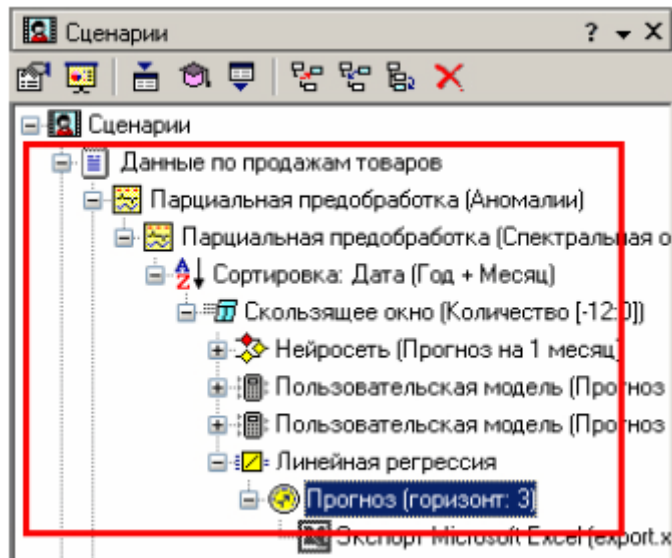


## ЛАБОРАТОРНАЯ РАБОТА 10 ПРОГНОЗИРОВАНИЕ С ПОМОЩЬЮ ЛИНЕЙНОЙ РЕГРЕССИИ

Линейная регрессия необходима тогда, когда предполагается, что зависимость между входными факторами и результатом линейная. Достоинством ее можно назвать быстроту обработки входных данных и простоту интерпретации полученных результатов.

Рассмотрим фрагмент проекта "Демопример анализа данных.ded".



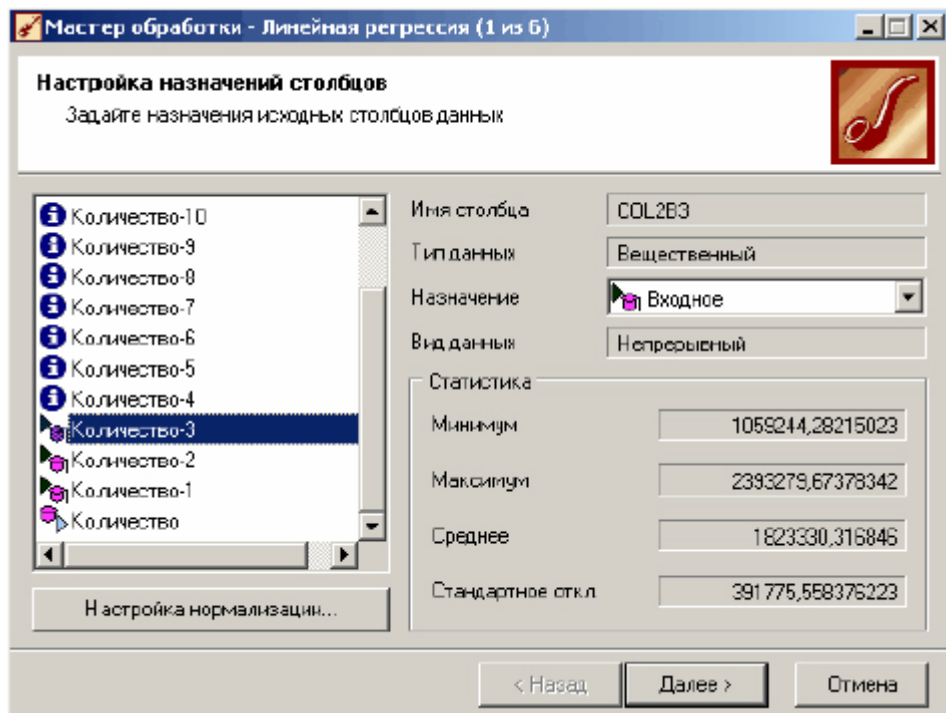
### **Исходные данные**

Рассмотрим применение линейной регрессии на примере данных по продажам, находящихся в файле "Trade.txt". Будем строить прогноз с помощью линейной регрессии от ветки импорта "Данные по продажам товаров" сразу после обработчика "Скользящее окно".

