# Package ‘MVar’

November 22, 2024

**Type** Package

**Title** Multivariate Analysis

**Version** 2.2.5

**Date** 2024-11-22

**Imports** graphics, grDevices, MASS, stats

**Description** Multivariate analysis, having functions that perform simple correspondence analysis (CA) and multiple correspondence analysis (MCA), principal components analysis (PCA), canonical correlation analysis (CCA), factorial analysis (FA), multidimensional scaling (MDS), linear (LDA) and quadratic discriminant analysis (QDA), hierarchical and non-hierarchical cluster analysis, simple and multiple linear regression, multiple factor analysis (MFA) for quantitative, qualitative, frequency (MFACT) and mixed data, biplot, scatter plot, projection pursuit (PP), grant tour method and other useful functions for the multivariate analysis.

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**NeedsCompilation** yes

**Author** Paulo Cesar Ossani [aut, cre] (<https://orcid.org/0000-0002-6617-8085>),
Marcelo Angelo Cirillo [aut] (<https://orcid.org/0000-0002-2647-439X>)

**Maintainer** Paulo Cesar Ossani <ossanipc@hotmail.com>

**Repository** CRAN

**Date/Publication** 2024-11-22 13:50:08 UTC

## Contents

---

`MVar-package` *Multivariate Analysis.*

---

## Description

Multivariate analysis, having functions that perform simple correspondence analysis (CA) and multiple correspondence analysis (MCA), principal components analysis (PCA), canonical correlation analysis (CCA), factorial analysis (FA), multidimensional scaling (MDS), linear (LDA) and quadratic discriminant analysis (QDA), hierarchical and non-hierarchical cluster analysis, simple and multiple linear regression, multiple factor analysis (MFA) for quantitative, qualitative, frequency (MFACT) and mixed data, biplot, scatter plot, projection pursuit (PP), grant tour method and other useful functions for the multivariate analysis.

## Details

| | |
|---|---|
| Package: | MVar |
| Type: | Package |
| Version: | 2.2.5 |
| Date: | 2024-11-22 |
| License: | GPL(>=2) |
| LazyLoad: | yes |

### Author(s)

Paulo Cesar Ossani and Marcelo Angelo Cirillo.

Maintainer: Paulo Cesar Ossani <ossanipc@hotmail.com>

### References

Abdessemed, L.; Escofier, B.; Analyse factorielle multiple de tableaux de frequencies: comparaison avec l'analyse canonique des correspondences. *Journal de la Societe de Statistique de Paris*, Paris, v. 137, n. 2, p. 3-18, 1996.

Abdi, H. Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition (GSVD). In: SALKIND, N. J. (Ed.). *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage, 2007. p. 907-912.

Abdi, H.; Valentin, D. Multiple factor analysis (MFA). In: SALKIND, N. J. (Ed.). *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage, 2007. p. 657-663.

Abdi, H.; Williams, L. Principal component analysis. *WIREs Computational Statatistics*, New York, v. 2, n. 4, p. 433-459, July/Aug. 2010.

Abdi, H.; Williams, L.; Valentin, D. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. *WIREs Computational Statatistics*, New York, v. 5, n. 2, p. 149-179, Feb. 2013.

Asimov, D. The Grand Tour: A Tool for Viewing Multidimensional Data. *SIAM Journal of Scientific and Statistical Computing*, 6(1), 128-143, 1985.

Asimov, D.; Buja, A. The grand tour via geodesic interpolation of 2-frames. in Visual Data Exploration and Analysis. *Symposium on Electronic Imaging Science and Technology*, IS&T/SPIE. 1994.

Becue-Bertaut, M.; Pages, J. A principal axes method for comparing contingency tables: MFACT. *Computational Statistics & Data Analysis*, New York, v. 45, n. 3, p. 481-503, Feb. 2004

Becue-Bertaut, M.; Pages, J. Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data. *Computational Statistics & Data Analysis*, New York, v. 52, n. 6, p. 3255-3268, Feb. 2008.

Benzecri, J. Analyse de l'inertie intraclasse par l'analyse d'un tableau de contingence: intra-classinertia analysis through the analysis of a contingency table. *Les Cahiers de l'Analyse des Donnees*, Paris, v. 8, n. 3, p. 351-358, 1983.

Buja, A.; Asimov, D. Grand tour methods: An outline. *Computer Science and Statistics*, 17:63-67. 1986.

Buja, A.; Cook, D.; Asimov, D.; Hurley, C. Computational Methods for High-Dimensional Rotations in Data Visualization, in C. R. Rao, E. J. Wegman & J. L. Solka, eds, *"Handbook of Statistics: Data Mining and Visualization"*, Elsevier/North Holland, http://www.elsevier.com, pp. 391-413. 2005.

Charnet, R., at al. *Analise de modelos de regressao lienar,* 2a ed. Campinas: Editora da Unicamp, 2008. 357 p.

Cook, D.; Lee, E. K.; Buja, A.; WickmamM, H. Grand tours, projection pursuit guided tours and manual controls. In Chen, Chunhouh, Hardle, Wolfgang, Unwin, e Antony (Eds.), *Handbook of Data Visualization*, Springer Handbooks of Computational Statistics, chapter III.2, p. 295-314. Springer, 2008.

Cook, D.; Buja, A.; Cabrera, J. Projection pursuit indexes based on orthonormal function expansions. *Journal of Computational and Graphical Statistics*, 2(3):225-250, 1993.

Cook, D.; Buja, A.; Cabrera, J.; Hurley, C. Grand tour and projection pursuit, *Journal of Computational and Graphical Statistics*, 4(3), 155-172, 1995.

Cook, D.; Swayne, D. F. Interactive and Dynamic Graphics for Data Analysis: With R and GGobi. Springer. 2007.

Escofier, B. Analyse factorielle en reference a un modele: application a l'analyse d'un tableau d'echanges. *Revue de Statistique Appliquee*, Paris, v. 32, n. 4, p. 25-36, 1984.

Escofier, B.; Drouet, D. Analyse des differences entre plusieurs tableaux de frequence. *Les Cahiers de l'Analyse des Donnees*, Paris, v. 8, n. 4, p. 491-499, 1983.

Escofier, B.; Pages, J. *Analyse factorielles simples et multiples.* Paris: Dunod, 1990. 267 p.

Escofier, B.; Pages, J. *Analyses factorielles simples et multiples:* objectifs, methodes et interpretation. 4th ed. Paris: Dunod, 2008. 318 p.

Escofier, B.; Pages, J. *Comparaison de groupes de variables definies sur le meme ensemble d'individus:* un exemple d'applications. Le Chesnay: Institut National de Recherche en Informatique et en Automatique, 1982. 121 p.

Escofier, B.; Pages, J. Multiple factor analysis (AFUMULT package). *Computational Statistics & Data Analysis*, New York, v. 18, n. 1, p. 121-140, Aug. 1994

Espezua, S.; Villanueva, E.; Maciel, C. D.; Carvalho, A. A projection pursuit framework for supervised dimension reduction of high dimensional small sample datasets. *Neurocomputing*, 149, 767-776, 2015.

Ferreira, D. F. *Estatistica multivariada.* 2. ed. rev. e ampl. Lavras: UFLA, 2011. 675 p.

Friedman, J. H., Tukey, J. W. A projection pursuit algorithm for exploratory data analysis. *IEEE Transaction on Computers*, 23(9):881-890, 1974.

Greenacre, M.; Blasius, J. *Multiple correspondence analysis and related methods.* New York: Taylor and Francis, 2006. 607 p.

Hastie, T.; Buja, A.; Tibshirani, R. Penalized discriminant analysis. *The Annals of Statistics*. 23(1), 73-102 . 1995.

Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, Arlington, v. 24, p. 417-441, Sept. 1933.

Huber, P. J. Projection pursuit. *Annals of Statistics*, 13(2):435-475, 1985.

Hurley, C.; Buja, A. Analyzing high-dimensional data with motion graphics, *SIAM Journal of Scientific and Statistical Computing*, 11 (6), 1193-1211. 1990.

Johnson, R. A.; Wichern, D. W. *Applied multivariate statistical analysis.* 6th ed. New Jersey: Prentice Hall, 2007. 794 p.

Jones, M. C.; Sibson, R. What is projection pursuit, (with discussion), *Journal of the Royal Statistical Society*, Series A 150, 1-36, 1987.

Lee, E.; Cook, D.; Klinke, S.; Lumley, T. Projection pursuit for exploratory supervised classification. *Journal of Computational and Graphical Statistics*, 14(4):831-846, 2005.

Lee, E. K., Cook, D. A projection pursuit index for large p small n data. *Statistics and Computing*, 20(3):381-392, 2010.

Martinez, W. L.; Martinez, A. R. *Computational Statistics Handbook with MATLAB*, 2th. ed. New York: Chapman & Hall/CRC, 2007. 794 p.

Martinez, W. L.; Martinez, A. R.; Solka, J. *Exploratory Data Analysis with MATLAB*, 2th. ed. New York: Chapman & Hall/CRC, 2010. 499 p.

Mingoti, S. A. *Analise de dados atraves de metodos de estatistica multivariada:* uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 297 p.

Ossani, P. C.; Cirillo, M. A.; Borem, F. M.; Ribeiro, D. E.; Cortez, R. M. Quality of specialty coffees: a sensory evaluation by consumers using the MFACT technique. *Revista Ciencia Agronomica (UFC. Online)*, v. 48, p. 92-100, 2017.

Ossani, P. C. *Qualidade de cafes especiais e nao especiais por meio da analise de multiplos fatores para tabelas de contingencias.* 2015. 107 p. Dissertacao (Mestrado em Estatistica e Experimentacao Agropecuaria) - Universidade Federal de Lavras, Lavras, 2015.

Pages, J. Analyse factorielle multiple appliquee aux variables qualitatives et aux donnees mixtes. *Revue de Statistique Appliquee*, Paris, v. 50, n. 4, p. 5-37, 2002.

Pages, J. Multiple factor analysis: main features and application to sensory data. *Revista Colombiana de Estadistica*, Bogota, v. 27, n. 1, p. 1-26, 2004.

Pena, D.; Prieto, F. Cluster identification using projections. *Journal of the American Statistical Association*, 96(456):1433-1445, 2001.

Posse, C. Projection pursuit exploratory data analysis, *Computational Statistics and Data Analysis*, 29:669-687, 1995a.

Posse, C. Tools for two-dimensional exploratory projection pursuit, *Journal of Computational and Graphical Statistics*, 4:83-100, 1995b

Rencher, A.C.; *Methods of Multivariate Analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

Young, F. W.; Rheingans P. Visualizing structure in high-dimensional multivariate data, *IBM Journal of Research and Development*, 35:97-107, 1991.

Young, F. W.; Faldowski R. A.; McFarlane M. M. *Multivariate statistical visualization, in Handbook of Statistics*, Vol 9, C. R. Rao (ed.), The Netherlands: Elsevier Science Publishers, 959-998, 1993.

Biplot *Biplot graph.*

## Description

Plots the Biplot graph.

