

Методичні рекомендації до домашньої роботи

Консолідація даних та аналітична звітність аптечної мережі

Опис бізнес-завдання

Постановка задачі. Компанія, що володіє невеликою аптечною мережею, займається роздрібним продажем лікарських препаратів. Керівництво компанії прийняло рішення про впровадження системи аналітичної OLAP-звітності, в якій їх цікавить інформація про динаміку продажів, завантаженості торгових точок, найбільші продажі товарів в різних розрізах. Так як існуюча облікова система відчуває навантаження (компанія постійно розширює свою мережу), було вирішено створити єдине консолідоване джерело - сховище даних, яке послужить базою для OLAP-звітності.

Попередньо програмісти компанії створили процедуру вивантаження даних з облікової системи в структуровані текстові файли (як проби сформована «пачка» даних за кілька місяців). Потрібно:

- спроектувати структуру реляційного сховища даних (СД);
- наповнити СД первинною інформацією;
- розробити процедури поповнення СД і контроль несуперечності даних, які містяться в сховищах даних.
- запропонувати набір OLAP -звітів.

Вихідні дані. Представлені в чотирьох файлах: Групи товарів.txt, Товари.txt, Відділи.txt, Продажі.txt.

Зобразимо послідовність виконання завдання в аналітичній платформі Deductor.

Deductor Warehouse

Сховище даних Deductor Warehouse- це спеціально організована база даних, орієнтована на вирішення завдань аналізу даних і підтримки прийняття рішень, що забезпечує максимально швидкий і зручний доступ до інформації. СД Deductor Warehouse відповідає моделі ROLAP (схема «сніжинка») і може бути розгорнута на одній з наступних СУБД: Firebird , SQL Server , Oracle (у версії Academic - тільки на Firebird). С Deductor Warehouse на базі Firebird є можливість працювати локально за допомогою динамічної бібліотеки fbclient.dll (поставляється разом з Deductor).

Сховище даних Deductor Warehouse включає в себе потоки даних, які поступають з різних джерел, і спеціальний семантичний шар, який містить метадані (дані про дані). Семантичний шар і самі дані зберігаються в одній СУБД.

Запит до сховища здійснюється безпосередньо крізь семантичний шар, який через внутрішню систему команд (приховану від користувача і аналітика) підбирає запитувану інформацію з різноманіття збережених даних. Роботу семантичного шару можна порівняти з діяльністю бібліотекаря, який на прохання читача дістає з різних полиць книги і розкриває їх на потрібних сторінках.

Всі дані в Deductor Warehouse зберігаються в структурах типу «сніжинка», де в центрі розташовані таблиці фактів, а «променями» є вимірювання, причому кожний вимір може посилатися на інший вимір. Саме ця схема найчастіше зустрічається в реляційних сховищах даних (рис. 1.1).



Рис.1.1. Структура Deductor Warehouse

У Deductor Warehouse маються такі типи об'єктів.

Вимірювання - послідовність значень одного з аналізованих параметрів. Наприклад, для параметра *Час* це послідовність календарних днів, для параметра *Region* - список міст. Кожне значення вимірювання може бути представлено координатою в багатовимірному просторі процесу, наприклад Товар, Клиент, Дата.

Атрибут - властивість вимірювання (тобто точки в просторі). Атрибут як би прихований всередині іншого виміру і допомагає користувачеві повніше описати досліджуване вимір. Атрибутами вимірювання *Товар* можуть виступати *Колір*, *Вага*, *Габарити*.

Факт - значення, відповідне виміру. Факти - це дані, які відображають сутність події. Як правило, фактами є чисельні значення, наприклад сума і кількість відвантаженого товару, знижка.

Посилання на вимір - встановлений зв'язок між двома і більше вимірами. Справа в тому, що деякі бізнес-поняття (відповідні вимірам в сховищі даних) можуть утворювати ієрархії, наприклад. Товари можуть включати *Продукти харчування* та *Лікарські препарати*, які, в свою чергу, поділяються на групи продуктів і ліків і т. д. У цьому випадку перший вимір містить посилання на друге, друге-на третє і т. д.

Процес - сукупність вимірів, фактів і атрибутів. По суті, процес і є « куб », « сніжинка ». Процес описує певну дію, наприклад продажу товару, відгрузки, надходження грошових коштів тощо.

Атрибут процесу - властивість процесу. Атрибут процесу, на відміну від виміру, не визначає координати в багатовимірному просторі. Це довідкове значення, що відноситься до процесу, наприклад *№ Накладної*, *Валюта документа* і т.д. Значення атрибута процесу, на відміну від вимірювання, не завжди може бути визначеним.

У Deductor Warehouse можливо одночасно зберігати безліч процесів («зірок» або «сніжинок»), що мають спільні вимірювання, наприклад вимір *Товар*, що фігурує в процесах *Постачання* і *Відвантаження*.

Всі завантаженні в СД дані обов'язково повинні бути визначені як вимір, атрибут або факт (рис. 1.2).



Рис.1.2. Проектирования структуры хранилища

Інформація про приналежність даних до того чи іншого типу (вимір, посилання на вимір, атрибут або факт) міститься в семантичному шарі сховища. Звернемо увагу на те, що:

- *таблиці вимірювань* містять тільки довідкову інформацію (коди, найменування т.д) й посилання на інші виміри при необхідності;
- *таблиця процесу* містить тільки факти і коди вимірювань (без їх атрибутів).

Проектирования хранилища «Фармація»

Перша підзадача – спроектувати структуру сховища нашої аптечної мережі. Всі дані представлені в чотирьох таблицях. Їх фрагменти зображені нижче.

Табл.1.1. Группы товаров (фрагмент)

Код группы	Наименование группы
33	Иммуномодуляторы
48	Общетонизирующие средства и адаптогены
50	Местные анестетики
108	Микро- и макроэлементы
198	Витамины и витаминоподобные средства
223	Желчегонные средства и препараты желчи
...	...

Табл.1.2. Товары (фрагмент)

Код товара	Наименование товара	Код группы
774	Альмагель	1
810	Иммунорм	33
824	Ревит	198
898	Настойка пустырника	48
...

Табл.1.3. Відділи

Код отдела	Наименование отдела
1	Аптека 1
2	Аптека 2
3	Аптека 3

Табл.1.4. Продажі (фрагмент)

Дата	Код отдела	Код товара	Час покупки	Количество	Сумма
01.01.2009	1	3382	15	1	293,92
01.01.2009	1	18346	17	1	22,15

Дата	Код отдела	Код товара	Час покупки	Количество	Сумма
01.01.2009	2	85600	16	1	32,16
01.01.2009	3	62535	14	4	202,72
01.01.2009	2	40315	15	3	47,52
...

Розглянемо, які дані є вимірами, які - атрибутами, а які - фактами і що являє собою процес.

У табл. 1.1 *Код групи* є виміром, а *Найменування групи* - його атрибутом.

У табл. 1.2 *Код товару* є виміром. *Найменування товару* - його атрибутом, а *Код групи* - посиланням на однойменне вимір.

У табл. 1.3 *Код відділу* є виміром, а *Найменування відділу* - його атрибутом.

У табл. 1.4 *Дата* є виміром, *Код відділу* і *Код товару*, як було сказано вище, - вимірювання, *Час покупки* - вимір, *Кількість* і *Сума* - факти. Тобто табл. 1.4 є описом процесу продажів в трьох аптеках.

При такій структурі СД ми припускаємо, що унікальність точки в просторі визначається сукупністю вимірів *Дата + Товар + Код відділу + Час покупки*. Тобто якщо в одній і тій же аптеці в один і той же день і годину буде скоєно кілька покупок, скажімо, препарату «анальгін», то в сховище даних буде відображена тільки один запис.

Взаємовідносини вимірювань, атрибутів і фактів всередині процесу продажів в трьох аптеках показано на рис. 1.3 (див. виділений рядок в табл. 1.4). В силу того, що візуально можна уявити тільки тривимірний простір, на рисунку показано взаємодію трьох вимірів (*Дата, Код відділу* і *Код товару*). У розглянутому прикладі вимірювань набагато більше. Кожне нове може бути представлено новою віссю.

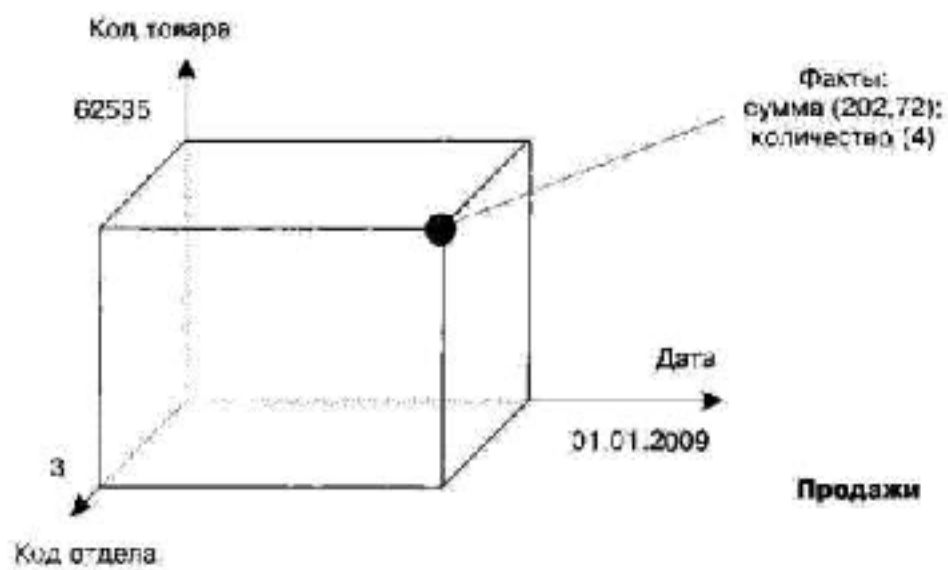


Рис.1.3. Виміри , атрибути й факти всередіні процесу продаж

Створення сховища

Запустіть програму Deductor Studio Academic. Для створення нового порожнього сховища даних або підключення до діючого перейдіть на вкладку Підключення меню Вид , натисніть правою клавішею миші й запустіть майстра підключення (рис.1.4).

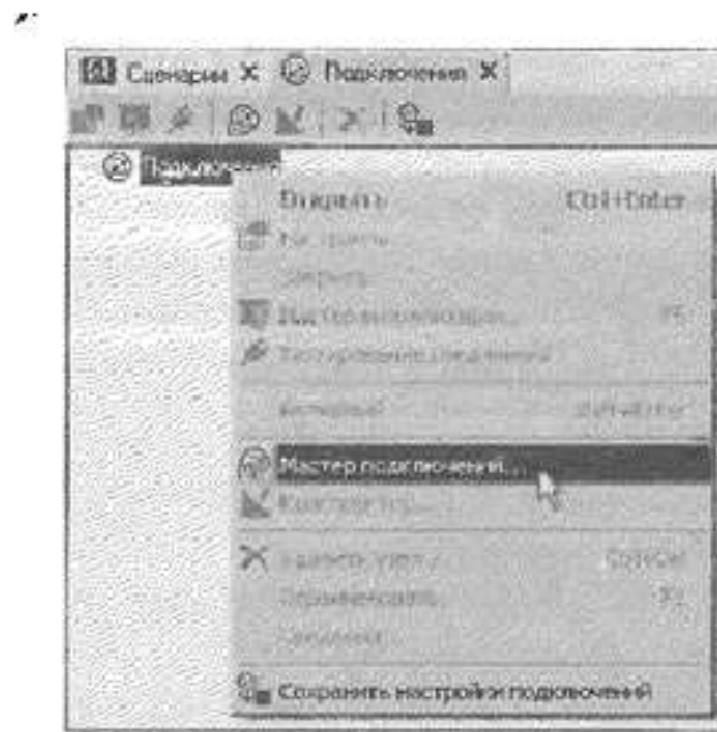


Рис.1.4. Створення (підключення) сховища даних

На першому кроці майстра необхідно обрати тип джерела (приемника) – Deductor Warehouse (рис.1.5).

