

# Select Rows

Selects data instances based on conditions over data features.

## Inputs

- Data: input dataset

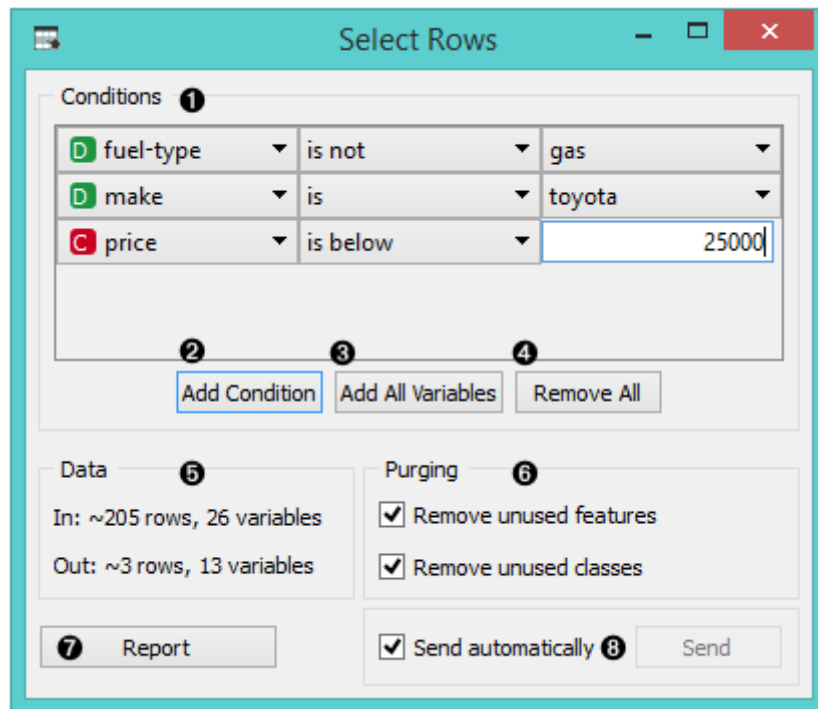
## Outputs

- Matching Data: instances that match the conditions
- Non-Matching Data: instances that do not match the conditions
- Data: data with an additional column showing whether a instance is selected

This widget selects a subset from an input dataset, based on user-defined conditions. Instances that match the selection rule are placed in the output *Matching Data* channel.

Criteria for data selection are presented as a collection of conjunct terms (i.e. selected items are those matching all the terms in '*Conditions*').

Condition terms are defined through selecting an attribute, selecting an operator from a list of operators, and, if needed, defining the value to be used in the condition term. Operators are different for discrete, continuous and string attributes.



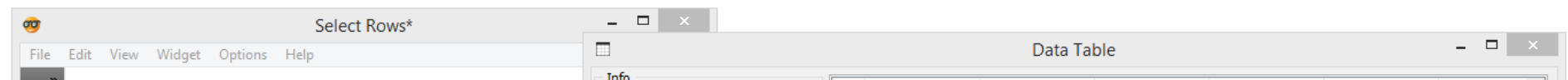
1. Conditions you want to apply, their operators and related values
2. Add a new condition to the list of conditions.
3. Add all the possible variables at once.
4. Remove all the listed variables at once.
5. Information on the input dataset and information on instances that match the condition(s)
6. Purge the output data.
7. When the *Send automatically* box is ticked, all changes will be automatically communicated to other widgets.
8. Produce a report.

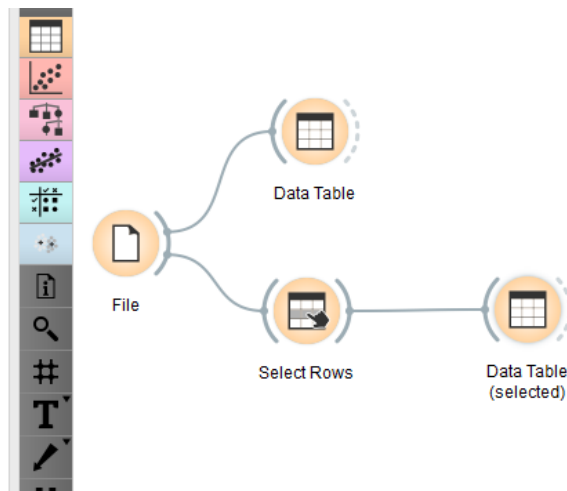
Any change in the composition of the condition will update the information pane (*Data Out*).

If *Send automatically* is selected, then the output is updated on any change in the composition of the condition or any of its terms.

## Example

In the workflow below, we used the Zoo data from the **File** widget and fed it into the **Select Rows** widget. In the widget, we chose to output only two animal types, namely fish and reptiles. We can inspect both the original dataset and the dataset with selected rows in the **Data Table** widget.





**Select Rows**

Conditions

☒ type is one of fish, reptile

Data

In: ~101 rows, 18 variables  
Out: ~18 rows, 12 variables

Purging

☒ Remove unused features  
☒ Remove unused classes  
☒ Send automatically

101 instances (no missing values)  
16 features (no missing values)  
Discrete class with 7 values (no missing values)  
1 meta attribute (no missing values)

Variables

☒ Show variable labels (if present)  
☒ Visualize continuous values  
☒ Color by instance classes

Selection

☒ Select full rows

	type	name	hair	feathers	eggs	milk
1	mammal	aardvark	1	0	0	1
2	mammal	antelope	1	0	0	1
3	fish	bass	0	0	1	0
4	mammal	bear	1	0	0	1
5	mammal	boar	1	0	0	1
6	mammal	buffalo	1	0	0	1
7	mammal	calf	1	0	0	1
8	fish	carp	0	0	1	0
9	fish	catfish	0	0	1	0
10	mammal	cavy	1	0	0	1
11	mammal	cheetah	1	0	0	1
12	bird	chicken	0	1	1	0
13	fish	chub	0	0	1	0
14	invertebrate	clam	0	0	1	0
15	invertebrate	crab	0	0	1	0
16	invertebrate	crayfish	0	0	1	0
17	bird	crow	0	1	1	0
18	mammal	deer	1	0	0	1

**Data Table (selected)**

Info

18 instances (no missing values)  
10 features (no missing values)  
Discrete class with 2 values (no missing values)  
1 meta attribute (no missing values)

Variables

☒ Show variable labels (if present)  
☒ Visualize continuous values  
☒ Color by instance classes

Selection

☒ Select full rows

☒

	type	name	eggs	aquatic	predator	toothed
1	fish	bass	1	1	1	1
2	fish	carp	1	1	0	1
3	fish	catfish	1	1	1	1
4	fish	chub	1	1	1	1
5	fish	dogfish	1	1	1	1
6	fish	haddock	1	1	0	1
7	fish	herring	1	1	1	1
8	fish	pike	1	1	1	1
9	fish	piranha	1	1	1	1
10	reptile	pitviper	1	0	1	1
11	fish	seahorse	1	1	0	1
12	reptile	seasnake	0	1	1	1
13	reptile	slowworm	1	0	1	1
14	fish	sole	1	1	0	1
15	fish	stingray	1	1	1	1
16	reptile	tortoise	1	0	0	0
17	reptile	tuatara	1	0	1	1
18	fish	tuna	1	1	1	1

In the next example, we used the data from the *Titanic* dataset and similarly fed it into the **Box Plot** widget. We first observed the entire dataset based on survival. Then we selected only first class passengers in the **Select Rows** widget and fed it again into the **Box Plot**. There we could see all the first class passengers listed by their survival rate and grouped by gender.

