

Table of Contents

| | |
|---|---|
| Preparation of data for analysis | 1 |
| <i>R functions</i> | 1 |

Preparation of data for analysis

Theory **R functions** Examples Exercise 

R functions

Simple EDA (Exploratory Data Analysis):

- `boxplot` - draws the boxplot of data; with argument `outline = TRUE` (default setting) draws also the outliers;
- `hist` - histogram with distribution of values in the variable;
- `qqnorm`, `qqline` - draws QQ plot (comparing the quantiles of values in variable `x` to quantiles of normal distribution) and adds the line for comparison.
- `pairs` - draws the pairwise relationships between variables in the matrix/data.frame; the original `pairs` function can be extended for panel functions from *Numerical Ecology with R*, which gives quite interesting tool for data overview (see details [here](#)).

Data transformation

- `sqrt`
- `log`, `log10`, `log1p`

Data standardization

- `scale`
- `decostand` (library `vegan`)

From:

<https://anadat-r.davidzeleny.net/> - **Analysis of community ecology data in R**

Permanent link:

https://anadat-r.davidzeleny.net/doku.php/en:data_preparation_r

Last update: **2019/02/22 18:00**