About the transformation - NA removal: use na.omit to drop instances with missing values. Useful for initial cleanup when imputation is not desired. - Use with care as it may bias results if missingness is not completely at random.

Environment setup.

# NA and Outlier analysis  
  
# installation   
#install.packages("daltoolbox")  
  
# loading DAL  
library(daltoolbox)

Demonstration: insert an artificial NA and remove rows with NA.

# NA removal  
  
iris <- datasets::iris  
head(iris)

## Sepal.Length Sepal.Width Petal.Length Petal.Width Species  
## 1 5.1 3.5 1.4 0.2 setosa  
## 2 4.9 3.0 1.4 0.2 setosa  
## 3 4.7 3.2 1.3 0.2 setosa  
## 4 4.6 3.1 1.5 0.2 setosa  
## 5 5.0 3.6 1.4 0.2 setosa  
## 6 5.4 3.9 1.7 0.4 setosa

nrow(iris)

## [1] 150

# introducing an NA to remove  
  
iris.na <- iris  
iris.na$Sepal.Length[2] <- NA  
head(iris.na)

## Sepal.Length Sepal.Width Petal.Length Petal.Width Species  
## 1 5.1 3.5 1.4 0.2 setosa  
## 2 NA 3.0 1.4 0.2 setosa  
## 3 4.7 3.2 1.3 0.2 setosa  
## 4 4.6 3.1 1.5 0.2 setosa  
## 5 5.0 3.6 1.4 0.2 setosa  
## 6 5.4 3.9 1.7 0.4 setosa

nrow(iris.na)

## [1] 150

# removing rows with NA  
  
iris.na.omit <- na.omit(iris.na)  
head(iris.na.omit)

## Sepal.Length Sepal.Width Petal.Length Petal.Width Species  
## 1 5.1 3.5 1.4 0.2 setosa  
## 3 4.7 3.2 1.3 0.2 setosa  
## 4 4.6 3.1 1.5 0.2 setosa  
## 5 5.0 3.6 1.4 0.2 setosa  
## 6 5.4 3.9 1.7 0.4 setosa  
## 7 4.6 3.4 1.4 0.3 setosa

nrow(iris.na.omit)

## [1] 149

References - Little, R. J. A., and Rubin, D. B. (2002). Statistical Analysis with Missing Data (2nd ed.). Wiley.