

РОЗДІЛ 10

МАТЕМАТИЧНІ МЕТОДИ, МОДЕЛІ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ЕКОНОМІЦІ

УДК 330.46

Роскладка А. А.

Київський національний торговельно-економічний університет

МОДЕЛЮВАННЯ ПРОЦЕСУ КОНСОЛІДАЦІЇ ДАНИХ АГРОПРОМИСЛОВОГО ПІДПРИЄМСТВА

На основі платформи *Deductor Studio* описано задачу підвищення ефективності роботи агрофірми, для розв'язування якої проведено моделювання процесу консолідації та аналізу даних агропромислового підприємства. У результаті створено і наповнено сховище даних; написаний сценарій завантаження або поповнення інформації з джерел у сховище даних; продуманий контроль непротиворіччя даних у сховищі. За результатами побудовані багатовимірні звіти у різних розрізах та проведено *ABC-XYZ*-аналіз діяльності агрофірми.

Ключові слова: *Data Mining*, консолідація, аналіз даних, сховище даних, *ABC-XYZ*-аналіз.

Постановка проблеми. Більшість підприємств роками реєструють та записують величезні обсяги різномірної інформації (кількісної, якісної, текстової, мультимедійної тощо), що стосується всіх аспектів їхньої діяльності. Однак сподіватися, що ці масиви фактичних даних суттєво допоможуть підприємствам в ухваленні якісних стратегічних рішень без відповідних технологій оброблення інформації, – марно. Адже для того щоб сформулювати коректні посилання для ухвалення ефективних управлінських рішень, потрібні об'єктивні знання про суть, зв'язки і закономірності досліджуваної предметної області. У цьому на допомогу приходять методи асоціації та консолідації даних, що є складовими методиками *Data Mining*.

Аналіз останніх досліджень і публікацій. Деякі методи моделювання процесів у галузі агропромислового виробництва досліджені в роботах [1–3]. Різні аспекти методології *Data Mining* розглянуті у джерелах [4–13]. З програмним забезпеченням процесу консолідації даних можна ознайомитися у роботах [13–15].

Виділення не вирішених раніше частин загальної проблеми. Важливість інформації для керівництва підприємством чи бізнесом сьогодні ні в кого не викликає сумніву. Але можна мати достатньо ресурсів, у тому числі й інформаційних, але не мати успіху в бізнесі. Треба знати і вміти раціонально використовувати ці ресурси, впізнавати нові течії і тенденції, вчасно корегувати або перебудовувати свою діяльність. Без консолідації та стандартизації даних з різних джерел неможливо отримати об'єктивної інформації.

Перевагою консолідації даних є те, що цей підхід дозволяє здійснювати трансформацію значних об'ємів даних (реструктуризацію, узгодження, очищення чи агрегацію) в процесі їх передачі від первинних систем до кінцевих місць зберігання.

Метою статті є дослідження процесу консолідації та аналіз бізнес-даних діяльності агропромислового підприємства за допомогою методів *Data Mining*.

Виклад основного матеріалу. Аналіз інформації є невід'ємною частиною ведення бізнесу і одним з важливих чинників підвищення його конкурентоспроможності. При цьому в переважній більшості випадків аналіз зводиться до використання одних і тих самих базових механізмів. Вони є універсальними і застосовуються до багатьох предметних областей, завдяки чому є можливість

створення уніфікованої програмної платформи, в якій реалізовані основні механізми аналізу.

Основою для створення закінчених прикладних рішень у галузі аналізу даних може служити аналітична платформа *Deductor* [15]. Реалізовані в *Deductor* технології дозволяють на базі єдиної архітектури пройти всі етапи побудови аналітичної системи – від консолідації даних до побудови моделей і візуалізації отриманих результатів.

Процес обробки даних діяльності агропромислового підприємства розіб'ємо на чотири етапи:

- 1) проектування структури реляційного сховища даних;
- 2) розробка процедури наповнення сховища даних та контроль їх несуперечливості;
- 3) побудова системи *OLAP*-звітів підприємства;
- 4) використання *ABC-XYZ*-аналізу.