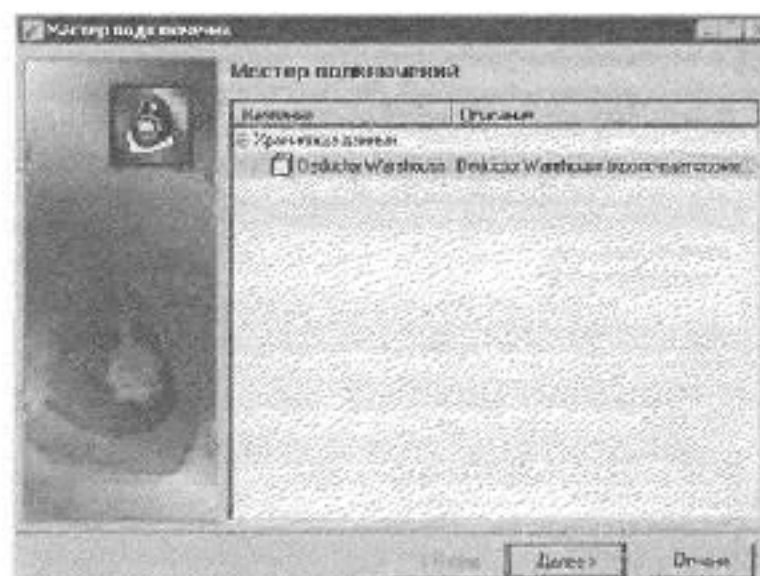


Рис.1.5. Вікно вибору типу підключення

На наступному кроці з єдиного доступного в списку типу бази даних виберіть Firebird. Задайте параметри бази даних, в якій буде створена фізична і логічна структура сховища даних (рис. 1.6) :

- база даних - D:\farma.gdb (або будь-який інший шлях);
- логін - sysdba, пароль – masterkey;
- встановіть прапорець Зберігати пароль.

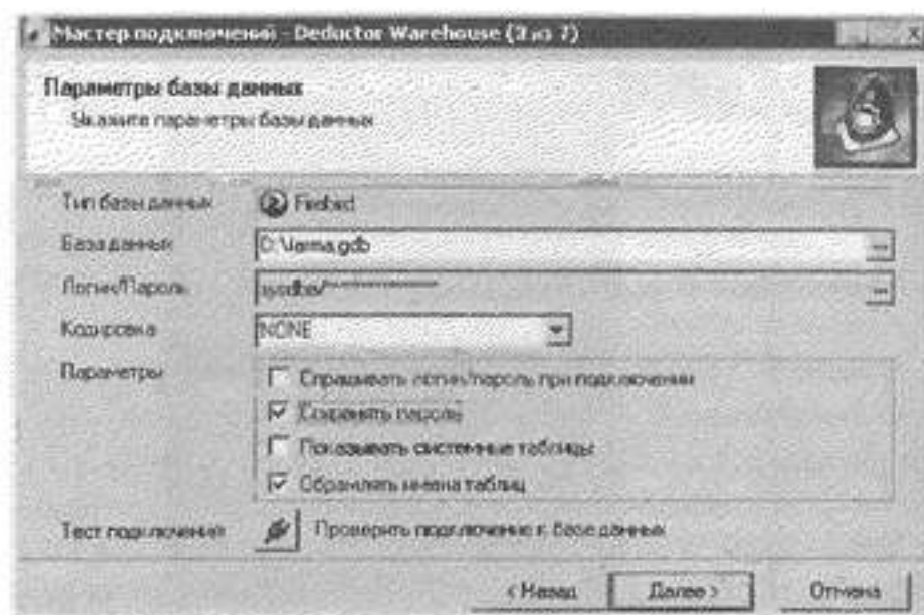


Рис.1.6. Встановлення параметрів бази даних

На наступній вкладці виберіть останню версію для роботи з СД Deductor Warehouse (попередні версії необхідні для сумісності з попередніми сховищами). Натисніть кнопку **Створити файл бази даних з необхідною структурою метаданих** і за вказаним раніше шляхом буде створений файл `farma.gdb` (з'явиться повідомлення про його успішне створення). Це і є пусте сховище даних.

Залишилося вибрати візуалізатор для підключення (в нашому випадку це Відомості та Метадані) і задати ім'я, мітку та опис нового сховища (рис. 1.7)

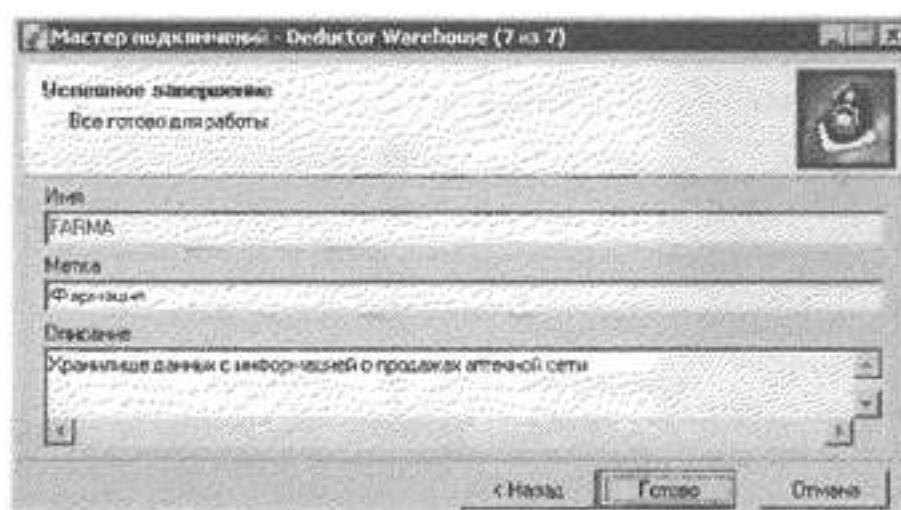


Рис.1.7. Налаштування семантики імен для вузла підключення

Ім'я сховища може бути введено тільки латинськими літерами.

Після натискання кнопки Готово на дереві вузлів підключення з'явиться мітка сховища (рис. 14.8).

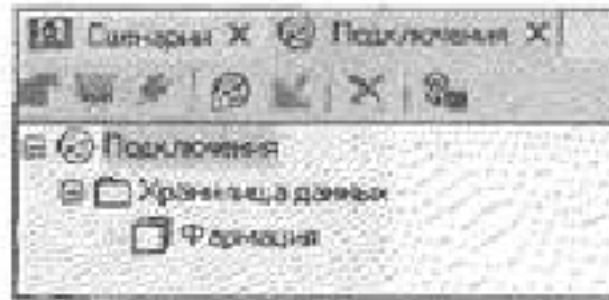







Рис.1.8. Сховище даних «Фармація»

Для перевірки доступу до нового сховища даних скористайтеся **кнопкою Тестування підключення** . Якщо через деякий час з'явиться повідомлення «Тестування з'єднання пройшло успішно», то сховище готове до роботи. Збережіть налаштування підключень, натиснувши відповідну кнопку .


Якщо з'єднання з якої-небудь причини встановити не вдалося, то буде видане повідомлення про помилку. В цьому випадку потрібно перевірити параметри підключення сховища даних і при необхідності ввести в них зміни (використовуйте для цього кнопку **Налаштувати підключення** ).

Таким чином, створено пусте сховище, в якому немає жодного об'єкта (процесу, вимірювання, факту). Раніше ми спроектували структуру сховища даних аптечної мережі. Залишилося відбити її в сховище. Для цього призначений редактор, який викликається натисканням кнопки **Відкрити редактор метаданих**  на вкладці **Підключення**.

Для переходу в режим внесення змін до структури сховища натисніть кнопку **Дозволити редагувати** . З'явиться діалогове вікно з попередженням про те, що це небезпечна операція. Виберіть вузол **Вимірювання**, клацніть правою кнопкою миші, потім натисніть кнопку **Додати** і створіть перший вимір **Код групи** з наступними параметрами:

- ім'я - *GR_ID*;
- мітка – *Група.Код*;
- тип даних - *цілий*.

Мітка - це семантична назва об'єкта сховища даних, яке бачить користувач, який працює з СД.

Проробіть аналогічні дії для створення всіх інших вимірів, взявши параметри з табл. 1.5. Зафіксуйте зміна структури СД, натиснувши кнопку **Прийняти зміни** .

Таблиця 1.5. Параметри зміни

Измерение	Имя	Метка	Тип данных
Код группы	GR_ID	Група.Код	Целый
Код товара	TV_ID	Товар.Код	Целый
Код отдела	PART_ID	Отдел.Код	Целый
Дата	S_DATE	Дата	Дата/время
Час покупки	S_HOUR	Час	Целый

В результаті структура метаданих сховища буде містити 5 вимірів (рис.1.9).

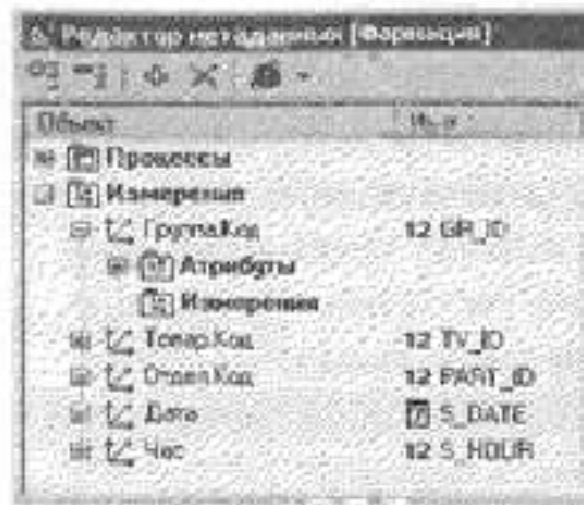



Рис.1.9. Структура метаданих сховища

До кожного виміру, крім *Дата* і *Час*, додайте текстовий атрибут (его також робиться натисканням кнопки Додати). Для вимірювання *Група.Код* це буде *Група.Найменування*, для вимірювання *Товар.Код* – *Товар.Найменування*, для виміру *Відділ.Код* - *Відділ.Найменування*. Намагайтеся не робити ім'я вимірювання зарезервованим ключовим словом мови SQL.

Кожен вимір може посилатися на інший вимір, реалізуючи тим самим ієрархію вимірів. У нашому випадку вимір *Товар.Код* посилається на *Група.Код*. Це посилання і встановіть шляхом простого додавання (посилання на вимір відображається значком ) , а ім'я посилання задайте CR_ID_1. Результат роботи ілюструє рис. 1.10.

