Package 'exposomex'

June 20, 2024

Title To expediate the discovery of "Exposure-Biology-Disease" nexus

Version 1.0.0

Description A new framework of exposomic analysis to expediate the discovery of "Exposure-Biology-Disease" nexus.

License GPL (>= 3)

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.1

Imports httr,vroom,readxl,ggplot2,writexl,gridExtra,GGally

NeedsCompilation no

Author bin Wang [aut, cph, cre]

Maintainer bin Wang <binwang@pku.edu.cn>

R topics documented:

`CrosAsso` *Association analysis*

Description

Association analysis for cross-sectional data

Usage

```
CrosAsso = function(PID, OutPath = "default",Linear = T, EpiDesign = "cross.sectional",
   VarsY, VarsX = "all.cx",VarsN = "single.factor",VarsSel = F,VarsSelThr = 0.1,
    Covariates = "all.c",Family,RepMsr = F,Corstr = "ar1")
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by StatCros
`OutPath`	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
`Linear`	lgl. T (or TRUE) and F (or FALSE). Whether the relationship between variables is linear

EpiDesign	chr. Epidemiological design of the study, including "cohort" "case.control" and "cross.sectional". It doesn't affect the modeling, but the format of the output file. For the three designs, the effect values are usually indicated by RR (relative risk) of cohort, OR (odds ratio) of case-control, and beta value of cross-sectional.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
VarsSel	lgl. T (or TRUE) and F (or FALSE). Whether to select the significant variable for the final model. Available options.
VarsSelThr	num. If "VarsSel" = T, provide the selection threshold of the P-value. Three value can be chosen, i.e. 0.05, 0.1, and 0.2.
Covariates	chr. Covariates used for modeling. The default option is "all.c" (All covariates are included).
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable
RepMsr	lgl. T (or TRUE) and F (or FALSE). Whether existing repeated measurement of the subjects. Available options.
Corstr	chr. If "RepMsr" = T, the generalized estimating equations (GEE) will be used. For GEE, three correlation structure options are "exchangeable" "ar1" "unstructured".

Value

A list containing the association analysis results.

Author(s)

Bin Wang

Examples

```
res <- InitStatCros()
   res1 = LoadStatCros(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatCros(PID = res$PID, OutPath = "default",TransDummyVars="default")
   res3 = CrosAsso(PID=res$PID, Linear = TRUE, EpiDesign = "cohort",
   VarsY = "Y1", VarsX = "X5,X6,X7,X8,X9,X10,X11", VarsN = "single.factor" ,
   VarsSel = FALSE, VarsSelThr = 0.1, Covariates = "all.c", Family = "gaussian",
   RepMsr = FALSE,Corstr = "ar1")
   FuncExit(PID = res$PID)
```

CrosPred *Build prediction models*

Description

Build prediction models

Usage

```
CrosPred = function(PID,OutPath = "default",VarsY,VarsX = "all.x",
   PredType = "response",VarsSel = FALSE,VarsSelThr = 0.1, Covariates = "all.c",
    RsmpMethod = "cv",Folds = 5,Ratio = 0.667,Repeats = 5)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatCros
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
PredType	chr. Prediction type of the outcome variable, including "response" for the actual values and "prob" for outcome with binary variable.
VarsSel	lgl. Whether to select the significant variable for the final model. Available options include T (or TRUE) and F (or FALSE).
VarsSelThr	num. If "VarsSel" = T, provide the selection threshold of the P-value. Three value can be chosen, i.e. 0.05, 0.1, and 0.2.
Covariates	chr. Covariates used for modeling. The default option is "all.c" (All covariates are included).
RsmpMethod	chr. Four resampling methods options for internal validation, including "cv" (i.e.,Cross validation) , "loo" (i.e., eave-one-out), "bootstrap", and "holdout".
Folds	num. Folds of Cross-validation resampling. It is ranging 2-10.
Ratio	num. Ratio of Bootstrap resampling. It is ranging 0.4-0.9.
Repeats	num. Number of Bootstrap resampling. It is ranging 2-20.

Value

A list containing the prediction performance evaluation.

Author(s)

Bin Wang

Examples

```
res <- InitStatCros()
   res1 = LoadStatCros(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatCros(PID = res$PID, OutPath = "default",TransDummyVars="default")
   res3 = CrosPred(PID=res$PID, VarsY = "Y1", VarsX = "X5,X6,X7,X8,X9,X10,X11",
   PredType = "response",VarsSel = FALSE,VarsSelThr = 0.1,Covariates = "all.c",
   RsmpMethod = "cv" ,Folds = 5,Ratio = 0.667,Repeats = 5)
   FuncExit(PID = res$PID)
```

DescComp *Size comparison for groups.*

Description

Perform size comparisons between groups.

Usage

```
DescComp(PID = res$PID, Task = "mean", Vars = "X5,X6,X7,X8,X9", VarsBy = "Y1" , Method = "wilcox",
   Layout = "density" , Brightness = "dark" , Palette = "default3").
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatDesc.
OutPath	chr. Output file directory, e.g. "D:/StatDesc/DescComp". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
Task	chr. Comparison task. At present, only the "mean" comparison is available.
Vars	chr. Target variables used for modelling. It should be noted that there is fixed format for the entering characters separated with "," and without space, e.g. "X5,X6,X7,X8,X9". The default values is "all" (all variables are included).
VarsBy	chr. Variable used to group the observation for size comparison.
Method	chr. Comparison method. At present, only "wilcox" (Wilcoxon rank sum test) is available.
Layout	chr. Visualization layout. Available values include "column.points", "density".
Brightness	chr. Visualization brightness. Available values include "light" and "dark".
Palette	chr. Visualization palette. Available values include "default1", "default2" and 6 journal option including "cell", "nature", "science", "lancet", "nejm" and "jama".

Value

A list object containing the results of size comparisons between groups for variables and visualization of the results.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStatDesc()
   res1 = LoadStatDesc(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatDesc(PID = res$PID)
  res3 = DescComp(PID=res$PID, OutPath="default",Task = "mean", Vars = "X5,X6,X7,X8,X9", VarsBy = "Y1",
  Method = "wilcox", Layout = "column.points",Brightness = "light" , Palette = "default1")
   FuncExit(PID = res$PID)
```

DescCorr *Correlation analysis.*

Description

Perform correlation analysis between variables.

Usage

```
DescCorr(PID = res$PID, VarsX = "X5,X6,X7,X8,X9", VarsY = "Y1", VarsBy = "Y1", Method = "pearson",
   Layout = "bubble", Brightness = "dark", Palette = "nature").
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatDesc.
OutPath	chr. Output file directory, e.g. "D:/StatDesc/DescCorr". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
VarsX	chr. Target variables used for modelling. It should be noted that there is fixed format for the entering characters separated with "," and without space, e.g. "X5,X6,X7,X8,X9". The default values is "all.x" (all variables are included).
VarsY	chr. Target outcome variables used for correlation analysis.
VarsBy	chr. Variable used to group the observation for correlation analysis.
Method	chr. Method for orrelation analysis. Available values include "spearman" (Spearman's rank correlation analysis) and "pearson" (Pearson correlation analysis).
Layout	chr. Visualization layout. Available values include "heatmap", "bubble", "matrix".
Brightness	chr. Visualization brightness. Available values include "light" and "dark".
Palette	chr. Visualization palette. Available values include "default1", "default2" , and 6 journal option including "cell", "nature", "science", "lancet", "nejm" and "jama".

Value

A list object containing the results of correlation analysis between variables and visualization of the results.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStatDesc()
   res1 = LoadStatDesc(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatDesc(PID = res$PID)
  res3 = DescCorr(PID = res$PID, VarsX = "X5,X6,X7,X8,X9", VarsY = "Y1", VarsBy = "Y1", Method = "spearman", Lay
   Brightness = "light", Palette = "default1")
   FuncExit(PID = res$PID)
```

`DescExtre` *Extreme value calculation.*

Description

Extreme value calculation for numeric variables.

Usage

```
DescExtre(PID=res$PID, Vars = "X5,X6,X7,X8,X9", LimitLow = 0.025, LimitUpper = 0.975,
   Layout = "column.points", Brightness = "light", Palette = "default2").
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by InitStatDesc.
`OutPath`	chr. Output file directory, e.g. "D:/StatDesc/DescExtre". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
`Vars`	chr. Target variables used for modelling. It should be noted that there is fixed format for the entering characters separated with "," and without space. The default values is "all" (all variables are included).
`LimitLow`	num. Lower limit ratio to screen the small extreme values located from 0 to this lower limit of the target variables.
`LimitUpper`	num. Upper limit ratio to screen the large extreme values located from this lower limit to 1 of the target variables.
`Layout`	chr. Visualization layout . Available values include "column.points", "heatmap".
`Brightness`	chr. Visualization brightness . Available values include "light" and "dark".
`Palette`	chr. Visualization palette . Available values include "default1", "default2" and 6 journal option including "cell", "nature", "science", "lancet", "nejm" and "jama".

Value

A list object containing the results of extremum for variables and visualization of the results.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStatDesc()
   res1 = LoadStatDesc(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatDesc(PID = res$PID)
  res3 = DescExtre(PID = res$PID, Vars = "X5,X6,X7,X8,X9", LimitLow = 0.025, LimitUpper = 0.975,
   Layout = "column.points", Brightness = "light", Palette = "default1")
   FuncExit(PID = res$PID)
```

`DescNorm` *Normality test.*

Description

Normality test for numeric variables.

Usage

```
DescNorm(PID = res$PID, Vars = 'X5,X6,X7,X8,X9', Method = "shapiro.test", Layout = "rose.chart" ,
   Brightness = "light", Palette = "default3").
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by InitStatDesc.
`OutPath`	chr. Output file directory, e.g. "D:/StatDesc/DescNorm". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
`Vars`	chr. Target variables used for modelling. It should be noted that there is fixed format for the entering characters separated with "," and without space, e.g. "X7,X8,X9,X10,X11". The default values is "all" (all variables are included).
`Method`	chr. Normality test method. Only "shapiro.test" method is available at present.
`Layout`	chr. Visualization layout. Available values include "column", "column.points", "rose.chart", and "density".
`Brightness`	chr. Visualization brightness. Available values include "light" and "dark".
`Palette`	chr. Visualization palette. Available values include "default1", "default2" and 6 journal option including "cell", "nature", "science", "lancet", "nejm" and "jama".

Value

A list object containing the results of normality test for variables and visualization of the results.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStatDesc()
   res1 = LoadStatDesc(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatDesc(PID = res$PID)
  res3 = DescNorm(PID=res$PID, Vars = 'X5,X6,X7,X8,X9', Method = "shapiro.test", Layout = "column" ,
   Brightness = "light", Palette = "default2")
   FuncExit(PID = res$PID)
```

DescSize *Variable description.*

Description

Calculate median, mean .etc for numeric variables,frequency and count for Discrete variable.

Usage

```
DescSize(PID=res$PID, Vars = "C1,C2,X5,X6,X7,X8,X9", VarsBy = "Y1", Layout = "box" , Brightness = "d
   Palette ="science").
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatDesc.
OutPath	chr. Output file directory, e.g. "D:/StatDesc/StatDesc". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
Vars	chr. Target variables used for modelling. It should be noted that there is fixed format for the entering characters separated with "," and without space, e.g. "X4,X5,X6,X7,X8". The default values is "all" (all variables are included).
VarsBy	chr. Variable used to group the observation for size description.
Layout	chr. Visualization layout. Available values include "box", "violin".
Brightness	chr. Visualization brightness. Available values include "light" and "dark".
Palette	chr. Visualization palette. Available values include "default1", "default2" and 5 journal option including "jama", "nature", "science", "lancet", "nejm".

Value

A list object containing the results of variable description for continuous and discrete variables respectively and visualization of the continuous variables.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStatDesc()
   res1 = LoadStatDesc(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatDesc(PID = res$PID)
   res3 = DescSize(PID = res$PID, Vars = "X5,X6,X7,X8,X9", VarsBy = "Y1", Layout = "box",
   Brightness = "light", Palette = "default1")
   FuncExit(PID = res$PID)
```

DescTable1 *Create Table 1 for StatDesc Module.*

Description

Create Table 1 for different epidemilogical study designs.

Usage

```
DescTable1(PID = res$PID,EpiDesign = "cohort", VarsY = "Y1", VarsC = "C1,C2,C3,C4,C5,C6",
   Missing = "ifany").
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatDesc.
OutPath	chr. Output file directory, e.g. "D:/StatDesc/DescTable1". It should be noted that the slash symbol is "/", not "\".If "default", the current working directory will be set.
EpiDesign	chr. Research types provided for users, include "cohort" , "case.control", "cross.sectional".
VarsY	chr. Outcome variable used for modelling. Only one variable can be entered.
VarsC	chr. Covariate variables needing further statistical test. It should be noted that there is fixed format for the entering characters separated with "," and without space, e.g. "C1,C2,C3,C4,C5". The defaults value is all covariate variables listed in the data file, which can be entered with "all.c".
Missing	chr. Counts of missing values in the table, available options include are "no" (never display missing values), "ifany" (only display if any missing values), and "always" (includes missing count row for all variables). Default is "ifany".

Value

A list object containing standardized table 1

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStatDesc()
   res1 = LoadStatDesc(PID = res$PID, UseExample = "example#1")
   res2 =  TidyStatDesc(PID = res$PID)
  res3 = DescTable1(PID = res$PID, EpiDesign = "cohort", VarsY = "Y1", VarsC = "C1,C2,C3,C4,C5,C6", Missing = "i
   FuncExit(PID = res$PID)
```

EBIOBiolink *Build the biological link*

Description

Build the biological link between the exposures and diseases

Usage

```
EBIOBiolink(PID, OutPath="default",Mode, MetBiospec="blood")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEBIO.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Mode	chr. Method to build the biological link between exposures and diseases. Available options include "EPPD" (protein-protein interaction), "EGoD" (gene ontoloty), "EPD" (protein), "EGeD" (Gene), "EPaD" (pathway).
MetBiospec	chr. Biological sample matrix for the metabolome analysis. Options include "blood" and "urine".

Value

A list object containing the edges and nodes of the biological link.

Author(s)

Mingliang Fang, Tianxiang Wu, Bin Wang,(corresponding author)

Examples

```
res = InitEBIO()
   res1 = LoadEBIO(PID = res$PID,UseExample = "example#1")
   res2 = EBIOConvToExpoID(PID = res$PID)
  res3 = EBIOBiolink(PID = res$PID, OutPath="default", Mode = "EGoD", MetBiospec = "blood")
   FuncExit(PID = res$PID)
```

EBIOConvToExpoID *Convert different IDs to the unified ExposomeX IDs*

Description

Convert the IDs of exposure, chemicals, metabolites, or proteins to the unified ExposomeX ID, i.e., unified identifier in ExposomeX platform

Usage

```
EBIOConvToExpoID(PID,OutPath)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEBIO.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.