## Usage

```
Biplot(data, alpha = 0.5, title = NA, xlabel = NA, ylabel = NA,
       size = 1.1, grid = TRUE, color = TRUE, var = TRUE,
       obs = TRUE, linlab = NA, class = NA, classcolor = NA,
       posleg = 2, boxleg = TRUE, axes = TRUE, savptc = FALSE,
       width = 3236, height = 2000, res = 300)
```

## Arguments

| | |
|---|---|
| `data` | Data for plotting. |
| `alpha` | Representativeness of the individuals (alpha), representativeness of the variables (1 - alpha), being 0.5 the default. |
| `title` | Titles of the graphics, if not set, assumes the default text. |
| `xlabel` | Names the X axis, if not set, assumes the default text. |
| `ylabel` | Names the Y axis, if not set, assumes the default text. |
| `size` | Size of the points in the graphs. |
| `grid` | Put grid on graphs (default = TRUE). |
| `color` | Colored graphics (default = TRUE). |
| `var` | Adds the variable projections to graph (default = TRUE). |
| `obs` | Adds the observations to graph (default = TRUE). |
| `linlab` | Vector with the labels for the observations. |
| `class` | Vector with names of data classes. |
| `classcolor` | Vector with the colors of the classes. |
| `posleg` | 0 with no caption,<br>1 for caption in the left upper corner,<br>2 for caption in the right upper corner (default),<br>3 for caption in the right lower corner,<br>4 for caption in the left lower corner. |
| `boxleg` | Puts the frame in the caption (default = TRUE). |
| `axes` | Plots the X and Y axes (default = TRUE). |
| `savptc` | Saves graphics images to files (default = FALSE). |
| `width` | Graphics images width when savptc = TRUE (defaul = 3236). |
| `height` | Graphics images height when savptc = TRUE (default = 2000). |
| `res` | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |

## Value

| | |
|---|---|
| `Biplot` | Biplot graph. |
| `Md` | Matrix eigenvalues. |
| `Mu` | Matrix U (eigenvectors). |
| `Mv` | Matrix V (eigenvectors). |
| `coorI` | Coordinates of the individuals. |
| `coorV` | Coordinates of the variables. |
| `pvar` | Proportion of the principal components. |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## References

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

## Examples

```
data(iris) # dataset

data <- iris[,1:4]

Biplot(data)


cls <- iris[,5]
res <- Biplot(data, alpha = 0.6, title = "Biplot of data valuing individuals",
              class = cls, classcolor = c("goldenrod3","gray56","red"),
              posleg = 2, boxleg = FALSE, axes = TRUE, savptc = FALSE,
              width = 3236, height = 2000, res = 300)
print(res$pvar)


res <- Biplot(data, alpha = 0.4, title = "Graph valuing the variables",
              xlabel = "", ylabel = "", color = FALSE, obs = FALSE,
              savptc = FALSE, width = 3236, height = 2000, res = 300)
print(res$pvar)
```

CA *Correspondence Analysis (CA).*

## Description

Performs simple correspondence analysis (CA) and multiple (MCA) in a data set.

## Usage

```
CA(data, typdata = "f", typmatrix = "I")
```

## Arguments

| | |
|---|---|
| `data` | Data to be analyzed (contingency table). |
| `typdata` | "f" for frequency data (default),<br>"c" for qualitative data. |
| `typmatrix` | Matrix used for calculations when typdata = "c".<br>"I" for indicator matrix (default),<br>"B" for Burt's matrix. |

## Value

| | |
|---|---|
| `depdata` | Verify if the rows and columns are dependent, or independent by the chi-square test, at the 5% significance level. |
| `typdata` | Data type: "F" frequency or "C" qualitative. |
| `numcood` | Number of principal components. |
| `mtxP` | Matrix of the relative frequency. |
| `vtrR` | Vector with sums of the rows. |
| `vtrC` | Vector with sums of the columns. |
| `mtxPR` | Matrix with profile of the rows. |
| `mtxPC` | Matrix with profile of the columns |
| `mtxZ` | Matrix Z. |
| `mtxU` | Matrix with the eigenvectors U. |
| `mtxV` | Matrix with the eigenvectors V. |
| `mtxL` | Matrix with eigenvalues. |
| `mtxX` | Matrix with the principal coordinates of the rows. |
| `mtxY` | Matrix with the principal coordinates of the columns. |
| `mtxAutvlr` | Matrix of the inertias (variances), with the proportions and proportions accumulated. |

### Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

### References

Mingoti, S. A. *Analise de dados atraves de metodos de estatistica multivariada:* uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 297 p.

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

### See Also

Plot.CA

### Examples

```
data(DataFreq) # frequency data set

data <- DataFreq[,2:ncol(DataFreq)]

rownames(data) <- as.character(t(DataFreq[1:nrow(DataFreq),1]))

res <- CA(data = data, "f") # performs CA

print("Is there dependency between rows and columns?"); res$depdata

print("Number of principal coordinates:"); res$numcood

print("Principal coordinates of the rows:"); round(res$mtxX,2)

print("Principal coordinates of the columns:"); round(res$mtxY,2)

print("Inertia of the principal components:"); round(res$mtxAutvlr,2)
```

---

## CCA *Canonical Correlation Analysis(CCA).*

---

### Description

Perform Canonical Correlation Analysis (CCA) on a data set.

### Usage

```
CCA(X = NULL, Y = NULL, type = 1, test = "Bartlett", sign = 0.05)
```

## Arguments

| | |
|---|---|
| `X` | First group of variables of a data set. |
| `Y` | Second group of variables of a data set. |
| `type` | 1 for analysis using the covariance matrix (default),<br>2 for analysis using the correlation matrix. |
| `test` | Test of significance of the relationship between the group X and Y:<br>"Bartlett" (default) or "Rao". |
| `sign` | Test significance level (default 5%). |

## Value

| | |
|---|---|
| `Cxx` | Covariance matrix or correlation Cxx. |
| `Cyy` | Covariance matrix or correlation Cyy. |
| `Cxy` | Covariance matrix or correlation Cxy. |
| `Cyx` | Covariance matrix or correlation Cyx. |
| `var.UV` | Matrix with eigenvalues (variances) of the canonical pairs U and V. |
| `corr.UV` | Matrix of the correlation of the canonical pairs U and V. |
| `coef.X` | Matrix of the canonical coefficients of the group X. |
| `coef.Y` | Matrix of the canonical coefficients of the group Y. |
| `corr.X` | Matrix of the correlations between canonical variables and the original variables of the group X. |
| `corr.Y` | Matrix of the correlations between the canonical variables and the original variables of the group Y. |
| `score.X` | Matrix with the scores of the group X. |
| `score.Y` | Matrix with the scores of the group Y. |
| `sigtest` | Returns the significance test of the relationship between group X and Y: "Bartlett" (default) or "Rao". |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## References

Mingoti, S. A. *Analise de dados atraves de metodos de estatistica multivariada:* uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 297 p.

Ferreira, D. F. *Estatistica Multivariada.* 2a ed. revisada e ampliada. Lavras: Editora UFLA, 2011. 676 p.

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

Lattin, J.; Carrol, J. D.; Green, P. E. *Analise de dados multivariados.* 1th. ed. Sao Paulo: Cengage Learning, 2011. 455 p.

### See Also

Plot.CCA

### Examples

```
data(DataMix) # data set

data <- DataMix[,2:ncol(DataMix)]

rownames(data) <- DataMix[,1]

X <- data[,1:2]

Y <- data[,5:6]

res <- CCA(X, Y, type = 2, test = "Bartlett", sign = 0.05)

print("Matrix with eigenvalues (variances) of the canonical pairs U and V:"); round(res$var.UV,3)

print("Matrix of the correlation of the canonical pairs U and V:"); round(res$corr.UV,3)

print("Matrix of the canonical coefficients of the group X:"); round(res$coef.X,3)

print("Matrix of the canonical coefficients of the group Y:"); round(res$coef.Y,3)

print("Matrix of the correlations between the canonical
       variables and the original variables of the group X:"); round(res$corr.X,3)

print("Matrix of the correlations between the canonical
       variables and the original variables of the group Y:"); round(res$corr.Y,3)

print("Matrix with the scores of the group X:"); round(res$score.X,3)

print("Matrix with the scores of the group Y:"); round(res$score.Y,3)

print("test of significance of the canonical pairs:"); res$sigtest
```

---

## Cluster *Cluster Analysis.*

---

### Description

Performs hierarchical and non-hierarchical cluster analysis in a data set.

### Usage

```
Cluster(data, titles = NA, hierarquic = TRUE, analysis = "Obs",
        cor.abs = FALSE, normalize = FALSE, distance = "euclidean",
        method = "complete", horizontal = FALSE, num.groups = 0,
```

```
lambda = 2, savptc = FALSE, width = 3236, height = 2000,
res = 300, casc = TRUE)
```

## Arguments

| | |
|---|---|
| `data` | Data to be analyzed. |
| `titles` | Titles of the graphics, if not set, assumes the default text. |
| `hierarquic` | Hierarchical groupings (default = TRUE), for non-hierarchical groupings (method K-Means), only for case 'analysis' = "Obs". |
| `analysis` | "Obs" for analysis on observations (default), "Var" for analysis on variables. |
| `cor.abs` | Matrix of absolute correlation case 'analysis' = "Var" (default = FALSE). |
| `normalize` | Normalize the data only for case 'analysis' = "Obs" (default = FALSE). |
| `distance` | Metric of the distances in case of hierarchical groupings: "euclidean" (default), "maximum", "manhattan", "canberra", "binary" or "minkowski". Case Analysis = "Var" the metric will be the correlation matrix, according to cor.abs. |
| `method` | Method for analyzing hierarchical groupings: "complete" (default), "ward.D", "ward.D2", "single", "average", "mcquitty", "median" or "centroid". |
| `horizontal` | Horizontal dendrogram (default = FALSE). |
| `num.groups` | Number of groups to be formed. |
| `lambda` | Value used in the minkowski distance. |
| `savptc` | Saves graphics images to files (default = FALSE). |
| `width` | Graphics images width when savptc = TRUE (defaul = 3236). |
| `height` | Graphics images height when savptc = TRUE (default = 2000). |
| `res` | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |
| `casc` | Cascade effect in the presentation of the graphics (default = TRUE). |

## Value

Several graphics.

| | |
|---|---|
| `tab.res` | Table with similarities and distances of the groups formed. |
| `groups` | Original data with groups formed. |
| `res.groups` | Results of the groups formed. |
| `R.sqt` | Result of the R squared. |
| `sum.sqt` | Total sum of squares. |
| `mtx.dist` | Matrix of the distances. |

## Author(s)

Paulo Cesar Ossani

## References

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

Mingoti, S. A. *analysis de dados atraves de metodos de estatistica multivariada:* uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 297 p.

Ferreira, D. F. *Estatistica Multivariada.* 2a ed. revisada e ampliada. Lavras: Editora UFLA, 2011. 676 p.

## Examples

```
data(DataQuan) # set of quantitative data

data <- DataQuan[,2:8]

rownames(data) <- DataQuan[1:nrow(DataQuan),1]

res <- Cluster(data, titles = NA, hierarquic = TRUE, analysis = "Obs",
               cor.abs = FALSE, normalize = FALSE, distance = "euclidean",
               method = "ward.D", horizontal = FALSE, num.groups = 2,
               savptc = FALSE, width = 3236, height = 2000, res = 300,
               casc = FALSE)

print("R squared:"); res$R.sqt
# print("Total sum of squares:"); res$sum.sqt
print("Groups formed:"); res$groups
# print("Table with similarities and distances:"); res$tab.res
# print("Table with the results of the groups:"); res$res.groups
# print("Distance Matrix:"); res$mtx.dist

write.table(file=file.path(tempdir(),"SimilarityTable.csv"), res$tab.res, sep=";",
            dec=",",row.names = FALSE)
write.table(file=file.path(tempdir(),"GroupData.csv"), res$groups, sep=";",
            dec=",",row.names = TRUE)
write.table(file=file.path(tempdir(),"GroupResults.csv"), res$res.groups, sep=";",
            dec=",",row.names = TRUE)
```

---

# CoefVar *Coefficient of variation of the data.*

---

## Description

Find the coefficient of variation of the data, either overall or per column.

## Usage

```
CoefVar(data, type = 1)
```

## Arguments

| | |
|---|---|
| `data` | Data to be analyzed. |
| `type` | 1 Coefficient of overall variation (default),<br>2 Coefficient of variation per column. |

## Value

Coefficient of variation, either overall or per column.

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## References

Ferreira, D. F.; *Estatistica Basica.* 2 ed. rev. Lavras: UFLA, 2009. 664 p.