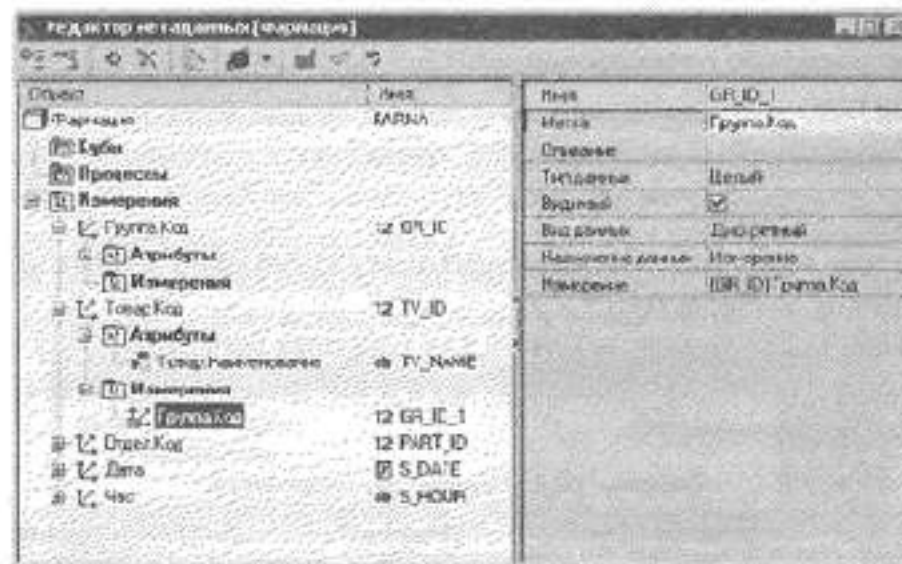


Рис.1.10. Формування посилання на вимірювання

Після того як всі виміри і посилання на вимірювання створені, приступайте до формування процесу («сніжинки»). Назвіть їх *Продажі* і додайте в нього посилання на чотири існуючих вимірювання: *Дата*, *Відділ.Код*, *Товар.Код*, *Час* (кнопка +). Крім них, в процесі беруть участь два факти: *Кількість* і *Сума*, причому перший - цілочисельний, а другий - речовинний (рис. 1.11).

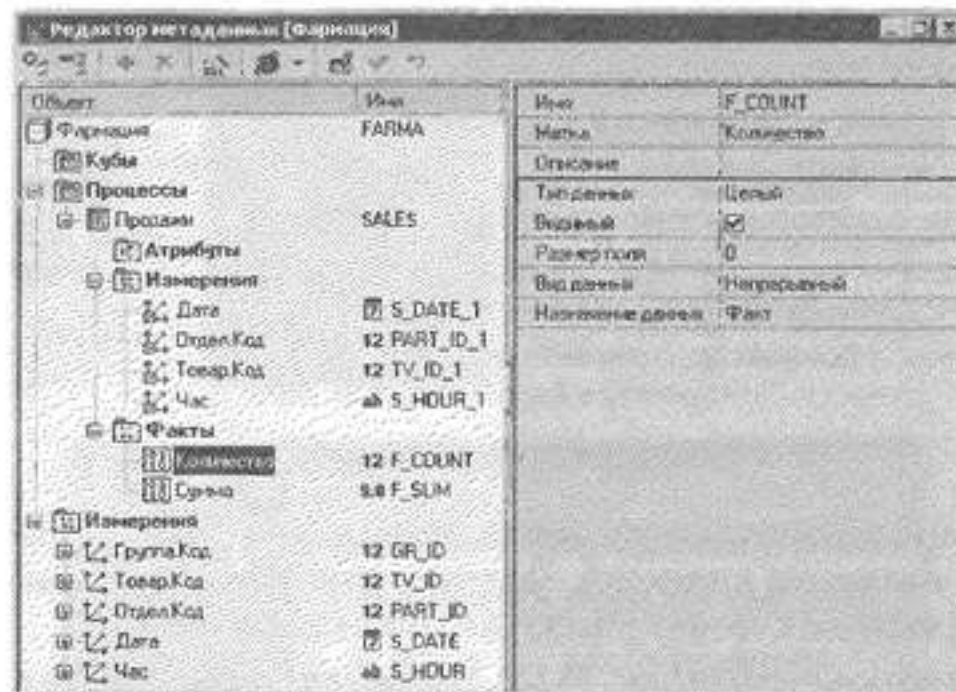


Рис.1.11. Створення метаданих процесу

На цьому проектуванні структури й метадані СД закінчені.

Наповнення сховища даних

Після створення структури сховища даних воно являє собою порожній файл з налаштованим семантичним шаром. У такому вигляді СД готове до завантаження в нього даних з зовнішніх структурованих джерел. Для цієї мети необхідно написати відповідний сценарій в Deductor Studio. Він повинен виконувати наступні функції:

- імпорт даних в Deductor Studio з бази даних, облікової системи або визначених файлів;
- опціональну преобробку даних, наприклад очищення або перетворення формату;
- завантаження даних в вимірювання і процеси сховища Deductor Warehouse.

У нашому прикладі вихідними даними для СД служать чотири текстових файли: Групи товарів.txt, Товари.txt, Відділи.txt, Продажі.txt. Тому сценарій завантаження повинен бути налаштований на використання цих файлів в якості джерел даних (рис. 1.12).

При створенні сценарію необхідно строго дотримуватися наступних правил.

Першими завантажуються всі вимірювання, що мають атрибути. Тільки після завантаження всіх вимірювань завантажуються дані в процес (-и).

Вимірювання потрібно завантажувати, починаючи з самого верхнього рівня ієрархії і спускатись нижче. Це вкрай важливо: в іншому випадку ієрархія не буде створена.

Допускається не завантажувати окремо вимірювання, які не мають атрибутів і не складаються в ієрархії вимірів. Значення таких вимірювань можна створювати під час завантаження в процес за допомогою спеціальної опції.

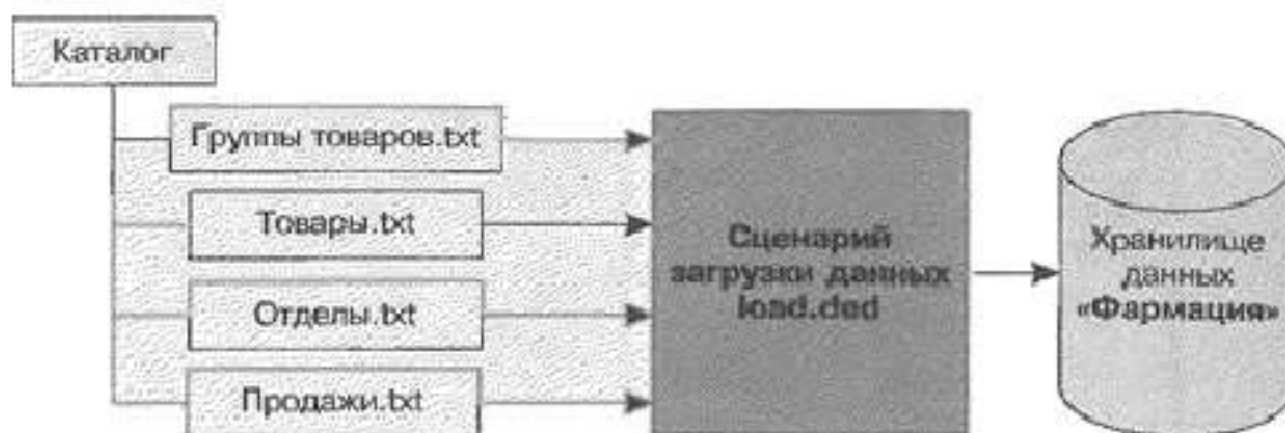


Рис.1.12. Схема сценарію завантаження

Пояснимо друге правило (рис. 1.13). Вимірювання *Група* знаходиться в ієрархії вище вимірювання *Товар*, тому послідовність завантаження вимірювань буде наступна: *Група*, *Товар*.

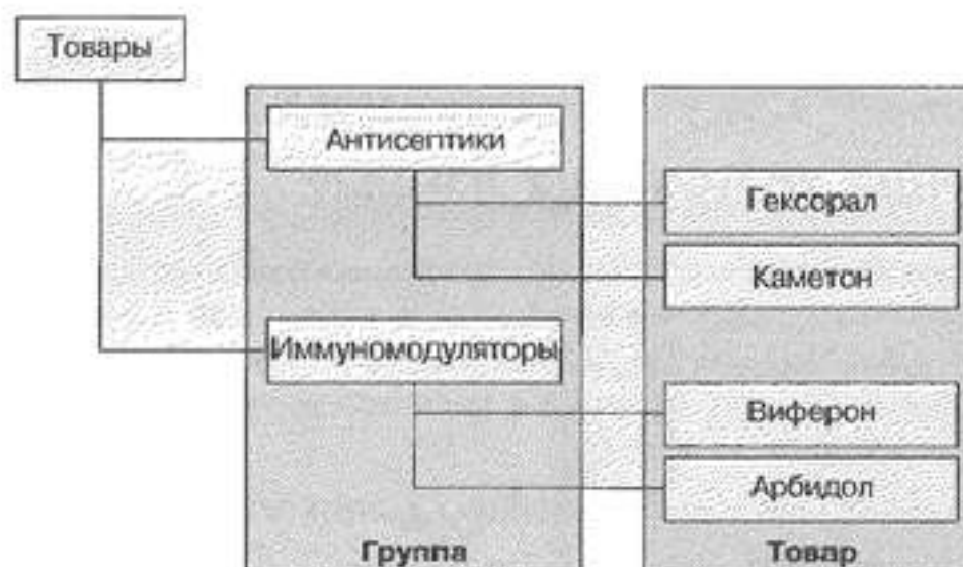
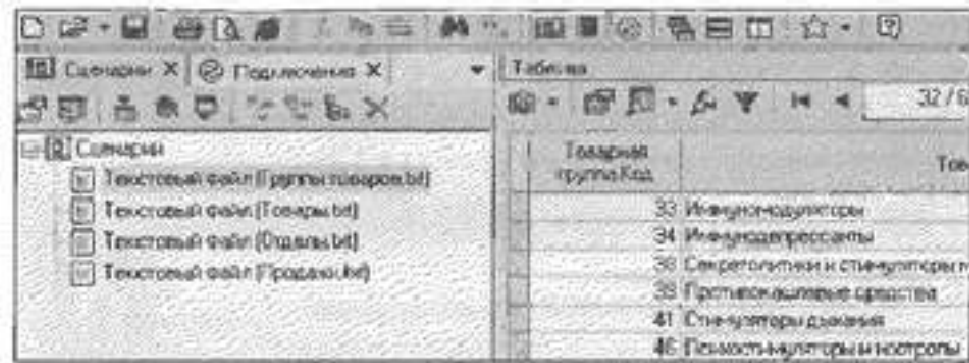


Рис.1.13. Ієрархія вимірів

Імпортуйте всі чотири текстових файлу в Deductor в тому порядку, як це показало на рис. 1.14. Для цього перейдіть на вкладку **Сценарії** і з контекстного меню або натисканням клавіші F6 викличте Майстер імпорту, виберіть тип джерела - текстовий файл і налаштуйте параметри імпорту. Послідовність створення вузлів імпорту повинна бути такою, щоб першими йшли вузли імпорту з файлів з таблицями вимірювань, і тільки в кінці - таблиця процесу Продажі.txt. Міняти порядок гілок сценарію можна за допомогою кнопок **CTRL + ↑** и **CTRL + ↓**.

Покажемо послідовність завантаження даних в вимір на прикладі першого виміру *Група.Код*. Виділивши перший вузол, викличте **Майстер експорту** (контекстне меню або клавіша F8). Зі списку типу приймачів виберіть Deductor Warehouse. (рис. 1.15).

На наступній вкладці зі списку доступних сховищ виберіть потрібне під назвою «Фармация». Далі потрібно вказати, в який саме вимір буде завантажуватися інформація. Це *Група.Код* (рис. 1.16).



