Rapport package team

Descriptive statistics

2011-04-26 20:25 CET

Contents

Description

This template will return descriptive statistics of a numerical or frequency table of a categorical variable.

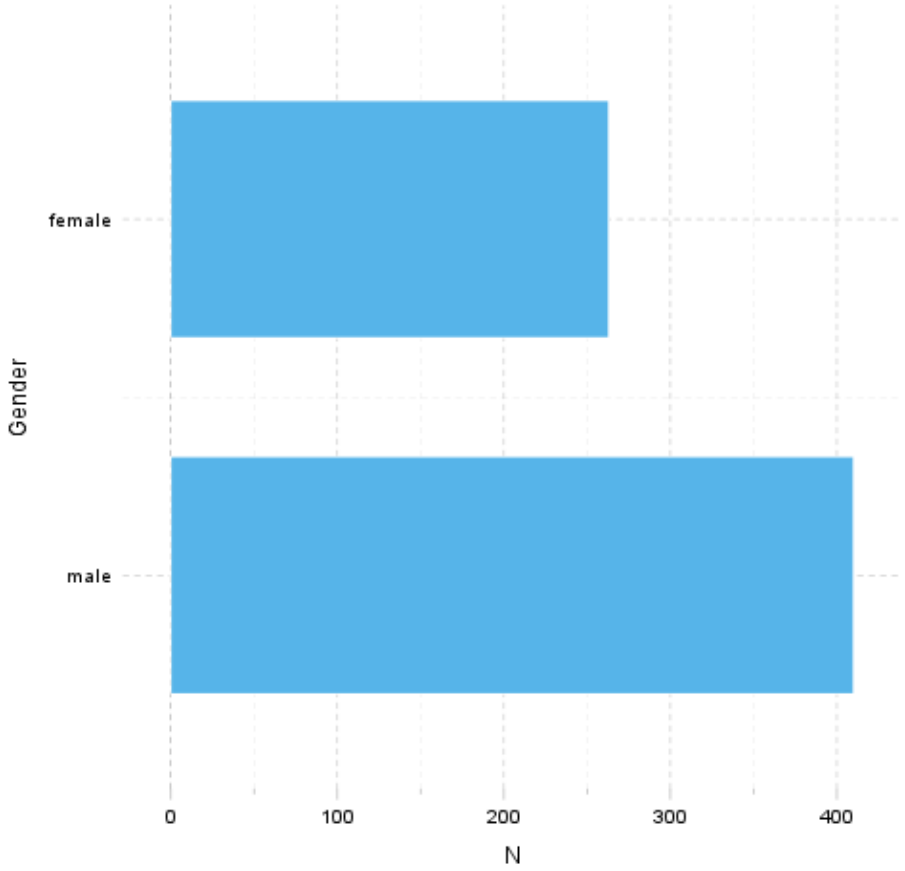
***gender* ("Gender")**

The dataset has *709* observations with *673* valid values (missing: *36*).

gender	N	%	Cumul. N	Cumul. %
male	410	60.92	410	60.92
female	263	39.08	673	100.00
Total	673	100.00	673	100.00

Table 1: Frequency table: Gender

The most frequent value is *male*.



Charts

It seems that the highest value is *2* which is exactly *2* times higher than the smallest value (*1*).

Description

This template will return descriptive statistics of a numerical or frequency table of a categorical variable.

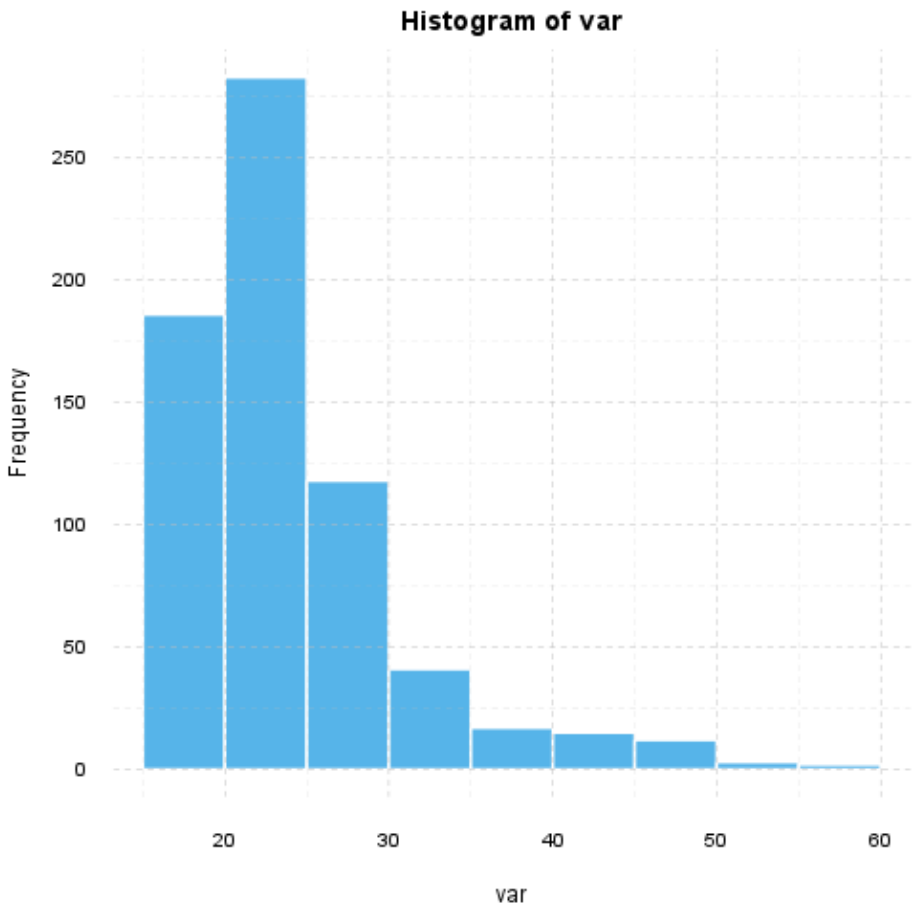
age ("Age")

