目录

[**回归分析**](#_一元线性回归预测模型)**1**

[**聚类分析**](#_1.1.1.4_聚类分析模型)**4**

[**时间序列分析**](#_1.1.1.5_时间序列分析模型)5

[**因子分析**](#_1.1.1.4_聚类分析模型)6

[**主成份分析**](#_1.1.1.4_聚类分析模型)7

#### 一元线性回归预测模型

一元线性回归分析中，只包括一个自变量和一个因变量，且二者的关系可用一条直线近似表示。

#### 1.1.1.2 多元线性回归分析模型

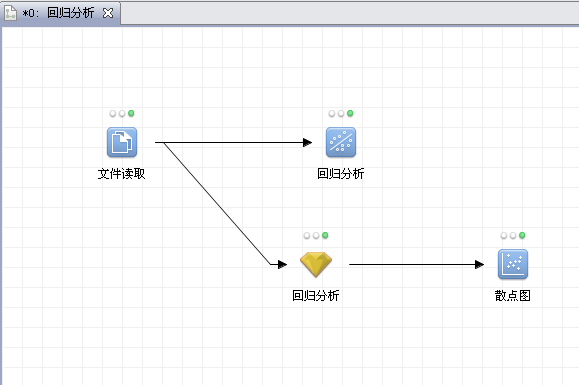
如果回归分析中包括两个或两个以上的自变量，且因变量和自变量之间是[线性关系](http://baike.baidu.com/view/91595.htm)，则称为[多元线性回归](http://baike.baidu.com/view/2362545.htm)分析。

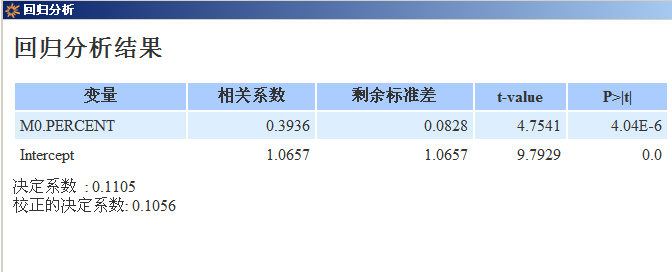
#### 1.1.1.3 非线性回归分析模型

[自变量](http://baike.baidu.com/view/379564.htm)和[因变量](http://baike.baidu.com/view/324030.htm)之间是非线性关系。

