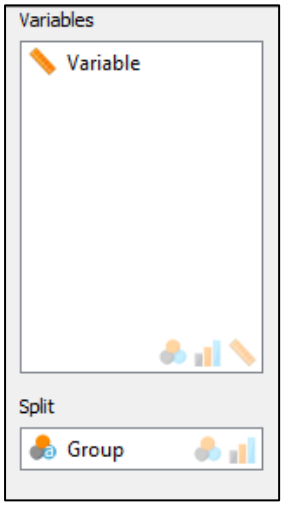
SPLITTING DATA FILES

If there is a grouping variable (categorical or ordinal), descriptive statistics and plots can be produced for each group. Using **Descriptive data.csv** with the variable data in the Variables box, now add Group to the Split box.



Descriptive Statistics

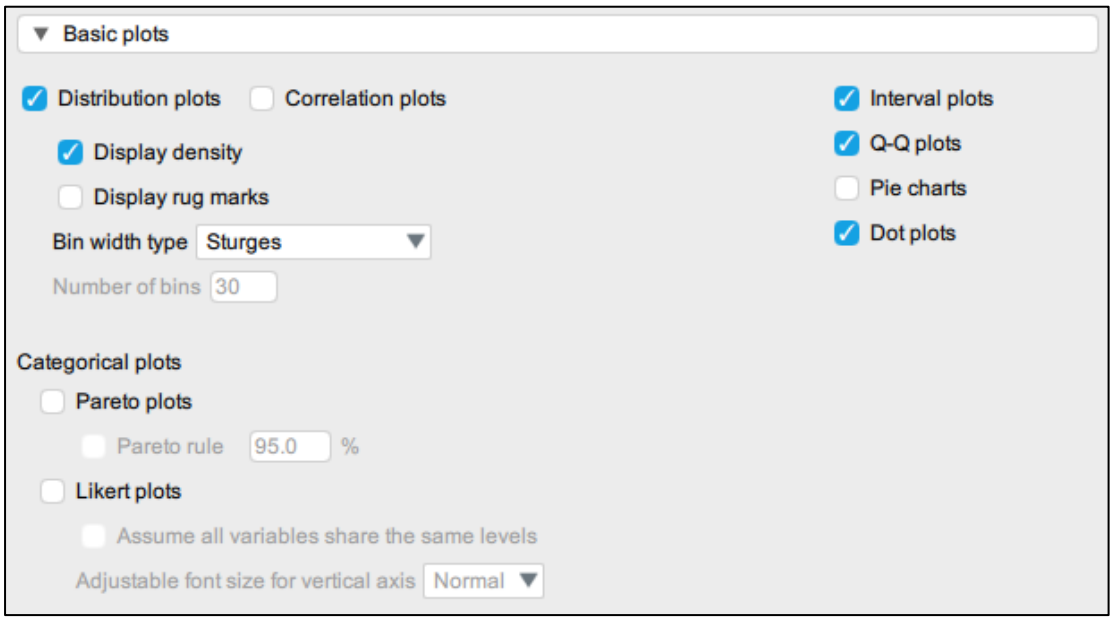
	Variable	
	Group 1	Group 2
Mean	16.021	18.787
Std. Error of Mean	0.362	0.316
Std. Deviation	6.424	7.040
Skewness	0.200	−0.176
Std. Error of Skewness	0.137	0.110
Kurtosis	−0.101	−0.397
Std. Error of Kurtosis	0.274	0.219
Shapiro-Wilk	0.993	0.993
P-value of Shapiro-Wilk	0.119	0.029

DESCRIPTIVE DATA VISUALISATION

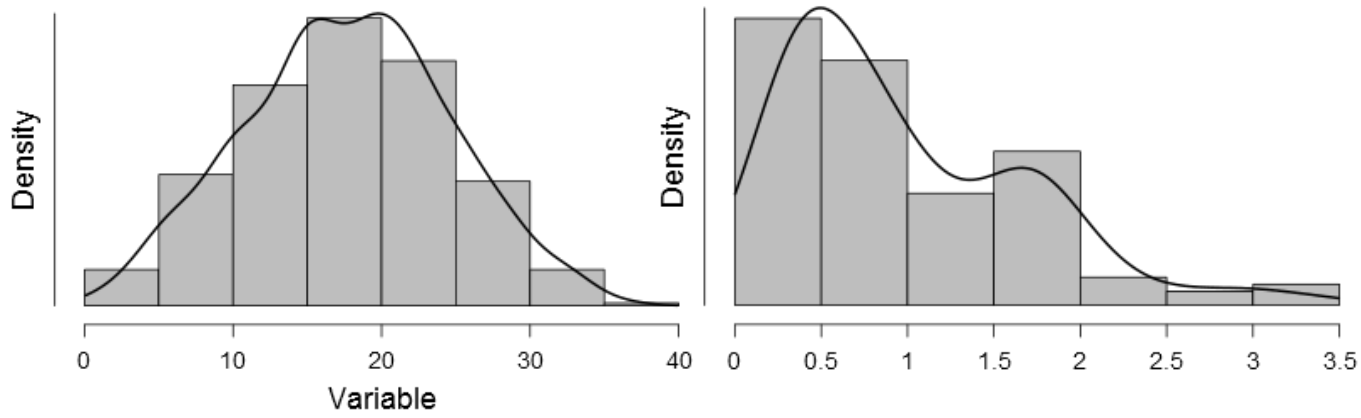
JASP produces a comprehensive range of descriptive and analysis-specific plots. These plots will be explained in their relevant chapters.

