

МНОГОМЕРНЫЙ АНАЛИЗ ДАННЫХ О ПРОДАЖАХ НА ОСНОВЕ ТЕХНОЛОГИИ OLAP

С. Сарсимбаева^{1, a}, В. Димитров^{2, b}

¹ АРГУ им. К. Жубанова, Акобе, Казахстан

² СУ им. Св. Кл. Охридского, София, Болгария

E-mail: ^a sarsi@mail.ru, ^b cht@fmi.uni-sofia.bg

В работе исследованы вопросы многомерного анализа данных на основе технологии семейства Business Intelligence и применение этой технологии для анализа продаж. Изучены OLAP технологии и требования к ним, способы реализации, на примере Business Intelligence. Рассмотрены основные положения технологии бизнес интеллекта в Visual Studio, внутренние интерфейсы Microsoft SQL Server. Разработана система управления базами данных для многомерного анализа данных в сфере продаж розничной сети бытовой и электронной техники, увеличивающая эффективность работы менеджеров, аналитиков компании. Представлен традиционный подход анализа данных со средствами Microsoft.

Ключевые слова: многомерная система управления базами данных, OLAP технология, OLAP куб, многомерный анализ, Business Intelligence, MS SQL Server, Microsoft Analysis Services, BI Dev Studio, базы данных.

© 2018 Сауле Мусаевна Сарсимбаева, Владимир Тодоров Димитров

1. Введение

Анализ информации был и остается одной из важнейших проблем сфер деятельности человека. Внедрение информационных технологий для анализа данных является ключевым фактором успешности компании. Для решения проблем требуется разносторонний анализ данных – по нескольким ключевым параметрам, поэтому многомерный анализ данных и применение компьютерных технологий для этого является одной из актуальных задач.

Однако, данные, которые доступны менеджерам и аналитикам компаний непосредственно из корпоративных информационных систем, не унифицированы, разрознены и в общем не подготовлены для анализа.

Существует множество средств для проведения многомерного анализа данных, но большинство из них не способны предоставить разработчикам инструменты для решения подавляющего большинства стоящих перед ним задач. При разработке приложения для многомерного анализа данных, необходимо было из всех современных технологий, выбрать наиболее подходящие для выполнения поставленных задач. Речь не идет о противопоставлении или взаимной конкуренции реляционного и многомерного подходов. Правильнее сказать, что эти два подхода взаимно дополняют друг друга. Как отметил Э. Кодд [1], реляционный подход никогда не предназначался для решения на его основе задач, требующих синтеза, анализа и консолидации данных. И изначально предполагалось, что такого рода функции должны реализовываться с помощью внешних по отношению к реляционным СУБД, инструментальных средств. Именно на решение таких задач и ориентированы многомерные СУБД. Область, где они наиболее эффективны, это хранение и обработка высоко агрегированных и стабильных во времени данных [7]. И их применение оправдано только при выполнении двух требований. Уровень агрегации данных в базе данных достаточно высок, и, соответственно, объем баз данных не очень велик, не более нескольких гигабайт. В качестве граней гиперкуба выбраны достаточно стабильные во времени Измерения, с точки зрения неизменности их взаимосвязей, и, соответственно, число несуществующих значений в ячейках гиперкуба относительно невелико. Поэтому сегодня многомерные СУБД все чаще используются не только как самостоятельный программный продукт, но и как аналитические средства переднего плана.

Системы Business Intelligence - это как раз тот класс информационных систем, который позволяет превратить сырые данные в полезную для бизнеса информацию и знания, используемые для успешного управления компанией, на основе которых можно принимать решения. Идея преобразования сырых данных в информацию востребована повсеместно, поэтому системы Business Intelligence абсолютно универсальны и применяются в различных отраслях для решения широкого круга задач.

