

NAME

InfoTextFiles.pl - List information about TextFile(s)

SYNOPSIS

InfoTextFiles.pl TextFile(s)...

InfoTextFiles.pl [-a, --all] [-c, --count] [--datacheck] [-d, --detail *infolevel*] [-e, --empty] [-h, --help] [--indelim *comma* | *semicolon*] [-m, --mode *colnum* | *collabel*] [-n, --numericaldatacols *colnum*,[*colnum*,...] | *collabel*,[*collabel*,...]] [-w, --workingdir *dirname*] TextFile(s)...

DESCRIPTION

List information about *TextFile(s)* contents: number of lines and columns, empty column values, and so on. The file names are separated by spaces. The valid file extensions are *.csv* and *.tsv* for comma/semicolon and tab delimited text files respectively. All other file names are ignored. All the text files in a current directory can be specified by **.csv*, **.tsv*, or the current directory name. The *--indelim* option determines the format of *TextFile(s)*. Any file which doesn't correspond to the format indicated by *--indelim* option is ignored.

OPTIONS

-a, --all

List all the available information.

-c, --count

List number of rows and columns. This is default behavior.

--datacheck

List number of numerical and non-numerical values for each column.

-d, --detail *infolevel*

Level of information to print about lines being ignored. Default: *1*. Possible values: *1, 2 or 3*.

-e, --empty

List number of empty row and column values.

-h, --help

Print this help message.

--indelim *comma* | *semicolon*

Input delimiter for CSV *TextFile(s)*. Possible values: *comma* or *semicolon*. Default value: *comma*. For TSV files, this option is ignored and *tab* is used as a delimiter.

-m, --mode *colnum* | *collabel*

Specify how to identify numerical data columns: using column number or column label. Possible values: *colnum* or *collabel*. Default value: *colnum*.

-n, --numericaldatacols *colnum*,[*colnum*,...] | *collabel*,[*collabel*,...]

This value is mode specific. It is a list of column number or labels to check for presence of numerical data only; otherwise, the value is flagged. Default value: *all;all;....*

For *colnum* mode, input value format is: *colnum,....;colnum,....;....*. Example:

```
1, 3, 5
"2, 4, 6"
```

For *collabel* mode, input value format is: *collabel,....;collabel,....;....*. Example:

```
"MW, SumNO, SumNHOH"
```

-w, --workingdir *dirname*

Location of working directory. Default: current directory.

EXAMPLES

To count number of lines and columns in Text file(s), type:

```
% InfoTextFiles.pl Sample1.csv
% InfoTextFiles.pl Sample1.csv Sample1.tsv
% InfoTextFiles.pl *.csv *.tsv
```

To count number of lines, columns and empty values in Sample1.csv file and print detailed information, type:

```
% InfoTextFiles.pl -d 3 -e Sample1.csv
```

To track all available information and non-numerical values for Mol_ID and MolWeight columns in Sample1.csv file and print detailed information, type:

```
% InfoTextFiles.pl -d 3 -a -m collabel -n Mol_ID,MolWeight Sample1.csv
```

AUTHOR

Manish Sud <msud@san.rr.com>

SEE ALSO

JoinTextFiles.pl, MergeTextFilesWithSD.pl, ModifyTextFilesFormat.pl, SplitTextFiles.pl, TextFilesToHTML.pl

COPYRIGHT

Copyright (C) 2024 Manish Sud. All rights reserved.

This file is part of MayaChemTools.

MayaChemTools is free software; you can redistribute it and/or modify it under the terms of the GNU Lesser General Public License as published by the Free Software Foundation; either version 3 of the License, or (at your option) any later version.