## Examples

```
data(DataQuan) # data set

data <- DataQuan[,2:8]

res <- CoefVar(data, type = 1) # Coefficient of overall variation
round(res,2)

res <- CoefVar(data, type = 2) # Coefficient of variation per column
round(res,2)
```

---

`DA` *Linear (LDA) and quadratic discriminant analysis (QDA).*

---

## Description

Perform linear and quadratic discriminant analysis.

## Usage

```
DA(data, class = NA, type = "lda", validation = "learning",
   method = "moment", prior = NA, testing = NA)
```

## Arguments

| | |
|---|---|
| `data` | Data to be classified. |
| `class` | Vector with data classes names. |
| `type` | "lda": linear discriminant analysis (default), or<br>"qda": quadratic discriminant analysis. |
| `validation` | Type of validation:<br>"learning" - data training (default), or<br>"testing" - classifies the data of the vector "testing". |
| `method` | Classification method:<br>"mle" to MLEs,<br>"mve" to use cov.mv,<br>"moment" (default) for standard mean and variance estimators, or<br>"t" for robust estimates based on the t distribution. |
| `prior` | Probabilities of occurrence of classes. If not specified, it will take the proportions of the classes. If specified, probabilities must follow the order of factor levels. |
| `testing` | Vector with indices that will be used in data as test. For validation = "learning", one has testing = NA. |

## Value

| | |
|---|---|
| `confusion` | Confusion table. |
| `error.rate` | Overall error ratio. |
| `prior` | Probability of classes. |
| `type` | Type of discriminant analysis. |
| `validation` | Type of validation. |
| `num.class` | Number of classes. |
| `class.names` | Class names. |
| `method` | Classification method. |
| `num.correct` | Number of correct observations. |
| `results` | Matrix with comparative classification results. |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## References

Ferreira, D. F. *Estatistica Multivariada.* 2a ed. revisada e ampliada. Lavras: Editora UFLA, 2011. 676 p.

Mingoti, S. A. *Analise de dados atraves de metodos de estatistica multivariada:* uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 297 p.

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

Ripley, B. D. *Pattern Recognition and Neural Networks.* Cambridge University Press, 1996.

Venabless, W. N.; Ripley, B. D. *Modern Applied Statistics with S.* Fourth edition. Springer, 2002.

## Examples

```
data(iris) # data set

data  = iris[,1:4] # data to be classified
class = iris[,5]   # data class
prior = c(1,1,1)/3 # a priori probability of the classs

res <- DA(data, class, type = "lda", validation = "learning",
          method = "mle", prior = prior, testing = NA)

print("confusion table:"); res$confusion
print("Overall hit ratio:"); 1 - res$error.rate
print("Probability of classes:"); res$prior
print("classification method:"); res$method
print("type of discriminant analysis:"); res$type
print("class names:"); res$class.names
print("Number of classess:"); res$num.class
print("type of validation:"); res$validation
print("Number of correct observations:"); res$num.correct
print("Matrix with comparative classification results:"); res$results


### cross-validation ###
amostra   = sample(2, nrow(data), replace = TRUE, prob = c(0.7,0.3))
datatrain = data[amostra == 1,] # training data
datatest  = data[amostra == 2,] # test data

dim(datatrain) # training data dimension
dim(datatest)  # test data dimension

testing  = as.integer(rownames(datatest)) # test data index

res <- DA(data, class, type = "qda", validation = "testing",
          method = "moment", prior = NA, testing = testing)

print("confusion table:"); res$confusion
print("Overall hit ratio:"); 1 - res$error.rate
print("Number of correct observations:"); res$num.correct
print("Matrix with comparative classification results:"); res$results
```

---

DataCoffee *Frequency data set.*

---

## Description

Set of data categorized by coffees, on sensorial abilities in the consumption of special coffees.

## Usage

```
data(DataCoffee)
```

## Format

Data set of a research done with the purpose of evaluating the concordance between the responses of different groups of consumers with different sensorial abilities. The experiment relates the sensorial analysis of special coffees defined by (A) Yellow Bourbon, cultivated at altitudes greater than 1200 m; (D) idem to (A) differing only in the preparation of the samples; (B) Acaia cultivated at an altitude of less than 1,100 m; (C) identical to (B) but differentiating the sample preparation. Here the data are categorized by coffees. The example given demonstrates the results found in OSSANI et al. (2017).

## References

Ossani, P. C.; Cirillo, M. A.; Borem, F. M.; Ribeiro, D. E.; Cortez, R. M.. Quality of specialty coffees: a sensory evaluation by consumers using the MFACT technique. *Revista Ciencia Agronomica (UFC. Online)*, v. 48, p. 92-100, 2017.

Ossani, P. C. *Qualidade de cafes especiais e nao especiais por meio da analise de multiplos fatores para tabelas de contingencias.* 2015. 107 p. Dissertacao (Mestrado em Estatistica e Experimentacao Agropecuaria) - Universidade Federal de Lavras, Lavras, 2015.

## Examples

```
data(DataCoffee) # categorized data set

data <- DataCoffee[,2:ncol(DataCoffee)]

rownames(data) <- as.character(t(DataCoffee[1:nrow(DataCoffee),1]))

group.names = c("Coffee A", "Coffee B", "Coffee C", "Coffee D")

mf <- MFA(data, c(16,16,16,16), c(rep("f",4)), group.names)

print("Principal components variances:"); round(mf$mtxA,2)

print("Matrix of the Partial Inertia / Score of the Variables:"); round(mf$mtxEV,2)

tit <- c("Scree-plot","Individuals","Individuals / Types coffees","Inercias Groups")

Plot.MFA(mf, titles = tit, xlabel = NA, ylabel = NA,
         posleg = 2, boxleg = FALSE, color = TRUE,
        namarr = FALSE, linlab = NA, casc = FALSE) # plotting several graphs on the screen
```

---

`DataFreq` *Frequency data set.*

---

## Description

Simulated data set with the weekly frequency of the number of coffee cups consumed weekly in some world capitals.

## Usage

```
data(DataFreq)
```

## Format

Set of data with 6 rows and 9 columns. There are 6 observations described by 9 variables: Group by sex and age, Sao Paulo - Cafe Bourbon, London - Cafe Bourbon, Athens - Cafe Bourbon, London - Cafe Acaia, Athens - Cafe Catuai, Sao Paulo - Cafe Catuai, Athens - Cafe Catuai.

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## Examples

```
data(DataFreq)
DataFreq
```

---

`DataInd` *Frequency data set.*

---

## Description

Set of data categorized by coffees, on sensorial abilities in the consumption of special coffees.

## Usage

```
data(DataInd)
```

## Format

Data set of a research done with the purpose of evaluating the concordance between the responses of different groups of consumers with different sensorial abilities. The experiment relates the sensorial analysis of special coffees defined by (A) Yellow Bourbon, cultivated at altitudes greater than 1200 m; (D) idem to (A) differing only in the preparation of the samples; (B) Acaia cultivated at an altitude of less than 1,100 m; (C) identical to (B) but differentiating the sample preparation. Here the data are categorized by coffees. The example given demonstrates the results found in OSSANI et al. (2017).

## References

Ossani, P. C.; Cirillo, M. A.; Borem, F. M.; Ribeiro, D. E.; Cortez, R. M.. Quality of specialty coffees: a sensory evaluation by consumers using the MFACT technique. *Revista Ciencia Agronomica (UFC. Online)*, v. 48, p. 92-100, 2017.

Ossani, P. C. *Qualidade de cafes especiais e nao especiais por meio da analise de multiplos fatores para tabelas de contingencias.* 2015. 107 p. Dissertacao (Mestrado em Estatistica e Experimentacao Agropecuaria) - Universidade Federal de Lavras, Lavras, 2015.

## Examples

```
data(DataInd) # categorized data set

data <- DataInd[,2:ncol(DataInd)]

rownames(data) <- as.character(t(DataInd[1:nrow(DataInd),1]))

group.names = c("Group 1", "Group 2", "Group 3", "Group 4")

mf <- MFA(data, c(16,16,16,16), c(rep("f",4)), group.names)

print("Principal components variances:"); round(mf$mtxA,2)

print("Matrix of the Partial Inertia / Score of the Variables:"); round(mf$mtxEV,2)

tit <- c("Scree-plot","Individuals","Individuals / Types coffees","Inercias Groups")

Plot.MFA(mf, titles = tit, xlabel = NA, ylabel = NA,
         posleg = 2, boxleg = FALSE, color = TRUE,
        namarr = FALSE, linlab = NA, casc = FALSE) # plotting several graphs on the screen
```

---

DataMix *Mixed data set.*

---

## Description

Simulated set of mixed data on consumption of coffee.

## Usage

```
data(DataMix)
```

## Format

Data set with 10 rows and 7 columns. Being 10 observations described by 7 variables: Cooperatives/Tasters, Average grades given to analyzed coffees, Years of work as a taster, Taster with technical training, Taster exclusively dedicated, Average frequency of the coffees Classified as special, Average frequency of the coffees as commercial.

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## Examples

```
data(DataMix)
DataMix
```

---

DataQuali *Qualitative data set*

---

## Description

Set simulated of qualitative data on consumption of coffee.

## Usage

```
data(DataQuali)
```

## Format

Data set simulated with 12 rows and 6 columns. Being 12 observations described by 6 variables: Sex, Age, Smoker, Marital status, Sportsman, Study.

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## Examples

```
data(DataQuali)
DataQuali
```

---

DataQuan *Quantitative data set*

---

## Description

Set simulated of quantitative data on grades given to some sensory characteristics of coffees.

## Usage

```
data(DataQuan)
```

## Format

Data set with 6 rows and 11 columns. Being 6 observations described by 11 variables: Coffee, Chocolate, Caramelised, Ripe, Sweet, Delicate, Nutty, Caramelised, Chocolate, Spicy, Caramelised.

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## Examples

```
data(DataQuan)
DataQuan
```

---

FA *Factor Analysis (FA).*

---

## Description

Performs factorial analysis (FA) in a data set.

## Usage

```
FA(data, method = "PC", type = 2, nfactor = 1, rotation = "None",
   scoresobs = "Bartlett", converg = 1e-5, iteracao = 1000,
   testfit = TRUE)
```

## Arguments

| | |
|---|---|
| `data` | Data to be analyzed. |
| `method` | Method of analysis:<br>"PC" - Principal Components (default),<br>"PF" - Principal Factor,<br>"ML" - Maximum Likelihood. |
| `type` | 1 for analysis using the covariance matrix,<br>2 for analysis using the correlation matrix (default). |
| `rotation` | Type of rotation: "None" (default), "Varimax" and "Promax". |
| `nfactor` | Number of factors (default = 1). |
| `scoresobs` | Type of scores for the observations: "Bartlett" (default) or "Regression". |
| `converg` | Limit value for convergence to sum of the squares of the residuals for Maximum likelihood method (default = 1e-5). |
| `iteracao` | Maximum number of iterations for Maximum Likelihood method (default = 1000). |
| `testfit` | Tests the model fit to the method of Maximum Likelihood (default = TRUE). |

## Value

| | |
|---|---|
| `mtxMC` | Matrix of correlation / covariance. |
| `mtxAutvlr` | Matrix of eigenvalues. |
| `mtxAutvec` | Matrix of eigenvectors. |
| `mtxvar` | Matrix of variances and proportions. |
| `mtxcarga` | Matrix of factor loadings. |
| `mtxvaresp` | Matrix of specific variances. |
| `mtxcomuna` | Matrix of commonalities. |
| `mtxresidue` | Matrix of residues. |
| `vlrsqrs` | Upper limit value for sum of squares of the residues. |
| `vlrsqr` | Sum of squares of the residues. |
| `mtxresult` | Matrix with all associated results. |
| `mtxscores` | Matrix with scores of the observations. |
| `coefscores` | Matrix with the scores of the coefficients of the factors. |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## References


Mingot, S. A. *Analise de dados atraves de metodos de estatistica multivariada:* uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 297 p.