1.14.Сценарій в Deductor

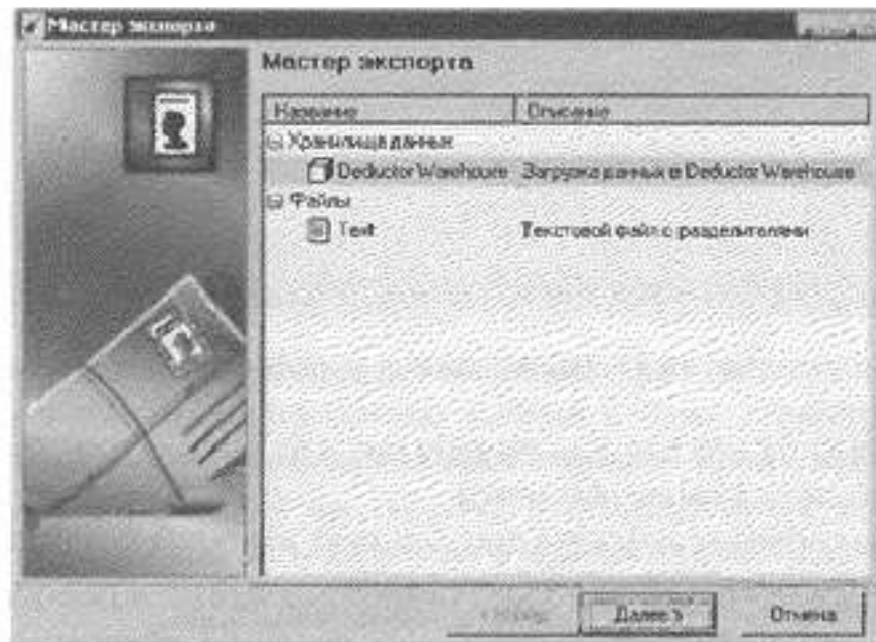


Рис.1.15. Экспорт в сховище даних

На наступній вкладці із списку доступних сховищ оберіть потрібне під назвою «Фармація». Далі необхідно вказати, в який саме вимір буде завантажуватись інформація. Це *Група.Код* (рис.1.16).



Рис.1.16. Вибір об'єкту для експорту

Залишилося встановити відповідність елементів об'єкта в сховище даних з полями вхідного джерела даних (тобто таблиці Групи товарів.txt). У разі, коли імена полів та (або) мітки в семантичному шарі сховища даних збігаються, робити нічого не потрібно (рис. 1.17).

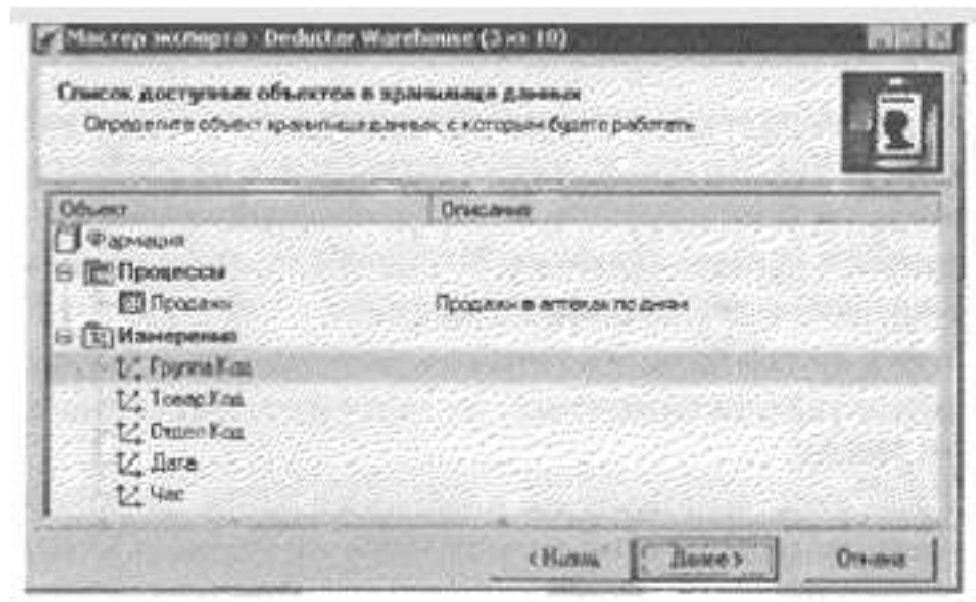


Рис.1.17. Налаштування відповідності полів

Натискання кнопки **Пуск** на наступному кроці завантажить в вимір дані. При цьому старі дані, якщо вони були, оновляться.

Проробивши аналогічні дії ще для двох вимірів – *Відділ.Код*, *Товар.Код*, отримаємо наступний сценарій (рис. 1.18).

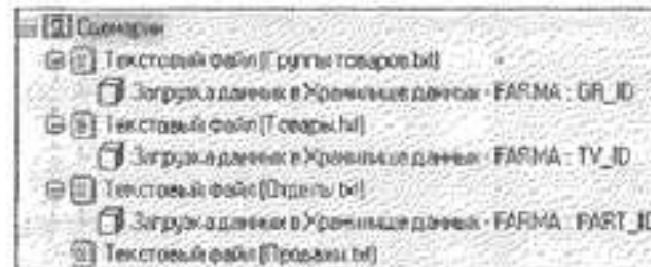


Рис.1.18. Незавершений сценарій завантаження даних в СД

Завантаження вимірювань на цьому закінчується, незважаючи на те що залишилися ще вимірювання - *Дата* і *Час*. Але, вони не мають атрибутів і не беруть участь в ієрархії, тому їх значення можна завантажити на етапі експорту в процес.

Завантажте дані в процес *Продажі*. На відміну від завантаження вимірювань, в **Майстері експорту** з'являються два специфічних кроки.

На одному з них потрібно задати параметри контролю несуперечності даних в сховище - вказати вимірювання, по яким слід видаляти дані зі сховища (рис. 1.19).

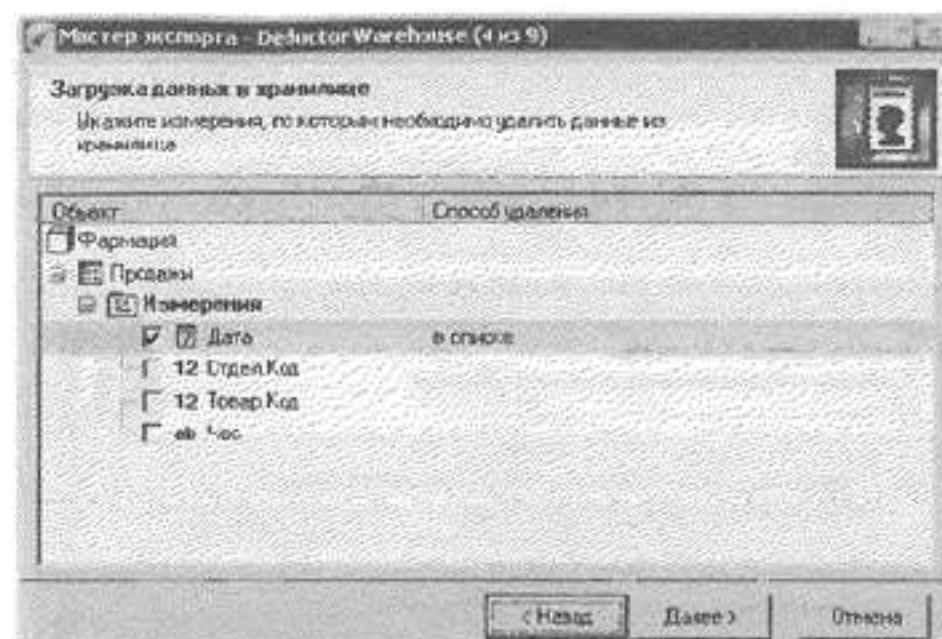


Рис.1.19. Параметри для контролю несуперечності інформації

Вибирається дія, яка виконується в ситуації, коли в процес завантажуються інформація, яка збігається зі значенням з кількох вимірів. Може бути два варіанти: видалити

старі дані і завантажити нові або заборонити видалення і залишити те, що було завантажено раніше.

Пояснимо операцію видалення на прикладі (рис. 1.20). Припустимо, в сховище є процес з двома вимірами: *Клієнт* і *Дата*. Необхідно завантажити в сховище дані про продажі за останні два дні. Якщо в наборі даних, який ми завантажуюємо, є всі відомості про продажі за ці два дні, то можна вказати: «Видаляти дані по виміру і вибрати таким виміром *Дата*. Програма визначить, що по вимірюванню *Дата* у вихідних даних всього два значення, а потім видалить зі сховища в процесі *Продажі* всю інформацію за ці два дні і завантажить нову.



Рис.1.20. Ілюстрація контролю несуперечності

Подібний спосіб завантаження зручний ще й тим, що дозволяє уникнути колізій, наприклад, коли в сховище є некоректні дані за якийсь період. У такому випадку краще усі дані за цей період видалити, а після завантажити нові коректні відомості.

На останній сторінці Майстра експорту краще залишити настройки за замовчуванням (рис. 1.21).

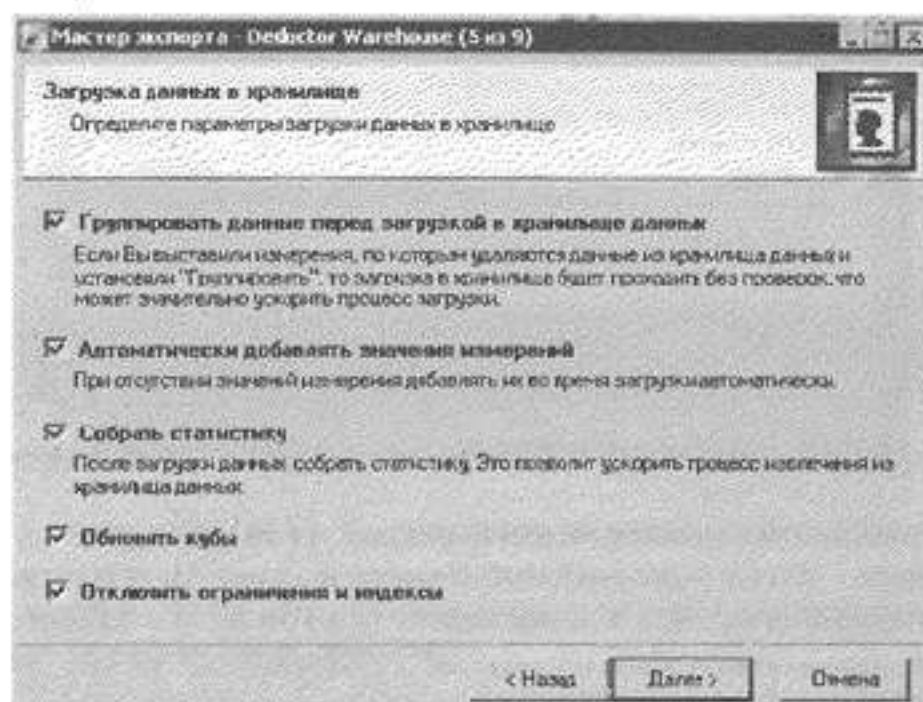


Рис.1.21. Додаткові параметри завантаження в процес

Прапорець **Автоматично додавати значення вимірювань** дозволяє «на льоту» додавати нові значення в існуючі вимірювання. Але користуватися опцією потрібно з обережністю. У разі бездумного її застосування можна дуже швидко засмітити сховище

даних, так як будь-яке значення вимірювання, навіть невірне, буде занесено як реально існуюче.

Прапорець **Групувати дані перед завантаженням в сховище даних** корисний в наступній ситуації: ви до кінця не впевнені, що сукупність вимірів процесу забезпечить унікальність точки в багатовимірному просторі, і одночасно такий рівень деталізації вас влаштовує. У нашій задачі, якщо в таблиці продажів зустрінуться два записи з однаковими значеннями вимірювань (табл. 1.6), то при відсутності встановленого прапорця **Групувати дані ... в сховище** потрапить тільки другий запис (останній, який зустрічається). Вийде, що один запис фактично загубиться, хоча потрібно підсумувати значення полів *Кількість* і *Сума*.

Табл.1.6. Випадок, при якому сукупність вимірів не дає унікальності

Дата	Код отдела	Код товара	Час покупки	Количество	Сумма
16.12.2008	1	3381	18	2	196,0
16.12.2008	1	3381	18	1	98,0

В Майстрі експорту можна задати будь який варіант агрегації даних (рис.1.22).