BASIC PLOTS

Firstly, to look at examples of the basic plots, open **Descriptive data.csv** with the variable data in the Variables box, go to Plots and tick Distribution plots, Display density, Interval plots, Q-Q plots, and dot plots.

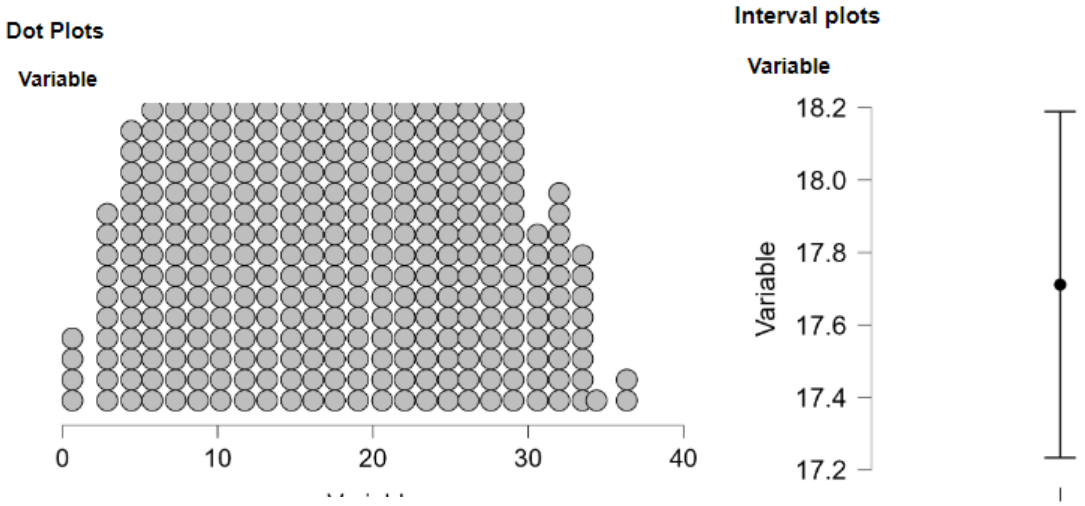


The Distribution plot is based on splitting the data into frequency bins, which are then overlaid with the distribution curve. As mentioned before, the highest bar is the mode (the most frequent value of the dataset. In this case, the curve looks approximately symmetrical, suggesting that the data is approximately normally distributed. The second distribution plot is from another dataset, which shows that the data is positively skewed.

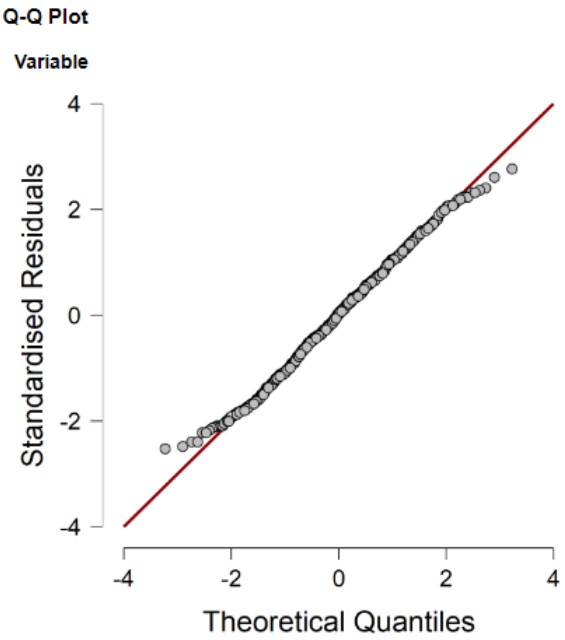


The dot plot displays the distribution where each dot represents a value. If a value occurs more than once, the dots are placed one above the other so that the height of the column of dots represents the frequency for that value.

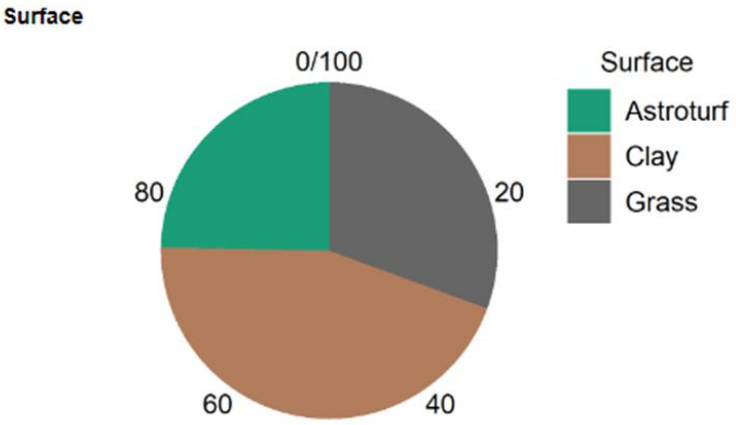
The interval plot shows a 95% confidence interval for the mean of each variable.



