factorMerger
Cheat Sheet

Agnieszka Sitko [aut, cre]
Przemysław Biecek [aut, ths]
University of Warsaw

MI DATA LAB

FactorMerger - set of tools to support results from post hoc testing

Basics

The aim of **factorMerger** is to provide set of tools to support results from post hoc comparisons. *Post hoc* testing is an analysis performed after running *ANOVA* to examine differences between group means (of some response numeric variable) for each pair of groups (groups are defined by a factor variable).

This project arose from the need to create a method of post hoc testing which gives the hierarchical interpretation of relations between groups means. Thereby, for a given significance level we may divide groups into nonoverlapping clusters.

1. TODO: here describe more details

Level fusing plot

The top-left plot shows level fusing paths (merging paths). With arguments **family=**, **show=**, **fuse=**, **spacing=**, one can select how to merge factors and what shall be presented on OX/OY axes.

Argument	Summary
panel = "all"	All panels
panel = "left"	Only left two panels
panel = "top"	Only top two panels
panel = "merging"	Merging path plot

Argument	Summary
show = "likelihood"	Plot likelihood on OX axis
show = "p-value"	Plot p-values on OX axis
fuse = "all2all"	Compare all pairs of groups
fuse = "nearby"	Compare nearby groups
fuse = "cluster"	DMR4glm algorithm
spacing = "equidistant"	Levels equidistant on OY scale
spacing = "effects"	Levels according to their effects
family = "gaussian"	For one-dimensional Gaussian
family = "mgaussian"	For multi dimensional Gaussian
family = "binomial"	For binomial regression
family = "survival"	For Cox regression

Group summaries

The top-right panel shows group characteristics.

Use the parameter **summary=** to select the most suitable presentation.

Argument	Summary
summary = "heatmap"	For mgaussian
summary = "profile"	For mgaussian
summary = "boxplot"	For gaussian
summary = "means"	For gaussian
summary = "survival"	For Cox regression
summary = "proportions"	For binomial regression

Merging path plot — Group success proportion — Optimal GIC partition — Model statistics