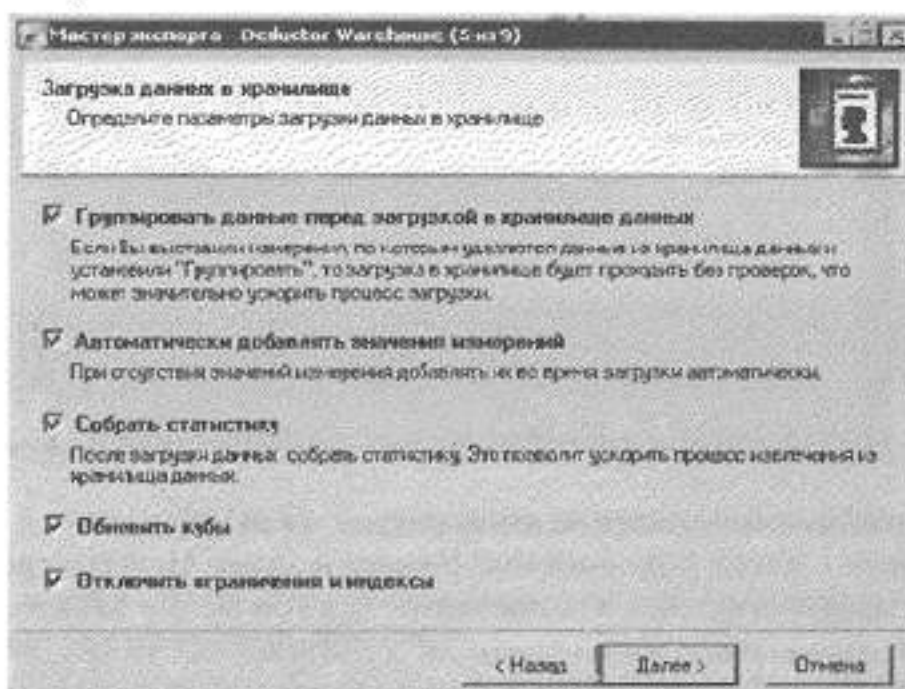


Рис.1.22. Додаткові параметри завантаження в процес

При завантаженні з встановленим прапорцем **Оновити куби** «перераховуються» всі куби, побудовані в сховище на основі даного процесу (куби «обраховуються» заздалегідь і зберігаються в окремих таблицях, тому операція імпорту з куба виконується значно швидше, ніж безпосередньо з процесу).

Включений прапорець **Відключити обмеження і індекси** корисний у разі експорту в процес великих обсягів даних (наприклад, при первинному завантаженні). Тоді під час виконання цієї дії при додаванні кожного нового рядка даних не буде витрачатися час на додаткові процедури, пов'язані з перебудовою індексів і перевіркою обмежень.

Остаточний сценарій завантаження наведено на рис. 1.23.

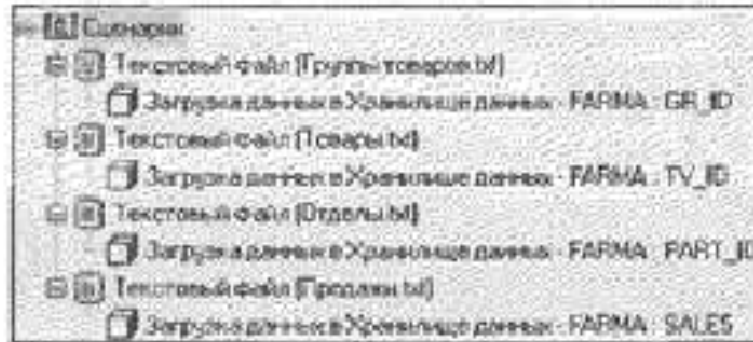


Рис.1.23. Остаточний сценарій завантаження

В результаті всіх вищеописаних дій буде:

- створено і наповнено сховище даних;
- написаний сценарій завантаження (поповнення) інформації з джерел в СД;
- продуманий контроль несуперечності даних в СД.

Зауважимо, що сценарій завантаження прив'язаний не до даних безпосередньо, а до їх структури, тобто в ньому змодельована послідовність дій, які необхідно виконати для завантаження інформації в СД: назви файлів-джерел , відповідність полів , тощо. Один раз створений сценарій згодом застосовується для поповнення сховища даних. Як правило, ці процедури проводяться по регламенту в неробочий час (наприклад, вночі) з використанням пакетного або серверного режиму.

Зрізи зі сховища даних та OLAP-куби

Процес отримання даних зі сховища здійснюється за допомогою Майстра імпорту (контекстне меню або клавіша F6). Побудуємо звіт, що відображає динаміку сум продажів за місяцями року в розрізі груп товарів і аптек. Для цього виконайте такі дії.

1. За допомогою Майстра імпорту виберіть тип джерела даних - Deductor Warehouse , на наступному кроці - сховище *Фармація*, а потім - процес *Продажі*. Далі задайте, які вимірювання і атрибути необхідно імпортувати (рис. 1.24). Зауважимо, що завдяки ієрархії всередині виміру *Товар.Код* з'явилася можливість доступу до виміру *Група.Код*.

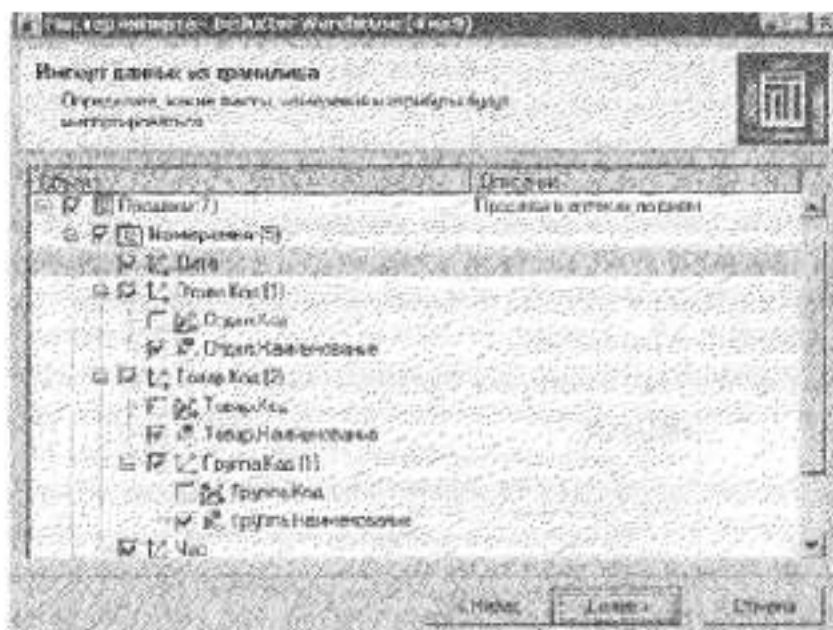


Рис.1.24. Вибір ілюстрованих вимірів та атрибутів

У цьому ж вікні задайте імпортовані факти і види їх агрегації (рис. 1.25).

У більшості випадків потрібно агрегація у вигляді суми.

2. Визначте зрізи для обраних вимірювань. Це доцільно робити при великій кількості значень виміру , так як дозволяє завантажувати з сервера, на якому розташовано СД, тільки

необхідні значення вимірів і тим самим економити час завантаження даних. Встановимо зріз по виміру *Дата* : «Всі продажі за останні 4 місяці від наявних даних» (рис. 1.26).

3. У цьому ж вікні внизу виберіть тип фільтра - **Фільтр користувача**. Це означає, що при кожному виконанні вузла імпорту буде виводитися вікно, аналогічне вікна налаштування зрізу, в якому він зможе вказати необхідні розрізи за цим виміром. Опція дозволяє будувати динамічні звіти, в яких користувачеві надається тільки цікава йому інформація, а конкретні умови фільтрації він вибирає в момент імпорту даних.

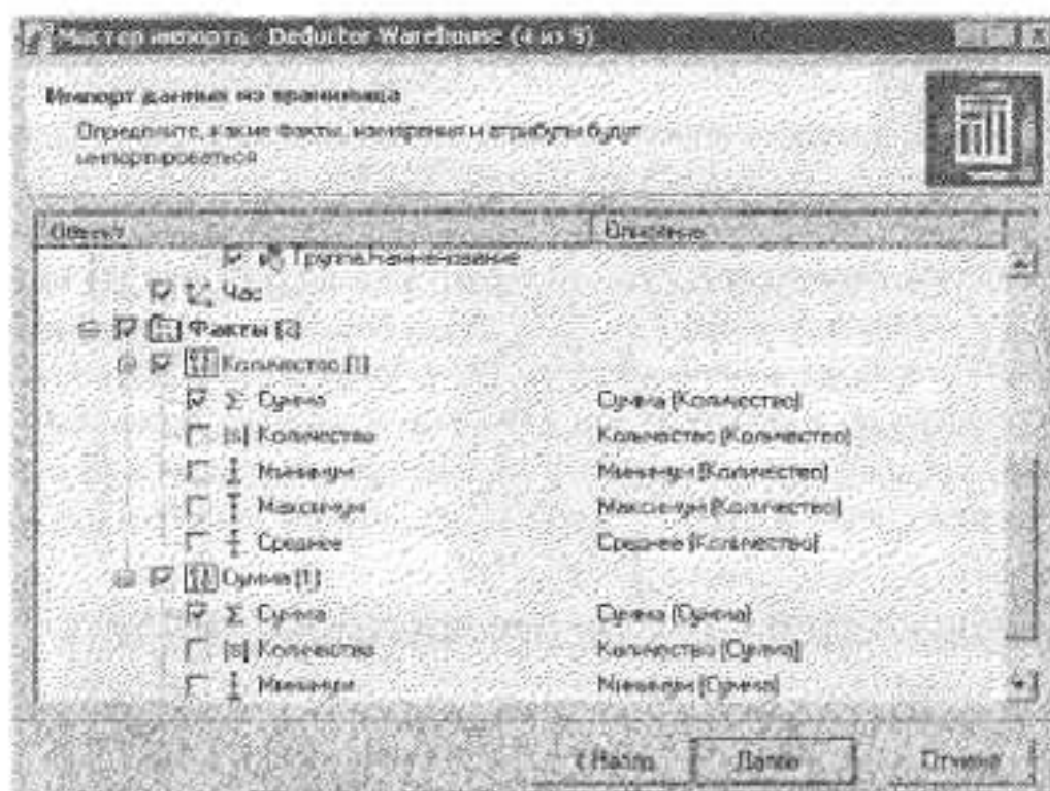


Рис.1.25. Вибір імпортованих файлів

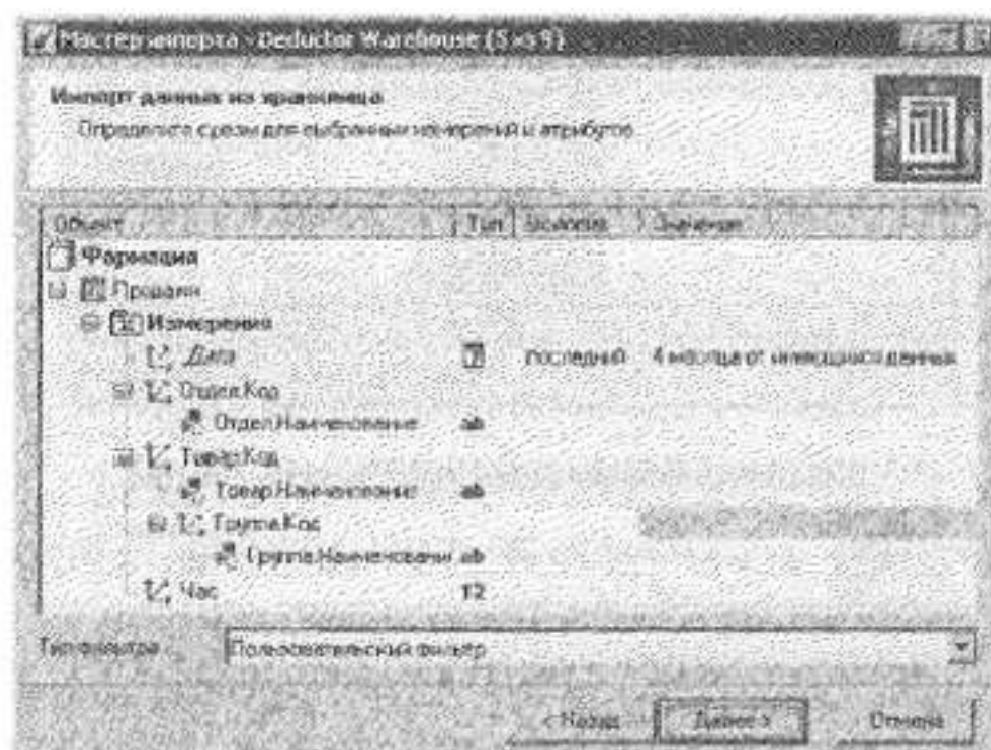


Рис.1.26. Вибір зрізу

Натисніть кнопку **Пуск**, дочекайтеся імпорту даних і виберіть візуалізатор **Таблиця**.

