Occupancy Matrix with Dendrograms

OccMat\_Dendro.py

There are two modes of running the script:

Command Line Mode – input options supplied to OS X prompt

Python Function Mode – called within another python script with options in function call

**Two modes of input:**

.csv file, with column labels in first row and row labels in first column (Example: test\_occmat.csv) as an option: –csvfile test\_occmat.csv. (Command Line or Python Function Modes)

OR

Supply occupancy matrix and row/column labels within a python script (Python Function mode only).

**Output:**

* Occupancy matrix with dendrograms is displayed on the screen
* Optional pdf output (i.e. test\_occmat.pdf)
* Optional output of .csv file with data ordered by clustering (i.e test\_occmat\_ordered.csv)

**Command Line Mode:**

>> python3 OccMat\_Dendro.py test\_occmat.csv

Full paths can be supplied, or the .py file can be copied to your local directory.

There are a number of command line arguments that can be supplied as ‘—opt value’. These can be listed as:

>>python3 OccMat\_Dendro.py -h

usage: OccMat\_Dendro.py [-h] [--csvfile [CSVFILE]] [--occmat [OCCMAT]]

[--row\_labels [ROW\_LABELS]] [--col\_labels [COL\_LABELS]] [--cluster\_rows [CLUSTER\_ROWS]]

[--cluster\_cols [CLUSTER\_COLS]] [--log\_row\_dendro [LOG\_ROW\_DENDRO]]

[--log\_col\_dendro [LOG\_COL\_DENDRO]] [--show\_row\_labels [SHOW\_ROW\_LABELS]]

[--show\_col\_labels [SHOW\_COL\_LABELS]] [--xsize XSIZE][--ysize YSIZE]

[--pixel\_size PIXEL\_SIZE] [--dendro\_linewidth DENDRO\_LINEWIDTH] [--font FONT]]

[--font\_path\_width FONT\_PATH\_WIDTH] [--row\_label\_x\_offset ROW\_LABEL\_X\_OFFSET]

[--row\_label\_y\_offset ROW\_LABEL\_Y\_OFFSET] [--col\_label\_x\_offset COL\_LABEL\_X\_OFFSET]

[--col\_label\_y\_offset COL\_LABEL\_Y\_OFFSET] [--row\_label\_size ROW\_LABEL\_SIZE]

[--col\_label\_size COL\_LABEL\_SIZE] [--show\_pdf [SHOW\_PDF]] [--hue HUE]

[--border\_color BORDER\_COLOR] [--save\_ordered\_csv [SAVE\_ORDERED\_CSV]]

Occupancy Matrix Clustering Arguments

**Note: Either csvfile or occmat/row\_labels/col\_labels arguments must be supplied**

optional arguments:

-h, --help show this help message and exit

--csvfile CSVFILE csv file containing occupancy matrix

--occmat OCCMAT numpy array containing values to be clustered

--row\_labels ROW\_LABELS

list of row labels

--col\_labels COL\_LABELS

list of col labels--cluster\_rows [CLUSTER\_ROWS]

Flag to cluster Rows, default = True

--cluster\_cols [CLUSTER\_COLS]

Flag to cluster Columns, default = True

--log\_row\_dendro [LOG\_ROW\_DENDRO]

Flag to take log2 of row dendrogram, default = True

--log\_col\_dendro [LOG\_COL\_DENDRO]

Flag to take log2 of column dendrogram, default = True

--show\_row\_labels [SHOW\_ROW\_LABELS]

Flag to show row labels, default = True

--show\_col\_labels [SHOW\_COL\_LABELS]

Flag to show column labels, default = True

--xsize XSIZE X size of output plot in inches, default = 8

--ysize YSIZE Y size of output plot in inches, default = 8

--pixel\_size PIXEL\_SIZE

number of pixels per element of the occupancy matrix, default = 40

--dendro\_linewidth DENDRO\_LINEWIDTH

linewidth to use for drawing dendrograms, default = 1.0

--font FONT font for dendrogram labels, default = Helvetica

--font\_path\_width FONT\_PATH\_WIDTH

linewidth to use for drawing dendrograms, default = 0.02

--row\_label\_x\_offset ROW\_LABEL\_X\_OFFSET

x offset for labels in fraction of row, default = 0.2

--row\_label\_y\_offset ROW\_LABEL\_Y\_OFFSET

y offset for labels in fraction of row, default = 0.3

--col\_label\_x\_offset COL\_LABEL\_X\_OFFSET

x offset for labels in fraction of column, default = 0.3

--col\_label\_y\_offset COL\_LABEL\_Y\_OFFSET

y offset for labels in fraction of column, default = -0.2

--row\_label\_size ROW\_LABEL\_SIZE

row label size as fraction of row, default = 0.6

--col\_label\_size COL\_LABEL\_SIZE

column label size as fraction of column, default = 0.6

--show\_pdf [SHOW\_PDF]

Flag to supress pdf output, default = True

--hue HUE hue (for HSV) for pixels, default = 0.62 (blue)

--border\_color BORDER\_COLOR

border color between pixels (y or w), default = y (yellow)

--save\_ordered\_csv [SAVE\_ORDERED\_CSV]

Flag to save ordered output in .csv

**Python Function Mode:**

The python file can be loaded as a module, and the occupancy matrix dendrogram is produced by calling the function. DrawOccMat(csvfile).

test\_dendro.py:

import OccMat\_Dendro as omd

file = "test\_occmat.csv"

omd.DrawOccMat(file)

OccMat\_Dendro.py can be either located in the same working directory as test\_dendro.py <OR> it can be installed in your local site-packages directory.

The site-packages directory on your machine can be found:

>> python -c 'import site; print(site.getsitepackages())'

Returning something like:

/Users/jrwill/opt/anaconda3/lib/python3.8/site-packages

To install, simply copy the file to this directory:

>> cp OccMat\_Dendro.py /Users/jrwill/opt/anaconda3/lib/python3.8/site-packages

*OPTIONS are different in this mode.* The same options are available, but must be specified in the python function mode ‘opt = value’, rather than the command line mode ‘—opt value’.

Example (test\_dendrol.py):

import OccMat\_Dendro as omd

file = "test\_occmat.csv"

omd.DrawOccMat(file, hue = 0.2, hide\_row\_labels = True)