NAME

RDKitDrawMoleculesAndDataTable.py - Generate a HTML data table

SYNOPSIS

RDKitDrawMoleculesAndDataTable.py [--alignmentSMARTS <SMARTS>] [--compute2DCoords <yes or no>] [--counterCol <yes or no>] [--colVisibility <yes or no>] [--colVisibilityCtrlMax <number>] [--footer <text>] [--footerClass <text>] [--freezeCols <yes or no>] [--header <text>] [--headerStyle <text>] [--highlightSMARTS <SMARTS,...>] [--highlightSMARTSDelim <text>] [--highlightValues <datalabel,datatype,criterion,value,...>] [--highlightValuesRanges <datalabel,datatype,criterion1,vaue1,criterion2,value2...>] [--highlightValuesClasses <RuleOf5,RuleOf3,...>] [--highlightColors <colortype,color1,color2>] [--highlightColorsRanges <colortype,color1,color2,color3>] [--highlightColorsRandom <colottype,color1,color2,...>] [--infileParams <Name,Value,...>] [--keysNavigation <yes or no>] [--molImageSize <width,height>] [--molImageEncoded <yes or no>] [--overwrite] [--paging <yes or no>] [--pagingType <numbers,simple, ...>] [--pageLength <number>] [--regexSearch <yes or no>] [--showMolName <yes or no>] [--scrollX <yes or no>] [--scrollY <yes or no>] [--scrollYSize <number>] [--tableStyle <table,table-striped,...>] [--tableFooter <yes or no>] [--tableHeaderStyle <thead-dark,thead-light,... >] [--wrapText <yes or no>] [--wrapTextWidth <number>] [-w <dir>] -i <infile> -o <outfile>

RDKitDrawMoleculesAndDataTable.py -h | --help | -e | --examples

DESCRIPTION

Generate an interactive HTML table with columns corresponding to molecules and available alphanumerical data in an input file. The drawing of molecules are embedded in the columns as in line SVG images.

The interactive HTML table may contain multiple columns with drawing of molecules. These columns are automatically generated for each data field in SD file or a column name in SMILES and CSV/TSV file containing SMILES string in their names. The first molecular drawing column in the HTML table represents primary molecular structure data available in an input file. It corresponds to MOL block is SD file or a first column containing SMILES string in its name in SMILES and CSV/TSV files.

The interactive table requires internet access for viewing in a browser and employs the following frameworks: JQuery, Bootstrap, and DataTable. It provides the following functionality: sorting by columns, page length control, page navigation, searching data with regular expressions, and horizontal/vertical scrolling, row highlighting during hovering, a counter column, freezing of primary structure and counter columns, and column visibility control.

The supported input file formats are: Mol (.mol), SD (.sdf, .sd), SMILES (.smi), CSV/TSV (.csv, .tsv, .txt)

The supported output file format is HTML (.html).

OPTIONS

-a, --alignmentSMARTS <SMARTS> [default: none]

SMARTS pattern for aligning molecules to a common template. This option is only used for primary molecular data in SD, SMILES and CSV/TSV files. It is ignored for all other molecular coordinates corresponding to data fields in SD file or columns in SMILES and CSV/TSV files containing SMILES string in their names.

-c, --compute2DCoords <yes or no> [default: yes]

Compute 2D coordinates of molecules before drawing. Default: yes for SMILES strings in SMILES, CSV/TSV, and SD file data fields. In addition, 2D coordinated are always calculated for molecules corresponding to data fields in SD file or columns in SMILES and CSV/TSV files containing SMILES string in their names.

--counterCol <yes or no> [default: yes]

Show a counter column as the first column in the table. It contains the position for each row in the table.

--colVisibility <yes or no> [default: yes]

Show a dropdown button to toggle visibility of columns in the table. The counter and primary structure columns are excluded from the list.

--colVisibilityCtrlMax <number> [default: 25]

Maximum number of columns to show in column visibility dropdown button. The rest of the data columns are not listed in the dropdown and are shown in the table. A word to the wise: The display of too many columns appear to hang interactive Javascript framework for Bootstrap and DataTables.