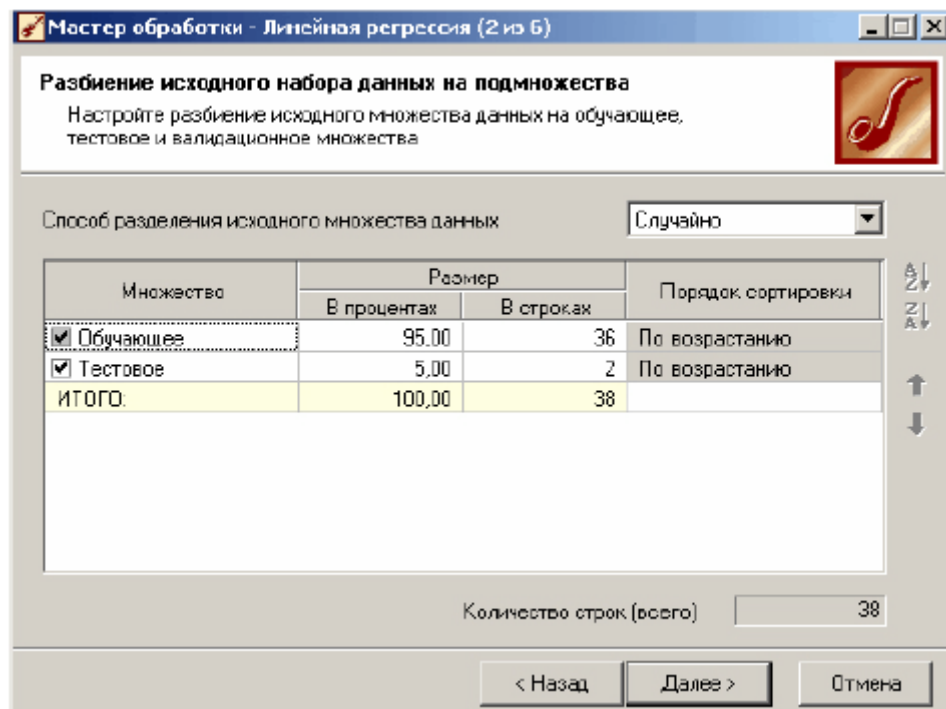
### **Обучение линейной регрессии**

Для построения линейной регрессии необходимо запустить Мастер обработки и выбрать в качестве обработки данных Линейную регрессию.

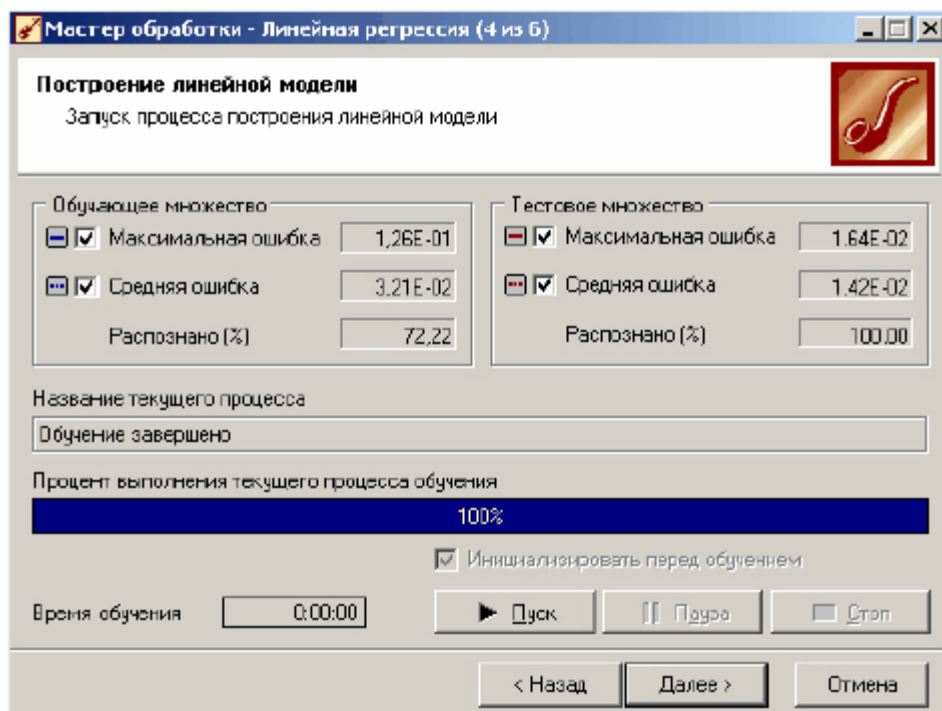
На первом шаге задаем назначение исходных столбцов. Предположим, что на прогноз влияет информация за 3 прошлых месяца, тогда укажем входными столбцами поля: "Количество - 3", "Количество - 2", и "Количество - 1". В качестве выходного поля укажем столбец "Количество".