Value

A data frame containing the converted ID information.

Author(s)

Mingliang Fang, Tianxiang Wu, Bin Wang,(corresponding author)

Examples

```
res = InitEBIO()
   res1 = LoadEBIO(PID = res$PID,UseExample = "example#1")
   res2 = EBIOConvToExpoID(PID = res$PID,OutPath = "default")
   FuncExit(PID = res$PID)
```

EDBLink *Find the nexuses between keywords*

Description

Find the nexuses between the keywords in ExposomeX platform. Nexus direction from LinkFrom (class A) to LinkTo (class B)

Usage

```
EDBLink(PID, OutPath = "default", ClassA, ClassB, LinkFrom, LinkTo)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by EDB
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
ClassA	chr. Find the Links between the keywords in ClassA and ClassB. Options include "Exposure","Protein","GO","Pathway","Gene"
ClassB	chr. Options include "Protein","GO","Disease","Pathway","Gene"
LinkFrom	chr. Name, alias, and ID of exposure, protein, GO, gene and pathway are all included. "Default" means using the values in the input data file.
LinkTo	chr. Name, alias, and ID of protein, GO, gene, pathwayand disease are all included. "Default" means using the values in the input data file.

Value

A data frame

Author(s)

Tianxiang Wu, Bin Wang (corresponding author)

Examples

```
res <- InitEDB()
   res1 = LoadEDB(PID = res$PID, UseExample = "example#1")
   res2 = EDBLink(PID=res$PID,OutPath ="default",ClassA = "Exposure",
       ClassB = "Protein",LinkFrom = "default",LinkTo = "default")
   FuncExit(PID = res$PID)
```

EDBNode *Explain keywords*

Description

Explain the keyword in ExposomeX platform

Usage

```
EDBNode(PID, OutPath = "default", Class, Nodes)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by EDB
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Class	chr. Choose the search range of the concerned keywords. Options include “Exposure” (In addition to containing chemicals, some exposure factors may be a mixture of chemicals, e.g. PM2.5, tobacco smoking.), "Metabolite" (the chemicals used for metabolome analysis in the KEGG database), "Protein", "Enzyme" (referring in particular to the enzymes in the KEGG database), "Disease", "GO" (gene ontology), and “Gene".
Nodes	chr. Any keywords to search. "default" means using the values in the input data file. Users can also enter the keywords here. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "7440-43-9,OMIM:619217,ARF5,GO:0010942".

Value

A data frame

Author(s)

Tianxiang Wu, Bin Wang (corresponding author)

Examples

```
res <- InitEDB()
   res1 = LoadEDB(PID = res$PID, UseExample = "example#1")
   res2 = EDBNode(PID=res$PID,OutPath = "default",Class = "GO",Nodes = "default")
   FuncExit(PID = res$PID)
```

EVIZCateDot *Plot category dot*

Description

Visualize data via dot plot.

Usage

```
EVIZCateDot(PID,OutPath,Vars, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEVIZ.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	chr. Specifying the variables. Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10"). No more than 50 variables be entered is recommended (< 50 variables).
Parameter	chr. Specifying which parameter of the data to be the ordinate of the output plot. Available options include: "mean", "median", "min", "max", "mad" or "sd".Default is "mean".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Details

The dot plot is used to display the relative position of two data points in the same time period, or compare the difference between the two categorical variables.

Value

A list containing the dot plot. "light_default1": the visualization result.

Author(s)

Ning Gao,Bin Wang(corresponding author)

Examples

```
res = InitEVIZ()
    res1 = LoadEVIZ(PID = res$PID, UseExample = "example#1")
  res2 = EVIZCateDot(PID=res$PID,Vars="X4,X5,X6,X7,X8,X9,X10",Parameter="mean",Brightness="light",Palette="
```

EVIZCompoDendrogram *Plot component dendrogram*

Description

Visualize data via dendrogram plot.

Usage

```
EVIZCompoDendrogram(Vars, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEVIZ.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	chr. Variables to be visualized(e.g."X4,X5,X6,X7,X8,X9,X10").Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10").
Parameter	chr. Specifying which parameter of the data to be the ordinate of the output plot. Available options include: "mean", "median", "min", "max", "mad" or "sd".Default is "mean".
DistMethod	chr.The distance measure. This must be one of "euclidean", "maximum" or "manhattan".Default is "euclidean".
ClusterMethod	chr.The agglomeration method. This should be one of "ward.D", "ward.D2" or "single".Default is "ward.D".
ClusterNum	num. The number of groups for cutting the tree.Default is 4.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Details

The dendrogram plot is used to plot beautiful dendrograms.

Value

A list containing one plot. (1) "light_default1": the whole dataset visualization result.

Author(s)

Ning Gao,Bin Wang(corresponding author)

Examples

```
res = InitEVIZ()
    res1 = LoadEVIZ(PID = res$PID, UseExample = "example#1")
  res2= EVIZCompoDendrogram(PID=res$PID,Vars = "X4,X5,X6,X7,X8,X61,X66,X67,X200",Parameter = "median",DistMe
```

EVIZDistrSierra *Plot distribution sierra*

Description

Visualize data via sierra plot.

Usage

```
EVIZDistrSierra(PID,OutPath,Vars, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEVIZ.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	chr. Variables to be visualized(e.g."X4,X5,X6,X7,X8,X9,X10").Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10"). No more than 50 variables be entered is recommended (< 20 variables).
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Details

The sierra plot is used to visualize the kernel density estimation of data.

Value

A list containing one plot. (1) "light_default1": the whole dataset visualization result.

Author(s)

Ning Gao,Bin Wang(corresponding author)

Examples

```
res = InitEVIZ()
    res1 = LoadEVIZ(PID = res$PID, UseExample = "example#1")
  res2 = EVIZDistrSierra(PID=res$PID,Vars = "X14,X15,X16,X17,X18,X19,X20",Brightness = "light",Palette = "def
```

EVIZRelatEdgeBundling *Plot elationship edge bundling*

Description

Visualize data via edge bundling plot.

Usage

```
EVIZRelatEdgeBundling(PID,OutPath,VarsY,VarsX, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEVIZ.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Dependent variables for visualization(e.g."Y2").
VarsX	chr. Independent variables.Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10").
Family	chr. Specifying the data distribution in order to determine the visualization method. Available options include:"gaussian","poisson" and "logistic".Notice that the family are determined by data type of an outcome, or the plot can not be visualized.
SizeFor	chr. Parameter to represent the size of the points in the output plot. Available options include "pvalue" and "beta".The default option is "pvalue".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).
CutOff	num. Partial outcomes to visualize which is determined by correlation coefficient r. The range must between 0 and 1.

Details

The edge bundling plot is used to bundle the edges closely in order to reduce complexity.

Value

A list containing one plot. (1) "light_default1": the visualization result;

Author(s)

Ning Gao,Bin Wang(corresponding author)

Examples

```
res = InitEVIZ()
    res1 = LoadEVIZ(PID = res$PID, UseExample = "example#1")
  res2 = EVIZRelatEdgeBundling(PID=res$PID,VarsY = "Y2",VarsX = "all.x",Family = "gaussian" ,SizeFor = "pvalue
```

`EVIZRelatHeatmap` *Plot relationship heatmep*

Description

Visualize data via heatmap plot.

Usage

```
EVIZRelatHeatmap(PID,OutPath,Vars, ...)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by InitEVIZ.
`OutPath`	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
`Vars`	chr. Independent variables.Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10").
`Method`	chr. Method to calculate the correlation. Default option is "spearman".
`Brightness`	chr. Visualization brightness. Available options include "light" and "dark".
`Palette`	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Details

The heatmap plot is used to display data in color changes as a matrix.

Value

A list containing one plot. (1) "light_default1": the whole dataset visualization result.

Author(s)

Ning Gao, Bin Wang (corresponding author)

Examples

```
res = InitEVIZ()
    res1 = LoadEVIZ(PID = res$PID, UseExample = "example#1")
  res2 = EVIZRelatHeatmap(PID=res$PID,Vars = "X1,X4,X5,X6,X7,X8,X9,X10",Method = "spearman",Brightness = "lig
```

EVIZRelatMatrix *Plot relationship matrix*

Description

Visualize data via matrix plot.

Usage

```
EVIZRelatMatrix(PID,OutPath,Vars, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEVIZ.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	chr. Independent variables.Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10").
Method	chr. Method to calculate the correlation. Default option is "spearman".

Details

The matrix plot is used to make a matrix of plots with a given data set.

Value

A list containing one plot. (1) "light_default1": the whole dataset visualization result.

Author(s)

Ning Gao, Bin Wang (corresponding author)

Examples

```
res = InitEVIZ()
    res1 = LoadEVIZ(PID = res$PID, UseExample = "example#1")
   res2 = EVIZRelatMatrix(PID=res$PID,Vars = "X4,X5,X6,X7,X8,X9,X10",Method = "spearman")
```

EVIZRelatNetwork *Plot relationship network*

Description

Visualize data via network plot.

Usage

```
EVIZRelatNetwork(PID,OutPath,VarsY,VarsX, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEVIZ.
OutPath	chr. Output file directory. e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Specifying the dependent variables(e.g."Y2").
VarsX	chr. Specifying the independent variables.Available options include: "all.x", all independent variables; or input a character string specifying the variables,separated by comma "," without space(e.g."X4,X5,X6,X7,X8,X9,X10").
Family	chr. Specifying the data distribution in order to determine the visualization method. Available options include:"gaussian","poisson" and "logistic".Notice that the family are determined by data type of an outcome, or the plot can not be visualized.
Layout	chr. Visualization layout. Available options include "force-directed" and "degree-circle".
CutOff	num. Partial outcomes to visualize which is determined by correlation coefficient r. The range must between 0 and 1.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Details

The network plot is used to visualize the relationship between the input variables by using ggraph package.

Value

A list containing one plot. (1) "light_default1": the visualization result;

Author(s)

Ning Gao,Bin Wang(corresponding author)

Examples

```
res = InitEVIZ()
    res1 = LoadEVIZ(PID = res$PID, UseExample = "example#1")
  res2 = EVIZRelatNetwork(PID=res$PID,VarsY="Y1",VarsX="all.x",Family="gaussian",Layout="force-directed",Cu
```

FuncExit *End the module analysis*

Description

End the module analysis

Usage

```
FuncExit(PID)
```

Arguments

PID	chr. Program ID. It should be the same with the PID generated by Init.

Value

Exit status

Author(s)

Bin Wang (corresponding author)

Examples

```
res = InitEBIO()
   FuncExit(PID = res$PID)
```

HEPPredBlood *Build StatHEP model*

Description

HEPPredBlood function is designed to integrate the multi-omic data to predict the incidence risk. It mainly aims to construct various stacked generalization models to predict the probability of outcome incidence, as well as providing the statistical explanation.

Usage

```
HEPPredBlood(PID, OutPath, MC, N)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatHEP.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
MC	lgl. T (or TRUE) and F (or FALSE). Whether to perform Mentocaro simulation.
N	num. Times of Mentocaro simulation.

Details

Users can easily get the modeling results and their visualization plots with high quality by following the detailed instructions in each step. Machine learning model is used to predict blood concentrations of chemicals and prioritize chemicals of health concern.

Value

A list containing StatHEP model results

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitStatHEP()
   res1 <- LoadStatHEP(PID=res$PID, UseExample = "example#1", DataPath = NULL)
   res2 <- HEPPredBlood(PID=res$PID, OutPath = "default", MC = "T", N = 100)
   FuncExit(PID = res$PID)
```

ImptM *Estimate the importance of mediators*

Description

Estimate the importance of mediators.

Usage

```
ImptM(PID, OutPath, VarsY, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
VarsY	chr. The outcome variable. Either continuous, binary or count type is permitted.
VarsX	chr. Exposure variables included in estimation procedure. When "default" is specified, all exposure variables in the data will be used. The shirnkaged exposure variables are also permitted. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "X1,X2,X3".
VarsC	chr. Covariates included in estimation procedure. When "default" is specified, all covariates in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "C1,C2".

Details

ImptM function estimates the importance of mediators among given exposures. The estimation procedure is mainly realized by the hima function in HIMA package. ImptM also provides another evaluation method by the bama function in bama package. The users can search for these packages for further information.

Value

A list containing four elements where the raw estimation results as well as the tidy tables for display are stored. The elements of that list include:

1. "MedtImptM_all": the raw estimation result where the importance of each mediator was evaluated among whole mediators.
2. "MedtImptM_list": the raw estimation result where the importance of each mediator was evaluated among each mediator group.
3. "MedtImptM_table1": the tidy table for MedtImptM_all and MedtImptM_list, where significant estimations (q value <0.2) are expressed with an asterisk mark (*).
4. "MedtImptM_table2": the tidy table for MedtImptM_all and MedtImptM_list, where only significant estimations (q value <0.2) are displayed in this table. MedtImptM_table2 is as same content as MedtImptM_table1.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMedt()
   res1 <- LoadStatMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   res2 = TidyStatMedt(PID = res$PID, OutPath = "default",LogTransM = "log10",
   RangeLow = 0,RangeUpp = 1)
   res3 <- RedX(PID=res$PID, VarsY = "Y1",
   VarsC = "default", VarsM = "default", Method = "mean", Folds = 10,
   Family = "linear", Iter = 500)
   res4 <- ImptM(PID = res$PID, VarsY = "Y1",
   VarsX = "default", VarsC = "default")
   FuncExit(PID = res$PID)
```

InitEBIO *Initialize EBIO module*

Description

Initialize EBIO module analysis. It can generate an R6 class object integrating all the analysis information.

Usage

```
InitEBIO()
```

Details

EBIO module is designed to find the biological relationships between exposure factors and health outcome. This module adopts the most frequently-used and authoritative databases, e.g., T3DB, CTD, ToxCast, StringDB, STITCH, KEGG, and GO.

Value

An R6 class object.

Author(s)

Mingliang Fang, Tianxiang Wu, Bin Wang,(corresponding author)

Examples

```
res <- InitEBIO()
   FuncExit(PID = res$PID)
```

InitEDB *Initialize EDB module*

Description

Initialize EDB module analysis. It can generate an R6 class object.

Usage

```
InitEDB()
```

Details

EDB module is designed as a convenient tool to explore the data, as well as facilitating to find the biological relationship between exposure and diseases from the perspective of bioinformatics. This module adopts the most frequently-used and authoritative databases, e.g., T3DB, CTD, ToxCast, StringDB, STITCH, KEGG, and GO.

Value

An R6 class object.

Author(s)

Tianxiang Wu, Bin Wang (corresponding author)

Examples

```
res <- InitEDB()
    FuncExit(PID = res$PID)
```

InitEMETA *Initialize EMETA Module*

Description

Initialize EMETA module, the first step to start EMETA Module.

Usage

```
InitEMETA()
```

Details

EMETA module mainly provides users preliminary information retrieval and screening for meta-analysis. Run InitEMETA to get a unique PID for following steps.

Value

An R6 class object.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res = InitEMETA()
   FuncExit(PID = res$PID)
```

InitEMS *Initialize EMS module*

Description

Initialize EMS module analysis. It can generate an R6 class object.

