

Correlations

Compute all pairwise attribute correlations.

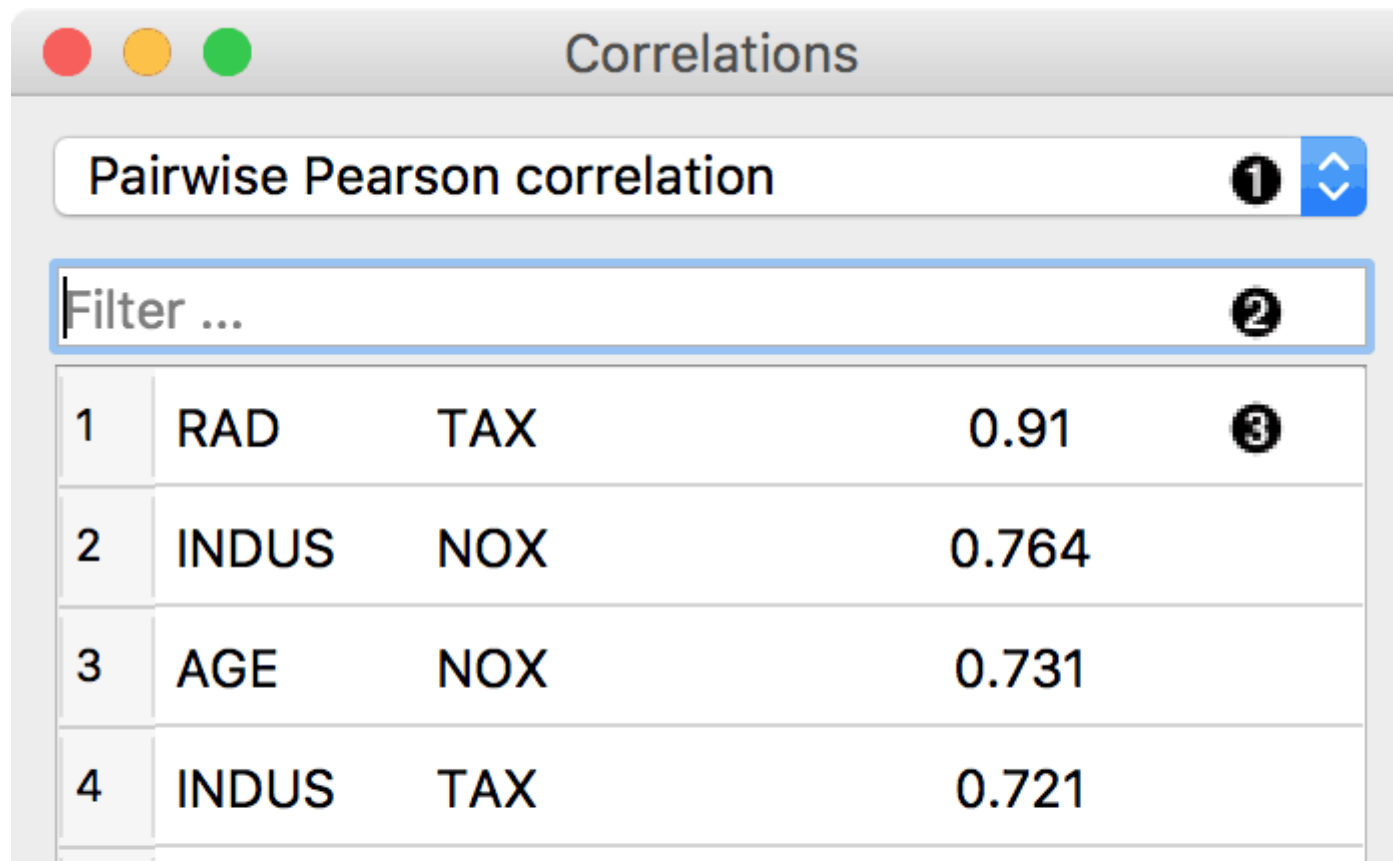
Inputs

- Data: input dataset

Outputs

- Data: input dataset
- Features: selected pair of features
- Correlations: data table with correlation scores




Correlations computes Pearson or Spearman correlation scores for all pairs of features in a dataset. These methods can only detect monotonic relationship.



1	RAD	TAX	0.91
2	INDUS	NOX	0.764
3	AGE	NOX	0.731
4	INDUS	TAX	0.721

5	NOX	TAX	0.668
6	DIS	ZN	0.664
7	AGE	INDUS	0.645
8	CRIM	RAD	0.626
9	NOX	RAD	0.611

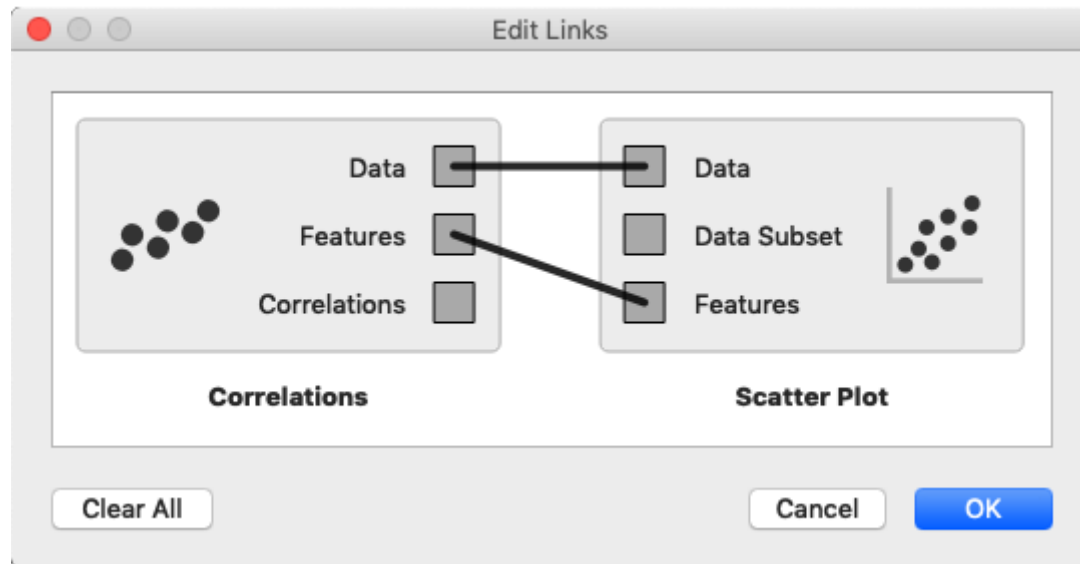
Finished



1. Correlation measure:
 - Pairwise **Pearson** correlation.
 - Pairwise **Spearman** correlation.
2. Filter for finding attribute pairs.
3. A list of attribute pairs with correlation coefficient. Press *Finished* to stop computation for large datasets.
4. Access widget help and produce report.

Example

Correlations can be computed only for numeric (continuous) features, so we will use *housing* as an example data set. Load it in the **File** widget and connect it to **Correlations**. Positively correlated feature pairs will be at the top of the list and negatively correlated will be at the bottom.



Go to the most negatively correlated pair, DIS-NOX. Now connect **Scatter Plot** to **Correlations** and set two outputs, Data to Data and Features to Features. Observe how the feature pair is immediately set in the scatter plot. Looks like the two features are indeed negatively correlated.