Проектування структури реляційного сховища даних. Для проектування структури сховища даних агрофірми необхідно представити вихідну інформацію, яка зберігається в базі даних, у вигляді таблиць. Дані таблиці знаходяться в окремих текстових файлах з розширенням *«.txt»*. Нижче наведено фрагменти таблиць (табл. 1–4), які містять початкові дані для заданої економічної системи агрофірми.

Таблиця 1
Групи товарів агрофірми (фрагмент)

Код групи	Назва групи
10	Жито озиме
20	Овес
30	Пшениця озима
35	Пшениця яра
...	...

Таблиця 2
Асортимент товарів агрофірми (фрагмент)

Код товару	Назва товару	Код групи
100	Жито озиме «Первісток»	10
101	Жито озиме «Слобожанець»	10
...
300	Пшениця озима «Альяс»	30
301	Пшениця озима «Астет»	30
...
500	Ячмінь «Аспект»	50
502	Ячмінь «Взірець»	50
...

Таблиця 3
Відділи реалізації агрофірми (фрагмент)

Код відділу	Назва відділу
1	Склад 1
2	Склад 2
...	...

Таблиця 4
Реалізація товарів агрофірми за 2012–2014 рр. (фрагмент)

Дата	Код товару	Код відділу	Час покупки, год.	Кількість, т	Сума, грн.
11.01.2012	100_WR	1	13	50	14000
11.02.2012	506_BA	1	12	60	19200
11.03.2012	102_WR	1	14	80	24000
...

У представленій системі таблиць визначається, які з даних є вимірами, які – атрибутами, а які – фактами, і що представлятиме собою процес:

- у таблиці 1 *Код групи* є виміром, а *Назва групи* – його атрибутом;
- у таблиці 2 *Код товару* є виміром, *Назва товару* – його атрибутом, а *Код групи* – посиланням на однойменний вимір;
- у таблиці 3 *Код відділу* є виміром, а *Назва відділу* – його атрибутом;
- у таблиці 4 *Дата* є виміром, *Код відділу*, *Код товару* і *Код групи* – виміри. *Час покупки* – вимір, *Кількість* і *Сума* – факти.

При такій структурі сховища даних передбачається, що унікальність точки в просторі визначається сукупністю вимірів *Дата* + *Товар* + *Код відділу* + *Час покупки*. Тобто якщо з одного і того ж складу в один і той самий день і час буде здійснено декілька відвантажень сільськогосподарських культур, то в сховищі даних буде відображений лише один запис.

Для того щоб відобразити структуру сховища даних агрофірми, необхідно використати *Редактор метаданих* і створити перший вимір *Код групи* з наступними параметрами: ім'я – *GR_ID*; мітка – *Група.Код*; тип даних – *цілий*. Аналогічні дії необхідно виконати для створення решти чотирьох вимірів, у результаті чого структура метаданих сховища буде складатися з таких вимірів, як *Група. Код*, *Товар. Код*, *Відділ. Код*, *Дата* і *Час* (рис. 1). До кожного виміру, крім *Дата* і *Час*, додано текстовий атрибут. Для виміру *Група. Код* це буде *Група. Найменування*, для виміру *Товар. Код* – *Товар. Найменування*, для виміру *Відділ. Код* – *Відділ. Найменування*.

Кожен вимір може посилатися на інший вимір, реалізуючи таким чином ієрархію вимірів.

Розробка процедури наповнення сховища даних та контроль їх несуперечливості. Створена структура сховища даних є порожнім файлом з налаштованим семантичним шаром. Далі потрібно створити відповідний сценарій в *Deductor Studio*, який повинен виконувати наступні функції:

- імпорт даних в *Deductor Studio* із бази даних, облікової системи або наперед визначених файлів;
- опціональну попередню обробку даних (очистку, перетворення формату тощо);
- завантаження даних у виміри та процеси сховища *Deductor Warehouse*.

