### percentile()

グループ内の値のうち、指定したパーセンタイルに該当する値を求める関数です。パーセンタイル基準で分位値を指定し、100は最大値、0は最小値を意味します。

percentile(x,0.0) の結果は [min(x)](https://docs.logpresso.comnull) と同じであり、percentile(x,100.0) の結果は [max(x)](https://docs.logpresso.comnull) と同じです。

#### 構文

percentile(EXPR, PERCENTILE, [INTERPOLATION])

必須パラメータ

**EXPR**

グループ値を返す式

**PERCENTILE**

求めたいパーセンタイル値

PERCENTILE値は0.0～100.0の範囲で指定する必要があり、範囲外の場合はクエリパースエラーとなります。

**INTERPOLATION**

2つの値の間の値を計算する方法

* interpolation値はオプションパラメータであり、指定したパーセンタイルに該当するデータが存在しない場合、両端の値をどのように処理して分位値を求めるかを決定します。例えば、1,2,3,4 の4つの値がある場合、percentile(x,50.0) を求めると2と3の間の値を計算する必要があります。単純に2つの平均値である2.5を返すのが一般的ですが、データ型によっては異なる計算方法が必要となる場合があります。デフォルト値は "linear" です。
* linear: 2つの値のPERCENTILEに応じた重み付けで計算
* midpoint: 2つの値の平均値で計算
* lower: 2つの値のうち小さい方を選択
* higher: 2つの値のうち大きい方を選択
* nearest: 2つの値のうちPERCENTILEにより近い方を選択
* 例）次の11個の値に対して37%パーセンタイルを求める場合、各INTERPOLATION方式ごとの結果は以下の通りです。

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 区分 | 1番 | 2番 | 3番 | 4番 | 5番 | 6番 | 7番 | 8番 | 9番 | 10番 | 11番 |
| 元データ | 97 | 18 | 2 | 85 | 81 | 85 | 65 | 14 | 96 | 22 | 42 |
| ソート | 2 | 14 | 18 | 22 | 42 | 65 | 81 | 85 | 85 | 96 | 97 |
| percentile | 0% | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |

* 値が11個なので、1番目は0%、2番目は10%、3番目は20%、…、11番目は100%となります。したがって、37%は4番目の値22と5番目の値42を基準に計算します。

|  |  |  |
| --- | --- | --- |
| オプション | 計算方法 | 結果 |
| linear | 小数点が0.7なので、42の比重を70%、22の比重を30%として計算 (22 \* 0.3) + (42 \* 0.7) | 36 |
| midpoint | (22+42)/2 | 32 |
| lower | min(22, 42) | 22 |
| higher | max(22, 42) | 42 |
| nearest | 3.7は3と4のうち4に近いため、その値を選択 | 42 |

#### 使用例

percentile(x,80.0)| # 上位80%の値を計算percentile(x,15.5, "midpoint")| # 上位15.5%の値を計算。該当する値がない場合は両端の平均値を計算。percentile(x,100.0)| # max(x)percentile(x,0.0)| # min(x)percentile(x,50.0)| # median(x)