

GEO Data Sets

Provides access to data sets from gene expression omnibus [GEO DataSets](#).

Inputs

- None

Outputs

- Expression data: Data set selected in the widget with genes or samples in rows.

[GEO DataSets](#) is a database of gene expression curated profiles maintained by [NCBI](#) and included in the [Gene Expression Omnibus](#). This Orange widget provides access to all its data sets and outputs a data set selected for further processing. For convenience, each downloaded data set is stored locally.

Info

4172 datasets
1 datasets cached
3 after filtering

Output

☐ Genes in rows
☒ Samples in rows
☒ Merge spots of same gene

Data set name
Caffeine effect: time course an...

Commit

☐

1 Filter

Caffeine Effect

ID	Title	Organism	Samples	Features	Genes	Subsets	PubMedID
GDS2914	Caffeine effect: time cour...	Saccharomyces cere...	36	15488	6370	8	16925551
GDS2336	Caffeine effect on TOR co...	Saccharomyces cere...	6	6257	6184	5	16923813
GDS2338	Caffeine and rapamycin ef...	Saccharomyces cere...	6	6257	6184	5	16923813

2

Description

Analysis of cells treated with a low dose of caffeine at 2 mM or a high one at 20 mM for 5, 15, or 90 minutes. Caffeine is a natural purine analogue that elicits pleiotropic effects leading to cell death. Results provide insight into the mechanism of action of caffeine.

Sample Annotations

Type (Sample annotations)	Sample count
<input checked="" type="checkbox"/> dose	
<input checked="" type="checkbox"/> high	18
<input checked="" type="checkbox"/> low	18
<input checked="" type="checkbox"/> time	
<input checked="" type="checkbox"/> 5 min	12
<input checked="" type="checkbox"/> 15 min	12

1. Information on the GEO data set collection. Cached data sets are the ones currently stored on the computer.
2. Output features. If *Samples in rows* is selected, genes (or spots) will be used as attributes. Alternatively samples will be used as attributes. *Merge spots of same gene* averages measures of the same gene. Finally, in the *Data set name* you can rename the output data.

GEO title will be used as a default name.

3. If *Auto commit is on*, then the selected data set will be automatically communicated to other widgets. Alternatively, click *Commit*.
4. *Filter* allows you to search for the data set. Below you see a list of GEO data sets with an ID number (link to the NCBI Data Set Browser), title, organism used in the experiment, number of samples, features, genes, subsets and a reference number for the PubMed journal (link to the article abstract).
5. Short description of the experiment from which the data set is sourced.
6. Select which *Sample Annotations* will be used in the output.

Example

GEO Data Sets is similar to the **File** widget, since it is used to load the data. In the example below we selected *Caffeine effect: time course and dose response* dataset from the GEO data base. Do not forget to press *Commit* to output the data. We can inspect the data in *Data Table*.

The screenshot displays the Orange Data Mining software interface. The main window shows a workflow with two widgets: 'GEO Data Sets' and 'Data Table'. The 'GEO Data Sets' widget is on the left, and the 'Data Table' widget is on the right. The 'Data Table' widget is currently selected, showing a detailed view of the data.

GEO Data Sets Widget Info:

- 4172 datasets
- 1 datasets cached
- 3 after filtering

Output:

- ☐ Genes in rows
- ☒ Samples in rows
- ☒ Merge spots of same gene

Data set name: Caffeine effect: time course an...

Commit: ☐

Filter: Caffeine Effect

ID	Title	Organism	Samples	Features	Genes	Subsets	PubMedID
GDS2914	Caffeine effect: time cour...	Saccharomyces cere...	36	15488	6370	8	16925551
GDS2336	Caffeine effect on TOR co...	Saccharomyces cere...	6	6257	6184	5	16923813
GDS2338	Caffeine and rapamycin ef...	Saccharomyces cere...	6	6257	6184	5	16923813

Description: Analysis of cells treated with a low dose of caffeine at 2 mM or a high one at 20 mM for 5, 15, or 90 minutes. Caffeine is a natural purine analogue that elicits pleiotropic effects leading to cell death. Results provide insight into the mechanism of action of caffeine.

Sample Annotations:

Type (Sample annotations)	Sample count
<input checked="" type="checkbox"/> dose	
<input checked="" type="checkbox"/> high	18
<input checked="" type="checkbox"/> low	18
<input checked="" type="checkbox"/> time	
<input checked="" type="checkbox"/> 5 min	12
<input checked="" type="checkbox"/> 15 min	12
<input checked="" type="checkbox"/> 90 min	12
<input checked="" type="checkbox"/> agent	

Data Table Widget Info:

- 36 instances
- 6370 features (no missing values)
- Discrete class with 18 values (no missing values)
- 3 meta attributes (38.9% missing values)

Variables:

- ☒ Show variable labels (if present)
- ☐ Visualize numeric values
- ☒ Color by instance classes

Selection:

- ☒ Select full rows

Data Table:

Entrez ID	class	dose	time	agent	-- EMPTY	AAC1 855078
1	caffeine low ...	low	5 min	caffeine	0.208	-0.31
2	caffeine low ...	low	5 min	caffeine	0.243	0.03
3	calcofluor w...	low	5 min	calcofluor w...	-0.117	0.13
4	calcofluor w...	low	5 min	calcofluor w...	-0.054	0.20
5	congo red lo...	low	5 min	congo red	-0.218	0.11
6	congo red lo...	low	5 min	congo red	0.148	0.02
7	caffeine low ...	low	15 min	caffeine	-0.580	0.20
8	caffeine low ...	low	15 min	caffeine	0.301	0.08
9	calcofluor w...	low	15 min	calcofluor w...	-0.410	0.01
10	calcofluor w...	low	15 min	calcofluor w...	0.251	0.00
11	congo red lo...	low	15 min	congo red	-0.177	-0.10
12	congo red lo...	low	15 min	congo red	0.247	0.07
13	caffeine low ...	low	90 min	caffeine	-0.232	-0.37
14	caffeine low ...	low	90 min	caffeine	0.283	0.00
15	calcofluor w...	low	90 min	calcofluor w...	0.033	-0.26
16	calcofluor w...	low	90 min	calcofluor w...	0.352	-0.03
17	congo red lo...	low	90 min	congo red	-0.184	-0.03
18	congo red lo...	low	90 min	congo red	-0.016	0.18
19	high	5 min	5 min	caffeine	-0.292	0.08
20	high	5 min	5 min	caffeine	0.273	-0.03
21	high	5 min	5 min	calcofluor w...	-0.151	0.07