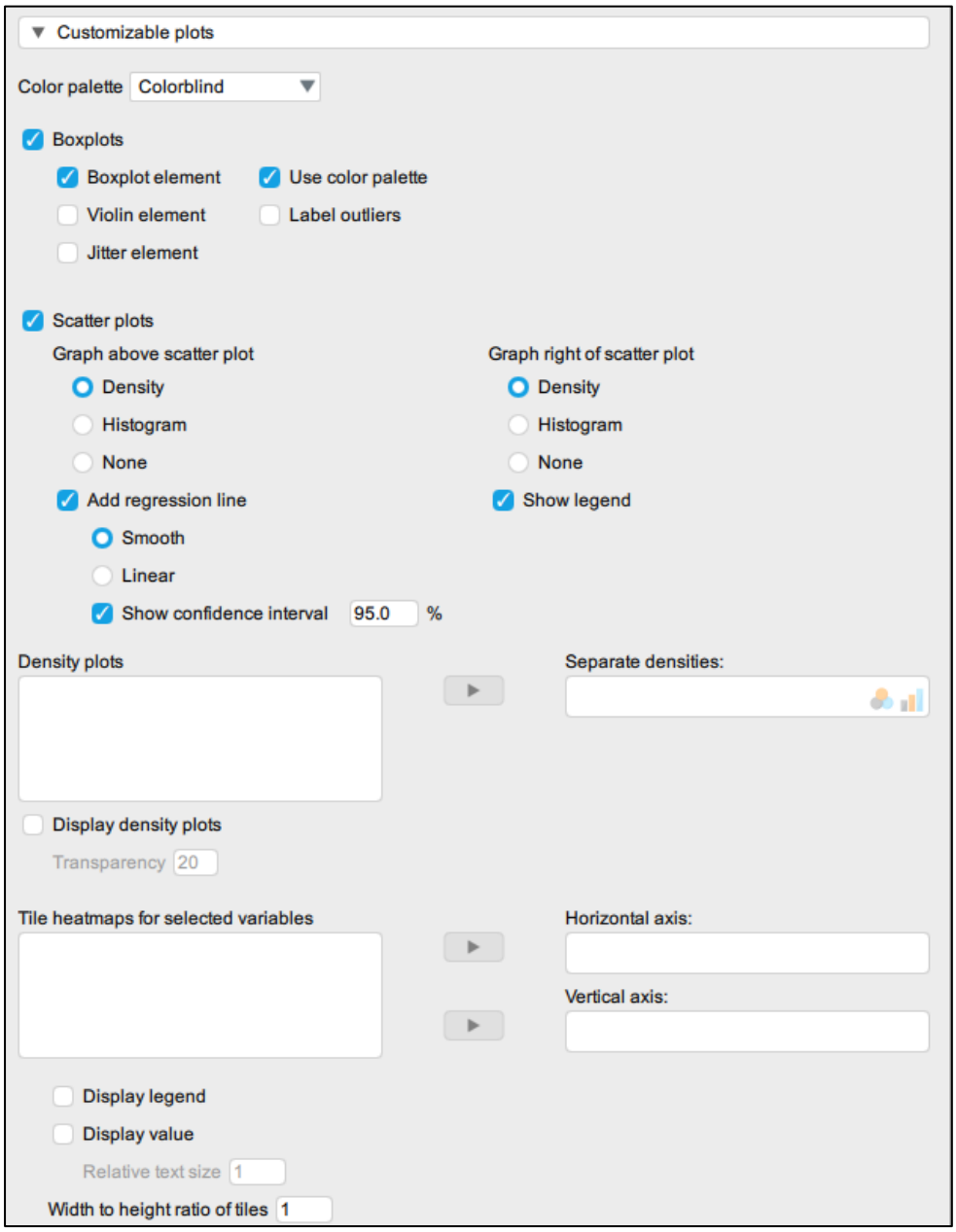
The Q-Q plot (quantile-quantile plot) can be used to visually assess if a set of data comes from a normal distribution. Q-Q plots take the sample data, sort it in ascending order, and then plot it against quantiles (percentiles) calculated from a theoretical distribution. If the data is normally distributed, the points will fall on or close to the 45-degree reference line. If the data is not normally distributed, the points will deviate from the reference line.



Depending on the data sets, basic correlation graphs and pie charts for non-scale data can also be produced, as can Pareto and Likert plots.



CUSTOMISABLE PLOTS



There are a variety of options depending on your datasets.

The boxplots visualise several statistics described above in one plot:

- Median value
- 25 and 75% quartiles
- Interquartile range (IQR) i.e., 75% - 25% quartile values
- Maximum and minimum values plotted with outliers excluded
- Outliers are shown if requested