У вашому розпорядженні є тільки вимір *Дата*, а для побудови OLAP-звіту потрібні окремі вимірювання *Місяць* і *Рік*. Їх можна отримати з дати, застосувавши до вузла імпорту зі сховища обробник *Дата і час* (він вибирається в Майстрі обробки, який можна викликати з контекстного меню). Суть цього обробника полягає в тому, що на основі стовпця з інформацією про дату / час формуються один або декілька стовпців, в яких вказується, до

якого заданому інтервалу часу належить рядок даних. Тип інтервалу задається єдиною вкладкою налаштувань вузла в залежності від того, що ви хочете виділити з дати (рис. 1.27).

В результаті у вихідному наборі буде створено два нових строкових стовпця з мітками *Дата (Рік)* і *Дата (Місяць)*.

4. Для результуючого набору даних визначте спосіб його відображення - куб і налаштуйте призначення полів куба, тобто вкажіть виміри і факти (рис. 1.28). Для нашого звіту вимірами будуть вимірювання *Дата (Місяць)*, *Дата (Рік)*, *Відділ.Найменування* і *Група.Найменування*, а фактами - *Кількість* і *Сума* проданих товарів (з агрегацією «Сума»).

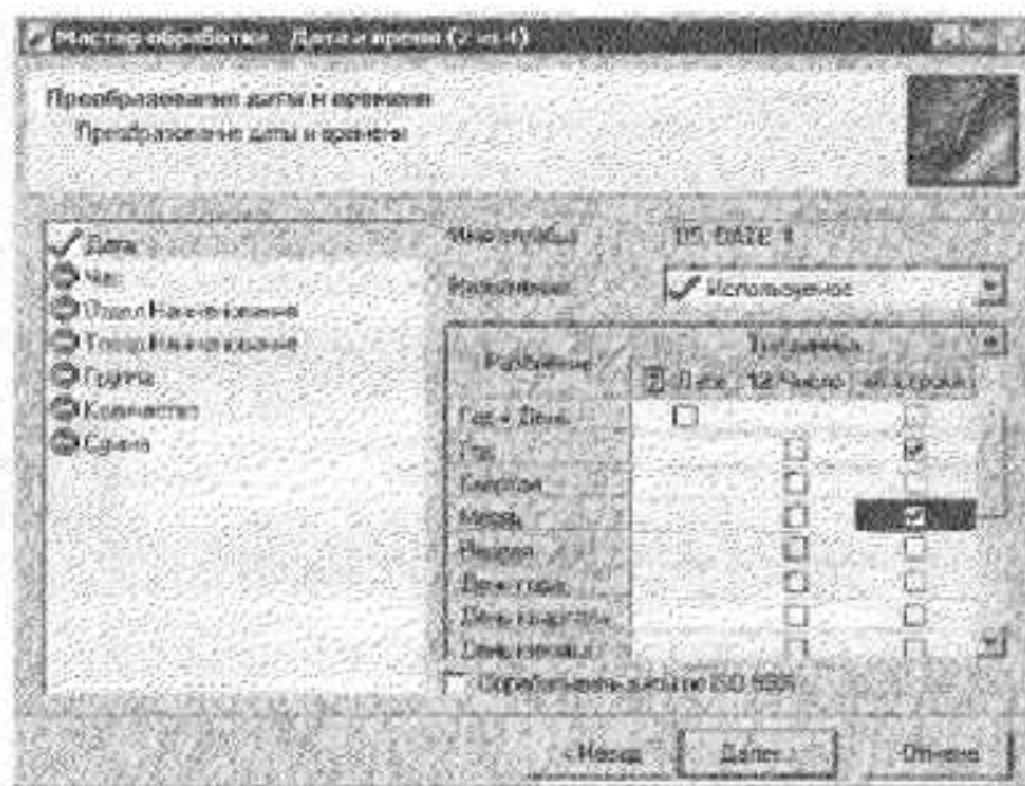


Рис.1.27. Витяг з дати місяця та року

При побудові куба інформаційне поле *Дата* не відобразиться, але буде доступно в деталізації.

5. На наступному кроці потрібно задати розміщення вимірювань по рядках / стовпцях (рис. 1.29).

6. На останньому кроці визначте, які факти потрібно відобразити в кубі на перетині вимірів, і їх агрегацію (рис. 1.30).

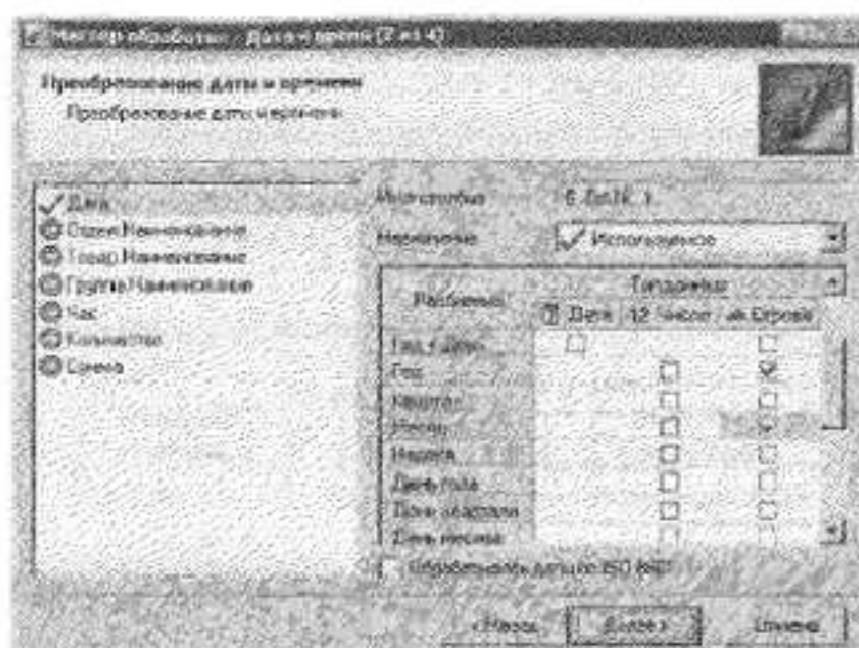



Рис.1.28. Налаштування призначень поля кубів

Фільтрація даних в кубі може здійснюватися двома способами:

- за значеннями фактів;
- за значеннями виміру.

Для фільтрації даних в кубі необхідно у спливаючому меню або на панелі інструментів натиснути кнопку Селектор ... , після чого відкриється діалогове вікно (рис. 1.33).

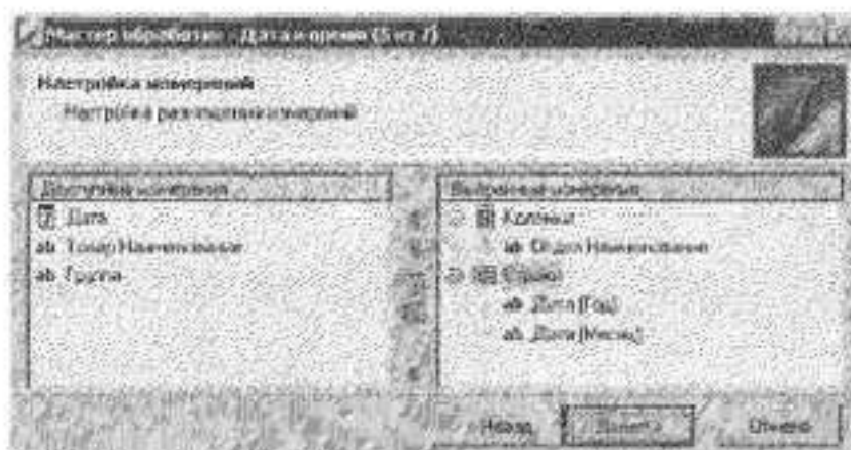


Рис.1.29. Налаштування розміщення полей куба

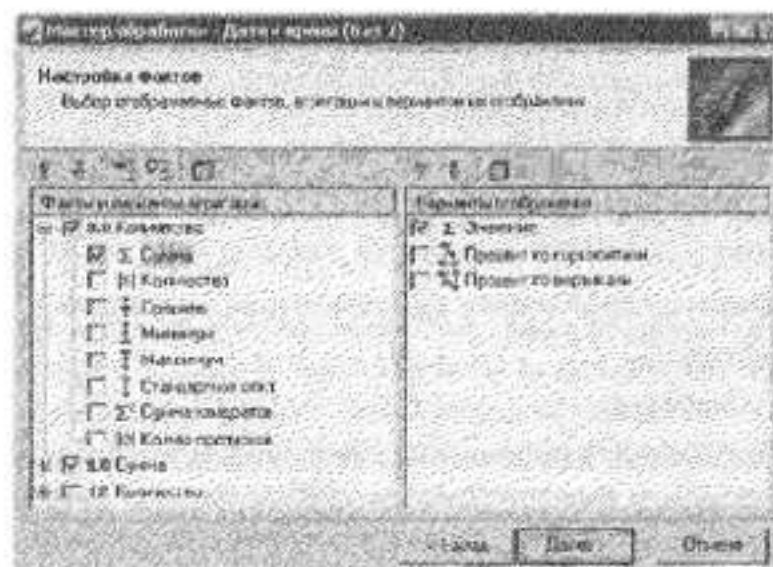


Рис.1.30. Налаштування відображення фактів

Таким чином, наш сценарій буде включати два вузли. (рис.1.31).



1.31. Фрагмент сценарію

В результаті отримуємо наступний багатомірний звіт (рис.1.32).



Рис.1.32. OLAP-звіт про продажі в розрізі місяця, року та аптеки

Нехай потрібно визначити товарні групи, які дають 80% виручки. Виберіть вимір *Група*, умова - *Частка від загального*, значення - 80 і налаштуйте в кубі одне активний вимір, додавши висновок відносних часток і відсортувавши по зменшенню (рис. 1.34).

