

Group By

Block Group:	Table Operations
Icon:	

The Group By block returns a new table that contains one row for each group of rows in the input table. A group is composed of rows that hold an identical value in the specified column.

For information on using dataflow blocks, see [Dataflow](#).

For answers to some common questions about working with tables, see [Tables](#).

Input/Output Properties

The following properties of the Group By block can take input and give output.

- input (*table*)
- baseColumn (*string*)
- nestedData (*boolean*)
- includeBlankValue (*boolean*)
- column *n* (*string*)
- method *n* (*string or enum*)
- outColumn *n* (*string*)

input receives the table whose values you want to group.

baseColumn specifies the name of the table column to use for creating groups. Rows that hold an identical value in this column will be grouped together.

nestedData specifies whether to keep the original rows and store them in a new column of the output table as nested tables. For more information about nested tables, see [How do I view the data in nested tables?](#)

includeBlankValue specifies whether to create a group for an empty cell in **baseColumn**, if any empty cells exist.

column *n* specifies the name of an input table column to include in the output table.

method *n* specifies the value that will be stored in **column *n***. **method *n*** can be one of the following:

- **first** returns the first value in the group.
- **last** returns the last value in the group.
- **average** returns the average of all number values in the group. Non-number values are treated as null and are not included in the calculation.
- **sum** returns the sum of all number values in the group. If there are no number values, returns 0.

- **max** returns the the largest number value in the group. If there are no number values, returns null.
- **min** returns the smallest number in the group. Non-number values are treated as null and are not included in the calculation.
- **count** returns the number of values in the group.
- **concat** returns all values in the group, separated by commas. Duplicate values are listed.
- **concatUnique** returns all unique values in the group, separated by commas. Duplicate values are not listed.
- **tableUnion** returns a table of values.

outColumn *n* specifies the output table column name for the column of grouped **column *n*** values.

Output Property

The following property of the Group By block can give output but cannot take input.

- output (*table*)

output returns the output table.

Example

The following image shows an example of the Group By block. In this example, rows are grouped using the values in the **v1** column. The output table displays a sum of **v2** values and a list of **v3** values for each group.

Dataflow

← > Stage > Group By > tableGroupBy

inputTable

data: inputTable

row	v1	v2	v3
0	4	1	v
1	2	3	a
2	6	3	k
3			j
4	6	5	h
5	5	6	2
6	5	6	f

tableGroupBy

- input: inputTable
- baseColumn: v1
- nestedData: false
- column 0: v2
- method 0: sum
- outColumn 0: sum of v2
- column 1: v3
- method 1: concat
- outColumn 1: list of v3
- output: Table

row	v1	sum of v2	list of v3
0	4	1	v
1	2	3	a
2	6	8	k,h
3		0	j
4	5	12	2,f

Previous: [Filter](#)

Next: [Join](#)

From:
<https://wiki.dglogik.com/> - **DGLogik**

Permanent link:
https://wiki.dglogik.com/dglux5_wiki:dataflow:dataflow_blocks_reference:table_operations:group_by

Last update: **2021/09/20 15:03**