Usage

```
InitEMS()
```

Details

ExpoNontarget module is designed to conduct the analysis of the features from the high-resolution mass spectrometry. It mainly aims to screen and annotate the significant features associated with the health outcomes.

Value

An R6 class object.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitEMS()
   FuncExit(PID = res$PID)
```

InitEVIZ *Initialize EVIZ module*

Description

Initialize EVIZ module analysis. It can generate an R6 class object integrating all the analysis information.

Usage

```
InitEVIZ()
```

Details

EVIZ module is designed for the data visualization of different statistical and biological analyses in a user friendly and easy way, including four typical classes of visualization.

Value

An R6 class object.

Author(s)

Ning Gao, Bin Wang (corresponding author)

Examples

```
res <- InitEVIZ()
```

InitStatCros *Initialize StatCros module*

Description

Initialize StatCros module analysis. It can generate an R6 class object.

Usage

```
InitStatCros()
```

Details

StatCros module was designed to analyze the cross-sectional data from exposome-wide association study (EWAS). This data structure can be obtained from the epidemiological designs of cross-section, case-control, and cohort.

Value

An R6 class object.

Author(s)

Bin Wang

Examples

```
res <- InitStatCros()
    FuncExit(PID = res$PID)
```

`InitStatDesc` *Initialize StatDesc Module.*

Description

The first step to start StatDesc Module.

Usage

```
InitStatDesc().
```

Value

An R6 class object.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStatDesc()
    FuncExit(PID = res$PID)
```

InitStatHEP *Initialize StatHEP module*

Description

Initialize StatHEP analysis. It can generate an R6 class object integrating all the analysis information.

Usage

```
InitStatHEP()
```

Details

Quantification of all trace organics in the biological fluids seems impossible and costly, regardless of the high individual exposure variability. We hypothesized that the blood concentration of organic pollutants could be predicted via their exposure and chemical properties. Developing a prediction model on the annotation of chemicals in human blood can provide new insight into the distribution and extent of exposures to a wide range of chemicals in humans. Our objective of this module is to develop a machine learning model (eg., random forest) to predict blood concentrations of chemicals and prioritize chemicals of health concern. Please see the details in Zhao et al., Environ Health Perspect. 2023 Mar;131(3):37009.

Value

An R6 class object.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitStatHEP()
   FuncExit(PID = res$PID)
```

InitStatIP *Initialize StatIP module*

Description

Initialize StatIP module analysis. It can generate an R6 class object.

Usage

```
InitStatIP()
```

Details

StatIP module is designed to find the statistical relationships between exposure factors and health outcome.

Value

An R6 class object.

Author(s)

Bin Wang

Examples

```
res <- InitStatIP()
  FuncExit(PID = res$PID)
```

`InitStatMedt` *Initialize StatMedt module*

Description

Initialize StatMedt module analysis. It can generate an R6 class object integrating all the analysis information.

Usage

```
InitStatMedt()
```

Details

InitStatMedt uses R6 package to generate an R6 class object where parameters to be used for the following mediation module program are initialized and save in that object. All executed function codes in the ExpoMediaiton packaged will be recorded in the form of log text in that object. Furthermore, a program ID (i.e., PID) is randomly created for the users to identify their own program.

Value

An R6 class object.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMedt()
   FuncExit(PID = res$PID)
```

InitStatMix *Initialize StatMix module*

Description

Initialize StatMix module analysis. It can generate an R6 class object.

Usage

```
InitStatMix()
```

Details

StatMix module is designed to analyze mixture effect of the various exposure factors. It mainly aims to screen the representative features with high contribution to the health outcome, as well as their potential interaction effect.

Value

An R6 class object.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   FuncExit(PID = res$PID)
```

InitStatMO *Initialize StatMO module*

Description

Initialize StatMO analysis. It can generate an R6 class object integrating all the analysis information.

Usage

```
InitStatMO()
```

Details

StatMO module is designed to integrate the multi-omic data to predict the incidence risk. It mainly aims to construct various stacked generalization(SG) models to predict the probability of outcome incidence, as well as providing the statistical explanation. In addition, the module can provide visualization plots with high quality of the final calculation results to make it easier for users to understand.

Value

An R6 class object.

Author(s)

Guohuan Zhang, Yuting Wang, Bin Wang (corresponding author)

Examples

```
res <- InitStatMO()
   FuncExit(PID = res$PID)
```

`InitStatPanel` *Initialize StatPanel module*

Description

Initialize StatPanel module analysis. It can generate an R6 class object.

Usage

```
InitStatPanel()
```

Details

StatPanel module is designed to conduct the analysis of the panel data. It mainly aims to evaluate the associations between exposure factors and the health outcome.

Value

An R6 class object.

Author(s)

Bin Wang (corresponding author)

Examples

```
res <- InitStatPanel()
    FuncExit(PID = res$PID)
```

InitStatSurv *Initialize StatSurv module*

Description

Initialize StatSurv module analysis. It can generate an R6 class object integrating all the analysis information.

Usage

```
InitStatSurv()
```

Details

InitStatSurv uses R6 package to generate an R6 class object where parameters to be used for the following mediation module program are initialized and save in that object. All executed function codes in the StatSurv packaged will be recorded in the form of log text in that object. Furthermore, a program ID (i.e., PID) is randomly created for the users to identify their own program.

Value

An R6 class object.

Author(s)

Changxin Lan, Bin Wang(corresponding author)

Examples

```
res <- InitStatSurv()
   FuncExit(PID = res$PID)
```

InitStatTidy *Initialize StatTidy module*

Description

Initialize StatTidy module analysis. It can generate an R6 class object.

Usage

```
InitStatTidy()
```

Details

It is designed to tidy the data for the target model analysis.

Value

An R6 class object.

Author(s)

Bin Wang

Examples

```
res <- InitStatTidy()
     FuncExit(PID = res$PID)
```

LoadEBIO *Load data file for EBIO module*

Description

Load data file for EBIO module.

Usage

```
LoadEBIO(PID, UseExample = "default", DataPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEBIO.
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input file directory, e.g. "D:/test/eg_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Mingliang Fang, Tianxiang Wu, Bin Wang,(corresponding author)

Examples

```
res <- InitEBIO()
   res = LoadEBIO(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

LoadEDB *Load data file for EDB module*

Description

Load data file for EDB module

Usage

```
LoadEDB(PID, UseExample = "default", DataPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by EDB
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Tianxiang Wu, Bin Wang (corresponding author)

Examples

```
res <- InitEDB()
   res1 = LoadEDB(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

LoadEMETA *Load Data for EMETA Module*

Description

Upload local data file for EMETA Module.

Usage

```
LoadEMETA(PID = NULL, UseExample = "default", DataPath = NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEMETA.
UseExample	chr. A character indicates whether uses example data for analyses, available option include "example#1" for using example data1 and "default" for using data uploaded.
DataPath	chr. Input file directory, e.g. "D:/test/eg_meta.xlsx". It should be noted that the slash symbol is "/", not "\".

Details

After initializing EMETA module, the second step is to upload local data file for EMETA Module. Set param "UseExample = 'example#1'" to use example data1. LoadEMETA can only run successfully after successfully running InitEMETA. Please attention, PID must be got from the return result of InitEMETA().

Value

An R6 class object containing the input data.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res = InitEMETA()
   res1 = LoadEMETA(PID = res$PID,
   UseExample = "example#1",
   DataPath = NULL)
   FuncExit(PID = res$PID)
```

LoadEMS *Load data file for EMS module*

Description

Load data file for EMS module

Usage

```
LoadEMS(PID, UseExample = "default", MS2, DataPath=NULL, VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by EMS
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1","example#2" provided by this module.
MS2	lgl.Whether to use MS2 information to annotate the target ion. If use example data, choose "T" to use "example#2" data with MS2 information, choose "F" to use "example#1" data without MS2 information.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitEMS()
   res1 = LoadEMS(PID = res$PID, UseExample = "example#1", MS2 = FALSE)
   FuncExit(PID = res$PID)
```

LoadEVIZ *Load data file for EVIZ module*

Description

Load data for visualization.

Usage

```
LoadEVIZ(PID, UseExample = "default", FileDirIn=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEVIZ.
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input data file directory, e.g. "D:/test/eg_expoviz_data.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input vocabulary file directory, e.g. "D:/test/eg_expoviz_voca.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Ning Gao, Bin Wang (corresponding author)

Examples

```
res = InitEVIZ()
    res1 = LoadEVIZ(PID = res$PID, UseExample = "example#1")
```

LoadStatCros *Load data file for StatCros module*

Description

Load data file for StatCros module

Usage

```
LoadStatCros(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatCros
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_StatCros.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_StatCros.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitStatCros()
   res1 = LoadStatCros(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

LoadStatDesc *Load data for StatDesc Module.*

Description

Upload data file for StatDesc Module.

Usage

```
LoadStatDesc(PID, UseExample = "default", DataPath = NULL, VocaPath = NULL).
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatDesc.
UseExample	chr. Whether uses example data for analyses, available option include "example#1" for using example data1 and "default" for using data.
DataPath	chr. Input file directory, e.g. "D:/test/StatDesc_data.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input file directory, e.g. "D:/test/StatDesc_voca.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

A list object containing imported data.

Author(s)

Yanqiu Feng, Bin Wang (corresponding author)

Examples

```
res = InitStatDesc()
   res1 = LoadStatDesc(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

LoadStatHEP *Load data file for StatHEP module*

Description

Upload data file for StatHEP module.

Usage

```
LoadStatHEP(PID, UseExample= "default", DataPath)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatHEP.
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input file directory, e.g. "D:/test/eg_data.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res <- InitStatHEP()
    res <- LoadStatHEP(PID=res$PID, UseExample = "example#1", DataPath = NULL)
   FuncExit(PID = res$PID)
```

LoadStatIP *Load data file for StatIP for module*

Description

Load data file for StatIP module

Usage

```
LoadStatIP(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatIP
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitStatIP()
   res = LoadStatIP(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

LoadStatMedt *Load data file for Mediation module*

Description

Load data file for Mediation module

Usage

```
LoadStatMedt(PID = PID, UseExample = "default", ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMedt.
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input data file directory, e.g. "D:/test/eg_Medt_data.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input vocabulary file directory, e.g. "D:/test/eg_Medt_data.xlsx". It should be noted that the slash symbol is "/", not "\".

Details

LoadStatMedt function loads the data file and the vocabulary file into the R6 object that InitStatMedt created. Noted that there are several data format requirments for the data and vocabulary file. For data file, the first three column must be named as "SampleID", "SubjectID" and "Group" in sequence. The "Gourp" variable should be a character variable to category data into two groups: "train" and "test" group. Outcome variables should be named as "Y*". e.g., Y1, Y2, Y3... Similarly, exposure variables should be named as "X*". e.g., X1, X2, X3..., and mediator variables should be named as "M*". e.g., M1, M2, M3... For vocabulary file, the first column should be a character variable named "SerialNo" indicating the names of outcome, exposure and mediator variables in the data file. These names should be consistent with the variable names in data file. The second column should be a character variable named "FullName" indicating the full names (labels) of the variables. The third column should be a character variable named "SubgroupName" indicating the groups the exposure or mediator variables belong to.

Value

An R6 class object containing the input data and vocabulary file.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMedt()
   res1 <- LoadStatMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   FuncExit(PID = res$PID)
```

LoadStatMix *Load data file for StatMix module*

Description

Load data file for StatMix module

Usage

```
LoadStatMix(PID, UseExample = "default", DataPath = NULL, VocaPath = NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatMix
UseExample	chr. Method of uploading data. If "default",user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   res = LoadStatMix(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

LoadStatMO *Load data file for multiomics module*

Description

Upload data file for multiomics module.

Usage

```
LoadStatMO(PID, UseExample= "default", DataPath, VocaPath)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMO.
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input file directory, e.g. "D:/test/eg_data.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input file vocabulary, e.g. "D:/test/eg_voca.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Guohuan Zhang, Yuting Wang, Bin Wang (corresponding author)

Examples

```
res <- InitStatMO()
   res1 <- LoadStatMO(PID=res$PID, UseExample = "example#1", DataPath = NULL, VocaPath = NULL)
    FuncExit(PID = res$PID)
```

LoadStatPanel *Load data file for StatPanel module*

Description

Load data file for StatPanel module

Usage

```
LoadStatPanel(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatPanel
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitStatPanel()
   res = LoadStatPanel(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

`LoadStatSurv` *Load data file for Survival module*

Description

Load data file for Survival module

Usage

```
LoadStatSurv(PID = PID, UseExample = "default", ...)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by InitStatSurv.
`UseExample`	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
`DataPath`	chr. Input data file directory, e.g. "D:/test/eg_Surv_data.xlsx". It should be noted that the slash symbol is "/", not "\".
`VocaPath`	chr. Input vocabulary file directory, e.g. "D:/test/eg_Surv_voca.xlsx". It should be noted that the slash symbol is "/", not "\".

Details

LoadStatSurv function loads the data file and the vocabulary file into the R6 object that InitStatSurv created. Noted that there are several data format requirments for the data and vocabulary file. For data file,the first three columns should be named as "SampleID", "SubjectID", and "Group", respectively. For the "Group" variable, only two values can be used, i.e. "train" and "test". If there is no data for test, all values should be set as "train". For outcome variables, their initials must be set as "Y" and serialized by adding Arabic numerals if needed, e.g., Y1, Y2, Y3. In this module, the survival time (Y1) and status (Y2) must be provided.For exposure variables, their initials must be set as "X" and serialized by adding Arabic numerals if needed, e.g., X1, X2, X3. For covariate variables, their initials must be set as "C" and serialized by adding Arabic numerals if needed, e.g., C1, C2, C3. It should be noted the covariates are not required if users don't have. For vocabulary file, the first two columns must be named as "SerialNo" and "FullName", respectively. The list of SerialNo of outcomes, exposure, and covariates should be the same with the column names of "Data file". The list of the FullName is prepared as users' like.