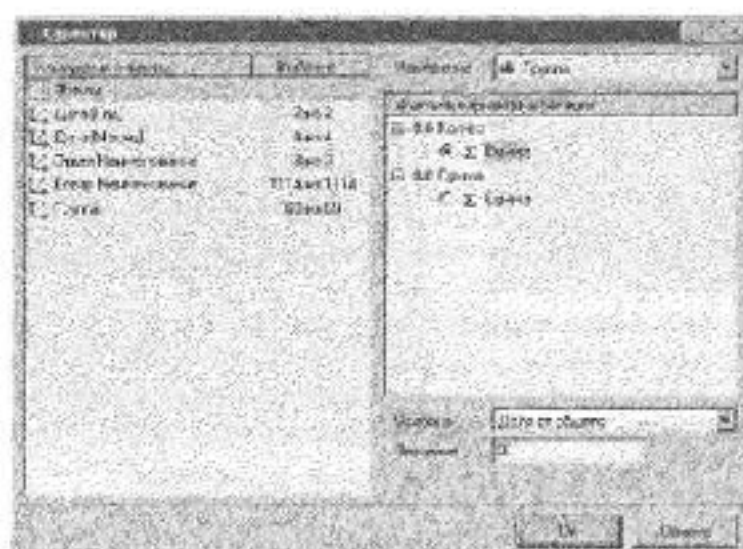


Рис.1.33. Вікно селектора

Група	Σ Частка		Σ Цена
	Σ Частка	Частка / по частоті	
Антибіотики	2 190,1	9,48%	252 482,0
Вітамінні препарати та інші добавки до їжі	2 180,1	4,17%	252 473,4
Середні цілі препарати, інгаляційні та інгаляційні	1 179,1	6,35%	24 700,8
Інші препарати та інші добавки до їжі	7 581,8	5,8%	279 150,3
Стероїди та інші препарати	7 735,8	15,83%	210 812,5
Спеціальні препарати та інші добавки до їжі	3 028,8	6,88%	198 707,6
Дітячі препарати та інші препарати	5 277,8	10,62%	25 649,5
Препарати та інші добавки до їжі	1 725,8	2,40%	14 402,0
Итого:	49 530,7%	100,00%	1 550 248,0

1.34.Куб з групами препаратів, які дають 80% виручки

Таку вибірку можна отримати з будь-якого факту. В даному прикладі це сума. Якщо вибрати вимір *Товар* і відфільтрувати але кількістю, то отримаємо лікарські препарати, що користуються найбільшим попитом.

Одночасно з кубом завжди будується крос-діаграма. Її відмінність від простої діаграми в тому, що вона однозначно відповідає поточному стану куба і при будь-яких його змінах (транспонування, обертанні) теж модифікується. Наприклад, побудуємо звіт і крос-діаграму завантаженості аптек (за кількістю проданих одиниць товарів) за останні 7 днів (рис. 1.35).

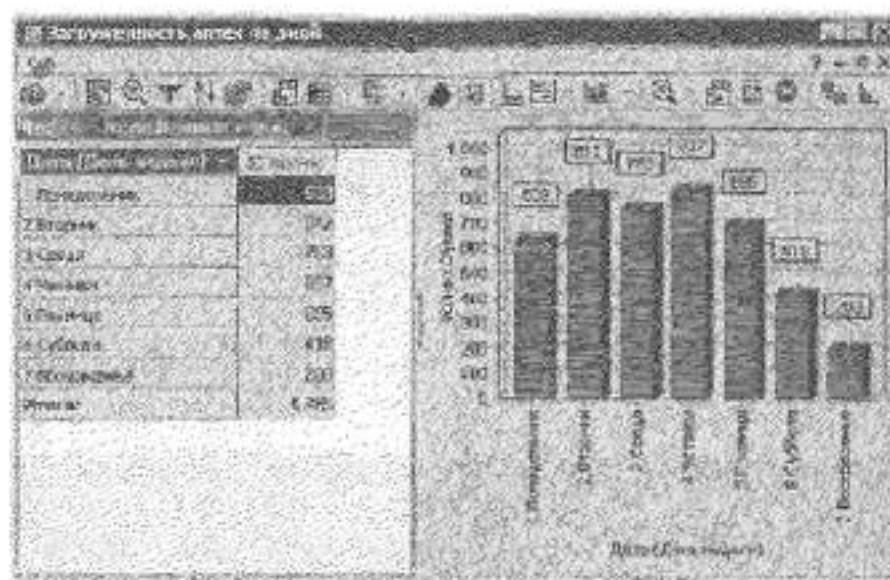


Рис.1.35

ABC-XYZ-аналіз

В аналітичній звітності дуже корисним часто виявляється ABC-XYZ-аналіз- розподіл об'єктів, наприклад товарів, клієнтів, постачальників, по стабільності продажів і прибутковості (це вже згадувалося в розділі 1.4). Якщо в якості об'єкта аналізу взяти лікарські препарати, а в якості критерію аналізу - обсяг продажів, то в результаті отримаємо:

- при ABC-аналізі продажів - найбільш і найменш користуються попитом товари;
- при XYZ-аналізі продажів - угруповання товарних позицій в залежності від стабільності продажів (кількість проданих одиниць).

Результатом ABC-аналізу є групування об'єктів за трьома категоріям:

- категорія А - найбільш цінні об'єкти, сума часток з накопичувальним підсумком яких становить перші 75% загальної суми параметрів;
- категорія В - проміжні об'єкти, такі за групою А, сума частин з накопичувальним підсумком яких становить від 75 до 90% загальної суми параметрів;
- категорія С - залишилися найменш цінні об'єкти, сума часток з накопичувальним підсумком яких становить від 90 до 100% загальної суми параметрів.

Існують і інші варіанти процентного розподілу об'єктів всередині категорій ABC, наприклад: *A - 80%, B - 15%, C-5%* або *A -50%, B - 30%, C-20%*.

Алгоритм ABC-аналізу складається з наступних кроків.

1. Визначити об'єкт аналізу: клієнт, постачальник, товарна група, товарна позиція. Для ABC-аналізу продажів це, як правило, товарна позиція.

2. Визначити параметр аналізу: обсяг продажів, дохід, кількість проданих одиниць, частота покупок товару, кількість замовлень, середній товарний запас, тощо. Для ABC-аналізу продажів це, як правило, обсяг продажів у грошовому вираженні.

3. Визначити період, за який буде проводитися ABC-аналіз, і підготувати агреговані дані по продажах за цей період. Так, якщо період дорівнює кварталу, то потрібно для кожної товарної позиції отримати суму квартальних продажів.

4. Відсортувати об'єкти аналізу в порядку зменшення значення параметра.

5. Щоб визначити приналежність обраного об'єкта до груп А, В і С, необхідно:

- розрахувати частку параметра від загальної суми параметрів обраних об'єктів;
- розрахувати цю частку з накопичувальним підсумком;
- привласнити значення груп вибраних об'єктів.

ABC-аналіз рекомендується проводити періодично, особливо якщо в асортименті є товарні позиції зі значними сезонними перепадами продажів. Періодичність аналізу визначається експертним шляхом і залежить від бізнес-процесів організації. Оптимальні періоди проведення ABC-аналізу: рік, півріччя, сезон (квартал), помісячно протягом року. Періодичне проведення ABC-аналізу дозволяє відстежити переміщення товарних позицій з групи в групу (наприклад нового товару з групи А в групу С, а також контролювати життєвий цикл товару).

Результат XYZ-аналізу - угруповання ресурсів за трьома категоріями на основі попередньо розрахованого коефіцієнта варіації:

- категорія X - об'єкти, коефіцієнт варіації яких не перевищує 10%. Характеризуються стабільною величиною споживання;
- категорія Y - об'єкти, коефіцієнт варіації яких становить 10-25%. Потреба в них залежить від відомих тенденцій, наприклад від сезонних коливань;
- категорія Z - об'єкти, коефіцієнт варіації яких перевищує 25%. Як і споживання ресурсів нерегулярне, будь-які тенденції відсутні. Існують і інші варіанти процентного розподілу об'єктів всередині категорій XYZ (наприклад, 15-50%). Коефіцієнт варіації розраховується за формулою:

$$v = \frac{\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}}{\bar{x}} \cdot 100\%$$

де x - значення параметра по оцінюваному об'єкту за 1-й період; \bar{x} - середнє значення параметра по оцінюваному об'єкту аналізу; n - число періодів.

Вдаватися до цього методу аналізу має сенс, якщо кількість аналізованих періодів більше трьох. Чим більше кількість періодів, тим більше ефективними будуть результати. При цьому сам період повинен бути не менше, ніж горизонт планування, прийнятий в компанії, інакше велика ймовірність того, що всі товари потраплять в категорію Z.

Алгоритм XYZ-аналізу також складається з декількох кроків.

1. Визначити об'єкт аналізу: клієнт, постачальник, товарна група, товарна позиція. Для XYZ-аналізу продажів це, як правило, товарна позиція.

2. Визначити параметр аналізу: обсяг продажів, дохід, кількість проданих одиниць, частота покупок товару, кількість замовлень, середній товарний запас і т.д. Для XYZ-аналізу продажів це, як правило, кількість проданих одиниць *товару* або *частота покупок*.

3. Визначити період, за який буде проводитися XYZ-аналіз, і підготувати агреговані дані по продажах за цей період.

4. Розрахувати коефіцієнт варіації.

5. Відсортувати об'єкти аналізу по зростанню значення коефіцієнта ва-РІАЦ.

6. Визначити групи X, Y і Z.

Результати ABC і XYZ-аналізу зазвичай подаються у вигляді матриці, яка складається з трьох рядків і трьох стовпців (табл. 1.7), і дозволяють дати практично повну характеристику ефективності асортиментної політики.

Таблиця 1.7. ABC-XYZ-матриця

Категорія	X	Y	Z
A	AX	AY	AZ
B	BX	BY	BZ
C	CX	CY	CZ

Прогнозування за групами такої асортиментної матриці дозволить диференціювати зусилля: одні позиції прогнозувати, інші планувати, чомусь приділяти максимум уваги, а щось розраховувати по жорстким методикам (рис. 1.36). Тому для торгових компаній матриця ABC-XYZ є ефективним інструментом при підборі стратегій управління товарними запасами.

Товарів груп AX, AY і BX слід приділяти підвищену увагу, оскільки відсутність цих товарів може призвести до серйозних збитків. У групи AX найвища ступінь надійності прогнозу, тому тут придатні складні моделі часових рядів з урахуванням сезонності, регресивні і нейромережеві моделі. Внаслідок стабільності споживання позиційні прогнози будуть забезпечувати гарну якість. Решта групи - AY і BX - можлива, слід прогнозувати за товарними групами.

Товари груп CX і CY характеризуються низькою споживчою вартістю і середньою надійністю прогнозу. Тут добре підходять нескладні розрахункові моделі, наприклад ковзне середнє з контролем страхового запасу. Взагалі, для товарів групи CX можна побудувати якісний прогноз, однак з огляду на низьку споживчу вартість витратити додатковий час на нього не потрібно.



1.36. Прогнозування по групам в ABC-XYZ-аналізі

Найменшу увагу приділяють позиціям, які потрапили в групи BZ і CZ. Якщо в цих товарах немає гострої необхідності, від них можна взагалі відмовитися. Товари групи C2 варто купувати в міру необхідності.

Сценарій ABC-аналізу продажів для сховища даних *Фармація* зображено на рис. 1.37. В якості періоду обрані останні 3 місяці; параметром обрана сума продажів в грошовому вираженні. Частки з накопичувальним підсумком розраховані в вузлі Калькулятор з використанням функції CumulativeSum()

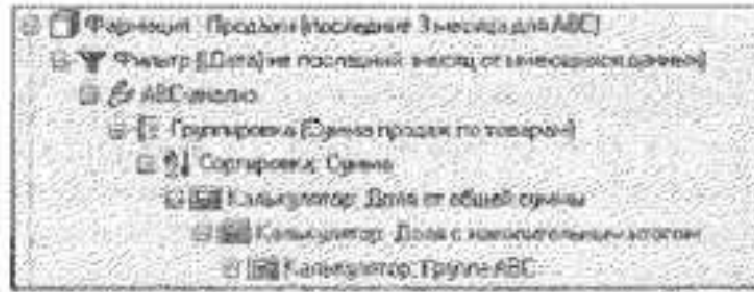


Рис.1.37. Сценарій ABC-аналізу для СД «Фармація»
 Результат ABC- аналізу в вигляді звіту с OLAP- кубом представлений на рис.1.38.

