

Лабораторная работа №6

Прогнозирование продаж товаров в оптовой компании

Описание бизнес-задачи

Постановка задачи. Оптовая компания занимается сбытом строительных материалов. Ассортимент насчитывает несколько тысяч товарных позиций, объединенных в группы (сухие смеси, грунтовка, напольные покрытия, плитка и т. д.). Менеджерам отдела логистики постоянно приходится решать задачи прогнозирования спроса, но большой ассортимент продукции не позволяет уделять каждой позиции должное внимание. Руководство поставило цели: снизить логистические затраты, повысить стабильность запасов и оптимизировать их структуру. Более конкретно требования формулируются следующим образом:

- ☐ автоматизировать ежемесячный расчет потоварного прогноза на следующие три периода;
- ☐ обеспечить аналитикам компании доступ к рассчитанным прогнозам для планирования запасов и формирования заказов.

Для этого было принято решение о внедрении системы автоматического расчета потребности в товарах.

Исходные данные. В наполненном хранилище данных о продажах товаров (начиная с 2004 г.) имеется единственный процесс *Продажи* с измерениями *Дата продажи*, *Товар*, *Товарная группа* и фактами *Количество* и *Сумма*.

Решение задачи

Нужно сразу оговориться, что задача прогнозирования спроса не имеет единой методики решения [1]. Не существует надежного способа количественной оценки наиболее существенных факторов, влияющих на спрос,

что еще сильнее проявляется в условиях рыночных отношений. Хороший аналитик, прогнозирующий спрос, использует все имеющиеся факты — количественные и качественные, а также интуицию, в основе которой лежит опыт работы на данном рынке. Но проблема заключается в том, что при большом ассортименте и частом обновлении спроса ручная обработка каждой товарной позиции не представляется возможной. Необходимо максимально автоматизировать эти процедуры. Например, можно построить несколько прогнозных моделей временных рядов для каждой товарной группы, выбрать из них оптимальную при помощи какого-либо критерия качества и перейти к позиционным прогнозам (рис. 18.1).