На следующем шаге происходит настройка обучающего и тестового множеств, способ разложения исходного множества данных.

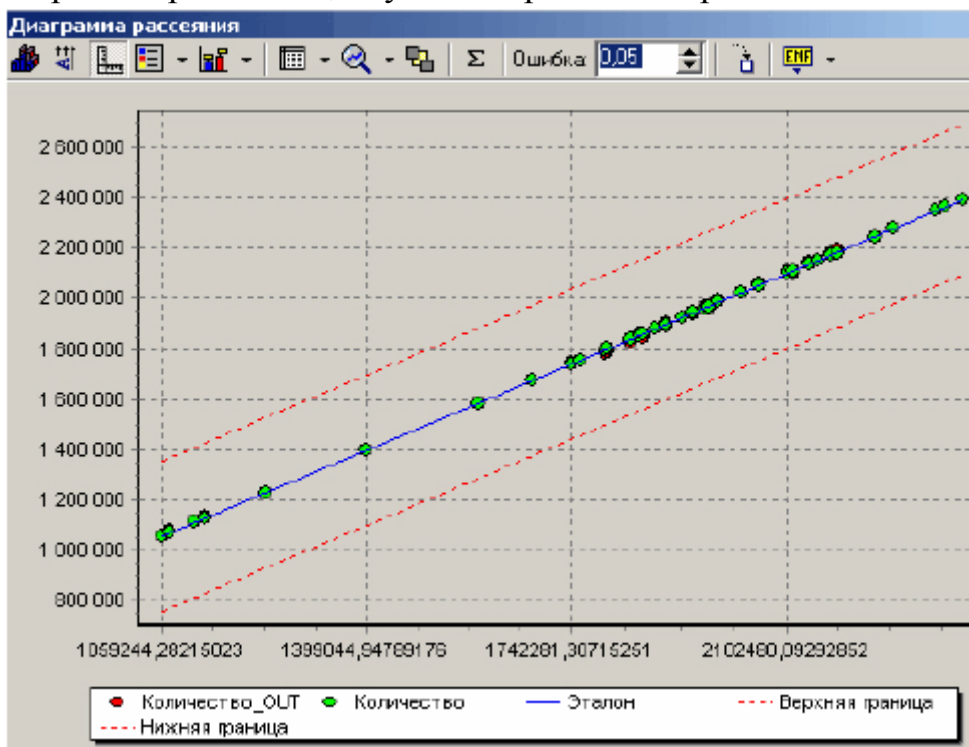


Третий шаг установки позволяет осуществить ограничение диапазона входных значений. Данный шаг оставим без изменений. При нажатии на кнопку "Далее" появляется окно запуска процесса обучения. В процессе выполнения видно, какая часть распознана на этапе обучения и теста.