Kaiser, H. F.*The varimax criterion for analytic rotation in factor analysis.* Psychometrika 23, 187-200, 1958.

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

Ferreira, D. F. *Estatistica Multivariada.* 2a ed. revisada e ampliada. Lavras: Editora UFLA, 2011. 676 p.


## See Also

`Plot.FA`

## Examples

```
data(DataQuan) # data set

data <- DataQuan[,2:ncol(DataQuan)]

rownames(data) <- DataQuan[,1]

res <- FA(data, method = "PC", type = 2, nfactor = 3, rotation = "None",
          scoresobs = "Bartlett", converg = 1e-5, iteracao = 1000,
          testfit = TRUE)

print("Matrix with all associated results:"); round(res$mtxresult,3)

print("Sum of squares of the residues:"); round(res$vlrsqr,3)

print("Matrix of the factor loadings.:"); round(res$mtxcarga,3)

print("Matrix with scores of the observations:"); round(res$mtxscores,3)

print("Matrix with the scores of the coefficients of the factors:"); round(res$coefscores,3)
```

---

`GrandTour` *Animation technique Grand Tour.*

---

## Description

Performs the exploration of the data through the technique of animation Grand Tour.

## Usage

```
GrandTour(data, method = "Interpolation", title = NA, xlabel = NA,
          ylabel = NA, size = 1.1, grid = TRUE, color = TRUE, linlab = NA,
          class = NA, classcolor = NA, posleg = 2, boxleg = TRUE,
          axesvar = TRUE, axes = TRUE, numrot = 200, choicerot = NA,
          savptc = FALSE, width = 3236, height = 2000, res = 300)
```

## Arguments

| | |
|---|---|
| `data` | Numerical data set. |
| `method` | Method used for rotations:<br>"Interpolation" - Interpolation method (default),<br>"Torus" - Torus method,<br>"Pseudo" - Pseudo Grand Tour method. |
| `title` | Titles of the graphics, if not set, assumes the default text. |
| `xlabel` | Names the X axis, if not set, assumes the default text. |
| `ylabel` | Names the Y axis, if not set, assumes the default text. |
| `size` | Size of the points in the graphs. |
| `grid` | Put grid on graphs (default = TRUE). |
| `color` | Colored graphics (default = TRUE). |
| `linlab` | Vector with the labels for the observations. |
| `class` | Vector with names of data classes. |
| `classcolor` | Vector with the colors of the classes. |
| `posleg` | 0 with no caption,<br>1 for caption in the left upper corner,<br>2 for caption in the right upper corner (default),<br>3 for caption in the right lower corner,<br>4 for caption in the left lower corner. |
| `boxleg` | Puts the frame in the caption (default = TRUE). |
| `axesvar` | Puts axes of rotation of the variables (default = TRUE). |
| `axes` | Plots the X and Y axes (default = TRUE). |
| `numrot` | Number of rotations (default = 200). If method = "Interpolation", numrot represents the angle of rotation. |
| `choicerot` | Choose specific rotation and display on the screen, or save the image if savptc = TRUE. |
| `savptc` | Saves graphics images to files (default = FALSE). |
| `width` | Graphics images width when savptc = TRUE (defaul = 3236). |
| `height` | Graphics images height when savptc = TRUE (default = 2000). |
| `res` | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |

### Value

Graphs with rotations.

| | |
|---|---|
| proj.data | Projected data. |
| vector.opt | Vector projection. |
| method | method used on Grand Tour. |

### Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

### References

Asimov, D. The Grand Tour: A Tool for Viewing Multidimensional data. *SIAM Journal of Scientific and Statistical Computing*, 6(1), 128-143, 1985.

Asimov, D.; Buja, A. The grand tour via geodesic interpolation of 2-frames. in Visual data Exploration and Analysis. *Symposium on Electronic Imaging Science and Technology*, IS&T/SPIE. 1994.

Buja, A.; Asimov, D. Grand tour methods: An outline. *Computer Science and Statistics*, 17:63-67. 1986.

Buja, A.; Cook, D.; Asimov, D.; Hurley, C. Computational methods for High-Dimensional Rotations in data Visualization, in C. R. Rao, E. J. Wegman & J. L. Solka, eds, *"Handbook of Statistics: data Mining and Visualization"*, Elsevier/North Holland, http://www.elsevier.com, pp. 391-413. 2005.

Hurley, C.; Buja, A. Analyzing high-dimensional data with motion graphics, *SIAM Journal of Scientific and Statistical Computing*, 11 (6), 1193-1211. 1990.

Martinez, W. L.; Martinez, A. R.; Solka, J.; *Exploratory data Analysis with MATLAB*, 2th. ed. New York: Chapman & Hall/CRC, 2010. 499 p.

Young, F. W.; Rheingans P. Visualizing structure in high-dimensional multivariate data, *IBM Journal of Research and Development*, 35:97-107, 1991.

Young, F. W.; Faldowski R. A.; McFarlane M. M. *Multivariate statistical visualization, in Handbook of Statistics*, Vol 9, C. R. Rao (ed.), The Netherlands: Elsevier Science Publishers, 959-998, 1993.

### Examples

```
data(iris) # database

res <- GrandTour(iris[,1:4], method = "Torus", title = NA, xlabel = NA, ylabel = NA,
                 color = TRUE, linlab = NA, class = NA, posleg = 2, boxleg = TRUE,
                 axesvar = TRUE, axes = FALSE, numrot = 10, choicerot = NA,
                 savptc = FALSE, width = 3236, height = 2000, res = 300)

print("Projected data:"); res$proj.data
print("Projection vectors:"); res$vector.opt
print("Grand Tour projection method:"); res$method
```

```
res <- GrandTour(iris[,1:4], method = "Interpolation", title = NA, xlabel = NA, ylabel = NA,
                 color = TRUE, linlab = NA, posleg = 2, boxleg = FALSE, axesvar = FALSE,
                 axes = FALSE, numrot = 10, choicerot = NA, class = iris[,5],
                 classcolor = c("goldenrod3","gray53","red"),savptc = FALSE,
                 width = 3236, height = 2000, res = 300)

print("Projected data:"); res$proj.data
print("Projection vectors:"); res$vector.opt
print("Grand Tour projection method:"); res$method
```

---

GSVD *Generalized Singular Value Decomposition (GSVD).*

---

## Description

Given the matrix $A$ of order $nxm$, the generalized singular value decomposition (GSVD) involves the use of two sets of positive square matrices of order $nxn$ and $mxm$ respectively. These two matrices express constraints imposed, respectively, on the lines and columns of $A$.

## Usage

```
GSVD(data, plin = NULL, pcol = NULL)
```

## Arguments

| | |
|---|---|
| data | Matrix used for decomposition. |
| plin | Weight for rows. |
| pcol | Weight for columns |

## Details

If plin or pcol is not used, it will be calculated as the usual singular value decomposition.

## Value

| | |
|---|---|
| d | Eigenvalues, that is, line vector with singular values of the decomposition. |
| u | Eigenvectors referring rows. |
| v | Eigenvectors referring columns. |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

### References

Abdi, H. Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition (GSVD). In: SALKIND, N. J. (Ed.). *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage, 2007. p. 907-912.

### Examples

```
data <- matrix(c(1,2,3,4,5,6,7,8,9,10,11,12), nrow = 4, ncol = 3)

svd(data)  # Usual Singular Value Decomposition

GSVD(data) # GSVD with the same previous results

# GSVD with weights for rows and columns
GSVD(data, plin = c(0.1,0.5,2,1.5), pcol = c(1.3,2,0.8))
```

---

## IM — *Indicator matrix.*

---

### Description

In the indicator matrix the elements are arranged in the form of *dummy* variables, in other words, 1 for a category chosen as a response variable and 0 for the other categories of the same variable.

### Usage

```
IM(data, names = TRUE)
```

### Arguments

| | |
|---|---|
| `data` | Categorical data. |
| `names` | Include the names of the variables in the levels of the Indicator Matrix (default = TRUE). |

### Value

| | |
|---|---|
| `mtxIndc` | Returns converted data in the indicator matrix. |

### Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

### References

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

## Examples

```
data <- matrix(c("S","S","N","N",1,2,3,4,"N","S","T","N"), nrow = 4, ncol = 3)

IM(data, names = FALSE)

data(DataQuali) # qualitative data set

IM(DataQuali, names = TRUE)
```

---

LocLab *Function for better position of the labels in the graphs.*

---

## Description

Function for better position of the labels in the graphs.

## Usage

```
LocLab(x, y = NULL, labels = seq(along = x), cex = 1,
       method = c("SANN", "GA"), allowSmallOverlap = FALSE,
       trace = FALSE, shadotext = FALSE,
       doPlot = TRUE, ...)
```

## Arguments

| | |
|---|---|
| x | Coordinate x |
| y | Coordinate y |
| labels | The labels |
| cex | cex |
| method | Not used |
| allowSmallOverlap | Boolean |
| trace | Boolean |
| shadotext | Boolean |
| doPlot | Boolean |
| ... | Other arguments passed to or from other methods |

## Value

See the text of the function.

MDS *Multidimensional Scaling (MDS).*

## Description

Performs Multidimensional Scaling (MDS) on a data set.

## Usage

```
MDS(data, distance = "euclidean", title = NA, xlabel = NA,
    ylabel = NA, posleg = 2, boxleg = TRUE, axes = TRUE,
    size = 1.1, grid = TRUE, color = TRUE, linlab = NA,
    class = NA, classcolor = NA, savptc = FALSE, width = 3236,
    height = 2000, res = 300)
```

## Arguments

| | |
|---|---|
| `data` | Data to be analyzed. |
| `distance` | Metric of the distance: "euclidean" (default), "maximum", "manhattan", "canberra", "binary" or "minkowski". |
| `title` | Titles of the graphics, if not set, assumes the default text. |
| `xlabel` | Names the X axis, if not set, assumes the default text. |
| `ylabel` | Names the Y axis, if not set, assumes the default text. |
| `posleg` | 0 with no caption,<br>1 for caption in the left upper corner,<br>2 for caption in the right upper corner (default),<br>3 for caption in the right lower corner,<br>4 for caption in the left lower corner. |
| `boxleg` | Puts the frame in the caption (default = TRUE). |
| `axes` | Plot the X and Y axes (default = TRUE). |
| `size` | Size of the points in the graphs. |
| `grid` | Put grid on graphs (default = TRUE). |
| `color` | Colored graphics (default = TRUE). |
| `linlab` | Vector with the labels for the observations. |
| `class` | Vector with names of data classes. |
| `classcolor` | Vector with the colors of the classes. |
| `savptc` | Saves graphics images to files (default = FALSE). |
| `width` | Graphics images width when savptc = TRUE (defaul = 3236). |
| `height` | Graphics images height when savptc = TRUE (default = 2000). |
| `res` | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |

### Value

Multidimensional Scaling.

`mtxD` Matrix of the distances.

### Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

### References

Mingoti, S. A. *Analise de dados atraves de metodos de estatistica multivariada:* uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 297 p.

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

### Examples

```
data(iris) # data set

data <- iris[,1:4]

cls <- iris[,5] # data class

md <- MDS(data = data, distance = "euclidean", title = NA, xlabel = NA,
          ylabel = NA, posleg = 2, boxleg = TRUE, axes = TRUE, color = TRUE,
          linlab = NA, class = cls, classcolor = c("goldenrod3","gray53","red"),
          savptc = FALSE, width = 3236, height = 2000, res = 300)

print("Matrix of the distances:"); md$mtxD
```

---

## MFA *Multiple Factor Analysis (MFA).*

---

### Description

Perform Multiple Factor Analysis (MFA) on groups of variables. The groups of variables can be quantitative, qualitative, frequency (MFACT) data, or mixed data.