Початковими даними для сховища служать чотири текстових файли: *Товари.txt*, *Групи товарів.txt*, *Відділи.txt*, *Реалізація.txt*.

Сценарій завантаження, що налаштований на використання текстових файлів в якості джерел даних, представлений на рис. 2.

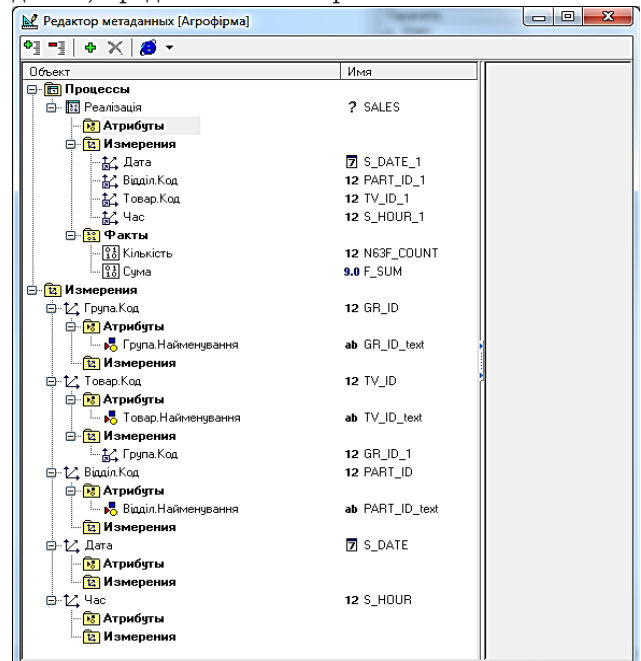


Рис. 1. Проектування структури метаданих у системі Deductor

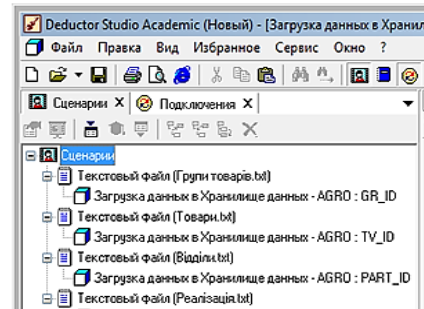


Рис. 2. Сценарій завантаження даних у сховище даних

Далі потрібно задати параметри непротиріччя даних у сховищі – вказати у вікні *Майстер експорту* виміри, за якими слід видалити дані зі сховища. У даному випадку таким виміром є *Дата* (рис. 3).

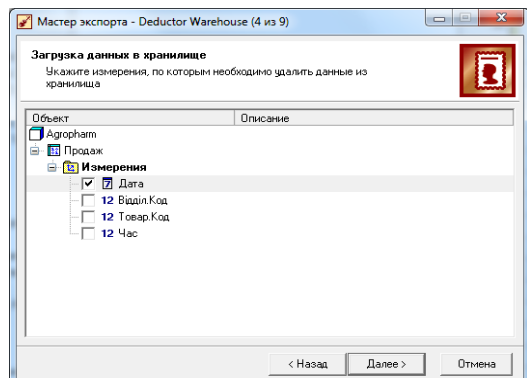


Рис. 3. Параметри для контролю непротиріччя даних

Побудова системи OLAP-звітів підприємства. Для того щоб побудувати OLAP-звіт, який відображає динаміку суми реалізації за місяцями року у розрізі груп товарів та відділів агропромислового підприємства, визначаються зрізи для обраних вимірів. Це має зміст при великій кількості значень

виміру, оскільки дозволяє завантажувати дані із сервера, на якому розміщено сховище даних.

Задаються також імпортуєчі факти і види їх агрегування, тобто об'єднання частин в одне ціле. У більшості випадків необхідне агрегування у вигляді суми.