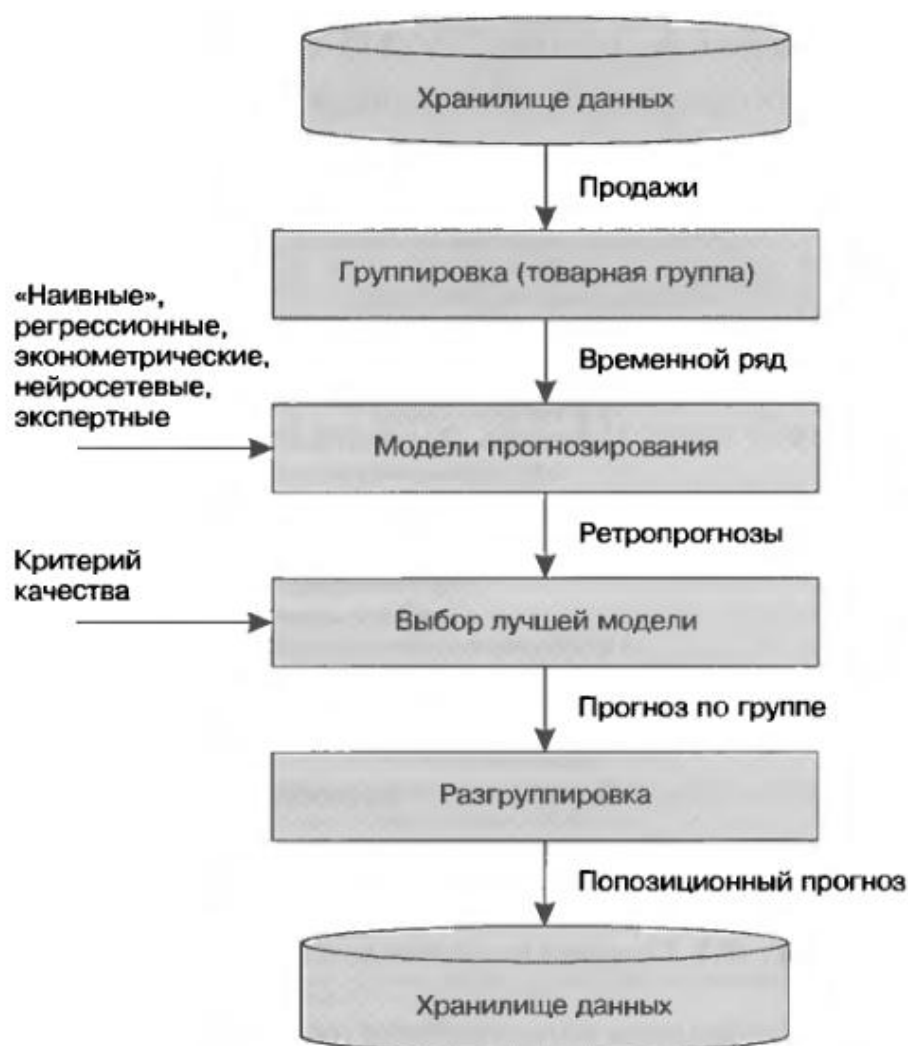


Рис. 18.1. Автоматизированный процесс получения прогноза спроса

Это не единственный вариант: можно отказаться от группировки и обрабатывать каждую позицию по отдельности или сгруппировать товары по ABC-XYZ-категориям и т. д.

Описанный общий подход не обеспечивает идеального результата и не гарантирует высокую точность прогнозов: это невозможно в принципе. Но гарантируется, что будет получен лучший из возможных вариантов прогноза. А поддерживаемые аналитическими платформами механизмы очистки, группировки и подстройки

под новые данные позволят постепенно повышать качество прогнозов, быстро добавлять или изменять прогностические модели.

С помощью сценариев в Deductor несложно реализовать такую схему. Покажем этапы создания сценария на примере товарной группы «Грунтовка». Построим несколько моделей временных рядов, затем выберем наилучшую и рассчитаем по ней прогноз, сделаем разгруппировку по товарным позициям, а результат прогноза сохраним в хранилище данных. В качестве моделей прогноза выберем:

- «наивную» модель скользящего среднего за предыдущие три месяца;
- линейную регрессию с периодами: (1, 2, 3), (1, 2, 6) и (1, 2, 12) месяцев.

Соответствующий сценарий обработки данных приведен на рис. 18.2.