### Usage

```
MFA(data, groups, typegroups = rep("n",length(groups)), namegroups = NULL)
```

## Arguments

| | |
|---|---|
| `data` | Data to be analyzed. |
| `groups` | Number of columns for each group in order following the order of data in 'data'. |
| `typegroups` | Type of group:<br>"n" for numerical data (default),<br>"c" for categorical data,<br>"f" for frequency data. |
| `namegroups` | Names for each group. |

## Value

| | |
|---|---|
| `vtrG` | Vector with the sizes of each group. |
| `vtrNG` | Vector with the names of each group. |
| `vtrplin` | Vector with the values used to balance the lines of the Z matrix. |
| `vtrpcol` | Vector with the values used to balance the columns of the Z matrix. |
| `mtxZ` | Matrix concatenated and balanced. |
| `mtxA` | Matrix of the eigenvalues (variances) with the proportions and proportions accumulated. |
| `mtxU` | Matrix U of the singular decomposition of the matrix Z. |
| `mtxV` | Matrix V of the singular decomposition of the matrix Z. |
| `mtxF` | Matrix global factor scores where the lines are the observations and the columns the components. |
| `mtxEFG` | Matrix of the factor scores by group. |
| `mtxCCP` | Matrix of the correlation of the principal components with original variables. |
| `mtxEV` | Matrix of the partial inertias / scores of the variables |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## References


Abdessemed, L.; Escofier, B. Analyse factorielle multiple de tableaux de frequencies: comparaison avec l'analyse canonique des correspondences. *Journal de la Societe de Statistique de Paris*, Paris, v. 137, n. 2, p. 3-18, 1996..

Abdi, H. Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition (GSVD). In: SALKIND, N. J. (Ed.). *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage, 2007. p. 907-912.

Abdi, H.; Valentin, D. Multiple factor analysis (MFA). In: SALKIND, N. J. (Ed.). *Encyclopedia of measurement and statistics.* Thousand Oaks: Sage, 2007. p. 657-663.

Abdi, H.; Williams, L. Principal component analysis. *WIREs Computational Statatistics*, New York, v. 2, n. 4, p. 433-459, July/Aug. 2010.

Abdi, H.; Williams, L.; Valentin, D. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. *WIREs Computational Statatistics*, New York, v. 5, n. 2, p. 149-179, Feb. 2013.

Becue-Bertaut, M.; Pages, J. A principal axes method for comparing contingency tables: MFACT. *Computational Statistics & data Analysis*, New York, v. 45, n. 3, p. 481-503, Feb. 2004

Becue-Bertaut, M.; Pages, J. Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data. *Computational Statistics & data Analysis*, New York, v. 52, n. 6, p. 3255-3268, Feb. 2008.

Bezecri, J. Analyse de l'inertie intraclasse par l'analyse d'un tableau de contingence: intra-classinertia analysis through the analysis of a contingency table. *Les Cahiers de l'Analyse des Donnees*, Paris, v. 8, n. 3, p. 351-358, 1983.

Escofier, B. Analyse factorielle en reference a un modele: application a l'analyse d'un tableau d'echanges. *Revue de Statistique Appliquee*, Paris, v. 32, n. 4, p. 25-36, 1984.

Escofier, B.; Drouet, D. Analyse des differences entre plusieurs tableaux de frequence. *Les Cahiers de l'Analyse des Donnees*, Paris, v. 8, n. 4, p. 491-499, 1983.

Escofier, B.; Pages, J. *Analyse factorielles simples et multiples.* Paris: Dunod, 1990. 267 p.

Escofier, B.; Pages, J. *Analyses factorielles simples et multiples:* objectifs, methodes et interpretation. 4th ed. Paris: Dunod, 2008. 318 p.

Escofier, B.; Pages, J. *Comparaison de groupes de variables definies sur le meme ensemble d'individus:* un exemple d'applications. Le Chesnay: Institut National de Recherche en Informatique et en Automatique, 1982. 121 p.

Escofier, B.; Pages, J. Multiple factor analysis (AFUMULT package). *Computational Statistics & data Analysis*, New York, v. 18, n. 1, p. 121-140, Aug. 1994

Greenacre, M.; Blasius, J. *Multiple correspondence analysis and related methods.* New York: Taylor and Francis, 2006. 607 p.

Ossani, P. C.; Cirillo, M. A.; Borem, F. M.; Ribeiro, D. E.; Cortez, R. M. Quality of specialty coffees: a sensory evaluation by consumers using the MFACT technique. *Revista Ciencia Agronomica (UFC. Online)*, v. 48, p. 92-100, 2017.

Pages, J. Analyse factorielle multiple appliquee aux variables qualitatives et aux donnees mixtes. *Revue de Statistique Appliquee*, Paris, v. 50, n. 4, p. 5-37, 2002.

Pages, J.. Multiple factor analysis: main features and application to sensory data. *Revista Colombiana de Estadistica*, Bogota, v. 27, n. 1, p. 1-26, 2004.

## See Also

Plot.MFA

## Examples

```
data(DataMix) # mixed dataset

data <- DataMix[,2:ncol(DataMix)]

rownames(data) <- DataMix[1:nrow(DataMix),1]
```

```
group.names = c("Grade Cafes/Work", "Formation/Dedication", "Coffees")

mf <- MFA(data = data, c(2,2,2), typegroups = c("n","c","f"), group.names) # performs MFA

print("Principal Component Variances:"); round(mf$mtxA,2)

print("Matrix of the Partial Inertia / Score of the Variables:"); round(mf$mtxEV,2)
```

---

NormData *Normalizes the data.*

---

## Description

Function that normalizes the data globally, or by column.

## Usage

```
NormData(data, type = 1)
```

## Arguments

| | |
|---|---|
| data | Data to be analyzed. |
| type | 1 normalizes overall (default),<br>2 normalizes per column. |

## Value

| | |
|---|---|
| dataNorm | Normalized data. |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## Examples

```
data(DataQuan) # set of quantitative data

data <- DataQuan[,2:8]

res  <- NormData(data, type = 1) # normalizes the data globally

res # Globally standardized data

sd(res)   # overall standard deviation

mean(res) # overall mean
```

```
res <- NormData(data, type = 2) # normalizes the data per column

res # standardized data per column

apply(res, 2, sd) # standard deviation per column

colMeans(res)     # column averages
```

---

NormTest *Test of normality of the data.*

---

## Description

Check the normality of the data, based on the asymmetry coefficient test.

## Usage

```
NormTest(data, sign = 0.05)
```

## Arguments

| | |
|---|---|
| `data` | Data to be analyzed. |
| `sign` | Test significance level (default 5%). |

## Value

| | |
|---|---|
| `statistic` | Observed Chi-square value, that is, the test statistic. |
| `chisquare` | Chi-square value calculated. |
| `gl` | Degree of freedom. |
| `p.value` | p-value. |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## References

Mingoti, S. A. *Analise de dados atraves de metodos de estatistica multivariada:* uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 297 p.

Rencher, A. C. *Methods of Multivariate Analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

Ferreira, D. F. *Estatistica Multivariada.* 2a ed. revisada e ampliada. Lavras: Editora UFLA, 2011. 676 p.

## Examples

```
data <- cbind(rnorm(100,2,3), rnorm(100,1,2))

NormTest(data)

plot(density(data))


data <- cbind(rexp(200,3), rexp(200,3))

NormTest(data, sign = 0.01)

plot(density(data))
```

---

PCA *Principal Components Analysis (PCA).*

---

## Description

Performs principal component analysis (PCA) in a data set.

## Usage

```
PCA(data, type = 1)
```

## Arguments

| | |
|---|---|
| data | Data to be analyzed. |
| type | 1 for analysis using the covariance matrix (default),<br>2 for analysis using the correlation matrix. |

## Value

| | |
|---|---|
| mtxC | Matrix of covariance or correlation according to "type". |
| mtxAutvlr | Matrix of eigenvalues (variances) with the proportions and proportions accumulated. |
| mtxAutvec | Matrix of eigenvectors - principal components. |
| mtxVCP | Matrix of covariance of the principal components with the original variables. |
| mtxCCP | Matrix of correlation of the principal components with the original variables. |
| mtxscores | Matrix with scores of the principal components. |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## References

Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, Arlington, v. 24, p. 417-441, Sept. 1933.

Mingoti, S. A. *Analise de dados atraves de metodos de estatistica multivariada:* uma abordagem aplicada. Belo Horizonte: UFMG, 2005. 297 p.

Ferreira, D. F. *Estatistica Multivariada.* 2a ed. revisada e ampliada. Lavras: Editora UFLA, 2011. 676 p.

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

## See Also

Plot.PCA

## Examples

```
data(DataQuan) # set of quantitative data

data <- DataQuan[,2:8]

rownames(data) <- DataQuan[1:nrow(DataQuan),1]

pc <- PCA(data, 2) # performs the PCA

print("Covariance matrix / Correlation:"); round(pc$mtxC,2)

print("Principal Components:"); round(pc$mtxAutvec,2)

print("Principal Component Variances:"); round(pc$mtxAutvlr,2)

print("Covariance of the Principal Components:"); round(pc$mtxVCP,2)

print("Correlation of the Principal Components:"); round(pc$mtxCCP,2)

print("Scores of the Principal Components:"); round(pc$mtxscores,2)
```

---

`Plot.CA` *Graphs of the simple (CA) and multiple correspondence analysis (MCA).*

---

## Description

Graphs of the simple (CA) and multiple correspondence analysis (MCA).

## Usage

```
Plot.CA(CA, titles = NA, xlabel = NA, ylabel = NA,
        size = 1.1, grid = TRUE, color = TRUE, linlab = NA,
        savptc = FALSE, width = 3236, height = 2000,
        res = 300, casc = TRUE)
```

### Arguments

| | |
|---|---|
| CA | Data of the CA function. |
| titles | Titles of the graphics, if not set, assumes the default text.. |
| xlabel | Names the X axis, if not set, assumes the default text. |
| ylabel | Names the Y axis, if not set, assumes the default text. |
| size | Size of the points in the graphs. |
| grid | Put grid on graphs (default = TRUE). |
| color | Colored graphics (default = TRUE). |
| linlab | Vector with the labels for the observations. |
| savptc | Saves graphics images to files (default = FALSE). |
| width | Graphics images width when savptc = TRUE (defaul = 3236). |
| height | Graphics images height when savptc = TRUE (default = 2000). |
| res | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |
| casc | Cascade effect in the presentation of the graphics (default = TRUE). |

### Value

Returns several graphs.

### Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

### See Also

CA

### Examples

```
data(DataFreq) # frequency data set

data <- DataFreq[,2:ncol(DataFreq)]

rownames(data) <- DataFreq[1:nrow(DataFreq),1]

res <- CA(data, "f") # performs CA

tit <- c("Scree-plot","Observations", "Variables", "Observations / Variables")

Plot.CA(res, titles = tit, xlabel = NA, ylabel = NA,
        color = TRUE, linlab = rownames(data), savptc = FALSE,
        width = 3236, height = 2000, res = 300, casc = FALSE)
```

```
data(DataQuali) # qualitative data set

data <- DataQuali[,2:ncol(DataQuali)]

res <- CA(data, "c", "b") # performs CA

tit <- c("","","Graph of the variables")

Plot.CA(res, titles = tit, xlabel = NA, ylabel = NA,
        color = TRUE, linlab = NA, savptc = FALSE,
        width = 3236, height = 2000, res = 300,
        casc = FALSE)
```

---

`Plot.CCA` *Graphs of the Canonical Correlation Analysis (CCA).*

---

## Description

Graphs of the Canonical Correlation Analysis (CCA).

