Package 'DescriptiveStats.OBeu'

November 10, 2016

Type Package

Title Descriptive Statistics OpenBudgets.eu

Version 1.1.5

Date 2016-09-18

Description Descriptive Statistics and other related analysis for OBeu datasets.

Author Kleanthis Koupidis

Contributor Aikaterini Chatzopoulou <mis16010@uom.edu.gr>

Maintainer Kleanthis Koupidis <koupidis.okfgr@gmail.com>

URL https://github.com/okgreece/DescriptiveStats.OBeu

BugReports https://github.com/okgreece/DescriptiveStats.OBeu/issues

License GPL-2 | file LICENSE

LazyData true

Imports grDevices,jsonlite,reshape,stats

RoxygenNote 5.0.1

R topics documented:

ds.analysis 2
ds.boxplot 3
ds.correlation 3
ds.frequency 4
ds.glm 5
ds.kurtosis 6
ds.skewness 6
ds.statistics 7
nums 8
sample_df_rudolf 8
sample_json_link_openspending 9
sample_json_link_rudolf 9
sample_json_openspending 10

Index **11**

`ds.analysis` *Read and Calculate the Basic Information for Basic Descriptive Tasks from Open Spending API*

Description

Extract and analyze the input data provided from Open Spending API, using the ds.analysis function.

Extract and analyze the input data provided from Open Spending API, using the ds.analysis function.

Usage

```
ds.analysis(data, box.out=1.5, corr.method= "pearson", fr.select=NULL)

open_spending.ds(json_data, box.outl=1.5, cor.method= "pearson", select=NULL)
```

Arguments

`data`	The input data
`box.out`	...
`corr.method`	The correlation coefficient method to compute: "pearson" (default), "kendall" or "spearman".
`fr.select`	One or more nominal variables to calculate their corresponding frequencies.
`json_data`	The json string, URL or file from Open Spending API
`box.outl`	...
`cor.method`	The correlation coefficient method to compute: "pearson" (default), "kendall" or "spearman".
`select`	One or more nominal variables to calculate their corresponding frequencies.

Details

This function is used to read data in json format from Open Spending API, in order to implement some basic descriptive tasks through `ds.analysis` function.

This function is used to read data in json format from Open Spending API, in order to implement some basic descriptive tasks through `ds.analysis` function.

Value

A json string with the resulted parameters of the `ds.analysis` function.

A json string with the resulted parameters of the `ds.analysis` function.

Author(s)

Kleanthis Koupidis

Kleanthis Koupidis

See Also

open_spending.ds

ds.analysis

`ds.boxplot` *Boxplot*

Description

Boxplot

Usage

```
ds.boxplot(data, out.level=1.5)
```

Arguments

`data`	The input matrix or data frame
`out.level`	...

Details

boxplot

Value

Description of the returns

Author(s)

Aikaterini Chatzopoulou, Kleanthis Koupidis

See Also

ds.analysis, open_spending.ds

`ds.correlation` *Correlation Coefficient*

Description

This functions calculates the correlation coefficient of the input vectors or datasets. By default, the correlation coefficient of pearson is computed.

Usage

```
ds.correlation(x, y=NULL, cor.method="pearson")
```

Arguments

x	A vector, matrix or data frame
y	A vector, matrix or data frame
cor.method	The correlation coefficient method to compute: "pearson" (default), "kendall" or "spearman".

Details

This function returns a symmetric matrix in json format with the correlation coefficients of the input data. The correlation coefficient of pearson is computed, by default. Other options are "kendall" or "spearman".

Author(s)

Aikaterini Chatzopoulou, Kleanthis Koupidis

See Also

ds.analysis, open_spending.ds

---

ds.frequency *Frequencies of nominal variables*

Description

This functions calculates frequencies of factors/characters of the input dataset.

Usage

```
ds.frequency(data,select=NULL)
```

Arguments

data	A vector, matrix or data frame which includes at least one factor/character.
select	One or more nominal variables to calculate their corresponding frequencies.

Details

This function returns a json output with the frequencies of factors/characters of the input dataset.

Author(s)

Kleanthis Koupidis

See Also

ds.analysis, open_spending.ds

ds.glm *Generalized Linear Models*

Description

ds.glm is used to fit generalized linear models through glm from stats package and return some results in json format.

Usage

```
ds.glm(x, dependent=NULL, independent=NULL, distr.family = "gaussian")
```

Arguments

x	The input matrix or data frame
dependent	The dependent variables of the model
independent	The independent variables of the model
distr.family	A character string naming a the error distribution and link function to be used in the model(See family for details of family functions.)