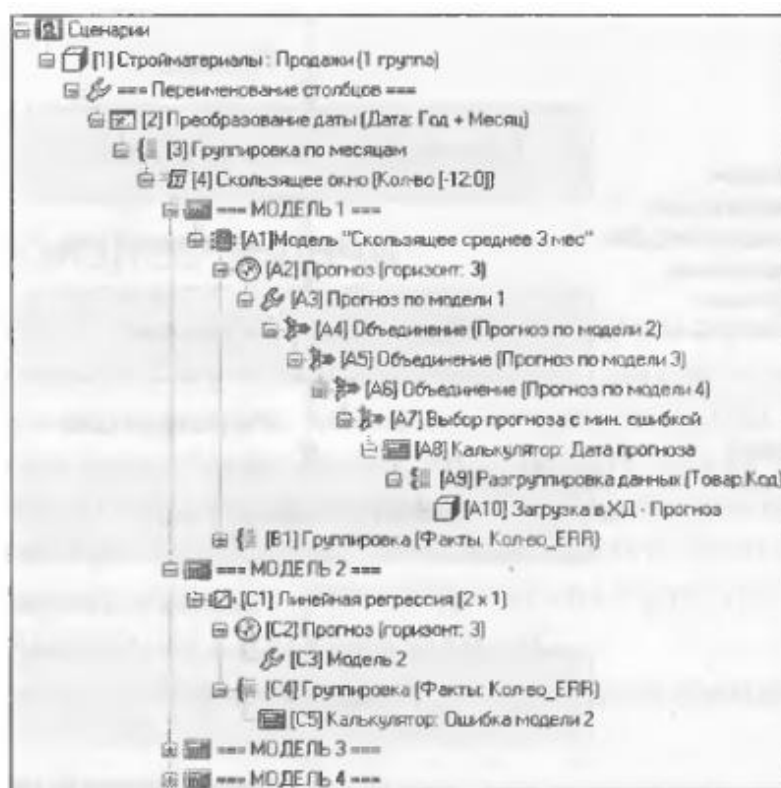


Рис. 18.2. Сценарий получения прогноза в Deductor

В первом узле из хранилища импортируются данные о продажах, после чего они группируются по месяцам и строится диаграмма, представляющая собой временной ряд продаж позиций товарной группы «Грунтовка» (рис. 18.3).

В узле под номером 4 ряд приводится к специальному виду для построения моделей при помощи скользящего окна (см. раздел 3.2). Далее сценарий делится на четыре практически идентичные ветви, в каждой из которых строится та или иная модель временного ряда.

Для «наивной» модели скользящего среднего используется обработчик Пользовательская модель, для остальных — Линейная регрессия. После каждого такого обработчика при помощи узла Прогнозирование (например, узлы A2, C1) рассчитывается прогноз на заданное число периодов.

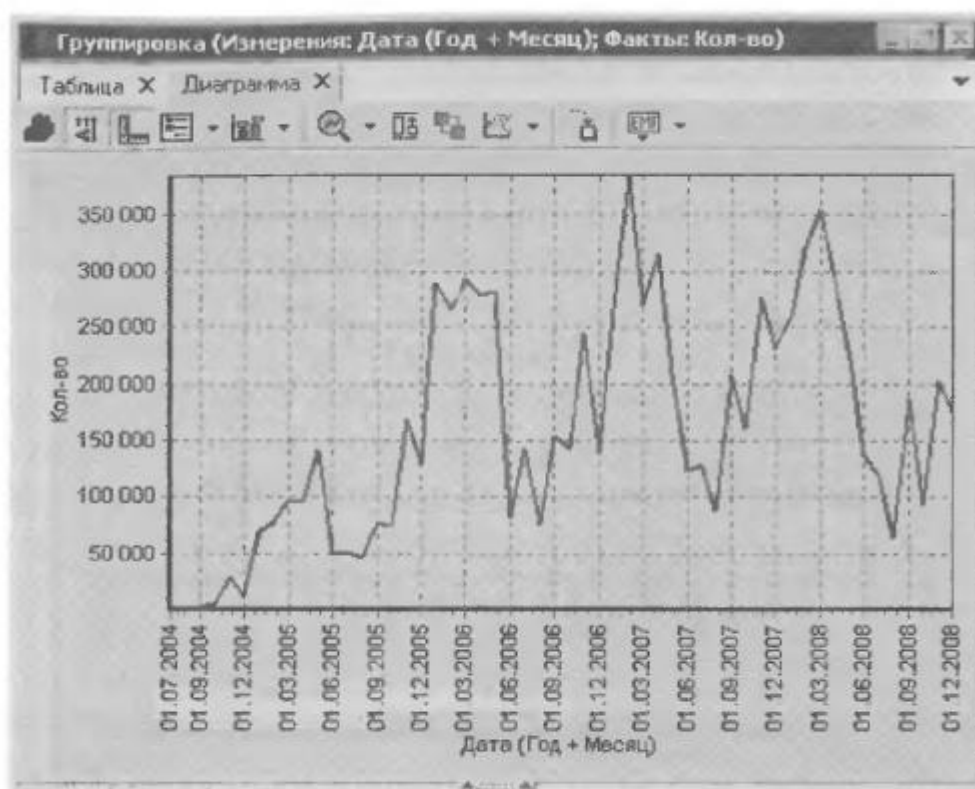


Рис. 18.3. Временной ряд продаж товаров из группы «Грунтовка»

В качестве критерия качества выберем среднеквадратическую ошибку, рассчитываемую на ретропрогнозе:

$$E = \frac{1}{n} \sum_{i=1}^n (p_i - x_i)^2,$$

где p_1, p_2, \dots, p_n — значения ряда, предсказанные моделью;
 x_1, x_2, \dots, x_n — фактические значения.