--freezeCols <yes or no> [default: yes]

Lock counter and primary structure columns in place during horizontal scrolling.

--footer <text> [default: none]

Footer text to insert at the bottom of the HTML page after the table.

--footerClass <text> [default: small text-center text-muted]

Footer class style to use with <p> tag.

-e, --examples

Print examples.

-h, --help

Print this help message.

--header <text> [default: none]

Header text to insert at the top of the HTML page before the table.

--headerStyle <text> [default: h5]

Header style to use. Possible values: h1 to h6.

--highlightSMARTS <SMARTS,...> [default: none]

SMARTS pattern for highlighting atoms and bonds in molecules. All matched substructures are highlighted.

The SMARTS string is used to highlight atoms and bonds in drawing of molecules present in a HTML table across multiple columns. These columns correspond to data field labels in SD file or a column name in SMILES and CSV/TSV file containing SMILES string in their names. The first molecular drawing column in HTML table corresponds to primary molecular structure data available in an input file. It is identified by a label 'Structure' across all input formats.

A single SMARTS string is used to highlight a common substructure across all columns containing drawing of molecules in HTML table.

Format:

```
SMARTS
Structure,SMARTS1,DataLabel,SMARTS2,...
Structure,SMARTS1,Collabel,SMARTS2,...
```

Example:

```
c1ccccc1
Structure,c1ccccc1,SMILESR1,c1ccccc1,SMILESR2,c1ccccc1
```

--highlightSMARTSDelim <text> [default: ,]

Delimiter for parsing SMARTS patterns specified using '--highlightSMARTS' option. Default: ',' comma character. Possible value: Any arbitrary text or a valid character. You may use arbitrary text as a delimiter to handle presence of special characters such as comma, semicolon, tilde etc. in SMARTS patterns.

--highlightValues <datalabel,datatype,criterion,value,...> [default: none]

Highlighting methodology to use for highlighting alphanumerical data corresponding to data fields in SD file or column names in SMILES and CSV/TSV text files.

Input text contains these quartets: DataLabel, DataType, Criterion, Value. Possible datatype values: numeric, text. Possible criterion values for numeric and text: gt, lt, ge, le.

The 'datalabel' corresponds to either data field label in SD file or column name in SMILES and CSV/TSV text files.

Examples:

```
MolecularWeight,numeric,le,500
MolecularWeight,numeric,le,450,SLogP,numeric,le,5
Name,text,eq,Aspirin
Name,regex,eq,acid|amine
```

--highlightValuesRanges <datalabel,datatype,...> [default: none]

Highlighting methodology to use for highlighting ranges of alphanumerical data corresponding to data

fields in SD file or column names in SMILES and CSV/TSV text files.

Input text contains these sextets: DataLabel, DataType, CriterionLowerBound, LowerBoundValue, CriterionUpperBound, UpperBoundValue.

Possible datatype values: numeric or text. Possible criterion values: Lower bound value - lt, le; Upper bound value: gt, ge.

The 'datalabel' corresponds to either data field label in SD file or column name in SMILES and CSV/TSV text files.

Examples:

```
MolecularWeight,numeric,lt,450,gt,1000
MolecularWeight,numeric,lt,450,gt,1000,SLogP,numeric,lt,0,gt,5
```

--highlightValuesClasses <RuleOf5,RuleOf3,...> [default: none]

Highlighting methodology to use for highlighting ranges of numerical data data corresponding to specific set of data fields in SD file or column names in SMILES and CSV/TSV text files. Possible values: RuleOf5, RuleOf3, DrugLike, Random.

The following value classes are supported: RuleOf5, RuleOf3, LeadLike, DrugLike. LeadLike is equivalent to RuleOf3.

Each supported class encompasses a specific set of data labels along with appropriate criteria to compare and highlight column values, except for 'Random' class. The data labels in these classes are automatically associated with appropriate data fields in SD file or column names in SMILES and CSV/TSV text files.

