App development

This app was created by Albert Xue in the Bagheri Lab at Northwestern University. Its original implementation was on an SNA dataset created in collaboration with the Mirkin and Mrksich labs. All code was developed by Albert Xue and is freely available for all to use. The github repo is intended to be published alongside our manuscript with the following authors: Gokay Yamankurt, Eric J. Berns, Albert Xue, Andrew Lee, Neda Bagheri, Milan Mrksich, Chad A. Mirkin. As of the last readme update, this work was submitted to Nature Nano.

App introduction

This app creates a dimensional stacking figure to visualize multiple dimensions of data. The app “stacks” the different dimensions on top of one another such that all datapoints can be visible simultaneously (See example\_output.pdf). I highly suggest viewing this pdf to understand what’s going on. This allows viewing of complicated relationships between multiple explanatory variables and a single response variable. In other words, this app visualizes how different explanatory variables affect the response variable. However, there are limitations: the explanatory variables must be discrete or categorical and continuous values will not work properly, datasets with more than ~10 variables and/or 1500 samples will not fit well, and the variable stacking order requires manual tuning.

This app is interactive meaning any change you make on the front should immediately update on the user interface, allowing real-time feedback on tuning parameters. If you make a change that breaks the code/figure, you can either reload the app (losing all progress) or revert the change.

App startup

To start the app, you need both R and RStudio installed. Then, double click on ui.R and RStudio should pop up (if not, configure it so all .R files run with RStudio; right click and try “run with Rstudio”). Press Ctrl-Shift-Enter to run the app and a user interface should pop up. To exit the app, simply close the pop-up app window.

Example data introduction

The example data comes from a spherical nucleic acid (SNA) dataset created by the Mirkin and Mrksich labs. The dataset was created to explore how SNAs can modify an immune system to retarget and attack a cancerous tumor. SNAs have millions of possible designs, so this experiment tested how different designs affect immune response. The experiment measured a macrophage’s immune response (measured by a response variable called SEAP) when exposed to SNA nanoparticles with different experimental designs (grouped together and called the explanatory variables). Higher response values correspond to higher immune activations and there are eight explanatory variables such as nanoparticle size or lipid type. The goal of this app is simply to visualize the different explanatory variables and their collective effect upon immune activation.

UI panel explanations

The following sections walk through each panel and option in the UI.

File Inputs and Outputs

File input is on the left. “Choose data file” loads in the explanatory variables while “Choose response file” loads in the response variable. The explanatory variable file must be in .csv format (comma separated values) with specified variable names in the first row. The response variable file should be a single column of numbers with the same number of rows as the explanatory variable file.

Checking “Specify value order?” brings up a file prompt in which you can load a file to customize ordering of values. For example, the example\_output.pdf has “OligoConc” as the highest column variable. This variable has four values composed of (1, 10, 100, 1000). If you wanted to change this order, say to (1000, 100, 10, 10), then you specify a value order file and use this option. An example is given in example\_value\_order.csv. The first word in each line must start with a variable name (spelled correctly) followed by the variable order (also spelled correctly).

If outputting the file to a pdf is needed, the output file name can be specified in “Choose output file name.” The pdf figure and raw values/orientation of the dimensional stacking visual can be output with the buttons below.

Plotting Parameters

This app allows basic adjustment of data before plotting. The first two bars represent the width and height in inches.

“Choose variables in columns” and “Choose variables in rows” allow reordering of variables within the plot. This changes the order that variables are stacked on each other and can drastically change the visual clarity. Variables must be specified exactly by name and separated by a comma (no space). If a variable name is spelled wrong, then the app should error out. Also, you may opt to omit some variables to see the effect on the visual; the app will ignore those variables accordingly. To recreate example\_output.pdf, the variable order is OligoConc,Backbone,Orientation,OligoDensity for the columns and Attachment,PeptideDensity,Lipid,Size for the rows.

“Normalize response?” is on by default and normalizes the response to be between 0 and 1. This option divides the response variable by the largest value.

“Log10 transform response?” is off by default and applies log10(response + 1) when active. This option exaggerates larger magnitudes and is good for visualizing small response changes.

“Diverging plot?” is off by default and will adjust the plot to account for negative values. When off, the app assumes that all values are positive (or zero) and two colors are used for response magnitudes (default color is white to red). When on, the app allows negative values and will use three colors for response: white is 0, red is positive values, and blue is negative values. The bubble sizes will be scaled according to absolute magnitude and colors scale with positive/negative.

The “Update plot” button is pretty self-explanatory.

Tinkering parameters

This app supports basic visual changes such as colors or tweaking the size/color of bubbles to fit an aesthetic.

“Legend title” customizes the title of the legend.

“Column label sizes” and “Row label sizes” tweaks the size of the labels on the sides of the figure. These parameters refer to the actual column/row labels (they should be underlined) and not the variable labels. There need to be as many values as variables in a column/row. For example, there are four column and four row variables in example\_output.pdf. The values are specified as numbers separated by commas (no spaces).

“Bubble size scaling factor” and “Variable label size” work the same way as column/row label sizes.

“Select color scale” allows customization of bubble colors. If “Diverging plot?” is on, then this option is ignored.

Author information

This app was created by Albert Xue of the Bagheri Lab at Northwestern University.

Version 0.1 updated on 04/27/18

You can reach him at [xuebert@gmail.edu](mailto:xuebert@gmail.edu)