В специальной ветви сценария (рис. 18.4) сливаются в один набор данных ошибки для каждой из моделей (узлы B1 — B5) и среди них выбирается минимальная (B6 — B7).

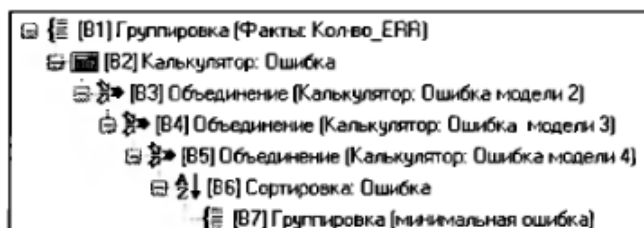


Рис. 18.4. Ветвь сценария, сравнивающая прогностические модели

В узлах A4 — A6 все прогнозы объединяются в один набор данных, и на основе информации из узла B7 о номере модели с минимальной ошибкой выбирается оптимальный прогноз. В нашем случае таким оказался прогноз на основе линейной регрессии с месяцами 1, 2 и 12 (рис. 18.5).

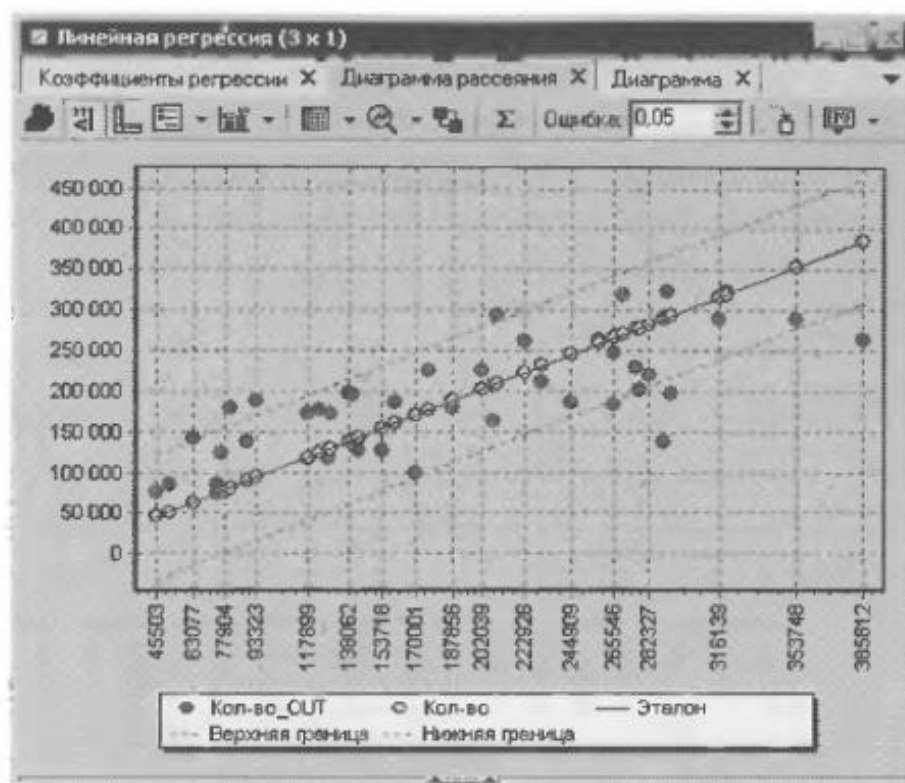


Рис. 18.5. Диаграмма рассеяния для модели 4

После получения прогнозной цифры по группе прогнозирования производится так называемая разгруппировка. Наиболее простой способ распределения квот — сделать предположение о том, что если продукция рассматриваемой прогнозной группы товаров в определенный период продавалась в известных пропорциях, то какое-то время данные товары будут продаваться в тех же пропорциях с поправкой на новый прогноз (этот период называется периодом актуальности). Для этой операции в Deductor имеется готовый обработчик Разгруппировка (рис. 18.6, см. также раздел 3.3).