Технологии Business Intelligence используются в таких отраслях, как: банковская деятельность, розничная торговля, страхование, телекоммуникации и другие, для которых на постоянной основе требуется аналитическая информация такого рода как анализ продаж, управление складскими запасами, оперативное управление поставками продукции, анализ эффективности маркетинговых акций, мониторинг вывода продукта на рынок, ежедневный контроль денежных потоков, анализ истории платежной дисциплины кредиторов, разработка новых продуктов, расчет сложных показателей.

Business Intelligence включает в себя процесс превращения данных в информацию и знания для поддержки принятия улучшенных и неформальных решений, аналитическую и управленческую отчетность, включающую в себя богатые возможности для построения запросов, визуализации данных, оперативный анализ информации OLAP предназначенный для менеджеров и аналитиков, которым требуется постоянное интерактивное взаимодействие с информацией, системы нетривиального анализа данных и получения знаний на основе технологии Data Mining, которые могут использоваться для решения сложно формализуемых задач, в том числе: прогнозирования, возможность выявления мошенничества, сегментацию клиентов, оценку риска и принятие решений в кредитовании, анализ покупательской корзины.

С использованием методов Business Intelligence разрабатываются программные приложения, которые являются мощным инструментом для разработки альтернативных вариантов действий, анализа последствий их применения и совершенствования навыков руководителя в столь важной области его деятельности как принятие решений

2. Разработка

На основе данной технологии разработано приложение для анализа данных магазина бытовой техники. Для разработки были использованы инструментальные средства SQL Server 2008, Visual Studio 2010 [2,3,4].

Основной задачей компонентов интеллектуальной обработки данных в SQL Server 2008 является поддержка разработки и использования систем интеллектуальной обработки данных на предприятиях любого размера и всеми сотрудниками - не только менеджерами и аналитиками, но и руководителями оперативных подразделений и партнерами. Для выполнения этой задачи SQL Server 2008 создана полноценная, интегрированная, легкая в использовании система, которая публикует данные в виде web-служб, обеспечивает высокую производительность на обычном аппаратном обеспечении и содержит множество возможностей, которые можно использовать для разработки инновационных аналитических приложений. В SQL Server 2008 входит компонент SQL Server Business Intelligence Dev Studio. Business Intelligence Dev Studio - это интегрированная среда разработки, созданная для разработчиков систем интеллектуальной обработки данных. Основанная на Visual Studio, Business Intelligence Dev Studio представляет собой мощную, интегрированную, профессиональную платформу разработки для разработчиков систем интеллектуальной обработки данных. Отладка, контроль исходного кода, разработка скриптов и кода доступны во всех компонентах приложения интеллектуальной обработки данных.

3. Анализ данных

Набор служб в SQL Server, связанных с бизнес-анализом и хранением данных называется Microsoft Analysis Services. Эти службы включают в себя службы интеграции - Integration Services и службы анализа - Analysis Services. Analysis Services, в свою очередь, включают в себя набор средств для работы с OLAP и интеллектуальным анализом данных. Главным компонентом Analysis Services является Business Intelligence Development Studio - инструмент управления, который предоставляет единую платформу разработки для Integration Services, извлечения данных Reporting Services и Analysis Services. Созданный в Visual Studio, Business Intelligence Development Studio поддерживает интегрированную платформу разработки для системных разработчиков в области бизнес-аналитики. Средства отладки, управление источниками данных, и разработка кода доступны во всех компонентах приложения бизнес-аналитики. Использование Business Intelligence Development Studio позволяет создавать и управлять многомерными кубами.

К примеру, в магазинах бытовой техники ведется постоянная работа по формированию информационной картины, используется база данных по приходу и реализации каждого отдельно взятого товара. Существует множество видов анализируемой информации, важнейшими из которых являются: данные об итогах и масштабах продаж конкретного товара определенного производителя.

