   --license ............ prints license terms for deviation.

DESCRIPTION

  deviation reads in multiple files and writes out the population standard deviation
  of numerical values that match string token positions across all input files.

  String tokens consist of contiguous non-whitespace and are delimited by whitespace.
  Whitespace is any continguous spaces, formfeeds, newlines, carriage returns, tabs,
  and the end of file. Numeric string tokens are tokens consisting entirely of numbers
  and optionally a decimal point. Non-numeric string tokens are tokens having any
  character that is not whitespace, a number, or a decimal point. A string token
  position is not the position in an absolute character offset, but rather the
  Nth occurrence of any string token (regardless of intervening whitespace).

  All string tokens in all lines are read and MUST MATCH associated string token
  positions in all input files. All non-numeric string tokens and delimiting white
  space will be output by line in order of command line specification. Numeric string
  tokens will be evaluated and replaced by their standard population deviation.
  The standard deviation values replacing numerical string tokens will only be
  printed out once, in the last line of each set of output lines.

  By default, the standard deviation values will be written with the same precision as
  that of the associated input numerical string token of the last line in the associated
  set of output lines. Option '--format=' ('-f') specifies the output format for all
  standard deviation values. The option '--format=' ('-f') must be followed by an option
  value specifying the output format as 'mmm.ddd', where 'mmm' is the minimum field width
  for the entire real number, and where 'ddd' is the precision. When the delimiting space
  in the existing files consists of tabs, then to suppress extraneous spaces, specify the
  format as '--format=X.Y' ('-f X.Y'), for example '--format=5.3' ('-f 5.3'), with the
  value of X being set to the value of Y+2; the abscissa will then be sized as required
  for the value and the mantissa will always be '#' digits. When the delimiting space
  preceding the associated input numerical string token consists of spaces, and the field
  width of the standard deviation being written out is different than the field width of
  the input numerical string token being replaced, then spaces will be added or removed to
  preserve existing input format; if the deviation field width is equal to or larger than
  the sum of the field width of the input token being replaced and the number of space
  characters preceding the existing input token, then the original format will be shifted
  and slightly distorted; to fix, create input with more leading spaces.

  By default, the output will consist of sets of lines where each set contains one line
  for each input file, with the last line of each set containing the standard deviation
  values substituted for numerical tokens in that line. The option '--condensed ('-c')
  produces condensed output where only the last line of each set of lines is written; only
  the labels associated with the last line will be written, i.e. the labels of the last
  specified input file.

  By default, a table with standard deviation values is output. Option '--mean=' ('-m')
  specifies that the output file will contain mean values instead of standard deviation values.

  By default, output goes to stdout. Option '--output=' ('-o') allows specification of the
  output filename. Errors and warnings go to stderr. Option '--help' ('-h') prints this help.
```

EXAMPLE

The three input files for the following example are named 'file1','file2', and 'file3'. In this example, the contents of the input files are the following labelled matrices.

file1 contents:

```
             alpha  bravo   charlie
    alpha     0.00  ------  ------
    bravo     1.00    0.00  ------
    charlie   2.00    3.00    0.00
```

file2 contents:

```
            point1  point2  point3
    pt1       0.00  ------  ------
    pt2      10.00    0.00  ------
    pt3      20.00   30.00    0.00
```

file3 contents:

```
            labelA  labelB  labelC
    labelA    0.00  ------  ------
    labelB  100.00    0.00  ------
    labelC  200.00  300.00    0.00
```

Entering the following command line would result in a new file named 'std_dev_out'.

With keyword options:

```
deviation --output=std_dev_out file1 file2 file3
```

With character options:

```
deviation -o std_dev_out file1 file2 file3
```

std_dev_out contents:

```
             alpha  bravo   charlie
            point1  point2  point3
            labelA  labelB  labelC
    alpha           ------  ------
    pt1             ------  ------
    labelA    0.00  ------  ------
    bravo                   ------
    pt2                     ------
    labelB   44.70    0.00  ------
    charlie
    pt3
    labelC   89.40  134.10    0.00
```

LICENSE INFORMATION

deviation is a software program from Arthur Weininger (www.weiningerworks.com). deviation is subject to a license; use the keyword option '--license' in order to view the license terms. Your use of this software contitutes an agreement to the license terms. Do not use this software if you do not agree to the license terms.

deviation Tutorial

The **dist and deviation Tutorial Page** gives examples of using **deviation.**