Value

An R6 class object containing the input data and vocabulary file.

Author(s)

Changxin Lan, Bin Wang(corresponding author)

Examples

```
res <- InitStatSurv()
   res1 <- LoadStatSurv(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   FuncExit(PID = res$PID)
```

LoadStatTidy *Load data file for StatTidy module*

Description

Load data file for StatTidy module

Usage

```
LoadStatTidy(PID, UseExample = "default", DataPath=NULL, VocaPath=NULL)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatTidy
UseExample	chr. Method of uploading data. If "default", user should upload their own data files, or use "example#1" provided by this module.
DataPath	chr. Input directory of data file, e.g. "D:/test/eg_data_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".
VocaPath	chr. Input directory of vocabulary file, e.g. "D:/test/eg_voca_biolink.xlsx". It should be noted that the slash symbol is "/", not "\".

Value

An R6 class object containing the input data.

Author(s)

Bin Wang

Examples

```
res <- InitStatTidy()
   res = LoadStatTidy(PID = res$PID, UseExample = "example#1")
   FuncExit(PID = res$PID)
```

MetaAsso *Pool Effect Value*

Description

Pool effect value in our meta database besed on chemical ID and disease ID.

Usage

```
MetaAsso(PID = NULL, CID = "default", DID = "default", OutPath = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEMETA.
CID	chr. "default" or a chemical ID character (separate different values by ","). If "default", the function will use the Chemical_ID values in the file loaded by LoadEMETA. If a character (separate different values by ","), the function will use the chemical ID in the character instead. Chemical_ID refers to the target chemical ID which can be inchikey (eg. JIAARYAFYJHUJI-UHFFFAOYSA-L), cas.rn (eg. 7784-42-1) or our EXC ID (eg. EX:C01631).
DID	chr. "default" or a disease ID character (separate different values by ","). If "default", the function will use the Disease_ID values in the file loaded by LoadEMETA. If a character (separate different values by ","), the function will use the disease ID in the character instead. Disease_ID refers to the target disease ID which can be MESH ID (format like MESH:D006973), OMIM ID (format like OMIM:182940) or our EXD ID (eg. EX:D16243)
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set. It should be noted that the slash symbol is "/", not "\".

Details

MetaAsso provides the functions of effect value pooling. In the publishd papers, what is the effect value between X and Y? This questions can be solved by MetaAsso function. It can provide the combined results of fixed effect model and random effect model. (It can only search the papers available in EMETA database DB_Meta) Please attention, PID must be got from the return result of InitEMETA(). MetaAsso can only run successfully after successfully running InitEMETA and LoadEMETA functions.

Value

A list object containing forest plots.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res = InitEMETA()
   res1 = LoadEMETA(PID = res$PID,
   UseExample = "example#1",
   DataPath = NULL)
   res2 = MetaAsso(PID=res$PID,
   OutPath = "default",
   CID = "default",
   DID = "default")
   FuncExit(PID = res$PID)
```

`MetaRef` *Search or Download Articles' Main Information*

Description

Search or download articles' main information based on keywords.

Usage

```
MetaRef(PID = NULL, Mode = "search", VarX, VarY, VarM, YearFrom, YearEnd, PMID, OutPath)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by InitEMETA.
`Mode`	chr. Two modes are provided. "Search" for paper retrieval by keywords VarX/VarY/VarM/YearFrom/Y and "Download" for downloading main informain (main information only) for specified PMID.
`VarX`	chr. "default" or a chemical names character (separate different values by ","). If "default", the function will use the VarX values in the file loaded by LoadEMETA. If a character (separate different values by ","), the function will use the chemical names in the character instead.
`VarY`	chr. "default" or a disease names character (separate different values by ","). If "default", the function will use the VarY values in the file loaded by LoadEMETA. If a character (separate different values by ","), the function will use the disease names in the character instead.
`VarM`	chr. "default" or a mediating factor names character (separate different values by ","). If "default", the function will use the VarM values in the file loaded by LoadEMETA. If a character (separate different values by ","), the function will use the mediating factor names in the character instead.
`YearFrom`	chr. "default" or a disease names character (separate different values by ","). Limits the time range searched for "search" mode. If "default", the function will use the YearFrom values in the file loaded by LoadEMETA. If a character (separate different values by ","), the function will use the YearFrom in the character instead.
`YearEnd`	chr. "default" or a disease names character (separate different values by ","). Limits the time range searched for "search" mode. If "default", the function will use the YearEnd values in the file loaded by LoadEMETA. If a character (separate different values by ","), the function will use the YearEnd in the character instead.

PMID	chr. “default” or a disease names character (separate different values by ","). Refs to the papers PMID for "download" mode. If “default”, the function will use the PMID values in the file loaded by LoadEMETA. If a character (separate different values by ","), the function will use the PMIDs in the character instead.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set. It should be noted that the slash symbol is "/", not "\".

Details

MetaRef provides the functions of paper retrieval and relevance sorting, returning the information to the user based on keywords. Please attention, PID must be got from the return result of InitEMETA(). MetaRef can only run successfully after successfully running InitEMETA and LoadEMETA functions.

Value

A list object containing dataframe of articles’ information.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res = InitEMETA()
   res1 = LoadEMETA(PID = res$PID,
   UseExample = "example#1",
   DataPath = NULL)
   res2 = MetaRef(PID = res$PID,
   OutPath = "default",
   Mode = "search",
   VarX = "default",
   VarY = "default",
   VarM = "default",
   YearFrom = "default",
   YearEnd = "default",
   PMID = "default")
   FuncExit(PID = res$PID)
```

MetaRev *Review Relationship Between Exposure and Outcome*

Description

Review the relationship between exposure and outcome besed on chemical ID and disease ID.

Usage

```
MetaRev(PID = NULL, CID = "default", DID = "default", OutPath = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitEMETA.
CID	chr. "default" or a chemical ID character (separate different values by ","). If "default", the function will use the Chemical_ID values in the file loaded by LoadEMETA. If a character (separate different values by ","), the function will use the chemical ID in the character instead. Chemical_ID refers to the target chemical ID which can be inchikey (eg. JIAARYAFYJHUJI-UHFFFAOYSA-L), cas.rn (eg. 7784-42-1) or our EXC ID (eg. EX:C01631).
DID	chr. "default" or a disease ID character (separate different values by ","). If "default", the function will use the Disease_ID values in the file loaded by LoadEMETA. If a character (separate different values by ","), the function will use the disease ID in the character instead. Disease_ID refers to the target disease ID which can be MESH ID (format like MESH:D006973), OMIM ID (format like OMIM:182940) or our EXD ID (eg. EX:D16243)
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set. It should be noted that the slash symbol is "/", not "\".

Details

MetaRev provides the functions of literature review. In the publishd papers, how many recorded that X is a protective/risky factor for Y? This questions can be solved by MetaRev function. (It can only search the papers available in our database) Please attention, PID must be got from the return result of InitEMETA(). MetaRev can only run successfully after successfully running InitEMETA and LoadEMETA functions.

Value

A list object containing relationship visualization.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res = InitEMETA()
   res1 = LoadEMETA(PID = res$PID,
   UseExample = "example#1",
   DataPath = NULL)
   res2 = MetaRev(PID = res$PID,
   OutPath = "default",
   CID = "default",
   DID = "default")
   FuncExit(PID = res$PID)
```

MixBKMR *Build the Bayesian Kernel Machine Regression (BKMR) model*

Description

Build the Bayesian Kernel Machine Regression (BKMR) model

Usage

```
MixBKMR(PID, OutPath = "default", VarsY, VarsX, IncCova,
    Family, Group = F,Iter = 2000, qfixed = 0.5, qsbivar = "default",
    qsoverall = "default", qsdiff = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatMix
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
Covariates	chr. To determine covariates in models. Default is "all.c", where all candidate covariates will be included in models. Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "C1,C2,C3".
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable
Group	lgl. T (or TRUE) and F (or FALSE). Whether to use group indicators for fitting hierarchical variable selection. If "TRUE", the group name (GroupName) should be provided in the vocabulary data file.
Iter	num. Number of iterations for modeling. The default is 500. For more accurate and stable results, a minimum of 10,000 iteration is recommended.
qfixed	num. Quantile at which to fix the other predictors. The default is 0.5.
qsbivar	chr. Quantiles at which to fix the second variable. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., the default character sequence is "0.1,0.5,0.9"
qsoverall	num. Quantiles at which to calculate the overall risk summary. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., the default character sequence is "0.25,0.30,0.35,0.40,0.45,0.50,0.55,0.60,0.6
qsdiff	chr. Indicating the two quantiles for computing their effect difference. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., the default character sequence is "0.25,0.75".

Value

A list containing the BKMR analysis results.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   res1 = LoadStatMix(PID = res$PID, UseExample = "example#1")
   res2 <- TidyStatMix(PID = res$PID, LogTransM = "log10", RangeLow = 0, RangeUpp = 1)
   res3 = MixBKMR(PID = res$PID, VarsY = "Y1", VarsX = "X4,X5,X6,X7,X8,X9,X10",
   Covariates = "all.c", Family = "gaussian", Group = 'F', Iter = 2000, qfixed = 0.5,
   qsbivar = "default", qsoverall = "default", qsdiff = "default")
   FuncExit(PID = res$PID)
```

MixGcomp *Build g-computation model*

Description

Build g-computation model

Usage

```
MixGcomp(PID, VarsY, VarsX, Covariates = "all.c", Family, q = 10)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatMix
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
Covariates	chr. To determine covariates in models. Default is "all.c", where all candidate covariates will be included in models. Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "C1,C2,C3".
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable.
q	num. Levels for ranking mixture variables, e.g. in quartiles (q = 4), deciles (q = 10), or percentiles (q = 100). The default is 10.

Value

A list containing the WQS analysis results.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   res1 = LoadStatMix(PID = res$PID, UseExample = "example#1")
   res2 <- TidyStatMix(PID = res$PID, LogTransM = "log10", RangeLow = 0, RangeUpp = 1)
   res3 = MixGcomp(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
   Covariates = "all.c", Family = "gaussian", q = 10)
   FuncExit(PID = res$PID)
```

MixMLR *Build multiple linear regression (MLR) model*

Description

Build multiple linear regression (MLR) model

Usage

```
MixMLR(PID,OutPath = "default",VarsY,VarsX,IncCova = "F",
    SelMethod = "lasso",PredType = "response",Family)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatMix
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
Covariates	chr. To determine covariates in models. Default is "all.c", where all candidate covariates will be included in models. Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "C1,C2,C3".
SelMethod	chr. Method to select the important features to the final model. Options include "stepwise" (stepwise regression), "lasso" (Regularization regression of least absolute shrinkage and selection operator), and "enet" (Regularization regression of elastic net).
PredType	chr. Prediction type of the outcome variable, including "response" for the actual values and "prob" for outcome with binary variable.
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable

Value

A list containing the MLR analysis.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   res1 = LoadStatMix(PID = res$PID, UseExample = "example#1")
   res2 <- TidyStatMix(PID = res$PID, LogTransM = "log10", RangeLow = 0, RangeUpp = 1)
   res3 = res3 = MixMLR(PID = res$PID, VarsY = "Y1", VarsX = "all.x",
   Covariates = "all.c", SelMethod = "lasso", PredType = "response",
   Family = "gaussian")
   FuncExit(PID = res$PID)
```

MixWQS *Build weighted quantile sum regression (WQS) model*

Description

Build weighted quantile sum regression (WQS) model

Usage

```
MixWQS(PID, OutPath = "default", VarsY, VarsX, IncCova = "F", Family,
    VarStrat = "none",RatioValidat = 0.3,q = 10,b = 100,b1_pos = F,b1_constr = F)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatMix
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
Covariates	chr. To determine covariates in models. Default is "all.c", where all candidate covariates will be included in models. Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "C1,C2,C3".
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable.
VarStrat	chr. A factor variable used for stratifying for the model.
RatioValidat	num. Percentage of the dataset to be used to validate the model. If validation = 0 then the test dataset is used as validation dataset too. The default is 0.3.
q	num. Levels for ranking mixture variables, e.g. in quartiles (q = 4), deciles (q = 10), or percentiles (q = 100). The default is 10.
b	num. Number of bootstrap samples used in parameter estimation. The default is 100.

b1_pos	lgl. T (or TRUE) and F (or FALSE). Whether the beta values were positive to derive weights from to build models.
b1_constr	lgl. T (or TRUE) and F (or FALSE). A logial value that determines whether to apply positive (if b1_pos = TRUE) or negative (if b1_pos = FALSE) constraints in the optimization function for the weight estimation.

Value

A list containing the WQS analysis results.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   res1 = LoadStatMix(PID = res$PID, UseExample = "example#1")
   res2 <- TidyStatMix(PID = res$PID, LogTransM = "log10", RangeLow = 0, RangeUpp = 1)
   res3 = MixWQS(PID=res$PID, VarsY = "Y1", VarsX = "all.x",Covariates = "all.c",
   Family = "gaussian", VarStrat = "none", RatioValidat = 0.3, q = 10, b=100,
   b1_pos = 'F', b1_constr = 'F')
   FuncExit(PID = res$PID)
```

MulOmicsCros *Build multiomics model*

Description

MulOmicsCros function is designed to integrate the multi-omic data to predict the incidence risk. It mainly aims to construct various stacked generalization models to predict the probability of outcome incidence, as well as providing the statistical explanation.

Usage

```
MulOmicsCros(PID, OutPath, OmicGroups,VarsY, VarsC, TuneMethod = "default", TuneNum, RsmpMethod, Fo
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMO.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
OmicGroups	chr. Groups to be integrated. The groups of outcome and covariates or confounders are not included. Note that separates different learners by "," and without space(e.g. OmicGroups = "immunome,metabolome,proteome").
VarsY	chr. Outcome variable for modelling. Only one variable can be entered.
VarsC	chr. Covariates needing further statistical test. "all.c" option refers to all covariate variables listed in the data file. Users can also select part of them by copying available vars. Note that separates different vars by "," and without space(e.g. VarsC = "C1,C2").