Учитывая набор перечисленных выше требований, было принято решение о разработке системы на принципах технологий семейства Business Intelligence - средств анализа и обработки данных масштаба определенной торговой сети, которая позволит эффективно решать широкий круг задач обработки информации и управления. Указанные возможности реализуются за счет использования методов и средств построения хранилищ данных, OLAP-технологий [5].

4. Заключение

Авторами разработана система многомерного анализа на основе OLAP технологии с целью унификации процесса управления торговой сетью, оптимизации процесса работы менеджеров, бизнес-аналитиков розничной сети. Решена задача внедрения системы Business

intelligence в розничные сети бытовой и электронной техники для многомерного анализа данных по продажам.

Разработка приложения по многомерному анализу данных по продажам бытовой и электронной техники позволило облегчить работу менеджеров, аналитиков компании по закупке товаров на основе анализа проданных товаров в определенный период (Рисунок 1).

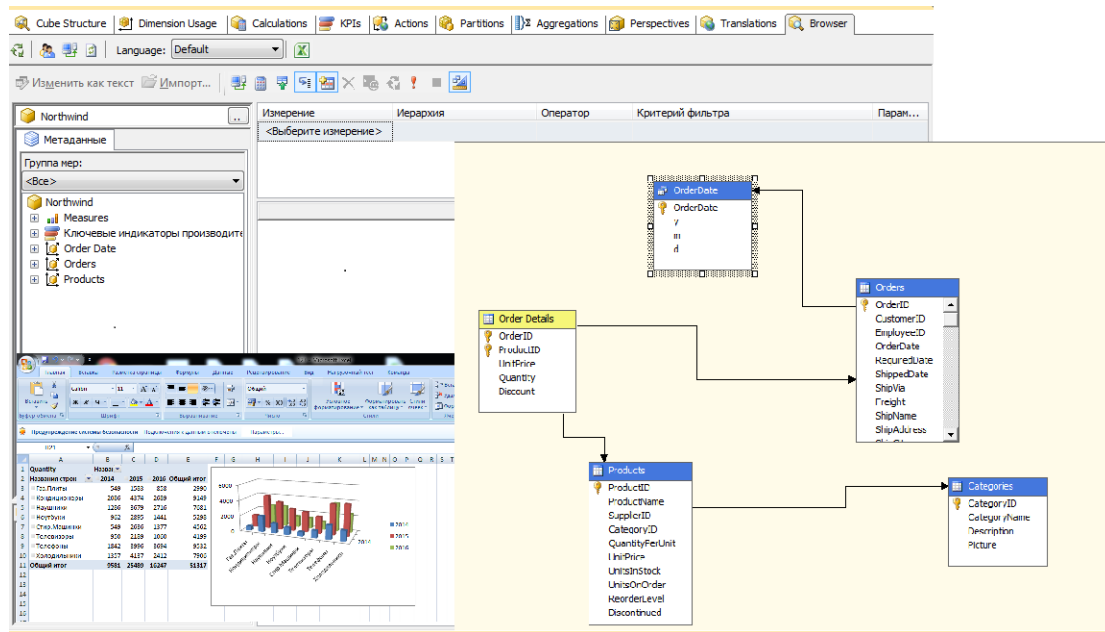


Рисунок 1. Многомерный куб и его виды через браузеры Visual Studio 2010 и MS Excel

На основе разработанного приложения был проведен анализ продаж сети магазинов бытовой и электронной техники по городу Актобе. Исходя из данных многомерного куба можем наблюдать, что наибольшей популярностью среди покупателей в 2016 году в категории кондиционеров лидерами продаж являются бренды Midea и Elenberg. (Рисунок 2.)