Для результуючого набору даних можна визначити спосіб його відображення – куб та (або) крос-діаграму. Її відмінність від звичайної діаграми у тому, що вона однозначно відповідає поточному стану куба, який показує кількість реалізованих товарів, і при довірливих його змінах (транспонуванні, обертанні) також модифікується. У результаті проведених операцій одержується наступний багатовимірний звіт про реалізацію товарів та крос-діаграма (рис. 4).

З діаграми (рис. 4) можна спостерігати, що пік реалізації досягається у вересні і становить 1950 т, що складає 10,62% від загального обсягу реалізованих сільськогосподарських культур за рік. Агрофірма за даний місяць отримала 1259,740 тис. грн. Натомість найменша кількість продажів спостерігається в січні і становить 590 т, що у грошовому еквіваленті складає 288540 тис. грн.

За допомогою *Майстра візуалізації* побудована діаграма, яка відображає кількість отриманої виручки за 2014 р. від реалізації товарів (рис. 5). З даної діаграми можна зробити висновок, що досліджуване підприємство за період аналізу отримувало виручку в розмірі більше 125 тис. грн. у травні, липні, вересні та листопаді 2014 р.

Використання *ABC-XYZ* аналізу для виявлення найбільш цінних товарів. В аналітичній платформі *Deductor* є можливість застосувати метод *ABC-XYZ*-аналізу для розподілу об'єктів за стабільністю продажів і доходності. У якості аналізу береться певний товар, а в якості критерію аналізу – обсяги його продажу. У результаті – при *ABC*-аналізі отримується групування товарних позицій за ступенем впливу на загальний результат – обсяг продажу у грошовому вираженні (рис. 6).

Результатом *ABC*-аналізу є групування сільськогосподарських культур за трьома категоріями:

- до категорії А потрапило 20 найбільш цінних культур, сума часток з накопичувальним підсумком яких становить 75% загальної суми реалізації;

- до категорії В віднесено 14 найменувань сільськогосподарських культур, які є проміжними, такими, що слідує за групою А, накопичувальна частка яких становить від 75 до 90% загальної суми параметрів;

- до категорії С потрапило 19 найменувань товарів – найменш цінні з тих, що залишилися, сума часток з накопичувальним підсумком яких становить від 90 до 100% загальної суми параметрів.

При *XYZ*-аналізі одержується групування товарних позицій залежно від стабільності продажу – кількість проданих одиниць (рис. 7).

Результатом *XYZ*-аналізу є групування культур за трьома категоріями на основі розрахованого коефіцієнта варіації.

Використання сумісного *ABC*- і *XYZ*-аналізів дозволяє підвищити ефективність системи управління товарними ресурсами; підвищити частку високоприбуткових товарів без порушення принципів асортиментної політики; виявити ключові товари та причини, що впливають на кількість товарів, які зберігаються на складі; перерозподілити зусилля персоналу залежно від кваліфікації та наявного досвіду.

Висновки і пропозиції. У статті на основі платформи *Deductor Studio* була описана задача підвищення ефективності роботи агрофірми, для розв'язування якої проведено моделювання процесу консолідації та аналізу даних агропромислового підприємства. У результаті створено і наповнено

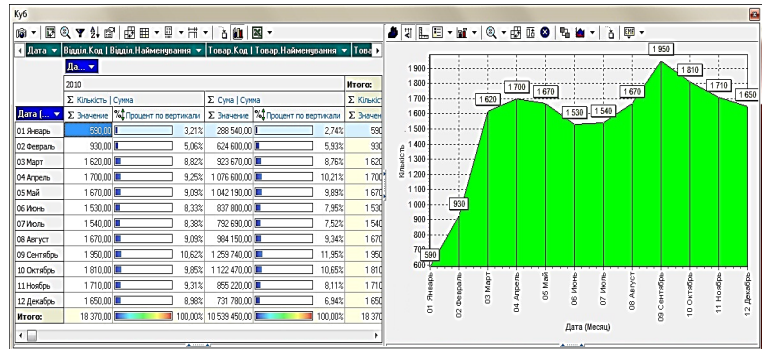


Рис. 4. Загальна кількість реалізованих товарів для усіх складів за 2012 р.