Details

Generalized linear models are used to fit data, using glm stats package, by specifying the dependent and independent variables and the description of the distribution error (default is gaussian). If user provides only the dependent variables, the rest variables are selected as independent. If user provides the independent variables, the rest variables are selected as dependent. If user provides the independent variables, the two variables with the highest correlation are selected.

Value

coefficients residuals fitted residuals.degfred qq.plot linear.predictors

Author(s)

Kleanthis Koupidis

See Also

ds.analysis, open_spending.ds

ds.kurtosis *Calculation of Kurtosis*

Description

This functions calculates kurtosis of the input dataset.

Usage

```
ds.kurtosis(x)
```

Arguments

x	A vector, matrix or data frame

Details

This function returns the kurtosis of numbers of the input data

Author(s)

Aikaterini Chatzopoulou

See Also

ds.skewness, ds.statistics, ds.analysis, open_spending.ds

ds.skewness *Calculation of Skewness*

Description

This functions calculates skewness of the input dataset.

Usage

```
ds.skewness(x)
```

Arguments

x	A vector, matrix or data frame which includes at least one number.

Details

This function returns the skewness of numbers of the input dataset.

Author(s)

Aikaterini Chatzopoulou

See Also

ds.kurtosis, ds.statistics, ds.analysis, open_spending.ds

`ds.statistics` *Calculation of the Statistic Measures*

Description

This functions calculates the basic descriptive measures of the input dataset.

Usage

```
ds.statistics(data)
```

Arguments

`data`	A numeric vector, matrix or data frame

Details

This function returns the min, max, range, mean, median, 0%,25%,50%,75%,100% quantiles variance, standartdeviation, skewness and kurtosis of the input data.

Value

A json file with the following components:

- Min: The minimum observed value of the input data
- Max: The maximum observed value of the input data
- Range: The range, defined as the difference of the maximum and the minimum value.
- Mean: The average value of the input data
- Median: The median value of the input data
- Quantiles: The 0%,25%,50%,75%,100% percentiles
- Variance: The variance of the input data
- StandartDeviation: The standard deviation of the input data
- Skewness: The Skewness of the input data
- Kurtosis: The Kurtosis of the input data

Author(s)

Katerina Chatzopoulou, Kleanthis Koupidis

See Also

`open_spending.ds`

nums *Select the numeric columns of a given dataset*

Description

Extract and return a dataframe with the columns that include only numeric values

Usage

```
nums(data)
```

Arguments

data	The json string, URL or file from Open Spending API

Value

This function returns a dataframe with the numeric columns of the input dataset.

Author(s)

Kleanthis Koupidis

sample_df_rudolf *Sample data from Rudolf*

Description

Sample data of Revised Budget phase amounts of Municipality of Athens

- The year (2016) of the recorded approved budget phase amounts
- The revised budget phase amounts of 2016
- The original amounts of this year
- The functional classification description
- The functional classification code

Format

A json format file

Source

Rudolf

sample_json_link_openspending
Sample data from Open Spending

Description

Sample data of Revised Budget phase amounts

- The year (2016) of the recorded approved budget phase amounts
- The revised budget phase amounts of 2016
- The original amounts of this year
- The functional classification description
- The functional classification code

Format

A link with the json format data

Source

OpenSpending

sample_json_link_rudolf
Sample data from Rudolf

Description

Sample data of Revised Budget phase amounts of Municipality of Athens

- The year (2016) of the recorded approved budget phase amounts
- The revised budget phase amounts of 2016
- The original amounts of this year
- The functional classification description
- The functional classification code

Format

A link to json format file

Source

Rudolf

`sample_json_openspending` *Sample data from Open Spending*

Description

Sample data of Revised Budget phase amounts

- The year (2016) of the recorded approved budget phase amounts
- The revised budget phase amounts of 2016
- The original amounts of this year
- The functional classification description
- The functional classification code

Format

A json format file

Source

OpenSpending

Index

ds.analysis, *2*, 2, *3–6*
ds.boxplot, 3
ds.correlation, 3
ds.frequency, 4
ds.glm, *5*, 5
ds.kurtosis, *6*, 6
ds.skewness, *6*, 6
ds.statistics, *6*, 7

nums, 8

open_spending.ds, *3–7*
open_spending.ds *(*ds.analysis*)*, 2

sample_df_rudolf, 8
sample_json_link_openspending, 9
sample_json_link_rudolf, 9
sample_json_openspending, 10