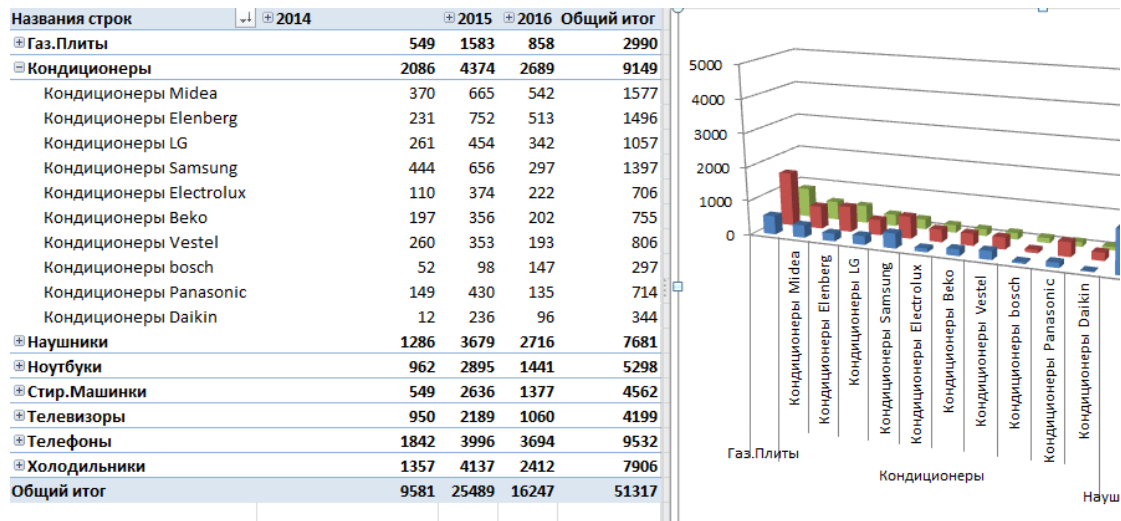


Рисунок 2. Анализ продаж кондиционеров по годам

При детальном анализе можно увидеть, что в последний месяц кондиционеры Elenberg опережают по продажам кондиционеры Midea. Исходя из этих данных при заказе товаров нужно сделать акцент на заказ кондиционеров Elenberg. (Рисунок 3).

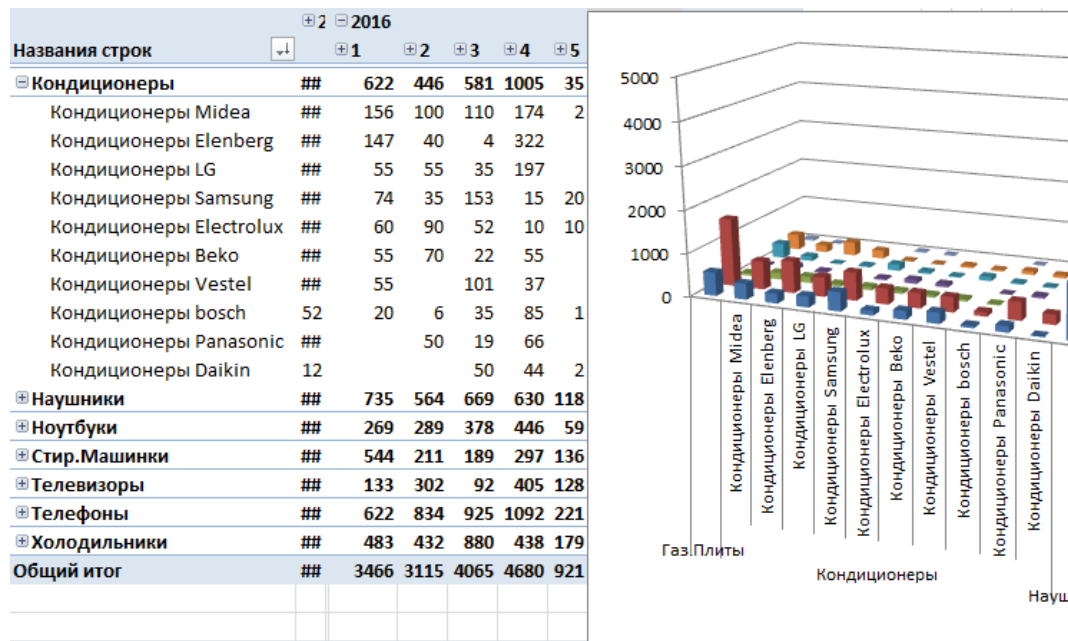


Рисунок 3. Анализ продаж кондиционеров по месяцам 2016 года

Такой анализ был проведен по товарам ноутбуки, газовые плиты, наушники и другим.

