FlowR

FlowR consists of a series of workflows, each representing a form of analysis. To date, these are; CutFlow, SurvFlow and CompFlow.

Installation

You can install the released version of FlowR from GITHUB with:

```
install_github("HHayman/FlowR")
```

You can load the FlowR package with:

```
library(FlowR)
```

CutFlow

Arrange your data to meet the following criteria;

- Carry out any exclusions
- Include an identifier variable
- Your status variable needs to be coded as 0 and 1, 0 for no event, 1 for an event
- Time variable must be included as a a continuous variable
- There is no limit on the number of variables you wish to generate a cutpoint for
- If coding multiple datasets, the respective variables must have the exact same name in all datasets
  - Additional datasets do not need to contain all variables from your training dataset, just those you wish to be coded in that dataset
- Save your dataset(s) as a CSV file
- Create a subdirectory in your R directory, and place your dataset files inside

To run CutFlow, simply fill in the blanks, as in the example below;

- Subdirectory is the name of the folder you placed your datasets in. It must be within your current directory
- TrainingData is the name of the dataset you wish to be used to generate the cutpoint
  - This cutpoint is then applied to all datasets within the subdirectory
- CutPointStatus is the status variable to be used (such as CSS)
- CutPointTime is the time variable to be used (such as Survival_2021)
- minprop is the minimum proportion of cases to be include either side of the cutpoint
  - Default is 0.1, exclude the argument if you don't want to change this
- Greyscale is an optional toggle to produce a greyscale variant of all plots
  - The default is colour, exclude the argument if you don't want to change this
- Variables is a list of your variables to generate cutpoints for

```
CutFlow(Subdirectory = "YourSubdirectory", TrainingData = "TrainingDataset",
        CutPointStatus = "CutpointVariable", CutPointTime = "TimeVariable", minprop = 0.1,
        Greyscale = TRUE, Variables = c( "Variable1",
                                         "Variable2",
                                         "Variable3",
                                         "Variable4",
                                         "Variable5",
                                         "Variable6",
                                         "Variable7",
                                         "Variable8",
                                         "Variable9",
                                         "Variable10",
                                         "Variable11",
                                         "Variable12"))
```

A new folder will be created in your R directory;

- Folder name format is CutFlow_SystemData_Number
- Three folders are contained within;
  - 0 - A copy of all datafiles fed into CutFlow, for record keeping
  - 1 - A copy of all cutpoint data, including a pdf list of cutpoints
  - 2 - A copy of all datasets, newly coded