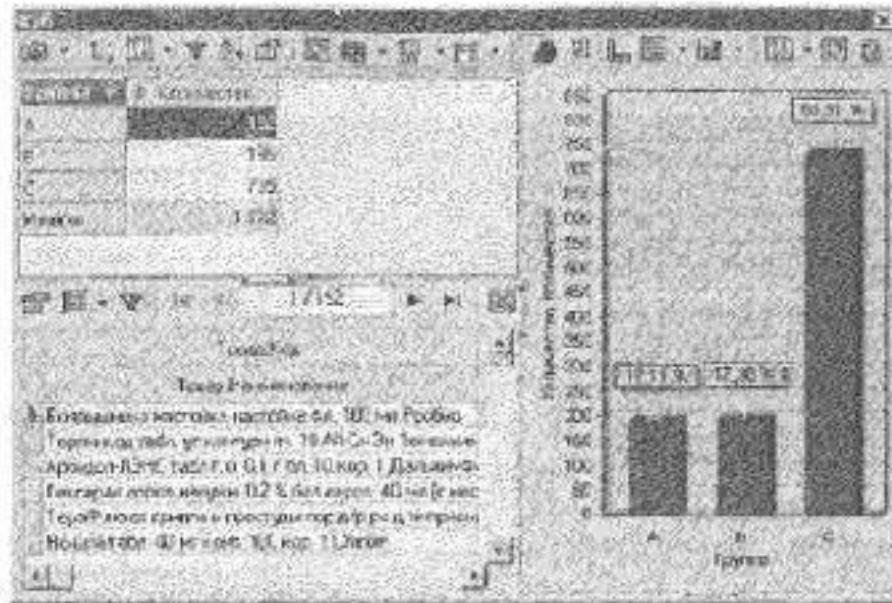


Рис.1.38. Результат ABC- аналізу в OLAP-кубі
 Сценарій XYZ- аналізу продаж для сховища даних Фармація зображений на рис.1.39.

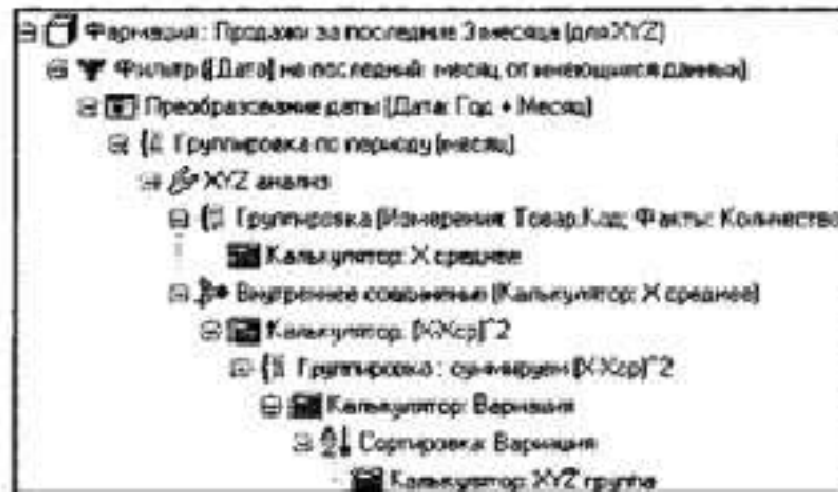


Рис.1.39. Сценарій XYZ- аналізу продажів для СД «Фармація»

Зауваження

Значення квадратного кореня у формулі для розрахунку s не що інше, як стандартне відхилення. Його (як і середня кількість продажів товару за аналізовані періоди) можна було б розрахувати, взявши функції **Середнє** і **Стандартне відхилення** у агрегацій факту в вузлі Групування. Але даний спосіб придатний тільки в разі, коли кожен товар продавався хоча б один раз в кожен період. У реальності цього може не бути. Тому в сценарії попередньо розраховується \bar{x} так, щоб кількість проданих одиниць завжди поділялося на $n = 3$. Потім воно внутрішнім сполученням «прикріплюється» до кожного товару і розраховуються суми квадратів відхилень від середнього $(x - \bar{x})$. Далі за формулою

обчислюються коефіцієнт варіації і групи XYZ. Якщо зміниться період аналізу n, то потрібно підправити формули розрахунку у відповідних вузлах.

Результат XYZ- аналізу у вигляді звіту з OLAP-кубом представлений на рис.1.40.

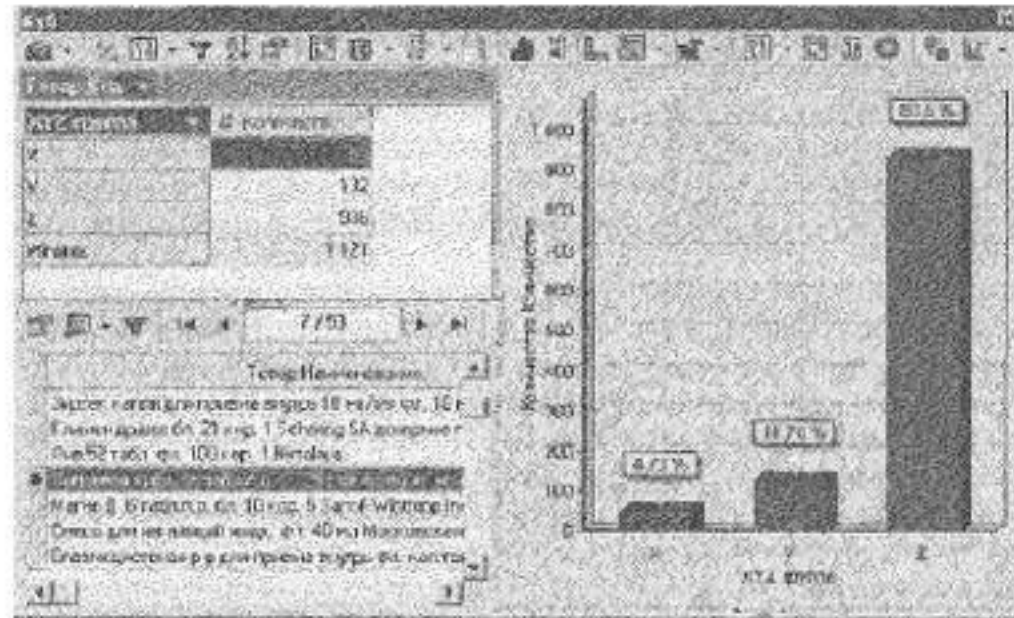


Рис.1.40. Результат XYZ- аналізу в OLAP-кубі

Тепер , щоб отримати ABC-XYZ- матрицю, залишилось додати вузол «Слияние» (дивіться розділ 3.4 «Керівництво аналітика » в поставці Deductor).

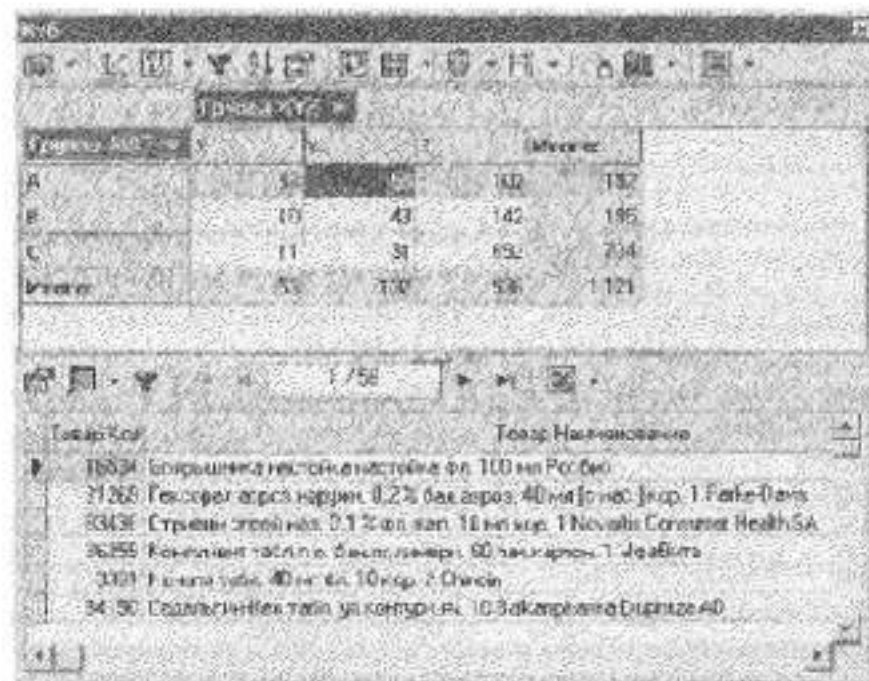


Рис.1.41. ABC-XYZ- матриця в OLAP-кубі

Налаштування звітів

В процесі роботи фахівець-аналітик виконує безліч операцій над аналізованими даними. Результати його роботи можуть бути цікавими широкому колу осіб - так званим кінцевим користувачам, яким не обов'язково вникати в послідовність дій аналітика, знати особливості математичного апарату і методів, що застосовуються при аналізі даних. Щоб

представити результати аналізу кінцевим користувачам, можна використовувати аналітичну звітність.

Аналітична звітність (звіти) - це один із засобів візуалізації і консолідації результатів аналізу даних для кінцевого користувача. Аналітична звітність забезпечує швидкий доступ до результатів аналізу, не вимагаючи від користувача навичок аналізу даних і роботи в пакеті Deductor. При роботі зі звітами користувач не бачить сценарію аналізу даних, йому доступні тільки кінцеві результати (витримки) з роботи аналітика.

Для створення аналітичної звітності в меню *Вид* виберіть пункт *Звіти* або натисніть відповідну кнопку на панелі інструментів. У робочій частині екрана з'явиться панель *Звіти*.

Звіти будуються у вигляді ієрархічного списку, кожним вузлом якого є окремий звіт або папка. Кожен вузол дерева звітності пов'язаний зі своїм вузлом в дереві сценарію. Для кожного звіту налаштовується свій спосіб відображення (таблиця, гістограма, куб, крос-діаграма, тощо). Це зручно, так як кілька звітів можуть бути пов'язані з одним вузлом дерева сценарію.

Для створення нового звіту необхідно вибрати команду *Додати вузол* з спливаючого меню або натиснути кнопку на панелі інструментів. В результаті відкриється вікно *Вибір вузла*, в якому слід виділити вузол дерева сценарію, де міститься потрібна вибірка даних, і клацнути по кнопці *Вибрати*.

В результаті виконання поданих операцій отримаємо дерево звітів (рис. 14.42).

На вкладеному до книги компакт-диску ви знайдете вихідні файли для наповнення сховища даних, готове сховище даних «Фармація» (farma.gdb), а також сценарії Deductor для завантаження і відображення звітів load.ded і reports.ded.

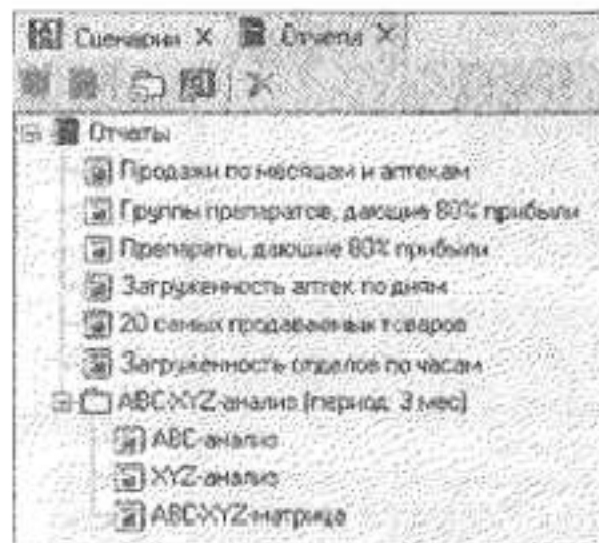


Рис.1.42.Панель звітів