Анализ данных традиционно осуществляется специальными средствами типа выше описанных. Данные анализа загружаются в многомерные кубы для обработки. Многомерные кубы могут быть многомерными хранилищами или просто таблицы.

При анализе Big Data тоже поддерживается вышеописанный подход, но кроме того рекомендуется анализ данных на месте хранения. К сожалению пока мало таких средств.

Представленная разработка еще используется для обучения студентов и аспирантов. Аналогичные средства «на примерах» представлены в обучении баз данных в [6].

Благодарности

Проекту 80-10-143/25.04.2018 ФНИ СУ им. Св. Кл. Охридского, за частичное финансирование работ.

Список литературы

- [1] Codd, E. F., Codd, S. B. & Salley, C. T. (1993). Providing OLAP (On-Line Analytical Processing) to User-Analysts: An IT Mandate. E. F. Codd and Associates.
- [2] Ponniah, P. (2011). Data Warehousing Fundamentals for IT Professionals. John Wiley & Sons, Inc.
- [3] Rainardi, V. (2014). Building a Data Warehouse: With Examples in SQL Server. Apress.
- [4] Сарка, Д., Йеркич, Г., Лах, М. (2014) Microsoft SQL Server 2012. Реализация хранилищ данных. Учебный курс Microsoft. Русская редакция.
- [5] Миронов, В., Макарова, Е. (2012) Агрегация показателей в OLAP-кубе при сведении по зависимым измерениям. Вестник УГАТУ, т. 16. – № 3 (48), стр. 180 - 186.
- [6] Orozova, D. Todorova, M. (2017) How to Follow Modern Trends in Courses in “Databases” - Introduction of Data Mining Techniques by Example, Proc. of XI INTED, 06-08.03.2017, Valencia, 2017, pp. 8186-8194, doi:10.21125/inted.2017.1929; ISBN: 978-84-617-8491-2.
- [7] Kaloyanova, K., Naydenova I. (2011). Regular Sparsity in OLAP system, Emerging Themes in Information Systems and Organization Studies, Springer- Verlag Berlin Heidelberg, pp. 115-123

MULTIDIMENSIONAL ANALYSIS OF DATA BASED ON OLAP TECHNOLOGY

S. Sarsimbayeva^{1, a}, V. Dimitrov^{2, b}

¹ *ARSU K. Zhubanov, Actobe, Kazakhstan*

² *SU St. Kliment Ohridski, Sofia, Bulgaria*

E-mail: ^a sarsi@mail.ru, ^b cht@fmi.uni-sofia.bg

The development problems of the database management system for multidimensional data analysis based on one of the modern OLAP technology are considered in article. This article reveals the necessity of the use of tools of the family Business Intelligence, such as Microsoft Analysis Services, which is part of MS SQL Server. Based on the conducted study authors developed a software - the system of multidimensional analysis based on OLAP technology to unify distribution network management process, optimization of the process of managers, business analysts, retail network. The developed application allows you to optimize the performance of managers, analysts, network of shops selling home appliances, electronic equipment based on multidimensional analysis of items sold in a given period.

Keywords: multidimensional database management system, the OLAP technology, multidimensional analysis, Business Intelligence, MS SQL Server, Microsoft Analysis Services, BI Dev Studio environment business intelligence, database, the OLAP cube.

© 2018 Saule Musaeвна Sarsimbaeva, Vladimir Todorov Dimitrov