

KTHLARGEST Function

Extracts the ranked value from the values in a column, where $k=1$ returns the maximum value. The value for k must be between 1 and 1000, inclusive. Inputs can be Integer, Decimal, or Datetime.

For purposes of this calculation, two instances of the same value are treated as separate values. So, if your dataset contains three rows with column values 10, 9, and 9, then `KTHLARGEST` returns 9 for $k=2$ and $k=3$.

When used in a `pivot` transform, the function is computed for each instance of the value specified in the `group` parameter. See *Pivot Transform*.

Input column can be of Integer, Decimal, or Datetime type. Other values column are ignored. If a row contains a missing or null value, it is not factored into the calculation.

Wrangle vs. SQL: This function is part of Wrangle, a proprietary data transformation language. Wrangle is not SQL. For more information, see *Wrangle Language*.

Basic Usage

```
kthlargest(myRating, 2)
```

Output: Returns the second highest value from the `myRating` column.

Syntax and Arguments

```
kthlargest(function_col_ref, k_integer) [ group:group_col_ref ] [ limit:limit_count ]
```

Argument	Required?	Data Type	Description
<code>function_col_ref</code>	Y	string	Name of column to which to apply the function
<code>k_integer</code>	Y	integer (positive)	The ranking of the value to extract from the source column

For more information on the `group` and `limit` parameters, see *Pivot Transform*.

For more information on syntax standards, see *Language Documentation Syntax Notes*.

`function_col_ref`

Name of the column the values of which you want to calculate the mean. Inputs must be Integer, Decimal, or Datetime values.

NOTE: If the input is in Datetime type, the output is in unixtime format. You can wrap these outputs in the `DATEFORMAT` function to generate the results in the appropriate Datetime format. See *DATEFORMAT Function*.

- Literal values are not supported as inputs.
- Multiple columns and wildcards are not supported.

Usage Notes:

Required?	Data Type	Example Value
Yes	String (column reference)	myValues

k_integer

Integer representing the ranking of the value to extract from the source column.

NOTE: The value for k must be an integer between 1 and 1,000 inclusive.

- $k=1$ represents the maximum value in the column.
- If k is greater than or equal to the number of values in the column, the minimum value is returned.
- Missing and null values are not factored into the ranking of k .

Usage Notes:

Required?	Data Type	Example Value
Yes	Integer (positive)	4

Examples

Tip: For additional examples, see *Common Tasks*.

This example explores how you can use aggregation functions to calculate rank of values in a column.

Functions:

Item	Description
KTHLARGEST Function	Extracts the ranked value from the values in a column, where $k=1$ returns the maximum value. The value for k must be between 1 and 1000, inclusive. Inputs can be Integer, Decimal, or Datetime.
KTHLARGESTUNIQUE Function	Extracts the ranked unique value from the values in a column, where $k=1$ returns the maximum value. The value for k must be between 1 and 1000, inclusive. Inputs can be Integer, Decimal, or Datetime.

Source:

You have a set of student test scores:

Student	Score
Anna	84
Ben	71
Caleb	76
Danielle	87
Evan	85
Faith	92
Gabe	87

Hannah	99
Ian	73
Jane	68

Transformation:

You can use the following transformations to extract the 1st through 4th-ranked scores on the test:

Transformation Name	New formula
Parameter: Formula type	Single row formula
Parameter: Formula	KTHLARGEST(Score, 1)
Parameter: New column name	'1st'

Transformation Name	New formula
Parameter: Formula type	Single row formula
Parameter: Formula	KTHLARGEST(Score, 2)
Parameter: New column name	'2nd'

Transformation Name	New formula
Parameter: Formula type	Single row formula
Parameter: Formula	KTHLARGEST(Score, 3)
Parameter: New column name	'3rd'

Transformation Name	New formula
Parameter: Formula type	Single row formula
Parameter: Formula	KTHLARGEST(Score, 4)
Parameter: New column name	'4th'

Transformation Name	New formula
Parameter: Formula type	Single row formula
Parameter: Formula	KTHLARGESTUNIQUE(Score, 3)
Parameter: New column name	'3rdUnique'

Transformation Name	New formula
Parameter: Formula type	Single row formula
Parameter: Formula	KTHLARGESTUNIQUE(Score, 4)

Parameter: New column name

'4thUnique'

Results:

When you reorganize the columns, the dataset might look like the following:

Student	Score	1st	2nd	3rd	4th	3rdUnique	4thUnique
Anna	84	99	92	87	87	87	85
Ben	71	99	92	87	87	87	85
Caleb	76	99	92	87	87	87	85
Danielle	87	99	92	87	87	87	85
Evan	85	99	92	87	87	87	85
Faith	92	99	92	87	87	87	85
Gabe	87	99	92	87	87	87	85
Hannah	99	99	92	87	87	87	85
Ian	73	99	92	87	87	87	85
Jane	68	99	92	87	87	87	85

Notes:

- The value 87 is both the third and fourth scores.
 - For the `KTHLARGEST` function, it is the output for the third and fourth ranking.
 - For the `KTHLARGESTUNIQUE` function, it is the output for the third ranking only.