No data labels are associated with 'Random' class. It is used to highlight available alphanumeric data by randomly selecting a highlight color from the list of colors specified using '--highlightColorsRandom' option. The 'Random' class value is not allowed in conjunction with '--highlightValues' or '--highlightValuesRanges'.

The rules to highlight values for the supported classes are as follows.

RuleOf5 [Ref 91]:

```
MolecularWeight,numeric,le,500 (MolecularWeight <= 500)
HydrogenBondDonors,numeric,le,5 (HydrogenBondDonors <= 5)
HydrogenBondAcceptors,numeric,le,10 (HydrogenBondAcceptors <= 10)
LogP,numeric,le,5 (LogP <= 5)
```

RuleOf3 or LeadLike [Ref 92]:

```
MolecularWeight,numeric,le,300 (MolecularWeight <= 300)
HydrogenBondDonors,numeric,le,3 (HydrogenBondDonors <= 3)
HydrogenBondAcceptors,numeric,le,3 (HydrogenBondAcceptors <= 3)
LogP,numeric,le,3 (LogP <= 3)
RotatableBonds,numeric,le,3 (RotatableBonds <= 3)
TPSA,numeric,le,60 (TPSA <= 60)
```

DrugLike:

```
MolecularWeight,numeric,le,500 (MolecularWeight <= 500)
HydrogenBondDonors,numeric,le,5 (HydrogenBondDonors <= 5)
HydrogenBondAcceptors,numeric,le,10 (HydrogenBondAcceptors <= 10)
LogP,numeric,le,5 (LogP <= 5)
RotatableBonds,numeric,le,10 (RotatableBonds <= 10)
TPSA,numeric,le,140 (TPSA <= 140)
```

The following synonyms are automatically detected for data labels used by MayaChemTools and RDKit packages during the calculation of physicochemical properties.

MayaChemTools: MolecularWeight, HydrogenBondDonors, HydrogenBondAcceptors, SLogP, RotatableBonds, TPSA.

RDKit: MolWt, NHOHCount, NOCount, MolLogP, NumRotatableBonds, TPSA

--highlightColors <colortype,color1,color2> [default: auto]

Background colors used to highlight column values based on criterion specified by '--highlightValues' and '--highlightColorsClasses' option. Default value: colorclass,table-success, table-danger.

The first color is used to highlight column values that satisfy the specified criterion for the column. The second color highlights the rest of the values in the column.

Possible values for colortype: colorclass or colorspec.

Any valid bootstrap contextual color class is supported for 'colorclass' color type. For example: table-primary (Blue), table-success (Green), table-danger (Red), table-info (Light blue), table-warning (Orange), table-secondary (Grey), table-light (Light grey), and table-dark (Dark grey).

The following bootstrap color classes may also used: bg-primary bg-success, bg-danger bg-info, bg-warning, bg-secondary.

Any valid color name or hexadecimal color specification is supported for 'colorspec' color type: For example: red, green, blue, #ff000, #00ff00, #0000ff.

--highlightColorsRanges <colortype,color1,color2,color3> [default: auto]

Background colors used to highlight column values using criteria specified by '--highlightValuesRanges' option. Default value: colorclass, table-success, table-warning, table-danger.

The first and third color are used to highlight column values lower and higher than the specified values for the lower and upper bound. The middle color highlights the rest of the values in the column.

The supported color type and values are explained in the section for '--highlightColors'.

--highlightColorsRandom <colortype,color1,color2,...> [default: auto]

Background color list to use for randomly selecting a color to highlight column values during 'Random" value of '--highlightValuesClasses' option.

Default value: colorclass,table-primary,table-success,table-danger,table-info, table-warning,table-secondary.

The supported color type and values are explained in the section for '--highlightColors'.

-i, --infile <infile>

Input file name.