TuneMethod	chr. Method for hyper-parameter autotuning. Options include "default", "random_search", "grid_search", "nloptr"(Non-linear optimization), and "gensa"(Generalized simulated annealing). The "default" option uses the simple training method for parameter optimization of mlr3 package.
TuneNum	num. Upper limit of model tuning times. It should be more than 20 times to search the appropriate parameters, but it takes more time. In theory, more time, better training results.
RsmpMethod	chr. Method for resampling. Options include "cv"(cross validation), "loo"(leave-one-out cross validation), "bootstrap"(bootstrapping), "holdout"(holdout).
Folds	num. Folds for cross validation resampling method. The default value is 5.
Ratio	num. Ratio for "Holdout" resampling method. The default value is 5.
Repeats	num. Repeats for "Bootstrap" resampling method.
VarsImpThr	num. Threshold for feature selection. It refers to the ratio of accumulated importance of all variables of the selected variables for building the final model.
SG_Lrns	chr. Learners for stacked generalization. Options include "lasso", "enet"(Elastic net), "rf"(Random forest), and "xgboost"(Xgboost). One or more arbitrary options can be selected at the same time. Note that separates different learners by "," and without space(e.g. SG_Lrns ="lasso,enet,rf,xgboost").

Details

The calculation time depends on the characteristics of your data, the number of learning methods, and the tuning method. For parameter "TuneMethod", the default option can provide faster calculations but less accurate results than other autotune methods. If you want to train a better model, choose other auto-tune method and increase the number of tuning times.

Value

An R6 class object containing eight elements. The elements of that object include: (1) "Importance": A list containing dataframes that contain the importance of features after modeling a single omic from different omicgroups. (2) "Feature": A list containing dataframes that contain the the coefficients or importance of selected features after modeling a single omic from different learners. (3) "Feature_select": A list containing dataframes that contain the selected features after modeling a single omic from different learners. (4) "ModelStat": A list containing dataframes that contain the r-square value of the single omic model built by different learners. (5) "Prediction_comp": A list containing dataframes that contain containing the prediction value of the SG model built by different combinations of learners. (6) "SGModel_summary": A dataframe containing the r-square value of the SG model built by different combinations of learners. (7) "NodeNum": The node number generated by different models. (8) "SGplot": A visualized plot for SG model summary.

Author(s)

Guohuan Zhang, Yuting Wang, Bin Wang (corresponding author)

Examples

```
res <- InitStatMO()
  res1 <- LoadStatMO(PID=res$PID, UseExample = "example#1", DataPath = NULL, VocaPath = NULL)
   res2 <- TidyStatMO(PID=res$PID, Vars = "X1,X2", To = "factor")
  res3 <- MulOmicsCros(PID=res$PID, OutPath = "default", OmicGroups = "immunome,metabolome",
  VarsY = "Y1", VarsC = "all.c", TuneMethod = "random_search", TuneNum = 2, RsmpMethod = "cv", Folds = 2,
   Ratio = 0.67, Repeats = 5, VarsImpThr = 0.85, SG_Lrns ="lasso,enet")
   FuncExit(PID = res$PID)
```

NtaAnno *Annotate the non-targeted features*

Description

Annotate the non-targeted features

Usage

```
NtaAnno(PID=res$PID,OutPath="default",VarsY = "Y1",VarsX = "default",VarsN = "single.factor",
  FdrCorrect = "F",AdductPos = "M+H",AdductNeg = "M-H",AccuracyMS1 = 5,MS2 = "F",
   AccuracyMS2 = 20,ScoreMS2 = 0.4,DiffRT = 5)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by EMS
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
FdrCorrect	lgl. T (or TRUE) and F (or FALSE). Whether to correct the multiple hypotheses by false positive rate (FDR) method.
AdductPos	chr. Adducts formed in the positive mode using LC-MS. The default is "all", i.e., all the adducts in positive mode will be chosen, including "M+3ACN+2H,M+ACN+H,M+NH4,M+2A H,M+ACN+Na,M+H+Na, M+2K-H,M+H+NH4,2M+K,M+K,2M+NH4,2M+Na,M+2Na,M+DMSO M+2H+Na,M+ACN+2H,M+H,2M+H,M+CH3OH+H,M+H+2Na,M+Na,2M+ACN+H,2M+ACN+N
AdductNeg	chr. Adducts formed in the negative mode using LC-MS. The default is "all", i.e., all the adducts in positive mode will be chosen, including "M+FA-H,M+Hac-H,M+Br,3M-H,2M+Hac-H,M+K-2H,2M+FA-H,M-H,M-H2O-H,M+Na-2H, M-2H,M+TFA-H,M+Cl,M-3H,2M-H"
AccuracyMS1	num. Accuracy threshold to match the target compounds of MS1 information. The unit is ppm. It is usually set ranging 1-20.
MS2	lgl. T (or TRUE) and F (or FALSE). Whether to use MS2 information to annotate the target ion.
AccuracyMS2	num. Accuracy threshold to match the target compounds of MS2 information. The unit is ppm. It is usually set ranging 1-20.
ScoreMS2	num. The lowest matching score ranging 0-1 using MS2 information. The default is 0.4.
DiffRT	num. Difference of the retention time between the measured and the theoretical prediction. The unit is second (s). It is usually set ranging 60-120.

Value

A list containing non-target analysis results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitEMS()
   res1 = LoadEMS(PID = res$PID, UseExample = "example#1", MS2 = FALSE)
   res2 = TidyEMS(PID = res$PID, OutPath = "default",TransDummyVars="default")
   res3 = NtaCros(PID=res$PID, OutPath="default",VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = "T", SelMethod = "lasso",StepwizeThr = 0.1,
  RF_ImpThr = 0.9, Covariates = "all.c", Family = "gaussian", RepMsr = "F", Corstr = "ar1")
  res4 = NtaAnno(PID=res$PID,OutPath="default",VarsY = "Y1",VarsX = "all.x",VarsN = "single.factor",
   FdrCorrect = "F",AdductPos = "M+H",AdductNeg = "M-H",AccuracyMS1 = 5,MS2 = "F",
   AccuracyMS2 = 20,ScoreMS2 = 0.4,DiffRT = 5)
   FuncExit(PID = res$PID)
```

NtaCros *Association analysis*

Description

Association analysis for non-targeted data

Usage

```
NtaCros(PID=res$PID, OutPath="default",VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = "T", SelMethod = "lasso",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, Covariates = "all.c", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by EMS
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
FdrCorrect	lgl. T (or TRUE) and F (or FALSE). Whether to correct the multiple hypotheses by false positive rate (FDR) method.
SelMethod	chr. Methods to select the significant features. Options include "stepwise" (multiple linear regress using stepwise algorithm), "lasso" (multiple linear regress using LASSO regularization algorithm), "random.forest" (random forest).

StepwizeThr	num. Threshold of the P value for stepwise regression to screen important variables. It ranges 0.05-0.25 with the default value of 0.1.
RF_ImpThr	num. Threshold of the total importance for the variables to a random forest model. It ranges 0.5-1.0 with the default value of 0.9.
Covariates	chr. Covariates used for modeling. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g. "X1,X2,X3".The default option is"all.c" (All covariates are included).
Family	chr. The link function for the regression model according the data type of outcomes, including "gaussian" for continuous variable, "binomial" for binary variable, and "poisson" for counting variable
RepMsr	lgl. T (or TRUE) and F (or FALSE). Whether existing repeated measurement of the subjects.
Corstr	chr. If "RepMsr" = T, the generalized estimating equations (GEE) will be used. For GEE, three correlation structure options are "exchangeable" "ar1" "unstructured".

Value

A list containing non-target analysis association results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitEMS()
   res1 = LoadEMS(PID = res$PID, UseExample = "example#1", MS2 = FALSE)
   res2 = TidyEMS(PID = res$PID, OutPath = "default",TransDummyVars="default")
   res3 = NtaCros(PID=res$PID, OutPath="default",VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = "T", SelMethod = "lasso",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, Covariates = "all.c", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
   FuncExit(PID = res$PID)
```

Pairwise *Implement pairwise mediation analyses*

Description

Implement pairwise mediation analyses for each pair of exposure and mediator.

Usage

```
Pairwise(PID, OutPath, VarsY, Family, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
VarsY	chr. The outcome variable. Either continuous, binary or count type is permitted.
VarsX	chr. Exposure variables included in pairwise mediation modelling. When "default" is specified, all exposure variables in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "X1,X2,X3".
VarsM	chr. Mediator variables included in pairwise mediation modelling. When "default" is specified, all mediator variables in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "M1,M2,M3".
VarsC	chr. Covariates included in pairwise mediation modelling. When "default" is specified, all covariates in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "C1,C2".
Family	chr. The error distribution and link function to be used in the model. Available options include "linear" and "gaussian" for continuous outcome variables, and "binomial" as well as "poisson" options for binary and count outcome variables.
Iter	num. The number of iteration times. Default is 500. To get a more stable result, a minimum iteration of 1,000 is recommended. To obtain more digits of P value of parameter estimation, a minimum iteration of 5,000 is suggested.

Details

Pairwise function implements mediation modelling for each pair of exposure and mediator. Given M exposures and N mediators, an exhaustive rule will be executed and M*N pair-wised mediation modelling are fitted. The modelling was realized using mediate function in mediation package.

Value

A list containing one dataframe that contains the pairwise mediation modelling results.

1. "MedtPairWise_Stats": pairwise mediation modelling results.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMedt()
   res1 <- LoadStatMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   res2 = TidyStatMedt(PID = res$PID, OutPath = "default",LogTransM = "log10",
   RangeLow = 0,RangeUpp = 1)
   res3 <- Pairwise(PID=res$PID, VarsY = "Y1",
   VarsX = "X1,X2,X3", VarsM = "M1,M2,M3", VarsC = "C1", Family = "linear",
   Iter = 500)
   FuncExit(PID = res$PID)
```

PanelAsso *Association analysis of panel data*

Description

Association analysis of panel data

Usage

```
PanelAsso = function(PID, OutPath = "default", VarsY, VarsX, VarsN = "single.factor",
    VarsRandomIpt = "SubjectID", VarsRandomSlp = "none",Covariates = "all.c")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatPanel
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
VarsRandomIpt	chr. Random intercept variable for the linear mixed-effect model. The default is "SubjectID".
VarsRandomSlp	chr. Random slope variable for the linear mixed-effect model. The default is "none". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
Covariates	chr. Covariates used for modeling. The default option is "all.c" (All covariates are included).

Value

A list containing the association analysis results.

Author(s)

Bin Wang

Examples

```
res <- InitStatPanel()
    res1 = LoadStatPanel(PID = res$PID, UseExample = "example#1")
    res2 = TidySataPanel(PID = res$PID, OutPath = "default",TransDummyVars="default")
    res3 = PanelAsso(PID=res$PID, OutPath = "default", VarsY = "Y1",
    VarsX = "X1,X2,X3,X4,X5,X6,X7,X8,X9,X10,X11,X12", VarsN = "single.factor",
    VarsRandomIpt = "SubjectID", VarsRandomSlp = "none", Covariates = "all.c")
    FuncExit(PID = res$PID)
```

RedM *Mediator dimension reduction*

Description

Implement mediator dimension reduction for mediators.

Usage

```
RedM(PID, OutPath, VarsY, Family, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
VarsY	chr. The outcome variable. Either continuous, binary or count type is permitted.
VarsX	chr. Exposure variables included in mediation modelling. When "default"(recommended) is specified, all exposure variables in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "X1,X2,X3".
VarsC	chr. Covariates included in mediation modelling. When "default" is specified, all covariates in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "C1,C2".
Method	chr. Dimension reduction method. Available options include "mean" and "pdm1"(default). "mean" option is recommended when outcome variable is a binary variable because in some situations error might occur in "pdm1" settings for binary outcome.
Family	chr. The error distribution and link function to be used in the model. Available options include "linear" and "gaussian" for continuous outcome variables, and "binomial" as well as "poisson" options for binary and count outcome variables.
Iter	num. The number of iteration times. Default is 500. To get a more stable result, a minimum iteration of 1,000 is recommended. To obtain more digits of P value of parameter estimation, a minimum iteration of 5,000 is suggested.

Details

RedM function provides two alternative methods for mediator dimension reduction, including sum method as well as pdm1 method via PDM package. By default, all mediators will be included for dimension reduction. Afterwards, mediation models will be built between given exposures as well as shrinkaged mediator variables.

Value

A list containing two elements where the mediator dimension reduction modelling results with exposures are stored. That list include:

1. "MedtRedM_all": a dataframe containing the mediator dimension reduction result where all mediators are considered as one group and that shrinkaged mediator is paired and modelled with each exposure.

2. "MedtRedM_list": a dataframe containing the mediator dimension reduction result where mediators are shrinkaged in their own subgroups, and these shrinkaged mediators are paired and modelled with each exposure.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMedt()
   res1 <- LoadStatMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   res2 = TidyStatMedt(PID = res$PID, OutPath = "default",LogTransM = "log10",
   RangeLow = 0,RangeUpp = 1)
   res4 <- RedM(PID=res$PID, VarsY = "Y1", VarsX = "X1,X2,X3",
   VarsC = "C1", Method = "mean", Family = "linear", Iter = 500)
   FuncExit(PID = res$PID)
```

RedX *Expoures dimension reduction*

Description

Implement dimension reduction for exposures.

Usage

```
RedX(PID, OutPath, VarsY, Family, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
VarsY	chr. The outcome variable. Either continuous, binary or count type is permitted.
VarsC	chr. Covariates included in mediation modelling. When "default" is specified, all covariates in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "C1,C2".
VarsM	chr. Mediator variables included in pairwise mediation modelling. When "default" is specified, all mediator variables in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "M1,M2,M3".
Method	chr. Dimension reduction method. Available options include "gcdnet" and "mean"(default). "mean" option is recommended when outcome variable is a binary variable because in some situations low variance of shrinkaged exposure variable might obtain in "gcdnet" settings for binary outcome.
Folds	num. Number of cross validation for gcdnet method. Default is 10.
Family	chr. The error distribution and link function to be used in the model. Available options include "linear" and "gaussian" for continuous outcome variables, and "binomial" as well as "poisson" options for binary and count outcome variables.

Iter	num. The number of iteration times. Default is 500. To get a more stable result, a minimum iteration of 1,000 is recommended. To obtain more digits of P value of parameter estimation, a minimum iteration of 5,000 is suggested.

Details

RedX function provides two alternative methods for exposure dimension reduction, including sum method as well as adaptive elastic net method via gcdnet package. By default, all exposures will be included for dimension reduction. Afterwards, mediation models will be built between given mediators as well as shrinkaged exposure variables.