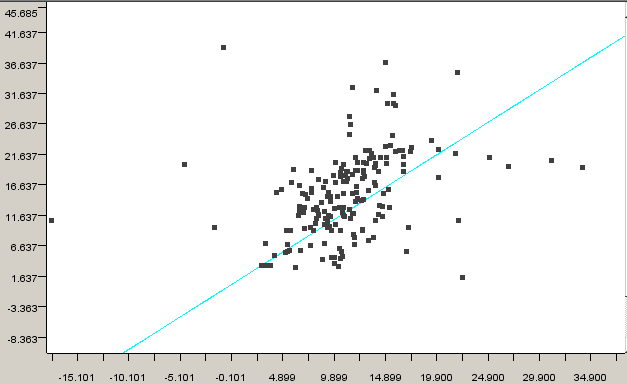
#### DataStudio中的回归分析实现

DataStudio中以上回归分析统一由回归分析算法节点实现，可以做多元线性，多元非线性回归分析。如下图所示：

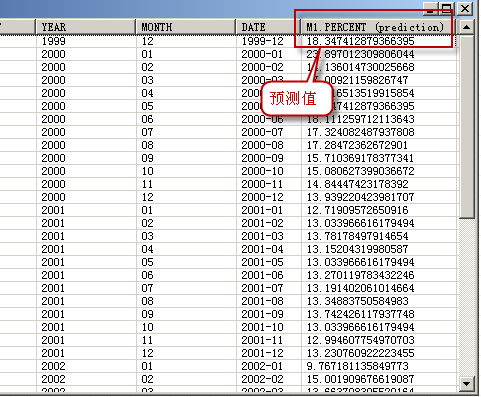
  
回归分析模型的结果展现：

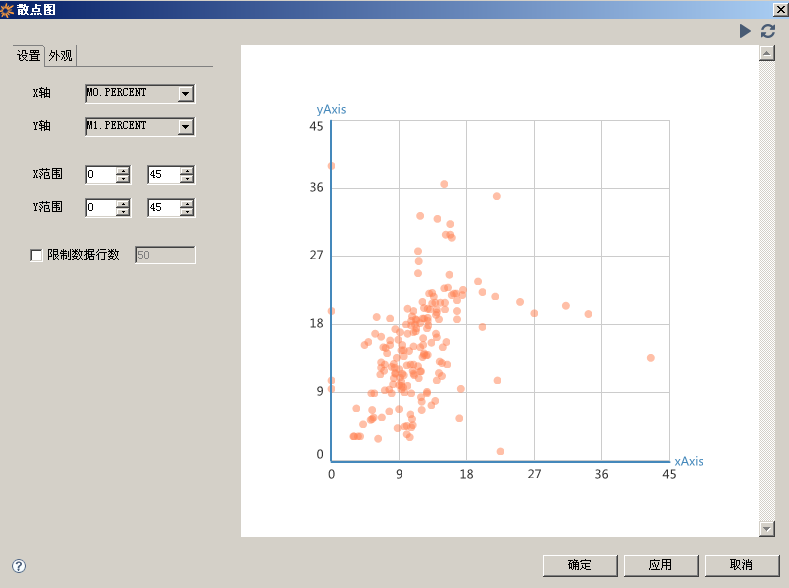


回归分析模型散点图展示：



回归分析预测结果的展现：





#### 1.1.1.4 聚类分析模型

聚类是将[数据分类](http://baike.baidu.com/view/3810806.htm" \t "_blank)到不同的类或者簇这样的一个过程，所以同一个簇中的对象有很大的相似性，而不同簇间的对象有很大的相异性。

#### DataStudio中的聚类分析实现

DataStudio中的聚类算法有多种，以Kmeans算法节点为例说明。



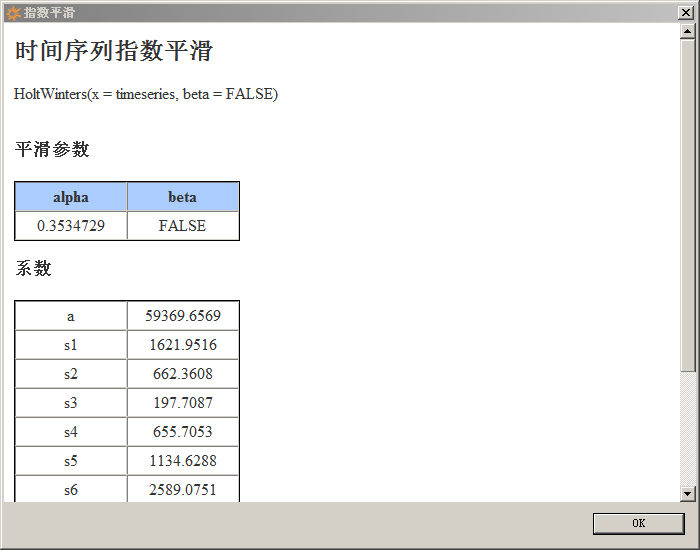
#### 1.1.1.5 时间序列分析模型

时间序列分析可以分析时间序列数据，可以区分出序列值中季节，趋势，随机部分。

#### DataStudio中的时间序列的实现

DataStudio中的时间序列算法节点包括指数平滑，ARIMA。这两种算法都可支持分析具有季节，趋势的时间序列数据。

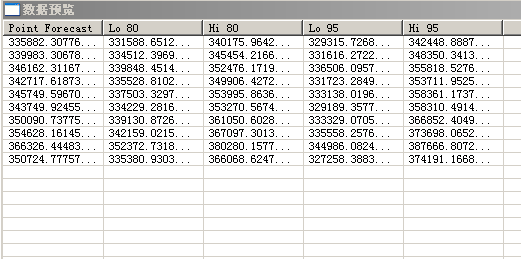
指数平滑算法生成模型的展现：



ARIMA模型的展现：



时间序列的预测结果展现（包括预测值，80%，95%的上下置信区间）：

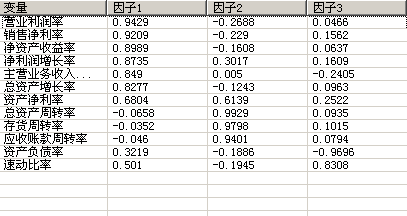


#### 1.1.1.6因子分析模型

在实际问题的分析中，有时需要对不能直接观测的潜在因素进行分析，如企业的形象，企业价值、创新能力等都是不可直接观测到的变量 ( 潜在变量) ，而利税总额、劳动生产率等都是显在变量。因子分析 (Factor Analysis) 是利用少数几个潜在变量或公共因子去解释多个显在变量或可观测变量中存在的复杂关系的分析方法，即将相关比较密切的几个变量归在同一类中，每一类变量就成为一个因子（之所以称其为因子，是因为它是不可观测的，即不是具体的变量），以较少的几个因子反映原资料的大部分信息。

#### DataStudio中因子分析的实现

使用因子分析节点，可以得到因子载荷矩阵，输出如下：

  
根据因子载荷矩阵，我们可以分析变量与哪些因子关系密切。

#### 1.1.1.7 主成分分析模型

主成分分析（Principal Component Analysis，PCA）， 将多个变量通过[线性变换](http://baike.baidu.com/view/325734.htm)以选出较少个数重要变量的一种[多元统计分析](http://baike.baidu.com/view/1010781.htm)方法。又称[主分量分析](http://baike.baidu.com/view/1364200.htm)。

#### DataStudio中主成分分析的实现

使用主成分分析节点，可以得到变量之间的交叉矩阵，模型输出如下：