heatmap / profile — box plot means — survival proportions

Example

```
library(factorMerger)
randSample <-
    generateMultivariateSample
    (N = 100, k = 10, d = 3)
fmAll <- mergeFactors
    (randSample$response,
     randSample$factor)
print(fmAll)
plotTree(fmAll)
```
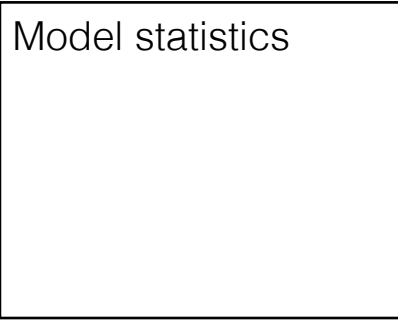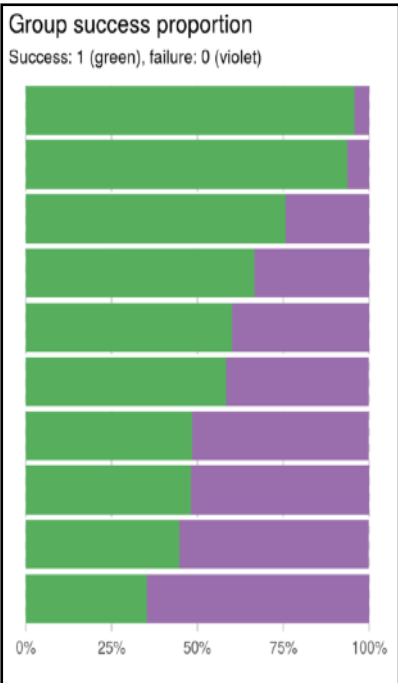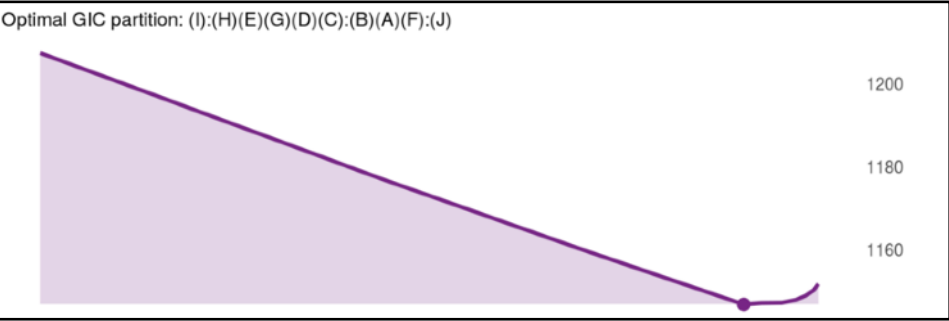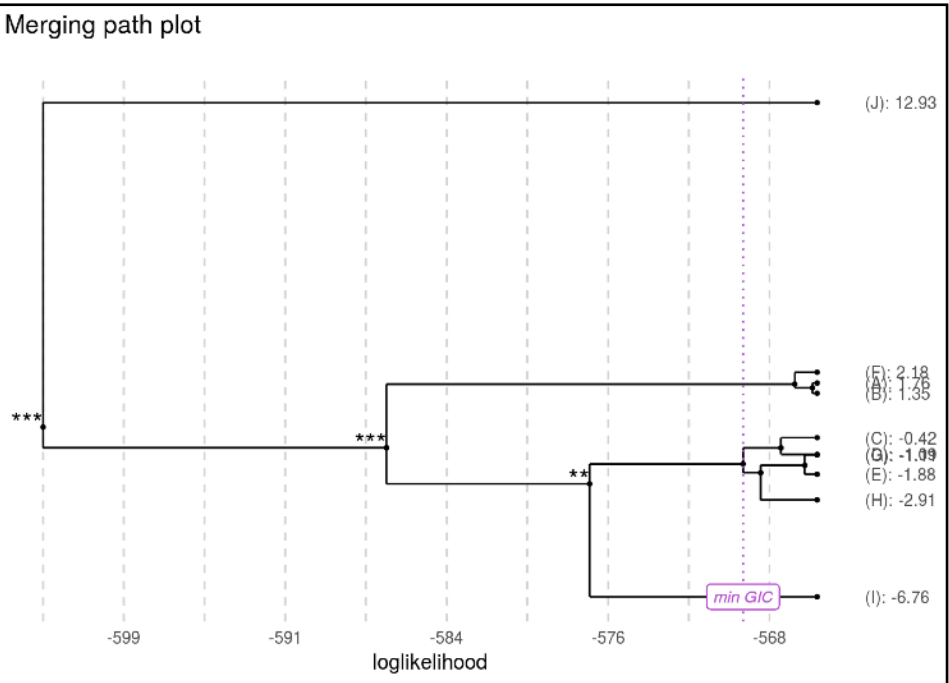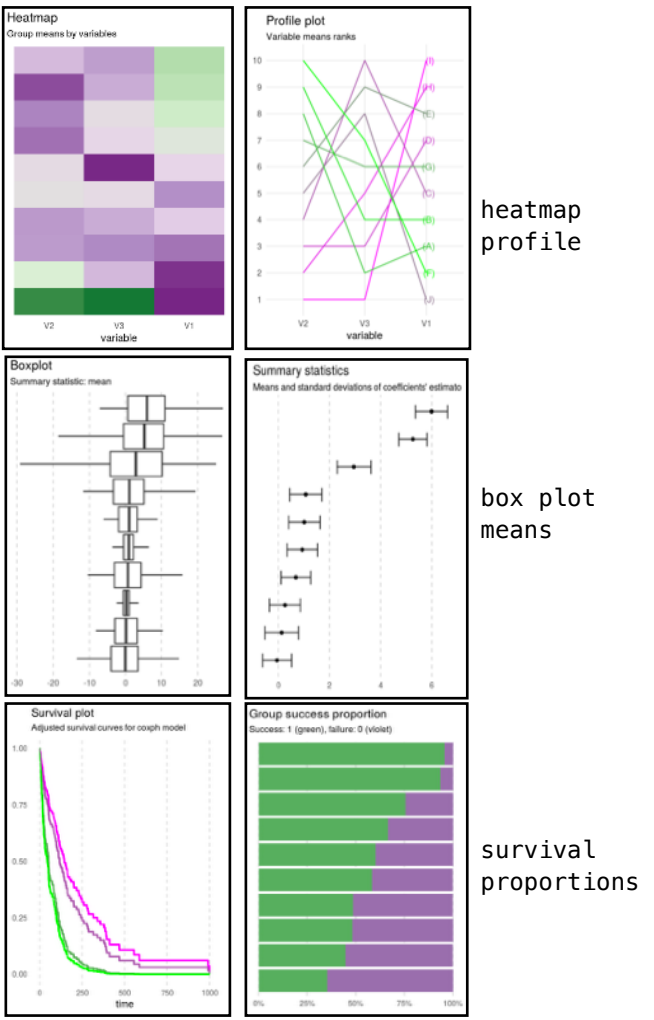
Optimal model selection

The bottom-left panel shows how the model selection criteria behaves for models on the merging path. Currently only criteria from the class of penalised likelihood are supported.

Set the **criterion=** to one from following

Argument	Criterion
criterion = "BIC"	Bayesian Information Criterion
criterion = "AIC"	Akaike Information Criterion
criterion = "GIC"	Generalized Information Criterion

Model statistics

TODO: Here put the description of model statistics…