Value

A list containing three elements where the exposure dimension reduction variables as well as their mediation modelling results with mediators are stored. That list include:

1. "MedtRedX_ERSall": a dataframe containing the exposure dimension reduction result where all exposures are considered as one group (ERS_All).
2. "MedtRedX_ERSlist": a dataframe containing the exposure dimension reduction results where exposures are shrinkaged in their own subgroups.
3. "MedtRedX_Stats": a dataframe containing the mediation modelling results where the shrinkaged exposure variables were paired with each mediator provided.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMedt()
   res1 <- LoadStatMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   res2 = TidyStatMedt(PID = res$PID, OutPath = "default",LogTransM = "log10",
   RangeLow = 0,RangeUpp = 1)
   res4 <- RedX(PID=res$PID, VarsY = "Y1",
   VarsC = "C1", VarsM = "M1,M2,M3", Method = "mean", Folds = 10,
   Family = "linear",Iter = 500)
   FuncExit(PID = res$PID)
```

RedXM *Exposure and mediator dimension reduction*

Description

Implement exposure and mediator dimension reduction. It should be noted that the exposure dimension reduction result has been built by RedX function prior to using it.

Usage

```
RedXM(PID, OutPath, VarsY, Family, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
VarsY	chr. The outcome variable. Either continuous, binary or count type is permitted.
VarsC	chr. Covariates included in mediation modelling. When "default" is specified, all covariates in the data will be used. It should be noted that there is fixed format for the characters separated with comma and without space, e.g., "C1,C2".
Method	chr. Dimension reduction method. Available options include "mean" and "pdm1"(default). "mean" option is recommended when outcome variable is a binary variable because in some situations error might occur in "pdm1" settings for binary outcome.
Family	chr. The error distribution and link function to be used in the model. Available options include "linear" and "gaussian" for continuous outcome variables, and "binomial" as well as "poisson" options for binary and count outcome variables.
Iter	num. The number of iteration times. Default is 500. To get a more stable result, a minimum iteration of 1,000 is recommended. To obtain more digits of P value of parameter estimation, a minimum iteration of 5,000 is suggested.

Details

RedXM function provides two alternative methods for mediator dimension reduction, including sum method as well as pdm1 method via PDM package. By default, all mediators will be included for dimension reduction. Afterwards, mediation models will be built between shrinkaged exposure variables as well as shrinkaged mediator variables. Prior to using RedXM, make sure that RedX function has been executed to obtain shrinkaged exposure variables.

Value

A list containing two elements where the mediator dimension reduction modelling results with shrinkaged exposures are stored. That list include:

1. "MedtRedXM_all": a dataframe containing the mediator dimension reduction result where all mediators are considered as one group and that shrinkaged mediator is paired and modelled with shrinkaged exposures.
2. "MedtRedXM_list": a dataframe containing the mediator dimension reduction result where mediators are shrinkaged in their own subgroups, and these shrinkaged mediators are paired and modelled with shrinkaged exposures.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMedt()
   res1 <- LoadStatMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   res2 = TidyStatMedt(PID = res$PID, OutPath = "default",LogTransM = "log10",
   RangeLow = 0,RangeUpp = 1)
   res3 <- RedX(PID=res$PID, VarsY = "Y1",
   VarsC = "default", VarsM = "default", Method = "mean", Folds = 10,
```

```
Family = "linear", Iter = 500)
res4 <- RedXM(PID=res$PID, VarsY = "Y1",
VarsC = "default", Method = "mean", Family = "linear", Iter = 500)
FuncExit(PID = res$PID)
```

`StatIPCros` *Build statistical link for cross-sectional data.*

Description

Build statistical link for cross-sectional data.

Usage

```
StatIPCros(PID, OutPath = "default", VarsY, VarsX, LinkModel = "ranger",
    ObsrPartType = "raw",ObsrPartNum = "50",ObsrProfType = "partial",
    ObsrProfNum = "100",ObsrProfVars = "all.x",ObsrProfGeom = "profiles",
    SubjPredSeq = "none",SubjPartType = "break_down")
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by StatIP
`OutPath`	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
`VarsY`	chr. Outcome variable used for modelling. Only one variable can be entered.
`VarsX`	chr. Exposure variable used for modelling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"
`LinkModel`	chr. Methods to interpret the model. Options include "ranger" (random forest), "glmnet" (elastic net), "svm" (support vector machine), "glm" (linear regression), "gam" (generalized additive model), and "xgboost" (eXtreme gradient boosting).
`ObsrPartType`	chr. Type of transformation that should be applied for dropout loss. Options include "raw" (drop losses), "ratio" (drop_loss/drop_loss_full_model), and "difference" (drop_loss - drop_loss_full_model)
`ObsrPartNum`	chr. Number of observations that should be sampled for calculation of variable importance. The default means variable importance will be calculated on whole dataset (no sampling). If "defult", use all Obsrrvations.
`ObsrProfType`	chr. Type of variable profile. Options include "partial", "conditional", and "accumulated"
`ObsrProfNum`	int. Number of observations used for calculation of aggregated profiles. By default 100.
`ObsrProfVars`	chr. Names of variables to be explained. If "all.x", all "X variable" above are chosen.
`ObsrProfGeom`	chr. Layout of the explanation profile in dataset level including "aggregates", "profiles" or "points".

SubjPredSeq	chr. Subjects which need explanation. Options include "all" (all the subjects), "none" (no subjects), and "other" (copy the subject list by clicking "Available vars").
SubjPartType	chr. Layout of the explanation profile in subject level. Options include "shap", "oscillations", "break_down", and "none".

Value

A list containing all the statistical explanation results.

Author(s)

Bin Wang

Examples

```
res <- InitStatIP()
   res1 = LoadStatIP(PID = res$PID, UseExample = "example#1")
   res2 = StatIPCros(PID=res$PID, OutPath = "default", VarsY = "Y1" ,VarsX = "all.x",
   LinkModel = "ranger",ObsrPartType = "raw" ,ObsrPartNum  = "50",
   ObsrProfType = "partial" ,ObsrProfNum = "100", ObsrProfVars = "all.x",
   ObsrProfGeom = "profiles",SubjPredSeq = "S1,S2,S3",SubjPartType = "break_down")
   FuncExit(PID = res$PID)
```

SurvAsso *Association analysis*

Description

Association analysis for survival data

Usage

```
SurvAsso = function(PID, OutPath = "default",
    TimeY, EventY,VarsX='all.x',VarsN = "single.factor",VarsSel = F,
    Covariates = "all.c")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatSurv()
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TimeY	chr. Outcome variable of survival time used for modelling. Only one variable can be entered.
EventY	chr. Outcome variable of status used for modelling. Only one variable can be entered.
VarsX	Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"

VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
VarsSel	lgl. T (or TRUE) and F (or FALSE). Whether to select the significant variable for the final model. Available options includes T and F
Covariates	chr. Covariates used for modeling. The default option is "all.c" (All covariates are included).

Value

A list containing the association analysis results.

Author(s)

Changxin Lan, Bin Wang(corresponding author)

Examples

```
res <- InitStatSurv()
    res1 = LoadStatSurv(PID = res$PID, UseExample = "example#1")
    res2 = TidyStatSurv(PID = res$PID, OutPath = "default",TransDummyVars="default")
    res3 = SurvAsso(PID=res$PID, TimeY = "Y1", EventY= 'Y2',VarsX='all.x',
    VarsN="single.factor",VarsSel="T",Covariates = "all.c")
    FuncExit(PID = res$PID)
```

SurvPred *Build prediction models*

Description

Build prediction models

Usage

```
SurvPred = function(PID,OutPath = "default",TimeY,EventY,VarsX = "all.x",
    Covariates = "all.c",RsmpMethod = "cv",Folds = 3,Ratio = 0.667,Repeats = 3)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatSurv()
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TimeY	chr. Outcome variable of survival time used for modelling. Only one variable can be entered.
EventY	chr. Outcome variable of status used for modelling. Only one variable can be entered.
VarsX	chr. Exposure variable used for modeling. The default option is "all.x" (All exposure variables are included). Users can also choose available variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3"

Covariates	chr. Covariates used for modeling. The default option is "all.c" (All covariates are included).
RsmpMethod	chr. Three resampling methods options for internal validation, including "cv" (i.e.,Cross validation) , "bootstrap", and "holdout".
Folds	num. Folds of Cross-validation resampling. It is ranging 2-10.
Ratio	num. Ratio of Bootstrap resampling. It is ranging 0.4-0.9.
Repeats	num. Number of Bootstrap resampling. It is ranging 2-20.

Value

A list containing the prediction performance evaluation.

Author(s)

Changxin Lan, Bin Wang(corresponding author)

Examples

```
res <- InitStatSurv()
   res1 = LoadStatSurv(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatSurv(PID = res$PID, OutPath = "default",TransDummyVars="default")
   res3 = SurvPred(PID=res$PID, TimeY = "Y1", EventY= 'Y2',VarsX='all.x',
   Covariates = "all.c",RsmpMethod="cv",Folds=3,Ratio=0.667,Repeats=3)
   FuncExit(PID = res$PID)
```

TidyDelMiss *Delete variables with missing values*

Description

Whether to delete missing variables with low variance. The default option is "yes". If skipped, it may result in failure during modeling.

Usage

```
TidyDelMiss(PID, OutPath = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.

Value

An R6 class object containing the variable(s) without missing values.

Author(s)

Bin Wang

Examples

```
res = InitStatTidy()
    res1 = LoadStatTidy(PID=res$PID, UseExample="example#1")
    res1$Expo$Data
    res2 = TidyDelMiss(PID=res$PID)
    res2$Expo$Data
    FuncExit(PID = res$PID)
```

TidyDelNearZeroVar *Delete variables with low variance*

Description

Whether to delete variables with low variance. The default option is "yes". If skipped, it may result in failure to build models.

Usage

```
TidyDelNearZeroVar(PID, OutPath = "default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.

Value

An R6 class object containing the variable(s) with acceptable variance.

Author(s)

Bin Wang

Examples

```
res = InitStatTidy()
    res1 = LoadStatTidy(PID=res$PID, UseExample="example#1")
    res2 = TidyDelNearZeroVar(PID=res$PID)
    FuncExit(PID = res$PID)
```

TidyEMS *Tidy data.*

Description

Tidy data file for EMS module

Usage

```
TidyEMS(PID, OutPath = "default",TransDummyVars="default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by EMS
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TransDummyVars	chr. Variables to be transformed as dummy variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3". If "default", all the factor variables will be transformed into dummy ones.

Value

An R6 class object containing the variable(s) after transforming the factor variables into dummy ones.

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitEMS()
    res1 = LoadEMS(PID = res$PID, UseExample = "example#1", MS2 = FALSE)
    res2 = TidyEMS(PID = res$PID, OutPath = "default",TransDummyVars="default")
    FuncExit(PID = res$PID)
```

TidySataPanel *Transform factor variables into dummy ones*

Description

Transform factor variables into dummy ones

Usage

```
TidySataPanel(PID, OutPath = "default",TransDummyVars="default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatPanel
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TransDummyVars	chr. Variables to be transformed as dummy variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3". If "default", all the factor variables will be transformed into dummy ones. These variables need to be transformed as factor ones in previous transform step using TransType function.

Value

An R6 class object containing the variable(s) after transforming the factor variables into dummy ones.

Author(s)

Bin Wang

Examples

```
res <- InitStatPanel()
    res1 = LoadStatPanel(PID = res$PID, UseExample = "example#1")
    res2 = TidySataPanel(PID = res$PID, OutPath = "default",TransDummyVars="default")
    res2$Expo$Data
    FuncExit(PID = res$PID)
```

TidyStatCros *Tidy data.*

Description

Tidy data file for StatCros module

Usage

```
TidyStatCros(PID, OutPath = "default",TransDummyVars="default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatCros
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TransDummyVars	chr. Variables to be transformed as dummy variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3". If "default", all the factor variables will be transformed into dummy ones. These variables need to be transformed as factor ones in previous transform step using TransType function.

Value

An R6 class object containing the variable(s) after transforming the factor variables into dummy ones.

Author(s)

Bin Wang

Examples

```
res <- InitStatCros()
    res1 = LoadStatCros(PID = res$PID, UseExample = "example#1")
    res2 = TidyStatCros(PID = res$PID, OutPath = "default",TransDummyVars="default")
    res2$Expo$Data
    FuncExit(PID = res$PID)
```

TidyStatDesc *Tidy data for StatDesc module*

Description

Tidy data file for StatDesc module

Usage

```
TidyStatDesc(PID = PID)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatDesc
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.

Details

TidyStatDesc function resembles several tidy steps for data pre-treatment before executing other functions in StatDesc module, including deleting vars with low variance and deleting missing values.

Value

An R6 class object containing the input data and vocabulary file.

Author(s)

Yanqiu Feng, Bin Wang(corresponding author)

Examples

```
res <- InitStatDesc()
   res1 <- LoadStatDesc(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   res2 <- TidyStatDesc(PID = res$PID)
   FuncExit(PID = res$PID)
```

TidyStatMix *Tidy data for Mix module*

Description

Tidy data file for Mix module

Usage

```
TidyStatMix(PID = PID, TransDummyVars = ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMix.
TransDummyVars	chr. A vector indicating factor variables to be transformed as dummy variables. Default value is "default" where all factor vars will be converted into dummies.
LogTransM	chr. Method of converting data distribution. Available options include "log10"(default), "log2", and "ln".
RangeLow	num. Lower range value for data scaling. Default value is 0. RangeLow value should be lower than that of RangeUpp.
RangeUpp	num. Upper range value for data scaling. Default value is 1. RangeUpp value should be larger than that of RangeLow.

Details

TidyStatMix function resembles several tidy steps for data pre-treatment before executing other functions in Mix module, including deleting vars with low variance, transforming categorical vars into factors, transferring data distribution and scaling for exposures.

Value

An R6 class object containing the input data and vocabulary file.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   res1 <- LoadStatMix(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   res2 <- TidyStatMix(PID = res$PID, LogTransM = "log10", RangeLow = 0, RangeUpp = 1)
   FuncExit(PID = res$PID)
```

TidyStatMO *Transform data type*

Description

Transform data type

Usage

```
TidyStatMO(PID,OutPath = "default",Vars,To)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; "all.c", all covariates; "all.cx", combination of All X and All C. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
To	chr. Indicate the type of the chosen variables to be transformed into. Available options include "integer", "numeric", "character", "factor", "logical", and "date".