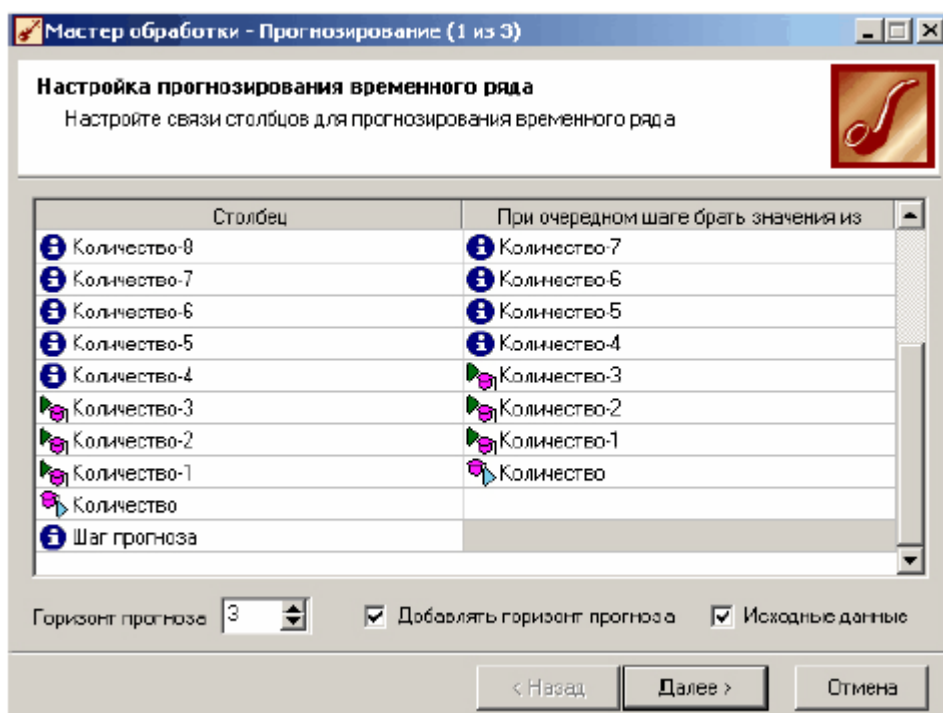
## Результат

После выполнения процесса выберем в качестве способа отображения диаграмму рассеяния и отображение результатов в виде диаграммы. Как видно из диаграммы рассеяния, обучение прошло с хорошей точностью.



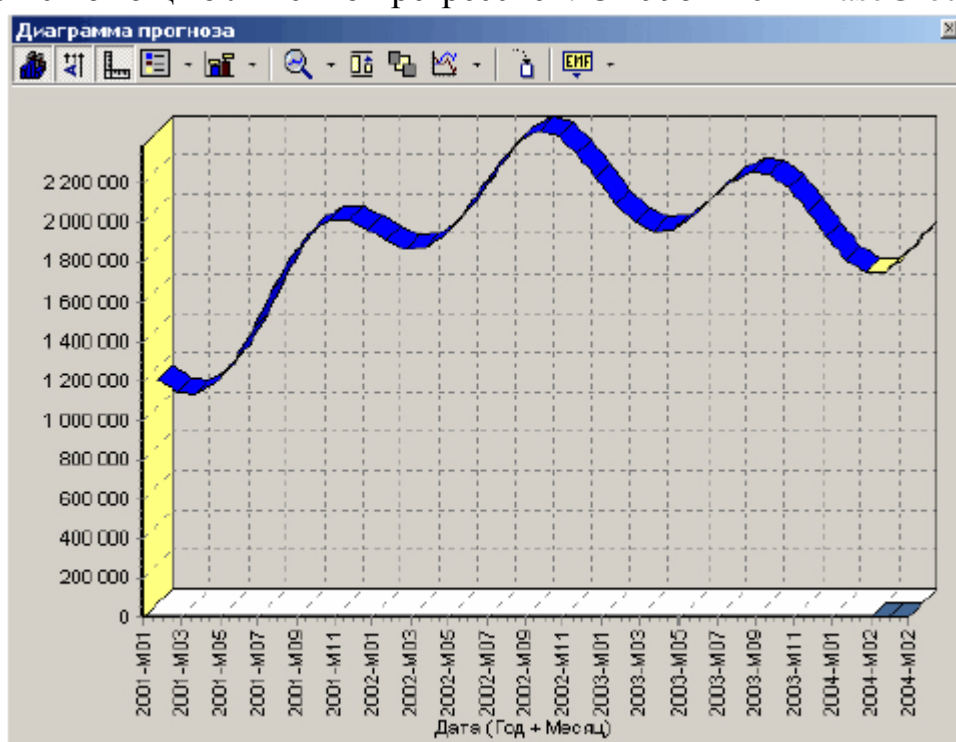
## Прогнозирование

Теперь для построения прогноза запустим Мастера обработки, в котором выберем прогнозирование. На первом шаге обработчика происходит настройка связи столбцов для прогнозирования. Укажем связь между столбцами и горизонт прогноза равный 3.



## Результат

На следующем шаге задаются параметры визуализации. Для данного примера выбираем отображение результатов в виде диаграммы прогноза. Теперь аналитик может дать прогноз о продажах, основываясь на модели, построенной с помощью линейной регрессии. © 1995 - 2011 BaseGroup Labs



## Выводы

Данный пример показал целесообразность применения линейного регрессионного анализа для прогнозирования линейных зависимостей.

Простота настроек и быстрота построения модели иногда бывают необходимы. Аналитику достаточно указать входные столбцы - факторы, выходные - результат, указать способ разбиения данных на тестовое и

обучающее множество и запустить процесс обучения. Причем после этого будут доступны все механизмы визуализации и анализа данных, позволяющие построить прогноз, провести эксперимент по принципу "Что-если", исследовать зависимость результата от значений входных факторов, оценить качество построенной модели по диаграмме рассеяния. Также по результатам работы этого алгоритма можно подтвердить или опровергнуть гипотезу о наличии линейной зависимости.