--infileParams <Name,Value,...> [default: auto]

A comma delimited list of parameter name and value pairs for reading molecules from files. The supported parameter names for different file formats, along with their default values, are shown below:

```
SD, MOL: removeHydrogens,yes,sanitize,yes,strictParsing,yes
SMILES: smilesColumn,1,smilesNameColumn,2,smilesDelimiter,space,
    sanitize,yes
```

Possible values for smilesDelimiter: space, comma or tab.

-k, --keysNavigation <yes or no> [default: yes]

Provide Excel like keyboard cell navigation for the table.

-m, --molImageSize <width,height> [default: 200,150]

Image size of a molecule in pixels.

--molImageEncoded <yes or no> [default: yes]

Base64 encode SVG image of a molecule for inline embedding in a HTML page. The inline SVG image may fail to display in browsers without encoding.

-o, --outfile <outfile>

Output file name.

--overwrite

Overwrite existing files.

-p, --paging <yes or no> [default: yes]

Provide page navigation for browsing data in the table.

--pagingType <numbers, simple, ...> [default: full_numbers]

Type of page navigation. Possible values: numbers, simple, simple_numbers, full, full_numbers, or first_last_numbers.

```
numbers - Page number buttons only
simple - 'Previous' and 'Next' buttons only
simple_numbers - 'Previous' and 'Next' buttons, plus page numbers
full - 'First', 'Previous', 'Next' and 'Last' buttons
full_numbers - 'First', 'Previous', 'Next' and 'Last' buttons, plus
    page numbers
first_last_numbers - 'First' and 'Last' buttons, plus page numbers
```

--pageLength <number> [default: 15]

Number of rows to show per page.

-r, --regexSearch <yes or no> [default: yes]

Allow regular expression search through alphanumerical data in the table.

-s, --showMolName <yes or no> [default: auto]

Show molecule names in a column next to the column corresponding to primary structure data in SD and SMILES file. The default value is yes for SD and SMILES file. This option is ignored for CSV/TSV text files.

--scrollX <yes or no> [default: yes]

Provide horizontal scroll bar in the table as needed.

--scrollY <yes or no> [default: yes]

Provide vertical scroll bar in the table as needed.

--scrollYSize <number> [default: 75vh]

Maximum height of table viewport either in pixels or percentage of the browser window height before providing a vertical scroll bar. Default: 75% of the height of browser window.

-t, --tableStyle <table,table-striped,...> [default: table,table-hover,table-sm]

Style of table. Possible values: table, table-striped, table-bordered, table-hover, table-dark, table-sm, none, or All. Default: 'table,table-hover'. A comma delimited list of any valid Bootstrap table styles is also supported.

--tableFooter <yes or no> [default: yes]

Show column headers at the end of the table.

--tableHeaderStyle <thead-dark,thead-light,...> [default: thead-dark]

Style of table header. Possible values: thead-dark, thead-light, or none. The names of the following contextual color classes are also supported: table-primary (Blue), table-success (Green), table-danger (Red), table-info (Light blue), table-warning (Orange), table-active (Grey), table-light (Light grey), and table-dark (Dark grey).

-w, --workingdir <dir>

Location of working directory which defaults to the current directory.

--wrapText <yes or no> [default: yes]

Wrap alphanumeric text using <br/> delimiter for display in a HTML table.

--wrapTextWidth <number> [default: 40]

Maximum width in characters before wraping alphanumeric text for display in a HTML table.

EXAMPLES

To generate a HTML table containing structure and alphanumeric data for molecules in a SD file along with all the bells and whistles to interact with the table, type:

```
% RDKitDrawMoleculesAndDataTable.py -i Sample.sdf -o SampleOut.html
```

To generate a HTML table containing structure and alphanumeric data for molecules in a SMILES file along with all the bells and whistles to interact with the table, type:

```
% RDKitDrawMoleculesAndDataTable.py  -i Sample.smi -o SampleOut.html
```

To generate a HTML table containing multiple structure columns for molecules in a CSV file along with all the bells and whistles to interact with the table, type:

```
% RDKitDrawMoleculesAndDataTable.py -i SampleSeriesRGroupsD3R.csv
  -o SampleSeriesRGroupsD3ROut.html
```