## Usage

```
Plot.CCA(CCA, titles = NA, xlabel = NA, ylabel = NA,
         size = 1.1, grid = TRUE, color = TRUE, savptc = FALSE,
         width = 3236, height = 2000, res = 300, casc = TRUE)
```

## Arguments

| | |
|---|---|
| `CCA` | Data of the CCA function. |
| `titles` | Titles of the graphics, if not set, assumes the default text. |
| `xlabel` | Names the X axis, if not set, assumes the default text. |
| `ylabel` | Names the Y axis, if not set, assumes the default text. |
| `size` | Size of the points in the graphs. |
| `grid` | Put grid on graphs (default = TRUE). |
| `color` | Colored graphics (default = TRUE). |
| `savptc` | Saves graphics images to files (default = FALSE). |
| `width` | Graphics images width when savptc = TRUE (defaul = 3236). |
| `height` | Graphics images height when savptc = TRUE (default = 2000). |
| `res` | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |
| `casc` | Cascade effect in the presentation of the graphics (default = TRUE). |

## Value

Returns several graphs.

### Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

### See Also

CCA

### Examples

```
data(DataMix) # database

data <- DataMix[,2:ncol(DataMix)]

rownames(data) <- DataMix[,1]

X <- data[,1:2]

Y <- data[,5:6]

res <- CCA(X, Y, type = 2, test = "Bartlett", sign = 0.05) # performs CCA

tit <- c("Scree-plot","Correlations","Scores of the group X","Scores of the group Y")

Plot.CCA(res, titles = tit, xlabel = NA, ylabel = NA,
         color = TRUE, savptc = FALSE, width = 3236,
         height = 2000, res = 300, casc = TRUE)
```

---

Plot.Cor *Plot of correlations between variables.*

---

### Description

It performs the correlations between the variables of a database and presents it in graph form.

### Usage

```
Plot.Cor(data, title = NA, grid = TRUE, leg = TRUE, boxleg = FALSE,
          text = FALSE, arrow = TRUE, color = TRUE, namesvar = NA,
          savptc = FALSE, width = 3236, height = 2000, res = 300)
```

### Arguments

| | |
|---|---|
| data | Numeric data set. |
| title | Title for the plot, if not defined it assumes standard text. |
| grid | Puts grid on plot (default = TRUE). |
| leg | Put the legend on the plot (default = TRUE) |

| | |
|---|---|
| `boxleg` | Put frame in the legend (default = FALSE). |
| `text` | Puts correlation values in circles (default = FALSE). |
| `arrow` | Positive (up) and negative (down) correlation arrows (default = TRUE). |
| `color` | Colorful plot (default = TRUE). |
| `namesvar` | Vector with the variable names, if omitted it assumes the names in 'date'. |
| `savptc` | Saves graphics images to files (default = FALSE). |
| `width` | Graphics images width when savptc = TRUE (defaul = 3236). |
| `height` | Graphics images height when savptc = TRUE (default = 2000). |
| `res` | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |

## Value

Plot with the correlations between the variables in 'date'.

## Author(s)

Paulo Cesar Ossani

## Examples

```
data(iris) # data set

Plot.Cor(data = iris[,1:4], title = NA, grid = TRUE, leg = TRUE, boxleg = FALSE,
         text = FALSE, arrow = TRUE, color = TRUE, namesvar = NA, savptc = FALSE,
         width = 3236, height = 2000, res = 300)

Plot.Cor(data = iris[,1:4], title = NA, grid = TRUE, leg = TRUE, boxleg = FALSE,
         text = TRUE, arrow = TRUE, color = TRUE, namesvar = c("A1","B2","C3","D4"),
         savptc = FALSE, width = 3236, height = 2000, res = 300)
```

---

# Plot.FA *Graphs of the Factorial Analysis (FA).*

---

## Description

Graphs of the Factorial Analysis (FA).

## Usage

```
Plot.FA(FA, titles = NA, xlabel = NA, ylabel = NA, size = 1.1,
        grid = TRUE, color = TRUE, linlab = NA, axes = TRUE, class = NA,
        classcolor = NA, posleg = 2, boxleg = TRUE, savptc = FALSE,
        width = 3236, height = 2000, res = 300, casc = TRUE)
```

## Arguments

| | |
|---|---|
| `FA` | Data of the FA function. |
| `titles` | Titles of the graphics, if not set, assumes the default text. |
| `xlabel` | Names the X axis, if not set, assumes the default text. |
| `ylabel` | Names the Y axis, if not set, assumes the default text. |
| `size` | Size of the points in the graphs. |
| `grid` | Put grid on graphs (default = TRUE). |
| `color` | Colored graphics (default = TRUE). |
| `linlab` | Vector with the labels for the observations. |
| `axes` | Plots the X and Y axes (default = TRUE). |
| `class` | Vector with names of data classes. |
| `classcolor` | Vector with the colors of the classes. |
| `posleg` | 0 with no caption,<br>1 for caption in the left upper corner,<br>2 for caption in the right upper corner (default),<br>3 for caption in the right lower corner,<br>4 for caption in the left lower corner. |
| `boxleg` | Puts the frame in the caption (default = TRUE). |
| `savptc` | Saves graphics images to files (default = FALSE). |
| `width` | Graphics images width when savptc = TRUE (defaul = 3236). |
| `height` | Graphics images height when savptc = TRUE (default = 2000). |
| `res` | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |
| `casc` | Cascade effect in the presentation of the graphics (default = TRUE). |

## Value

Returns several graphs.

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## See Also

FA

## Examples

```
data(iris) # conjunto de dados

data <- iris[,1:4]

cls <- iris[,5] # classe dos dados

res <- FA(data, method = "PC", type = 2, nfactor = 3)

tit <- c("Scree-plot","Scores of the Observations","Factorial Loadings","Biplot")

cls <- as.character(iris[,5])

Plot.FA(FA = res, titles = tit, xlabel = NA, ylabel = NA,
        color = TRUE, linlab = NA, savptc = FALSE, size = 1.1,
        posleg = 1, boxleg = FALSE, class = cls, axes = TRUE,
        classcolor = c("blue3","red","goldenrod3"),
        width = 3236, height = 2000, res = 300, casc = FALSE)
```

---

`Plot.MFA` *Graphics of the Multiple Factor Analysis (MFA).*

---

## Description

Graphics of the Multiple Factor Analysis (MFA).

## Usage

```
Plot.MFA(MFA, titles = NA, xlabel = NA, ylabel = NA,
         posleg = 2, boxleg = TRUE, size = 1.1, grid = TRUE,
         color = TRUE, groupscolor = NA, namarr = FALSE,
         linlab = NA, savptc = FALSE, width = 3236,
         height = 2000, res = 300, casc = TRUE)
```

## Arguments

| | |
|---|---|
| `MFA` | Data of the MFA function. |
| `titles` | Titles of the graphics, if not set, assumes the default text. |
| `xlabel` | Names the X axis, if not set, assumes the default text. |
| `ylabel` | Names the Y axis, if not set, assumes the default text. |
| `posleg` | 1 for caption in the left upper corner,<br>2 for caption in the right upper corner (default),<br>3 for caption in the right lower corner,<br>4 for caption in the left lower corner. |
| `boxleg` | Puts frame in legend (default = TRUE). |

| | |
|---|---|
| size | Size of the points in the graphs. |
| grid | Put grid on graphs (default = TRUE). |
| color | Colored graphics (default = TRUE). |
| groupscolor | Vector with the colors of the groups. |
| namarr | Puts the points names in the cloud around the centroid in the graph corresponding to the global analysis of the Individuals and Variables (default = FALSE). |
| linlab | Vector with the labels for the observations, if not set, assumes the default text. |
| savptc | Saves graphics images to files (default = FALSE). |
| width | Graphics images width when savptc = TRUE (defaul = 3236). |
| height | Graphics images height when savptc = TRUE (default = 2000). |
| res | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |
| casc | Cascade effect in the presentation of the graphics (default = TRUE). |

## Value

Returns several graphs.

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## See Also

MFA

## Examples

```
data(DataMix) # set of mixed data

data <- DataMix[,2:ncol(DataMix)]

rownames(data) <- DataMix[1:nrow(DataMix),1]

group.names = c("Grade Cafes/Work", "Formation/Dedication", "Coffees")

mf <- MFA(data, c(2,2,2), typegroups = c("n","c","f"), group.names) # performs MFA

tit <- c("Scree-Plot","Observations","Observations/Variables",
         "Correlation Circle","Inertia of the Variable Groups")

Plot.MFA(MFA = mf, titles = tit, xlabel = NA, ylabel = NA,
         posleg = 2, boxleg = FALSE, color = TRUE,
         groupscolor = c("blue3","red","goldenrod3"),
         namarr = FALSE, linlab = NA, savptc = FALSE,
         width = 3236, height = 2000, res = 300,
```

```
         casc = TRUE) # plotting several graphs on the screen

Plot.MFA(MFA = mf, titles = tit, xlabel = NA, ylabel = NA,
         posleg = 2, boxleg = FALSE, color = TRUE,
         namarr = FALSE, linlab = rep("A?",10),
         savptc = FALSE, width = 3236, height = 2000,
         res = 300, casc = TRUE) # plotting several graphs on the screen
```

---

Plot.PCA *Graphs of the Principal Components Analysis (PCA).*

---

## Description

Graphs of the Principal Components Analysis (PCA).

## Usage

```
Plot.PCA(PC, titles = NA, xlabel = NA, ylabel = NA, size = 1.1,
         grid = TRUE, color = TRUE, linlab = NA, axes = TRUE, class = NA,
         classcolor = NA, posleg = 2, boxleg = TRUE, savptc = FALSE,
         width = 3236, height = 2000, res = 300, casc = TRUE)
```

## Arguments

| | |
|---|---|
| PC | Data of the PCA function. |
| titles | Titles of the graphics, if not set, assumes the default text. |
| xlabel | Names the X axis, if not set, assumes the default text. |
| ylabel | Names the Y axis, if not set, assumes the default text. |
| size | Size of the points in the graphs. |
| grid | Put grid on graphs (default = TRUE). |
| color | Colored graphics (default = TRUE). |
| linlab | Vector with the labels for the observations. |
| axes | Plots the X and Y axes (default = TRUE). |
| class | Vector with names of data classes. |
| classcolor | Vector with the colors of the classes. |
| posleg | 0 with no caption,<br>1 for caption in the left upper corner,<br>2 for caption in the right upper corner (default),<br>3 for caption in the right lower corner,<br>4 for caption in the left lower corner. |
| boxleg | Puts the frame in the caption (default = TRUE). |
| savptc | Saves graphics images to files (default = FALSE). |
| width | Graphics images width when savptc = TRUE (defaul = 3236). |

| | |
|---|---|
| `height` | Graphics images height when savptc = TRUE (default = 2000). |
| `res` | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |
| `casc` | Cascade effect in the presentation of the graphics (default = TRUE). |

## Value

Returns several graphs.

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## See Also

PCA

## Examples

```
data(iris) # data set

data <- iris[,1:4]

cls <- iris[,5] # data class

pc <- PCA(data, 2)

tit <- c("Scree-plot","Observations","Correlations")

cls <- as.character(iris[,5])

Plot.PCA(PC = pc, titles = tit, xlabel = NA, ylabel = NA,
         color = TRUE, linlab = NA, savptc = FALSE, size = 1.1,
         posleg = 2, boxleg = FALSE, class = cls, axes = TRUE,
         classcolor = c("blue3","red","goldenrod3"),
         width = 3236, height = 2000, res = 300, casc = FALSE)
```

---

`Plot.PP` *Graphics of the Projection Pursuit (PP).*

---

## Description

Graphics of the Projection Pursuit (PP).