Value

An R6 class object containing the variable(s) after transforming data type.

Author(s)

Guohuan Zhang, Yuting Wang, Bin Wang (corresponding author)

Examples

```
res = InitStatMO()
   res1 <- LoadStatMO(PID=res$PID, UseExample = "example#1", DataPath = NULL, VocaPath = NULL)
    res1$Expo$Data
    res2 = TidyStatMO(PID=res$PID, Vars = "X1,X2", To = "factor")
    res2$Expo$Data
    FuncExit(PID = res$PID)
```

TidyStatSurv *Transform factor variables into dummy ones*

Description

Transform factor variables into dummy ones

Usage

```
TidyStatSurv(PID, OutPath = "default",TransDummyVars="default")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by ExpoSurv
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TransDummyVars	chr. Variables to be transformed as dummy variables. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3". If "default", all the factor variables will be transformed into dummy ones. These variables need to be transformed as factor ones in previous transform step using TransType function.

Value

An R6 class object containing the variable(s) after transforming the factor variables into dummy ones.

Author(s)

Bin Wang

Examples

```
res <- InitStatSurv()
    res1 = LoadStatSurv(PID = res$PID, UseExample = "example#1")
    res2 = TidyStatSurv(PID = res$PID, OutPath = "default",TransDummyVars="default")
    res2$Expo$Data
    FuncExit(PID = res$PID)
```

TidyTransClass *Classify variables into various groups*

Description

Classify variables into various groups

Usage

```
TidyTransClass(PID, OutPath = "default", Vars, LevelTo)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; Users can also choose available variables by selecting "Other" option, and copy the variables by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
LevelTo	The number of levels to convert variables to.

Value

An R6 class object containing the variable(s) after classifying data into various levels.

Author(s)

Bin Wang

Examples

```
res = InitStatTidy()
    res1 = LoadStatTidy(PID=res$PID, UseExample="example#1")
    res1$Expo$Data
    res2 = TidyTransClass(PID=res$PID, Vars="X1", LevelTo="4")
    res2$Expo$Data
    FuncExit(PID = res$PID)
```

TidyTransDistr *Transform variable distribution*

Description

Transform variable distribution

Usage

```
TidyTransDistr(PID,OutPath = "default", Vars, Method)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; Users can also choose available variables by selecting "Other" option, and copy the variables by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".

Value

An R6 class object containing the variable(s) after transforming distribution.

Author(s)

Bin Wang

Examples

```
res = InitStatTidy()
    res1 = LoadStatTidy(PID=res$PID, UseExample="example#1")
    res1$Expo$Data
    res2 = TidyTransDistr(PID=res$PID, Vars="X6,X7", Method="log10")
    res2$Expo$Data
    FuncExit(PID = res$PID)
```

TidyTransImput *Missing data imputation.*

Description

Missing data imputation.

Usage

```
TidyTransImput(PID,OutPath = "default",Vars,Method)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; Users can also choose available variables by selecting "Other" option, and copy the variables by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
Method	Methods used for imputation. Available options include "lod" or "cart" methods. For "lod" method, limit of detection (LOD) should be included in the "Vocabulary" file.

Value

An R6 class object containing variable(s) with imputation.

Author(s)

Bin Wang

Examples

```
res = InitStatTidy()
    res1 = LoadStatTidy(PID=res$PID, UseExample="example#1")
    res2 = TidyTransImput(PID=res$PID, Vars="all.x", Method="lod")
    FuncExit(PID = res$PID)
```

`TidyTransScale` *Scale variables*

Description

Scale variables

Usage

```
TidyTransScale(PID,OutPath = "default",Vars,Method,...)
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by initial functions.
`OutPath`	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
`Vars`	Variables to be imputed. Available options include: "all.x", all exposure variables; Users can also choose available variables by selecting "Other" option, and copy the variables by clicking "Available vars". It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
`Method`	chr. Scaling methods. Available options include "normal" and "range".
`Direct`	chr. Direction to be transformed, Available options include "positive" and "negative".
`RangeLow`	num. Lower limit for range method.
`RangeUpper`	num. Upper limit for range method. It should be greater than the lower limit.

Value

An R6 class object containing the variable(s) after scaling data.

Author(s)

Bin Wang

Examples

```
res = InitStatTidy()
    res1 = LoadStatTidy(PID=res$PID, UseExample="example#1")
    res1$Expo$Data
    res2 = TidyTransScale(PID=res$PID, Vars="all.x", Method="normal")
    res2$Expo$Data
    FuncExit(PID = res$PID)
```

TidyTransType *Transform data type*

Description

Transform data type

Usage

```
TidyTransType(PID,OutPath = "default",Vars,To)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by initial functions.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Vars	Variables to be imputed. Available options include: "all.x", all exposure variables; It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., "X1,X2,X3".
To	chr. Indicate the type of the chosen variables to be transformed into. Available options include "integer", "numeric", "character", "factor", "logical", and "date".

Value

An R6 class object containing the variable(s) after transforming data type.

Author(s)

Bin Wang

Examples

```
res = InitStatTidy()
    res1 = LoadStatTidy(PID=res$PID, UseExample="example#1")
    res1$Expo$Data
    res2 = TidyTransType(PID=res$PID, Vars = "X1,X2", To = "character")
    res2$Expo$Data
    FuncExit(PID = res$PID)
```

urlhead *Exposomex Link*

Description

Exposomex Link

Usage

```
urlhead
```

Format

An object of class character of length 1.

VizCrosAsso *Visualize association analysis*

Description

Visualize association analysis

Usage

```
VizCrosAsso = function(PID, OutPath = "default",VarsN = "single.factor",
    Layout = "volcano",Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatCros
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
Layout	chr. Visualization layout. Available options include "forest" and "volcano".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).
ColorFor	chr. #' @param SizeFor chr

Value

An R6 class object containing the results' plot.

Author(s)

Bin Wang

Examples

```
res <- InitStatCros()
   res1 = LoadStatCros(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatCros(PID = res$PID, OutPath = "default",TransDummyVars="default")
   res3 = CrosAsso(PID=res$PID, Linear = TRUE, EpiDesign = "cohort",
   VarsY = "Y1", VarsX = "X5,X6,X7,X8,X9,X10,X11", VarsN = "single.factor" ,
   VarsSel = FALSE, VarsSelThr = 0.1, Covariates = "all.c", Family = "gaussian",
   RepMsr = FALSE,Corstr = "ar1")
  res4 = VizCrosAsso(PID=res$PID, VarsY = "Y1", VarsN="single.factor", Layout = "volcano",
  Brightness = "dark",Palette = "default1", ColorFor = "beta.value", SizeFor = "p.value")
   FuncExit(PID = res$PID)
```

VizCrosPred *Visualize the prediction performance*

Description

Visualize the prediction performance

Usage

```
VizCrosPred = function(PID,OutPath = "default",Layout = "bar",
    Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatCros
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
Layout	chr. Visualization layout. Available options include "bar" and "roc"
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the results' plot.

Author(s)

Bin Wang

Examples

```
res <- InitStatCros()
   res1 = LoadStatCros(PID = res$PID, UseExample = "example#1")
   res3 = CrosPred(PID=res$PID, VarsY = "Y1", VarsX = "X5,X6,X7,X8,X9,X10,X11",
   PredType = "response",VarsSel = FALSE,VarsSelThr = 0.1, Covariates = "all.c",
   RsmpMethod = "cv" ,Folds = 5,Ratio = 0.667,Repeats = 5)
   res4 = VizCrosPred(PID=res$PID, VarsY = "Y1", Layout = "bar",
   Brightness = "light", Palette = "science")
   FuncExit(PID = res$PID)
```

VizHEPPredBlood *Visualize StatHEP results*

Description

VizHEPPredBlood function is mainly aimed to visualize the modeling results calculated by HEPPredBlood function. It can provide plots with high quality of the final results to make it easier for users to understand.

Usage

```
VizHEPPredBlood(PID, OutPath, MC,Layout,Brightness,Palette)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatHEP.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
MC	lgl. T (or TRUE) and F (or FALSE). Whether to perform Mentocaro simulation.
Layout	chr. Visualization layout. Available options include "forest" and "boxplot".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and other options about some journal preference styles including "cell", "nature", "science", "lancet", "nejm", etc.

Details

You can get different styles of images by selecting different parameters.

Value

An R6 class object.

Author(s)

Weinan Lin, Ning Gao, Bin Wang (corresponding author)

Examples

```
res <- InitStatHEP()
   res1 <- LoadStatHEP(PID=res$PID, UseExample = "example#1", DataPath = NULL)
   res2 <- HEPPredBlood(PID=res$PID, OutPath = "default", MC = "T", N = 100)
   res3 <- VizHEPPredBlood(PID=res$PID, OutPath = "default", MC = "T",
   Layout = "forest", Brightness = "light", Palette = 'default1')
   FuncExit(PID = res$PID)
```

VizMedtPair *Visualize pairwise mediation modelling result*

Description

Visualize the pairwise mediation result. It should be noted that the pairwise mediation modelling result has been built by Pairwise function prior to using it.

Usage

```
VizMedtPair(PID, OutPath, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Pallette	chr. Visualization palette. Available options include "default1", "default2" and "Journal". The "Journal" option provides several journal preference styles including cell, nature, science, lancet, nejm, and jama.

Details

VizMedtPair draws a visualized display for pairwise modelling results. Piror to using VizMedtPair, make sure that the users have built mediation models by Pairwise function.

Value

A list containing two elements where the visualized pairwise modelling plot as well as the organized plotting data are stored. The elements of that list include:

1. "plotdata": a dataframe containing the organized plotting data extracted from #' pairwise mediation modelling result.
2. "plot": a visualized plot for pairwise mediation result.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMedt()
   res1 <- LoadStatMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   res2 = TidyStatMedt(PID = res$PID, OutPath = "default",LogTransM = "log10",
   RangeLow = 0,RangeUpp = 1)
   res3 <- Pairwise(PID=res$PID, VarsY = "Y1",
   VarsX = "X1,X2,X3", VarsM = "M1,M2,M3", VarsC = "C1", Family = "linear",
   Iter = 500)
   res4 <- VizMedtPair(PID=res$PID, Brightness = "light",
   Pallette = "default1")
   FuncExit(PID = res$PID)
```

VizMixBKMR *Visualize the model results of Bayesian Kernel Machine Regression (BKMR)*

Description

Visualize the model results of Bayesian Kernel Machine Regression (BKMR)

Usage

```
VizMixBKMR(PID, OutPath = "default", VarsY, Brightness = "dark", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatMix
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the BKMR analysis results' plot.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   res1 = LoadStatMix(PID = res$PID, UseExample = "example#1")
   res2 <- TidyStatMix(PID = res$PID, LogTransM = "log10", RangeLow = 0, RangeUpp = 1)
   res3 = MixBKMR(PID = res$PID, VarsY = "Y1", VarsX = "X4,X5,X6,X7,X8,X9,X10",
   Covariates = "all.c", Family = "gaussian", Group = 'F', Iter = 2000, qfixed = 0.5,
   qsbivar = "default", qsoverall = "default", qsdiff = "default")
   res4 = VizMixBKMR(PID=res$PID, VarsY = "Y1", Brightness = "dark",
   Palette = "default1")
   FuncExit(PID = res$PID)
```

VizMixGcomp *Visualize results of g-computation model*

Description

Visualize results of g-computation model

Usage

```
VizMixGcomp(PID, OutPath = "default", Brightness  = "dark", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatMix
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the g-computation analysis results' plot.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   res1 = LoadStatMix(PID = res$PID, UseExample = "example#1")
   res2 <- TidyStatMix(PID = res$PID, LogTransM = "log10", RangeLow = 0, RangeUpp = 1)
   res3 = MixGcomp(PID=res$PID, VarsY = "Y1", VarsX = "all.x",
   Covariates = "all.c", Family = "gaussian", q = 10)
   res4 = VizMixGcomp(PID = res$PID, VarsY = "Y1", Brightness = "dark",
   Palette = "default1")
   FuncExit(PID = res$PID)
```

VizMixMLR *Visualize the model results*

Description

Visualize the results of multiple linear regression (MLR) model

Usage

```
VizMixMLR(PID,OutPath = "default",VarsY,SelMethod,Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatMix
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
SelMethod	chr. Method to select the important features to the final model. Options include "stepwise" (stepwise regression), "lasso" (Regularization regression of least absolute shrinkage and selection operator), and "enet" (Regularization regression of elastic net).
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the MLR analysis plot.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   res1 = LoadStatMix(PID = res$PID, UseExample = "example#1")
   res2 <- TidyStatMix(PID = res$PID, LogTransM = "log10", RangeLow = 0, RangeUpp = 1)
   res3 = res3 = MixMLR(PID = res$PID, VarsY = "Y1", VarsX = "all.x",
   Covariates = "all.c", SelMethod = "lasso", PredType = "response",
   Family = "gaussian")
   res4 = VizMixMLR(PID=res$PID, VarsY = "Y1", SelMethod = 'lasso',
   Brightness = "light", Palette = "default1")
   FuncExit(PID = res$PID)
```

VizMixWQS *Visualize results of weighted quantile sum regression (WQS) model*

Description

Visualize results of weighted quantile sum regression (WQS) model

Usage

```
VizMixWQS(PID, OutPath = "default", VarsY, Brightness = "dark", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatMix
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the WQS analysis results' plot.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMix()
   res1 = LoadStatMix(PID = res$PID, UseExample = "example#1")
   res2 <- TidyStatMix(PID = res$PID, LogTransM = "log10", RangeLow = 0, RangeUpp = 1)
   res3 = MixWQS(PID=res$PID, VarsY = "Y1", VarsX = "all.x",Covariates = "all.c",
   Family = "gaussian", VarStrat = "none", RatioValidat = 0.3, q = 10, b=100,
   b1_pos = 'F', b1_constr = 'F')
   res4 = VizMixWQS(PID = res$PID, VarsY = "Y1", Brightness = "dark",
   Palette = "default1")
   FuncExit(PID = res$PID)
```

VizMultOmic *Visualize multiomics model results*

Description

VizMultOmic function is mainly aimed to visualize the modeling results calculated by MulOmicsCros function. It can provide plots with high quality of the final results to make it easier for users to understand.

Usage

```
VizMultOmic(PID, OutPath, VarsY, NodeNum, EdgeThr, Layout, Brightness, Palette)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMO.
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable for modeling. Only one variable can be entered.