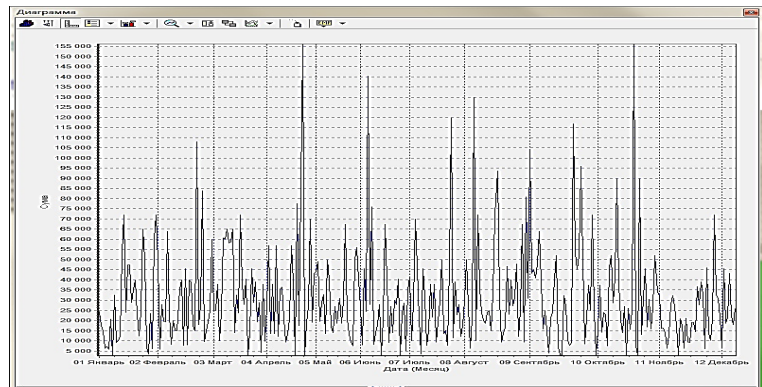


Рис. 5. Виручка агрофірми за 2014 р., грн.

Товар.Код	Товар.Найменування	Кількість	Сума	Доля	Накопительная доля	ABC-група
803	Соняшник Український F 1	310	279000	0,102972167988573	10,297	A
1302	Горюх Ефектний	330	214500	0,0791667743137957	18,214	A
713	Куркураза Явдошин 200 СВ	110	171600	0,0633334194510365	24,547	A
701	Куркураза Василея 273 СВ	130	169000	0,062338221866269	30,785	A
705	Куркураза Марлін 190 СВ	140	169000	0,0620047463157001	36,985	A
1005	Соя Скеля	180	136800	0,05048959791427844	42,034	A
1600	Люпин	190	95000	0,0350622077380447	45,54	A
800	Соняшник Заграва	120	90000	0,0332168283834108	48,862	A

Рис. 6. Визначення ABC-груп (фрагмент)

Товар.Код	Кількість	X _{середнє}	[X _{середнє}] ²	Варіація	Доля V	Накопительная доля V	XYZ-категорія
355	2	55	4050	0,688042657122685	0,109889254311965	10,369	X
713	2	55	4050	0,688042657122685	0,109889254311965	21,938	Y
105	2	65	2688,88888888889	0,460587814882307	0,0756262095256284	29,5	Y
100	2	65	2450	0,439652005114929	0,0721888545471908	36,719	Y
200	2	20	200	0,408248290463853	0,067032220795343	43,423	Y
357	2	35	450	0,349927106111883	0,0574562760681722	49,168	Y
508	3	66,6666666666667	1400	0,324037034920393	0,0532052562076895	54,489	Z
500	2	30	200	0,272165526979909	0,0446882147196895	58,959	Z
502	2	15	50	0,272165526979909	0,0446882147196895	63,426	Z

Рис. 7. Визначення XYZ-категорій товарів (фрагмент)

сховище даних; написаний сценарій завантаження або поповнення інформації з джерел у сховище даних; продуманий контроль непротиворіччя даних у сховищі. За результатами побудовані багатовимірні звіти у різних розрізах та проведено *ABC-XYZ*-аналіз діяльності агрофірми.

Для підвищення ефективності управління бізнес-процесами підприємства були запропоновані наступні рекомендації:

1) впровадити систему аналітичної *OLAP*-звітності, в якій містилася б інформація про товари та динаміку завантаженості відділів реалізації, а також найбільш популярних товарів у різних розрізах;

2) оскільки існуюча облікова система зазнає навантаження, запропоновано створити єдине консолідоване джерело – сховище даних, яке б слугувало базою для *OLAP*-звітності.