## Usage

```
Plot.PP(PP, titles = NA, xlabel = NA, ylabel = NA, posleg = 2, boxleg = TRUE,
        size = 1.1, grid = TRUE, color = TRUE, classcolor = NA, linlab = NA,
       axesvar = TRUE, axes = TRUE, savptc = FALSE, width = 3236, height = 2000,
        res = 300, casc = TRUE)
```

## Arguments

| | |
|---|---|
| `PP` | Data of the PP_Optimizer function. |
| `titles` | Titles of the graphics, if not set, assumes the default text. |
| `xlabel` | Names the X axis, if not set, assumes the default text. |
| `ylabel` | Names the Y axis, if not set, assumes the default text. |
| `posleg` | 0 with no caption,<br>1 for caption in the left upper corner,<br>2 for caption in the right upper corner (default),<br>3 for caption in the right lower corner,<br>4 for caption in the left lower corner. |
| `boxleg` | Puts the frame in the caption (default = TRUE). |
| `size` | Size of the points in the graphs. |
| `grid` | Put grid on graphs (default = TRUE). |
| `color` | Colored graphics (default = TRUE). |
| `classcolor` | Vector with the colors of the classes. |
| `linlab` | Vector with the labels for the observations. |
| `axesvar` | Puts axes of rotation of the variables, only when dimproj > 1 (default = TRUE). |
| `axes` | Plots the X and Y axes (default = TRUE). |
| `savptc` | Saves graphics images to files (default = FALSE). |
| `width` | Graphics images width when savptc = TRUE (defaul = 3236). |
| `height` | Graphics images height when savptc = TRUE (default = 2000). |
| `res` | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |
| `casc` | Cascade effect in the presentation of the graphics (default = TRUE). |

## Value

Graph of the evolution of the indices, and graphs whose data were reduced in two dimensions.

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## See Also

`PP_Optimizer` and `PP_Index`

## Examples

```
data(iris) # dataset

# Example 1 - Without the classes in the data
data <- iris[,1:4]

findex <- "kurtosismax" # index function

dim <- 1 # dimension of data projection

sphere <- TRUE # spherical data

res <- PP_Optimizer(data = data, class = NA, findex = findex,
                    optmethod = "GTSA", dimproj = dim, sphere = sphere,
                    weight = TRUE, lambda = 0.1, r = 1, cooling = 0.9,
                    eps = 1e-3, maxiter = 500, half = 30)

Plot.PP(res, titles = NA, posleg = 1, boxleg = FALSE, color = TRUE,
        linlab = NA, axesvar = TRUE, axes = TRUE, savptc = FALSE,
        width = 3236, height = 2000, res = 300, casc = FALSE)


# Example 2 - With the classes in the data
class <- iris[,5] # data class

res <- PP_Optimizer(data = data, class = class, findex = findex,
                    optmethod = "GTSA", dimproj = dim, sphere = sphere,
                    weight = TRUE, lambda = 0.1, r = 1, cooling = 0.9,
                    eps = 1e-3, maxiter = 500, half = 30)

tit <- c(NA,"Graph example") # titles for the graphics

Plot.PP(res, titles = tit, posleg = 1, boxleg = FALSE, color = TRUE,
        classcolor = c("blue3","red","goldenrod3"), linlab = NA,
        axesvar = TRUE, axes = TRUE, savptc = FALSE, width = 3236,
        height = 2000, res = 300, casc = FALSE)


# Example 3 - Without the classes in the data, but informing
#             the classes in the plot function
res <- PP_Optimizer(data = data, class = NA, findex = "Moment",
                    optmethod = "GTSA", dimproj = 2, sphere = sphere,
                    weight = TRUE, lambda = 0.1, r = 1, cooling = 0.9,
                    eps = 1e-3, maxiter = 500, half = 30)

lin <- c(rep("a",50),rep("b",50),rep("c",50)) # data class

Plot.PP(res, titles = NA, posleg = 1, boxleg = FALSE, color = TRUE,
        linlab = lin, axesvar = TRUE, axes = TRUE, savptc = FALSE,
        width = 3236, height = 2000, res = 300, casc = FALSE)
```

```
# Example 4 - With the classes in the data, but not informed in plot function
class <- iris[,5] # data class

dim <- 2 # dimension of data projection

findex <- "lda" # index function

res <- PP_Optimizer(data = data, class = class, findex = findex,
                    optmethod = "GTSA", dimproj = dim, sphere = sphere,
                    weight = TRUE, lambda = 0.1, r = 1, cooling = 0.9,
                    eps = 1e-3, maxiter = 500, half = 30)

tit <- c("",NA) # titles for the graphics

Plot.PP(res, titles = tit, posleg = 1, boxleg = FALSE, color = TRUE,
        linlab = NA, axesvar = TRUE, axes = TRUE, savptc = FALSE,
        width = 3236, height = 2000, res = 300, casc = FALSE)
```

---

`Plot.Regr` *Graphs of the linear regression results.*

---

## Description

Graphs of the linear regression results.

## Usage

```
Plot.Regr(Reg, typegraf = "Scatterplot", title = NA, xlabel = NA,
          ylabel = NA, namevary = NA, namevarx = NA, size = 1.1,
          grid = TRUE, color = TRUE, intconf = TRUE, intprev = TRUE,
          savptc = FALSE, width = 3236, height = 2000, res = 300,
          casc = TRUE)
```

## Arguments

| | |
|---|---|
| `Reg` | Regression function data. |
| `typegraf` | Type of graphic:<br>"Scatterplot" - Scatterplot 2 to 2,<br>"Regression" - Graph of the linear regression,<br>"QQPlot" - Graph of the normal probability of the residues,<br>"Histogram" - Histogram of the residues,<br>"Fits" - Graph of the adjusted values versus residuals,<br>"Order" - Graph of the order of the observations versus the residuals. |
| `title` | Titles of the graphics, if not set, assumes the default text. |
| `xlabel` | Names the X axis, if not set, assumes the default text. |
| `ylabel` | Names the Y axis, if not set, assumes the default text. |

| | |
|---|---|
| namevary | Variable name Y, if not set, assumes the default text. |
| namevarx | Name of the variable X, or variables X, if not set, assumes the default text. |
| size | Size of the points in the graphs. |
| grid | Put grid on graphs (default = TRUE). |
| color | Colored graphics (default = TRUE). |
| intconf | Case typegraf = "Regression". Graphics with confidence interval (default = TRUE). |
| intprev | Case typegraf = "Regression". Graphics with predictive interval (default = TRUE). |
| savptc | Saves graphics images to files (default = FALSE). |
| width | Graphics images width when savptc = TRUE (defaul = 3236). |
| height | Graphics images height when savptc = TRUE (default = 2000). |
| res | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |
| casc | Cascade effect in the presentation of the graphics (default = TRUE). |

## Value

Returns several graphs.

## Author(s)

Paulo Cesar Ossani

## See Also

Regr

## Examples

```
data(DataMix)

Y <- DataMix[,2]

X <- DataMix[,7]

name.y <- "Medium grade"

name.x <- "Commercial coffees"

res <- Regr(Y, X, namevarx = name.x , intercept = TRUE, sigf = 0.05)

tit <- c("Scatterplot")
Plot.Regr(res, typegraf = "Scatterplot", title = tit,
          namevary = name.y, namevarx = name.x, color = TRUE,
          savptc = FALSE, width = 3236, height = 2000, res = 300)

tit <- c("Scatterplot with the adjusted line")
Plot.Regr(res, typegraf = "Regression", title = tit,
```

```
          xlabel = name.x, ylabel = name.y, color = TRUE,
          intconf = TRUE, intprev = TRUE, savptc = FALSE,
          width = 3236, height = 2000, res = 300)

dev.new() # necessary to not overlap the following graphs to the previous graph

par(mfrow = c(2,2))

Plot.Regr(res, typegraf = "QQPlot", casc = FALSE)
Plot.Regr(res, typegraf = "Histogram", casc = FALSE)
Plot.Regr(res, typegraf = "Fits", casc = FALSE)
Plot.Regr(res, typegraf = "Order", casc = FALSE)
```

---

PP_Index *Function to find the Projection Pursuit indexes (PP).*

---

## Description

Function used to find Projection Pursuit indexes (PP).

## Usage

```
PP_Index(data, class = NA, vector.proj = NA,
         findex = "HOLES", dimproj = 2, weight = TRUE,
         lambda = 0.1, r = 1, ck = NA)
```

## Arguments

| | |
|---|---|
| `data` | Numeric dataset without class information. |
| `class` | Vector with names of data classes. |
| `vector.proj` | Vector projection. |
| `findex` | Projection index function to be used:<br>"lda" - LDA index,<br>"pda" - PDA index,<br>"lr" - Lr index,<br>"holes" - Holes index (default),<br>"cm" - Central Mass index,<br>"pca" - PCA index,<br>"friedmantukey" - Friedman Tukey index,<br>"entropy" - Entropy index,<br>"legendre" - Legendre index,<br>"laguerrefourier" - Laguerre Fourier index,<br>"hermite" - Hermite index,<br>"naturalhermite" - Natural Hermite index,<br>"kurtosismax" - Maximum kurtosis index,<br>"kurtosismin" - Minimum kurtosis index,<br>"moment" - Moment index, |

| | |
|---|---|
| | "mf" - MF index,<br>"chi" - Chi-square index. |
| `dimproj` | Dimension of data projection (default = 2). |
| `weight` | Used in index LDA, PDA and Lr to weight the calculations for the number of elements in each class (default = TRUE). |
| `lambda` | Used in the PDA index (default = 0.1). |
| `r` | Used in the Lr index (default = 1). |
| `ck` | Internal use of the CHI index function. |

## Value

| | |
|---|---|
| `num.class` | Number of classes. |
| `class.names` | Class names. |
| `findex` | Projection index function used. |
| `vector.proj` | Projection vectors found. |
| `index` | Projection index found in the process. |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## References


Ossani, P. C.; Figueira, M. R.; Cirillo, M. A. Proposition of a new index for projection pursuit in the multiple factor analysis. *Computational and Mathematical Methods*, v. 1, p. 1-18, 2020.

Cook, D.; Buja, A.; Cabrera, J. Projection pursuit indexes based on orthonormal function expansions. *Journal of Computational and Graphical Statistics*, 2(3):225-250, 1993.

Cook, D.; Buja, A.; Cabrera, J.; Hurley, C. Grand tour and projection pursuit, *Journal of Computational and Graphical Statistics*, 4(3), 155-172, 1995.

Cook, D.; Swayne, D. F. Interactive and Dynamic Graphics for data Analysis: With R and GGobi. Springer. 2007.

Espezua, S.; Villanueva, E.; Maciel, C. D.; Carvalho, A. A projection pursuit framework for supervised dimension reduction of high dimensional small sample datasets. *Neurocomputing*, 149, 767-776, 2015.

Friedman, J. H., Tukey, J. W. A projection pursuit algorithm for exploratory data analysis. *IEEE Transaction on Computers*, 23(9):881-890, 1974.

Hastie, T., Buja, A., Tibshirani, R. Penalized discriminant analysis. *The Annals of Statistics*. 23(1), 73-102 . 1995.

Huber, P. J. Projection pursuit. *Annals of Statistics*, 13(2):435-475, 1985.

Jones, M. C.; Sibson, R. What is projection pursuit, (with discussion), *Journal of the Royal Statistical Society*, Series A 150, 1-36, 1987.

Lee, E. K.; Cook, D. A projection pursuit index for large p small n data. *Statistics and Computing*, 20(3):381-392, 2010.

Lee, E.; Cook, D.; Klinke, S.; Lumley, T. Projection pursuit for exploratory supervised classification. *Journal of Computational and Graphical Statistics*, 14(4):831-846, 2005.

Martinez, W. L., Martinez, A. R.; *Computational Statistics Handbook with MATLAB*, 2th. ed. New York: Chapman & Hall/CRC, 2007. 794 p.

Martinez, W. L.; Martinez, A. R.; Solka, J. *Exploratory data Analysis with MATLAB*, 2th. ed. New York: Chapman & Hall/CRC, 2010. 499 p.

Pena, D.; Prieto, F. Cluster identification using projections. *Journal of the American Statistical Association*, 96(456):1433-1445, 2001.

Posse, C. Projection pursuit exploratory data analysis, *Computational Statistics and data Analysis*, 29:669-687, 1995a.

Posse, C. Tools for two-dimensional exploratory projection pursuit, *Journal of Computational and Graphical Statistics*, 4:83-100, 1995b.