NodeNum	num. Number of nodes in the network plot. The maximum number is generated by Multiomics function. User can set a smaller value if needed.
EdgeThr	num. Threshold of correlation coefficient ranging 0-1 for generating the concerned edges of the network plot.
Layout	chr. Visualization layout. Available options include "force-directed" and "degree-circle".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and other options about some journal preference styles including "cell", "nature", "science", "lancet", "nejm", etc.

Details

You can get different styles of images by selecting different parameters.

Value

An R6 class object containing seven elements. The elements of that object include: (1) "Importance_plot": Plots for the importance of features after modeling a single omic from different learners. (2) "Measures_boxplot": Plots for the r-square value of the single omic model built by different learners. (3) "NetWork_State": A list containing dataframes that contain nodes and edges used to draw interOmic network from models built by different learners. (4) "Nodeplot": Interomic node plots from models built by different models. (5) "Networkplot": Interomic network plots from models built by different models. (6) "Prediction_train_plot": Plots for the prediction value of the train set in the SG model built by different combinations of learners. (7) "Prediction_test_plot": Plots for the prediction value of the test set in the SG model built by different combinations of learners.

Author(s)

Guohuan Zhang, Yuting Wang, Ning Gao, Bin Wang (corresponding author)

Examples

```
res <- InitStatMO()
  res1 <- LoadStatMO(PID=res$PID, UseExample = "example#1", DataPath = NULL, VocaPath = NULL)
   res2 <- TidyStatMO(PID=res$PID, Vars = "X1,X2", To = "factor")
  res3 <- MulOmicsCros(PID=res$PID, OutPath = "default", OmicGroups = "immunome,metabolome",
  VarsY = "Y1", VarsC = "all.c", TuneMethod = "random_search", TuneNum = 2, RsmpMethod = "cv", Folds = 2,
   Ratio = 0.67, Repeats = 5, VarsImpThr = 0.85, SG_Lrns ="lasso,enet")
  res4 <- VizMultOmic(PID=res$PID, OutPath = "default", VarsY = "Y1", NodeNum=100, EdgeThr= 0.45,
   Layout = "force-directed", Brightness = "light", Palette = 'default1')
   FuncExit(PID = res$PID)
```

VizNtaAnno *Visualize annotation results*

Description

Visualize annotation results of non-targeted data

Usage

```
VizNtaAnno(PID=res$PID,OutPath= "default",VarsY = "Y1",VarsN = "single.factor",
  AccuracyMS1 = 5,MS2 = "F",AccuracyMS2 = 20,DiffRT = 100,Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by EMS
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
AccuracyMS1	num. Accuracy threshold to match the target compounds of MS1 information. The unit is ppm. It is usually set ranging 1-20.
MS2	lgl. T (or TRUE) and F (or FALSE). Whether to use MS2 information to annotate the target ion.
AccuracyMS2	num. Accuracy threshold to match the target compounds of MS2 information. The unit is ppm. It is usually set ranging 1-20.
DiffRT	num. Difference of the retention time between the measured and the theoretical prediction. The unit is second (s). It is usually set ranging 60-120.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing plots of non-target annotation analysis results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitEMS()
   res1 = LoadEMS(PID = res$PID, UseExample = "example#1", MS2 = "F")
   res2 = TidyEMS(PID = res$PID, OutPath = "default",TransDummyVars="default")
   res3 = NtaCros(PID=res$PID, OutPath="default",VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = "T", SelMethod = "lasso",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, Covariates = "all.c", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
  res4 = NtaAnno(PID=res$PID,OutPath="default",VarsY = "Y1",VarsX = "all.x",VarsN = "single.factor",
   FdrCorrect = "F",AdductPos = "M+H",AdductNeg = "M-H",AccuracyMS1 = 5,MS2 = "F",
   AccuracyMS2 = 20,ScoreMS2 = 0.4,DiffRT = 5)
   res5 = VizNtaAnno(PID=res$PID,OutPath="default",VarsY = "Y1",VarsN = "single.factor",
  AccuracyMS1 = 5,MS2 = "F",AccuracyMS2 = 20,DiffRT = 100,Brightness = "light",Palette = "default1")
   FuncExit(PID = res$PID)
```

VizNtaCros *Visualize the results of Association analysis*

Description

Visualize the results of Association analysis for non-targeted data

Usage

```
VizNtaCros(PID, OutPath = "default", VarsY,VarsN, Layout = "volcano",
    EffectThr = 0.5, Brightness = "light", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by EMS
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor".
Layout	chr. Visualization layout. Available options include "volcano" and "forest".
EffectThr	num. Threshold of the total importance for the variables to a random forest model. It ranges 0.5-1.0 with the default value of 0.9.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing plots of non-target analysis association results

Author(s)

Mingliang Fang, Bin Wang (corresponding author)

Examples

```
res <- InitEMS()
   res1 = LoadEMS(PID = res$PID, UseExample = "example#1", MS2 = FALSE)
   res2 = TidyEMS(PID = res$PID, OutPath = "default",TransDummyVars="default")
   res3 = NtaCros(PID=res$PID, OutPath="default",VarsY = "Y1", VarsX = "all.x",
   VarsN = "single.factor", FdrCorrect = "T", SelMethod = "lasso",
   StepwizeThr = 0.1, RF_ImpThr = 0.9, Covariates = "all.c", Family = "gaussian",
   RepMsr = "F", Corstr = "ar1")
   res4 = VizNtaCros(PID=res$PID,OutPath="default",VarsY = "Y1",VarsN = "single.factor",
   Layout = "volcano",EffectThr = 0.5,Brightness = "light",Palette = "default1")
   FuncExit(PID = res$PID)
```

VizPanelAsso *Visualize the results of association analysis for panel data*

Description

Visualize the results of association analysis for panel data

Usage

```
VizPanelAsso = function(PID, OutPath, VarsY, VarsN, Layout = "volcano",
    Brightness = "dark", Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by StatPanel
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsY	chr. Outcome variable used for modeling. Only one variable can be entered.
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
Layout	chr. Visualization layout. Available options include "forest" and "volcano".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., cell, nature, science, lancet, nejm, and jama).

Value

A list containing the plots of association analysis results.

Author(s)

Bin Wang

Examples

```
res <- InitStatPanel()
    res1 = LoadStatPanel(PID = res$PID, UseExample = "example#1")
    res2 = TidySataPanel(PID = res$PID, OutPath = "default",TransDummyVars="default")
    res3 = PanelAsso(PID=res$PID, OutPath = "default", VarsY = "Y1",
    VarsX = "X1,X2,X3,X4,X5,X6,X7,X8,X9,X10,X11,X12", VarsN = "single.factor",
    VarsRandomIpt = "SubjectID", VarsRandomSlp = "none", Covariates = "all.c")
    res4 = VizPanelAsso(PID = res$PID, OutPath = "default", VarsY = "Y1",
    VarsN = "single.factor", Layout = "forest", Brightness = "dark",Palette = "default1")
    FuncExit(PID = res$PID)
```

VizRedXM *Visualize RedXM mediation modelling result*

Description

Visualize the RedXM mediation result. It should be noted that the RedXM mediation modelling result has been built by RedXM function prior to using it.

Usage

```
VizRedXM(PID, OutPath, ...)
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatMedt.
OutPath	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set.It should be noted that the slash symbol is "/", not "\".
Brightness	chr. Visualization brightness. Available options include "Light" and "Dark".
Pallette	chr. Visualization palette. Available options include "default1", "default2" and "Journal". The "Journal" option provides several journal preference styles including cell, nature, science, lancet, nejm, and jama.

Details

VizRedXM draws a visualized display for RedXM results. Piror to using VizMedtPair, make sure that the users have built mediation models by RedXM function.

Value

A list containing two elements where the visualized exposure and mediator dimension reduction plot as well as the organized plotting data are stored. The elements of that list include:

1. "MedtRedXM_plotadta": a dataframe containing the organized plotting data extracted from exposure and mediator dimension reduction result.
2. "MedtRedXM_plot": a visualized plot for exposure and mediator dimension reduction result.

Author(s)

Mengyuan Ren, Bin Wang(corresponding author)

Examples

```
res <- InitStatMedt()
   res1 <- LoadStatMedt(PID = res$PID, UseExample = "example#1", DataPath = NULL,
   VocaPath = NULL)
   res2 = TidyStatMedt(PID = res$PID, OutPath = "default",LogTransM = "log10",
   RangeLow = 0,RangeUpp = 1)
   res3 <- RedX(PID=res$PID, VarsY = "Y1",
   VarsC = "default", VarsM = "default", Method = "mean", Folds = 10,
   Family = "linear", Iter = 500)
   res4 <- RedXM(PID=res$PID, VarsY = "Y1",
   VarsC = "default", Method = "mean", Family = "linear", Iter = 500)
```

```
res5 <- VizRedXM(PID=res$PID, Brightness = "light",
Pallette = "nature")
FuncExit(PID = res$PID)
```

`VizRefer` *Visualize the Articles' Main Information*

Description

Visualize the articles' main information after MetaRef function.

Usage

```
VizRefer(PID = NULL, OutPath = "default")
```

Arguments

`PID`	chr. Program ID. It must be the same with the PID generated by InitEMETA.
`OutPath`	chr. Output file directory, e.g. "D:/output". If "default", the current working directory will be set. It should be noted that the slash symbol is "/", not "\".

Details

VizRefer provides the visual function of articles' main information, showing the year and region distribution directly. Please attention, PID must be got from the return result of InitEMETA(). VizRefer can only run successfully after successfully running InitEMETA, LoadEMETA and MetaRef functions.

Value

A list object containing plots of article information visualization.

Author(s)

Weinan Lin, Bin Wang (corresponding author)

Examples

```
res = InitEMETA()
  res1 = LoadEMETA(PID = res$PID,
  UseExample = "example#1",
  DataPath = NULL)
  res2 = MetaRef(PID = res$PID,
  OutPath = "default",
  Mode = "search",
  VarX = "default",
  VarY = "default",
  VarM = "default",
  YearFrom = "default",
  YearEnd = "default",
  PMID = "default")
  res3 = VizRefer(PID=res$PID,
  OutPath = "default")
  FuncExit(PID = res$PID)
```

VizSurvAsso *Visualize association analysis*

Description

Visualize association analysis

Usage

```
VizSurvAsso = function(PID, OutPath = "default",VarsN = "single.factor",
  Layout = "volcano",Brightness = "light",Palette = "default1",ColorFor= "p.value",SizeFor= "p.val
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatSurv()
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
VarsN	chr. Choose the single factor or multiple factor model. Available options include "single.factor" and "multiple.factor"
Layout	chr. Visualization layout. Available options include "forest" and "volcano".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., "cell", "nature", "science", "lancet", "nejm", and "jama").
ColorFor	chr. Volcano plot dot color. Available options include "p.value" and "hr".
SizeFor	chr. Volcano plot dot size. Available options include "p.value" and "hr".

Value

A list containing the results' plot.

Author(s)

Changxin Lan, Ning Gao, Bin Wang(corresponding author)

Examples

```
res <- InitStatSurv()
   res1 = LoadStatSurv(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatSurv(PID = res$PID, OutPath = "default",TransDummyVars="default")
   res3 = SurvAsso(PID=res$PID, TimeY = "Y1", EventY= 'Y2',VarsX='all.x',
   VarsN="single.factor",VarsSel="T",Covariates = "all.c")
  res4 = VizSurvAsso(PID=res$PID,VarsN="single.factor",Layout="volcano",Brightness= "light",
   Palette = "default1",ColorFor= "p.value",SizeFor= "p.value")
   FuncExit(PID = res$PID)
```

VizSurvCompGroup *Compare the survival curves of two groups*

Description

Compare the survival curves of two groups

Usage

```
VizSurvCompGroup = function(PID,OutPath = "default",TimeY,EventY,
   VarsG,Model='km',VarsAdj,AdjMethod='average',Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatSurv()
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
TimeY	chr. Outcome variable of survival time used for modelling. Only one variable can be entered.
EventY	chr. Outcome variable of status used for modelling. Only one variable can be entered.
VarsG	chr. Grouping variable, must be a binary variable.
Model	chr. Methods to depict the survival curve. Options include 'km' (Kaplan-Meier estimate) and "coxph" (Cox proportional hazards regression mode).
VarsAdj	If you choose the cox model, co-variables used for modelling. It should be noted that there is fixed format for the entering characters separated with comma and without space, e.g., X1,X2,X3.
AdjMethod	If you choose the cox model, method for adjusting model, include: "average","single","margin" and "conditional".
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., "cell", "nature", "science", "lancet", "nejm", and "jama").

Value

A list containing the results' plot.

Author(s)

Changxin Lan, Ning Gao, Bin Wang(corresponding author)

Examples

```
res <- InitStatSurv()
   res1 = LoadStatSurv(PID = res$PID, UseExample = "example#1")
   res2 = VizSurvCompGroup(PID=res$PID,TimeY="Y1",EventY="Y2",VarsG="C3",
   Model="km",Brightness="light",Palette='default1')
   FuncExit(PID = res$PID)
```

VizSurvPred *Visualize the prediction performance*

Description

Visualize the prediction performance

Usage

```
VizSurvPred = function(PID,OutPath = "default",Layout = "curve",
    Brightness = "light",Palette = "default1")
```

Arguments

PID	chr. Program ID. It must be the same with the PID generated by InitStatSurv()
OutPath	chr. Output file directory, e.g. "D:/test". It should be noted that the slash symbol is "/", not "\". If "default", the current working directory will be set.
Layout	chr. Visualization layout. Available options include "curve" , "bar" and 'all'.
Brightness	chr. Visualization brightness. Available options include "light" and "dark".
Palette	chr. Visualization palette. Available options include "default1", "default2" and several journal preference styles (i.e., "cell", "nature", "science", "lancet", "nejm", and "jama").

Value

A list containing the results' plot.

Author(s)

Changxin Lan, Ning Gao, Bin Wang(corresponding author)

Examples

```
res <- InitStatSurv()
   res1 = LoadStatSurv(PID = res$PID, UseExample = "example#1")
   res2 = TidyStatSurv(PID = res$PID, OutPath = "default",TransDummyVars="default")
   res3 = SurvPred(PID=res$PID, TimeY = "Y1", EventY= 'Y2',VarsX='all.x',
   Covariates = "all.c",RsmpMethod="cv",Folds=3,Ratio=0.667,Repeats=3)
   res4 = VizSurvPred(PID=res$PID,Layout="curve",Brightness="light",Palette='default1')
   FuncExit(PID = res$PID)
```

Index