<https://www.dataiku.com/learn/guide/tutorials/lab.html>

Aggregate 合计

能够合并两张dataset,连接方式有左连接、右连接、外连接等。可以选择要合并的列。

内连接：保留两张表中都匹配的数据

外连接：保留两张表中所有的数据，包括不匹配的

Lab 实验室

Preliminary 初步的

dedicated 专用

Drafting 制图

Complementarily 互补

Overcrowding 人满为患

Steady 稳定的

Reused 重用

Assess 评估

Satisfied 满意

Lab：实验室，模拟数据的处理等，该实验室中可以有repie,机器学习的算法等，不会生成dataset等，但可以把模拟的dataset等转换为实际的dataset。可以进行可视化分析。

go through 经过

Distribution 分配

be prompted to 被提示

Interactive 互动

Topic 话题

couple 一对

可以用两个列的值相减得到新的一列

distribution 分配

turns out 事实证明

Outliers 离群

Indicative 指示

InterQuartile（**四分位数） 把数值从小到大分成四等份，分割点的数值就是四分位数，第一个分割点与第三个分割点的距离就是四分位距（InterQuartile Range）**

**Clip 清除**

**可以查看一个列的数据的分布，清除掉不合理数据，可以删除掉列属性**

**Leveraging 利用**

**Agent 代理人**

**Classify 分类**

**可以对一个列的数据进行分类**

**Recognize 认识**

**Extract 提取**

**deselect 取消**

**可以解析ip的所在的城市，国家等**

**Label 标签**

**revenue 收入**

**Formula 式**

**Validate 验证**

**可以对某一列添加表达式，并且产生一列判断它是true或false(true和false是可以自己设定的)**

**graphing 制图**

**Productive 生产的，丰富的**

**Loop 循环 套**

**a pie 一块**

**Donut 甜甜圈**

**可以以一个圈的方式看一个属性的分布（一个值随另一个值的影响）**

**relative share 相对份额**

**Enlightening 启发**

**approximately 大约**

**或在是一个值比如总数随着两个值的影响**

**Campaign 运动**

**Insight 眼光**

**Glean 收集**

**Moderately 适当**

**Valued 价值**

**可以画出一个属性和另一个属性的X-Y的散列表，还可以看到该点的其他细节**