To generate a HTML table containing structure and alphanumeric data for molecules in a SD file along without any bells and whistles to interact with the table, type:

```
% RDKitDrawMoleculesAndDataTable.py --colVisibility no --freezeCols no
```

```
    --keysNavigation no --paging no --regexSearch no --scrollX no
    --scrollY no -i Sample.sdf -o SampleOut.html
```

To generate a HTML table containing structure and alphanumeric data for molecules in a SD file along with highlighting molecular weight values using a specified criterion, type:

```
% RDKitDrawMoleculesAndDataTable.py  --highlightValues
  "MolecularWeight,numeric,le,500" -i Sample.sdf -o SampleOut.html
```

To generate a HTML table containing structure and alphanumeric data for molecules in a SD file along with highlighting range of molecular weight values using a specified criterion, type:

```
% RDKitDrawMoleculesAndDataTable.py  --highlightValuesRanges
  "MolecularWeight,numeric,lt,400,gt,500" -i Sample.sdf -o SampleOut.html
```

To generate a HTML table containing structure and alphanumeric data for molecules in a SD file along with highlighting molecular weight values and ranges of SLogP values using a specified criterion and color schemes, type:

```
% RDKitDrawMoleculesAndDataTable.py  --highlightValues
  "MolecularWeight,numeric,le,500" --highlightValuesRanges
  "SLogP,numeric,lt,0,gt,5" --highlightColors "colorclass,table-success,
  table-danger" --highlightColorsRanges "colorclass,table-danger,
  table-success,table-warning" -i Sample.sdf -o SampleOut.html
```

To generate a HTML table containing structure and alphanumeric data for molecules in a SD file along with highlighting RuleOf5 physicochemical properties using a pre-defined set of criteria, type:

```
% RDKitDrawMoleculesAndDataTable.py  --highlightValuesClasses RuleOf5
  -i Sample.sdf -o SampleOut.html
```

To generate a HTML table containing structure and alphanumeric data for molecules in a SD file along with all the bells and whistles to interact with the table and highlight a specific SMARTS pattern in molecules, type:

```
% RDKitDrawMoleculesAndDataTable.py  --highlightSMARTS "c1ccccc1"
  -i Sample.sdf -o SampleOut.html
```

To generate a HTML table containing structure and alphanumeric data for molecules in a SD file along with highlighting of values using random colors from a default list of colors, type:

```
% RDKitDrawMoleculesAndDataTable.py --highlightValuesClasses Random
  -i Sample.sdf -o SampleOut.html
```

To generate a HTML table containing structure and alphanumeric data for molecules in a SD file along with highlighting of values using random colors from a specified list of colors, type:

```
% RDKitDrawMoleculesAndDataTable.py --highlightValuesClasses Random
  --highlightColorsRandom "colorspec,Lavendar,MediumPurple,SkyBlue,
  CornflowerBlue,LightGreen,MediumSeaGreen,Orange,Coral,Khaki,Gold,
  Salmon,LightPink,Aquamarine,MediumTurquoise,LightGray"
  -i Sample.sdf -o SampleOut.html
```

To generate a HTML table containing structure and alphanumeric data for molecules in a SMILES file specific columns, type:

```
% RDKitDrawMoleculesAndDataTable.py --infileParams "smilesDelimiter,
  comma, smilesColumn,1,smilesNameColumn,2"
  -i SampleSMILES.csv -o SampleOut.html
```

AUTHOR

Manish Sud(msud@san.rr.com)

SEE ALSO

RDKitConvertFileFormat.py, RDKitDrawMolecules.py, RDKitRemoveDuplicateMolecules.py, RDKitSearchFunctionalGroups.py, RDKitSearchSMARTS.py

COPYRIGHT

Copyright (C) 2026 Manish Sud. All rights reserved.

The functionality available in this script is implemented using RDKit, an open source toolkit for cheminformatics developed by Greg Landrum.

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.