### kmeans

ユークリッド距離に基づき、入力レコードをK個のクラスターに分類します。

#### 構文

kmeans [k=INT] [iter=INT] FIELD, ...

必須パラメーター

**FIELD, ...**

計算対象フィールドのリスト。カンマ（,）で区切ります。フィールド値は数値型である必要があり、指定したフィールド値が数値でない入力レコードは無視されます。最大10万件の入力レコードを許容し、有効な入力レコードが10万件を超える場合はクエリを終了します。クラスタリング結果は\*\*\_cluster\*\*フィールドに1から始まる番号で出力されます。

オプションパラメーター

**k=INT**

クラスター数（デフォルト: 3）

**iter=INT**

計算反復回数（デフォルト: 10000）

#### 使用例

機械学習でよく利用されるirisデータを用いて試すことができます。長さと幅を用いて分類を行い、実際の種（species）名と比較します（ダウンロード: <https://github.com/illinois-cse/data-fa14/blob/gh-pages/data/iris.csv>）。

csvfile /opt/logpresso/iris.csv| eval sepal\_length = double(sepal\_length), sepal\_width = double(sepal\_width)| kmeans k=4 iter=100000 sepal\_length, sepal\_width