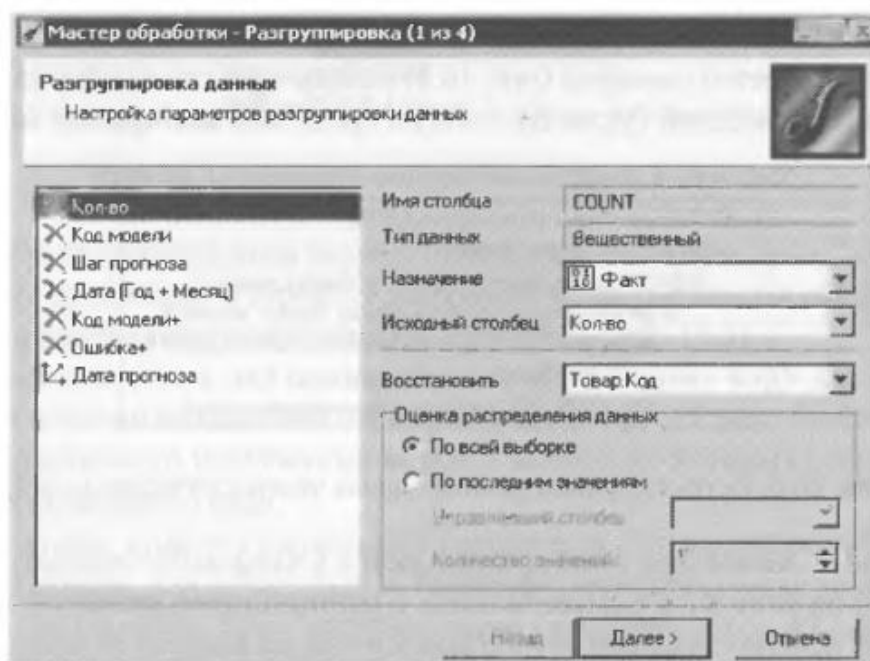


Рис. 18.6. Разгруппировка

Рассчитанный прогноз необходимо сохранить. Для этого откроем Редактор метаданных и создадим новый процесс *Прогноз* с измерениями *Дата прогноза* и *Товар.Код* и единственным фактом *Количество* (рис. 18.7).