The dataset has *709* observations with *677* valid values (missing: *32*).

Variable	mean	sd	var
Age	24.57	6.849	46.91

Table 2: Descriptives: Age

Base statistics The standard deviation is *6.849* (variance: *46.91*). The expected value is around *24.57*, somewhere between *24.06* and *25.09* with the standard error of *0.2632*.



Charts

It seems that the highest value is *58* which is exactly *3.625* times higher than the smallest value (*16*).

If we *suppose* that *Age* is not near to a normal distribution (skewness: *1.925*, kurtosis: *4.463*), checking the median (*23*) might be a better option instead of

the mean. The interquartile range (*6*) measures the statistics dispersion of the variable (similar to standard deviation) based on median.

Description

This template will return descriptive statistics of a numerical or frequency table of a categorical variable.

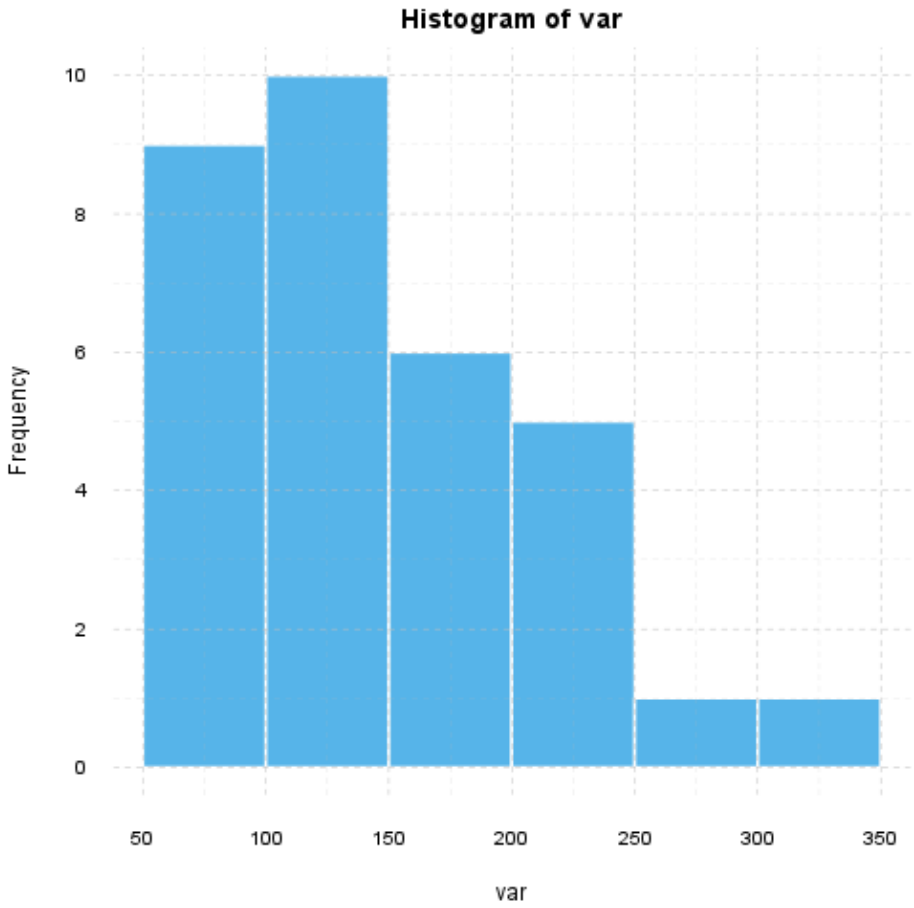
hp

The dataset has *32* observations with *32* valid values (missing: *0*).

Variable	mean	sd	var
hp	146.7	68.56	4701

Table 3: Descriptives: hp

Base statistics The standard deviation is *68.56* (variance: *4701*). The expected value is around *146.7*, somewhere between *122.9* and *170.4* with the standard error of *12.12*.



Charts

It seems that the highest value is *335* which is exactly *6.442* times higher than the smallest value (*52*).

If we *suppose* that *hp* is not near to a normal distribution (skewness: *0.726*, kurtosis: *-0.1356*), checking the median (*123*) might be a better option instead of the mean. The interquartile range (*83.5*) measures the statistics dispersion of the variable (similar to standard deviation) based on median.

This report was generated with R (2.15.1) and rapport (0.4) in *1.258* sec on x86_64-unknown-linux-gnu platform.

Figure 1: