

图 118




图 119

预览数据				
species	sepal_width	petal_length	sepal_length	petal_width
setosa	4.4	1.9	5.8	0.6
versicolor	3.4	5.1	7	1.8
virginica	3.8	6.9	7.9	2.5

图 120

3.4.2.16 数据标准化

图标:  数据标准化

描述: 数据标准化处理是将数据按比例缩放，使之落入一个小的特定区间。

字段属性

特征列: 选择进行标准化的列，请选择数值型数据，如果勾选了非数值类型数据，则会自动过滤，下个组件可能无法获取所有列，如图 121 所示。



图 121

参数设置

标准化方式：标准化方式包括极差标准化、零均值标准化和小数定标标准化，默认零均值标准化。

最小值：选择极差标准化时有效。

最大值：选择极差标准化时有效。如图 122 所示



图 122

输出

表结果：标准化结果。

报告：无。

示例

下面对某数据进行标准化处理。原数据如图 123 所示。

预览数据			
id	r	f	m
1	27	6	232.61
2	3	5	1507.11
3	4	16	817.62
4	3	11	232.81
5	14	7	1913.05
6	19	6	220.07
7	5	2	615.83
8	26	2	1059.66

共 940 条 25 条/页 < 1 2 3 4 5 6 ... 38 > 前往 1 页

图 123

- 勾选需要进行数据标准化的数据。如图 124 所示。
- 选择标准化方式为零均值标准化，如图 125 所示。
- 结果如图 126 所示。

The screenshot shows a data processing workflow in a software interface. The main workspace contains two components: '输入源' (Input Source) and '数据标准化' (Data Standardization), connected by a flow line. The right-hand panel, titled '字段属性' (Field Properties), displays a table of selected features for standardization.

特征列	字段	类型	取值
<input type="checkbox"/>	id	数值	-
<input checked="" type="checkbox"/>	m	数值	1535
<input checked="" type="checkbox"/>	r	数值	1-29
<input checked="" type="checkbox"/>	f	数值	1-24

Below the table, there are sections for '参数设置' (Parameter Settings) and '组件描述' (Component Description).

图 124

∨ 参数设置

标准化方式

零均值标准化

最小值 ?

最大值 ?

图 125

预览数据		
m	r	f
-1.159328122362407	0.7645931616651589	-0.49384157634634734
0.62285862671049	-1.025302213157502	-0.6304144159360991
-0.34128412858042545	-0.9507232392065578	0.8718868195511712
-1.1590484539827959	-1.025302213157502	0.18902262160241196
1.19050153680751	-0.20493349969711572	-0.3572687367565955
-1.1768633297640345	0.16796137005760528	-0.49384157634634734
-0.6234555401892288	-0.8761442652556135	-1.0401329347053547
-0.00282945557485135	0.6900141877142147	-1.0401329347053547

共 940 条 25 条/页 < 1 2 3 4 5 6 ... 38 > 前往 页

图 126

3.4.2.17 衍生变量

图标: 衍生变量

描述: 衍生变量是指将一系列或多列通过基本运算生成新列。

字段属性

特征列: 必选。选择进行衍生变量的列。请选择数值型数据,增加的序列会在勾选的字段