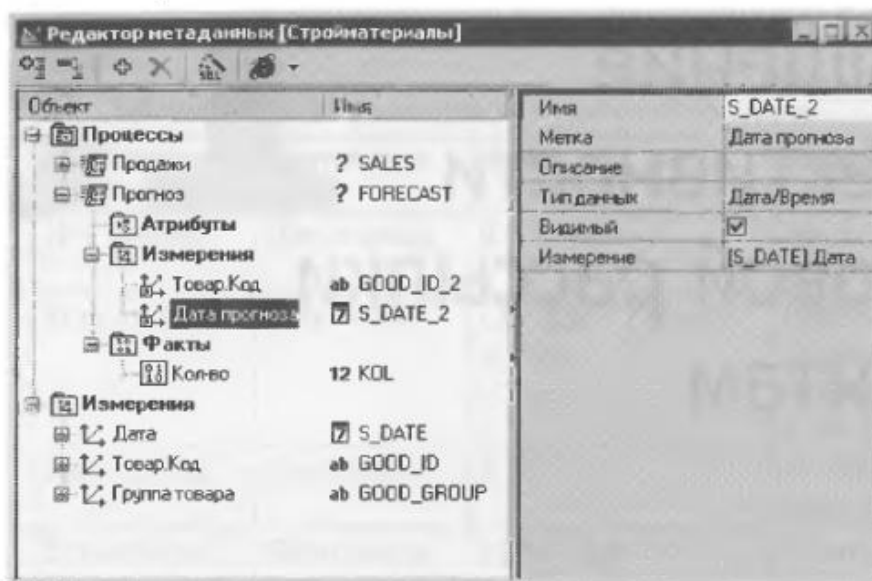
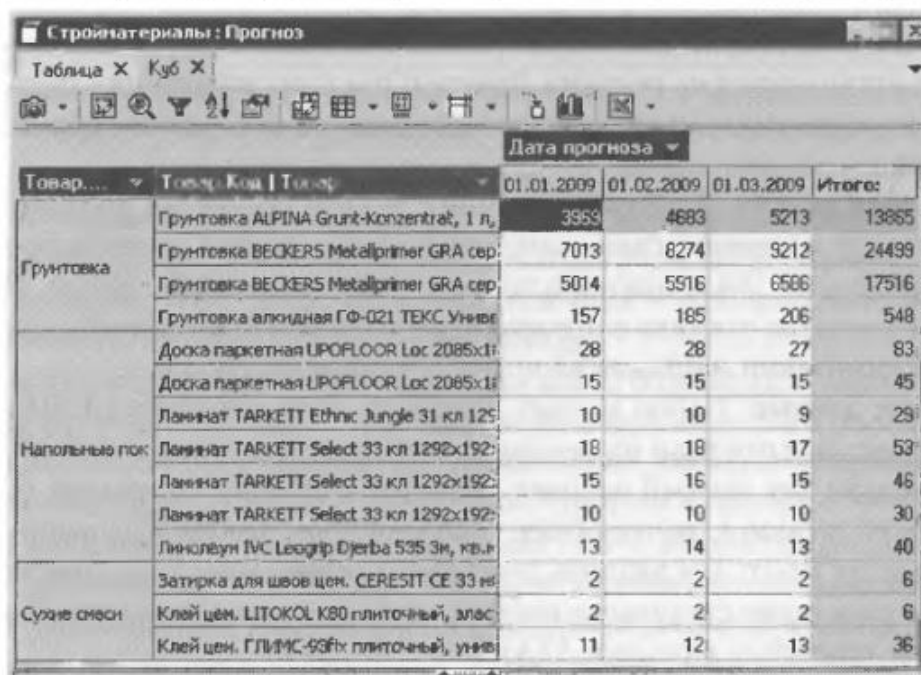


Рис. 18.7. Добавление процесса «Прогноз» в хранилище

Если запустить выполнение этого сценария в пакетном или серверном режиме для всех товарных групп, то мы получим прогноз, рассчитываемый автоматически без участия человека, который удобно просматривать в виде OLAP-куба (рис. 18.8).



		Дата прогноза ▾			
Товар...	Товар. Код Товар	01.01.2009	01.02.2009	01.03.2009	Итого:
Грунтовка	Грунтовка ALPINA Grunt-Konzentrat, 1 л,	3953	4683	5213	13865
	Грунтовка BECKERS Metallprimer GRA сер:	7013	8274	9212	24499
	Грунтовка BECKERS Metallprimer GRA сер:	5014	5516	6586	17516
	Грунтовка алкидная ГФ-021 ТЕКС Униве	157	185	206	548
Напольные пок:	Доска паркетная UPOFLOOR Loc 2085x11	28	28	27	83
	Доска паркетная UPOFLOOR Loc 2085x11	15	15	15	45
	Ламинат TARKETT Ethnic Jungle 31 кл 125	10	10	9	29
	Ламинат TARKETT Select 33 кл 1292x192	18	18	17	53
	Ламинат TARKETT Select 33 кл 1292x192	15	16	15	46
	Ламинат TARKETT Select 33 кл 1292x192	10	10	10	30
	Линолеум IVC Leogrif Djerba 535 3м, кл	13	14	13	40
Сухие смеси	Затирка для швов цен. CERESIT CE 33 не	2	2	2	6
	Клей цен. LITOKOL K80 плиточный, алас	2	2	2	6
	Клей цен. ГЛИМС-93fix плиточный, уни	11	12	13	36

Рис. 18.8. Прогноз в виде OLAP-куба

ЗАДАНИЕ.

Реализовать рассмотренный проект (см. постановку задачи). В отчет по лабораторной работе включить принт-скрины хода выполнения работы, сформировать и предоставить файлы проекта.