## See Also

PP_Optimizer and Plot.PP

## Examples

```
data(iris) # data set

data <- iris[,1:4]

# Example 1 - Without the classes in the data
ind <- PP_Index(data = data, class = NA, vector.proj = NA,
                findex = "moment", dimproj = 2, weight = TRUE,
                lambda = 0.1, r = 1)

print("Number of classes:"); ind$num.class
print("class Names:"); ind$class.names
print("Projection index function:"); ind$findex
print("Projection vectors:"); ind$vector.proj
print("Projection index:"); ind$index


# Example 2 - With the classes in the data
class <- iris[,5] # data class

findex <- "pda" # index function

sphere <- TRUE # spherical data

res <- PP_Optimizer(data = data, class = class, findex = findex,
                    optmethod = "SA", dimproj = 2, sphere = sphere,
                    weight = TRUE, lambda = 0.1, r = 1, cooling = 0.9,
                    eps = 1e-3, maxiter = 1000, half = 30)

# Comparing the result obtained
if (match(toupper(findex),c("LDA", "PDA", "LR"), nomatch = 0) > 0) {
  if (sphere) {
```

```
    data <- apply(predict(prcomp(data)), 2, scale) # spherical data
  }
} else data <- as.matrix(res$proj.data[,1:Dim])

ind <- PP_Index(data = data, class = class, vector.proj = res$vector.opt,
                findex = findex, dimproj = 2, weight = TRUE, lambda = 0.1,
                r = 1)

print("Number of classes:"); ind$num.class
print("class Names:"); ind$class.names
print("Projection index function:"); ind$findex
print("Projection vectors:"); ind$vector.proj
print("Projection index:"); ind$index
print("Optimized Projection index:"); res$index[length(res$index)]
```

---

`PP_Optimizer` *Optimization function of the Projection Pursuit index (PP).*

---

## Description

Optimization function of the Projection Pursuit index (PP).

## Usage

```
PP_Optimizer(data, class = NA, findex = "HOLES",
             dimproj = 2, sphere = TRUE, optmethod = "GTSA",
             weight = TRUE, lambda = 0.1, r = 1, cooling = 0.9,
             eps = 1e-3, maxiter = 3000, half = 30)
```

## Arguments

| | |
|---|---|
| `data` | Numeric dataset without class information. |
| `class` | Vector with names of data classes. |
| `findex` | Projection index function to be used:<br>"lda" - LDA index,<br>"pda" - PDA index,<br>"lr" - Lr index,<br>"holes" - Holes index (default),<br>"cm" - Central Mass index,<br>"pca" - PCA index,<br>"friedmantukey" - Friedman Tukey index,<br>"entropy" - Entropy index,<br>"legendre" - Legendre index,<br>"laguerrefourier" - Laguerre Fourier index,<br>"hermite" - Hermite index,<br>"naturalhermite" - Natural Hermite index,<br>"kurtosismax" - Maximum kurtosis index,<br>"kurtosismin" - Minimum kurtosis index, |

| | |
|---|---|
| | "moment" - Moment index,<br>"mf" - MF index,<br>"chi" - Chi-square index. |
| `dimproj` | Dimension of the data projection (default = 2). |
| `sphere` | Spherical data (default = TRUE). |
| `optmethod` | Optimization method GTSA - Grand Tour Simulated Annealing or SA - Simulated Annealing (default = "GTSA"). |
| `weight` | Used in index LDA, PDA and Lr to weight the calculations for the number of elements in each class (default = TRUE). |
| `lambda` | Used in the PDA index (default = 0.1). |
| `r` | Used in the Lr index (default = 1). |
| `cooling` | Cooling rate (default = 0.9). |
| `eps` | Approximation accuracy for cooling (default = 1e-3). |
| `maxiter` | Maximum number of iterations of the algorithm (default = 3000). |
| `half` | Number of steps without incrementing the index, then decreasing the cooling value (default = 30). |

## Value

| | |
|---|---|
| `num.class` | Number of classes. |
| `class.names` | Class names. |
| `proj.data` | Projected data. |
| `vector.opt` | Projection vectors found. |
| `index` | Vector with the projection indices found in the process, converging to the maximum, or the minimum. |
| `findex` | Projection index function used. |

## Author(s)

Paulo Cesar Ossani

Marcelo Angelo Cirillo

## References

Cook, D.; Lee, E. K.; Buja, A.; Wickmam, H. Grand tours, projection pursuit guided tours and manual controls. In Chen, Chunhouh, Hardle, Wolfgang, Unwin, e Antony (Eds.), *Handbook of data Visualization*, Springer Handbooks of Computational Statistics, chapter III.2, p. 295-314. Springer, 2008.

Lee, E.; Cook, D.; Klinke, S.; Lumley, T. Projection pursuit for exploratory supervised classification. *Journal of Computational and Graphical Statistics*, 14(4):831-846, 2005.

## See Also

`Plot.PP` and `PP_Index`

### Examples

```
data(iris) # data set

# Example 1 - Without the classes in the data
data <- iris[,1:4]

class <- NA # data class

findex <- "kurtosismax" # index function

dim <- 1 # dimension of data projection

sphere <- TRUE # spherical data

res <- PP_Optimizer(data = data, class = class, findex = findex,
                    optmethod = "GTSA", dimproj = dim, sphere = sphere,
                    weight = TRUE, lambda = 0.1, r = 1, cooling = 0.9,
                    eps = 1e-3, maxiter = 1000, half = 30)

print("Number of classes:"); res$num.class
print("class Names:"); res$class.names
print("Projection index function:"); res$findex
print("Projected data:"); res$proj.data
print("Projection vectors:"); res$vector.opt
print("Projection index:"); res$index


# Example 2 - With the classes in the data
class <- iris[,5] # classe dos dados

res <- PP_Optimizer(data = data, class = class, findex = findex,
                    optmethod = "GTSA", dimproj = dim, sphere = sphere,
                    weight = TRUE, lambda = 0.1, r = 1, cooling = 0.9,
                    eps = 1e-3, maxiter = 1000, half = 30)

print("Number of classes:"); res$num.class
print("class Names:"); res$class.names
print("Projection index function:"); res$findex
print("Projected data:"); res$proj.data
print("Projection vectors:"); res$vector.opt
print("Projection index:"); res$index
```

---

## Regr *Linear regression.*

---

### Description

Performs linear regression on a data set.

## Usage

```
Regr(Y, X, namevarx = NA, intercept = TRUE, sigf = 0.05)
```

## Arguments

| | |
|---|---|
| Y | Variable response. |
| X | Regression variables. |
| namevarx | Name of the variable, or variables X, if not set, assumes the default text. |
| intercept | Consider the intercept in the regression (default = TRUE). |
| sigf | Level of significance of residue tests(default = 5%). |

## Value

| | |
|---|---|
| Betas | Regression coefficients. |
| CovBetas | Covariance matrix of the regression coefficients. |
| ICc | Confidence interval of the regression coefficients. |
| hip.test | Hypothesis test of the regression coefficients. |
| ANOVA | Regression analysis of the variance. |
| R | Determination coefficient. |
| Rc | Corrected coefficient of determination. |
| Ra | Adjusted coefficient of determination. |
| QME | Variance of the residues. |
| ICQME | Confidence interval of the residue variance. |
| prev | Prediction of the regression fit. |
| IPp | Predictions interval |
| ICp | Interval of prediction confidence |
| error | Residuals of the regression fit. |
| error.test | It returns to 5% of significance the test of independence, normality and homogeneity of the variance of the residues. |

## Author(s)

Paulo Cesar Ossani

## References

Charnet, R.; at al.. *Analise de modelos de regressao lienar,* 2a ed. Campinas: Editora da Unicamp, 2008. 357 p.

Rencher, A. C.; Schaalje, G. B. *Linear models in statisctic.* 2th. ed. New Jersey: John & Sons, 2008. 672 p.

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

## See Also

Plot.Regr

## Examples

```
data(DataMix)

Y <- DataMix[,2]

X <- DataMix[,6:7]

name.x <- c("Special Coffees", "Commercial Coffees")

res <- Regr(Y, X, namevarx = name.x , intercept = TRUE, sigf = 0.05)

print("Regression Coefficients:"); round(res$Betas,4)
print("Analysis of Variance:"); res$ANOVA
print("Hypothesis test of regression coefficients:"); round(res$hip.test,4)
print("Determination coefficient:"); round(res$R,4)
print("Corrected coefficient of determination:"); round(res$Rc,4)
print("Adjusted coefficient of determination:"); round(res$Ra,4)
print("Tests of the residues"); res$error.test
```

---

`Scatter` *Scatter plot.*

---

## Description

Performs the scatter plot.

## Usage

```
Scatter(data, ellipse = TRUE, ellipse.level = 0.95, rectangle = FALSE,
        title = NA, xlabel = NA, ylabel = NA,  posleg = 2, boxleg = TRUE,
        axes = TRUE, size = 1.1, grid = TRUE, color = TRUE, linlab = NA,
        class = NA, classcolor = NA, savptc = FALSE, width = 3236,
        height = 2000, res = 300)
```

## Arguments

| | |
|---|---|
| `data` | Data with x and y coordinates. |
| `ellipse` | Place an ellipse around the classes (default = TRUE). |
| `ellipse.level` | Significance level of the ellipse (defaul = 0.95). |
| `rectangle` | Place rectangle to differentiate classes (default = FALSE). |
| `title` | Titles of the graphics, if not set, assumes the default text. |
| `xlabel` | Names the X axis, if not set, assumes the default text. |

| | |
|---|---|
| ylabel | Names the Y axis, if not set, assumes the default text. |
| posleg | 0 with no caption,<br>1 for caption in the left upper corner,<br>2 for caption in the right upper corner (default),<br>3 for caption in the right lower corner,<br>4 for caption in the left lower corner. |
| boxleg | Puts the frame in the caption (default = TRUE). |
| axes | Plots the X and Y axes (default = TRUE). |
| size | Size of the points in the graphs. |
| grid | Put grid on graphs (default = TRUE). |
| color | Colored graphics (default = TRUE). |
| linlab | Vector with the labels for the observations. |
| class | Vector with names of data classes. |
| classcolor | Vector with the colors of the classes. |
| savptc | Saves graphics images to files (default = FALSE). |
| width | Graphics images width when savptc = TRUE (defaul = 3236). |
| height | Graphics images height when savptc = TRUE (default = 2000). |
| res | Nominal resolution in ppi of the graphics images when savptc = TRUE (default = 300). |

## Value

Scatter plot.

## Author(s)

Paulo Cesar Ossani

## References

Rencher, A. C. *Methods of multivariate analysis.* 2th. ed. New York: J.Wiley, 2002. 708 p.

Anton, H.; Rorres, C. *Elementary linear algebra: applications version.* 10th ed. New Jersey: John Wiley & Sons, 2010. 768 p.

## Examples

```
data(iris) # data set

data <- iris[,3:4]

cls <- iris[,5] # data class

Scatter(data, ellipse = TRUE, ellipse.level = 0.95, rectangle = FALSE,
        title = NA, xlabel = NA, ylabel = NA,  posleg = 1, boxleg = FALSE,
        axes = FALSE, size = 1.1, grid = TRUE, color = TRUE, linlab = NA,
        class = cls, classcolor = c("goldenrod3","blue","red"),
```

```
        savptc = FALSE, width = 3236, height = 2000, res = 300)

Scatter(data, ellipse = FALSE, ellipse.level = 0.95, rectangle = TRUE,
        title = NA, xlabel = NA, ylabel = NA,  posleg = 1, boxleg = TRUE,
        axes = FALSE, size = 1.1, grid = TRUE, color = TRUE, linlab = NA,
        class = cls, classcolor = c("goldenrod3","blue","red"),
        savptc = FALSE, width = 3236, height = 2000, res = 300)
```

# Index