Список літератури:

1. Леньков И.И. Экономико-математическое моделирование экономических систем и процессов в сельском хозяйстве / И.И. Леньков. – М.: Дизайн-ПРО, 1997. – 304 с.
2. Вітлінський В.В. Математичне програмування: [навч. посіб.] / [В.В. Вітлінський, С.І. Наконечний, Т.О. Терещенко]. – К.: КНЕУ, 2001. – 248 с.
3. Томас Р. Количественные методы анализа хозяйственной деятельности / Р. Томас. – М.: Изд-во «Дело и Сервис», 1999. – 432 с.
4. Petra Perner. *Advances in Data Mining: Applications and Theoretical Aspects*. Springer, 2010. – 662 p.
5. Rud O. P. – *Data Mining Cookbook, Modeling Data for Marketing, Risk, and Customer Relationship Management*, 2001, John Wiley & Sons, Inc. – 429 p.
6. Барсегян А.А. и др. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP / А.А. Барсегян. – СПб.: БХВ-Петербург, 2007. – 384 с.
7. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining / [А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод]. – СПб.: БХВ-Петербург, 2004. – 331 с.
8. Дюк В. *Data Mining: [учебный курс]* / В. Дюк, А. Самойленко. – СПб.: Питер, 2001. – 368 с.
9. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining / [А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод] [Электронный ресурс]. – Режим доступа: <http://www.kodges.ru/14938-metody-i-modeli-analiza-dannykh-olap-i-data-mining.html>.
10. Ханк Д.Э., Уичерн Д.У., Райтс А.Дж. Бизнес-прогнозирование / [Д.Э. Ханк, Д.У. Уичерн, А.Дж. Райтс]. – М.: Издательский дом «Вильямс», 2003. – 457 с.
11. Чубукова И.А. Курс лекций по Data Mining [Электронный ресурс]. – Режим доступа: <http://www.kodges.ru/10257-kurs-lekcijj-intuit-po-data-mining.html>.
12. Чубукова И.А. *Data Mining: [учеб. пособ.]* / И.А. Чубукова. – М.: БИНОМ: Лаборатория знаний, 2006. – 382 с.
13. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям / Н.Б. Паклин, В.И. Орешков. – СПб.: Питер, 2009. – 624 с.
14. Seidman, C. *Data Mining with Microsoft SQL Server 2000: Technical Reference*. – Microsoft Press, 2001. – P. 360.
15. Анализ бизнес информации – основные принципы. [Электронный ресурс]. – Режим доступа: <http://www.basegroup.ru/library/methodology/>.

Роскладка А. А.

Киевский национальный торгово-экономический университет

МОДЕЛИРОВАНИЕ ПРОЦЕССА КОНСОЛИДАЦИИ ДАННЫХ АГРОПРОМЫШЛЕННОГО ПРЕДПРИЯТИЯ

Резюме

На основе платформы *Deductor Studio* описана задача повышения эффективности работы агрофирмы, для решения которой проведено моделирование процесса консолидации и анализа данных агропромышленного предприятия. В результате создано и наполнено хранилище данных; написан сценарий загрузки или пополнения информации из источников в хранилище данных; продуман контроль непротиворечивости данных в хранилище. По результатам построены многомерные отчеты в различных разрезах и проведен *ABC-XYZ*-анализ деятельности агрофирмы.

Ключевые слова: *Data Mining*, консолидация, анализ данных, хранилище данных, *ABC-XYZ*-анализ.

Roskladka A. A.

Kyiv National University of Economics and Trade

MODELLING OF THE DATA'S CONSOLIDATION PROCESS OF THE AGRO-INDUSTRIAL ENTERPRISE

Summary

In article on the basis of the *Deductor Studio* platform the problem of increase of overall performance of agro-industrial firm for which decision modeling of process of consolidation and the analysis of data of the agro-industrial enterprise is constructed. The warehouse is as a result created and filled; the scenario of loading or replenishment of information from sources in warehouse is written; control of consistency of warehouse is exercised. By results of researches multidimensional reports in various cuts are constructed and the *ABC-XYZ*-analysis of activity of agro-industrial firm are created.

Keywords: *Data Mining*, consolidation, data analysis, warehouse, *ABC-XYZ*-analysis.