ОНЛАЙН КУРС:

УПРАВЛЕНИЕ ДАННЫМИ

ОНЛАЙН КУРС:

УПРАВЛЕНИЕ ДАННЫМИ

ТЕМА 6.1

Информационные хранилища и склады данных

Вопрос 1

OLAP-технология

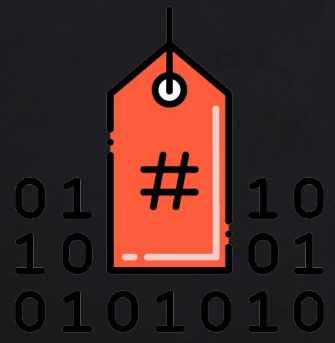
OLAP - технология

набор технологий для оперативной обработки информации, включающих динамическое построение отчётов в различных разрезах, анализ данных, мониторинг и прогнозирование ключевых показателей бизнеса

Сферы применения

~~OLAP~~

- Упрощения работы с многоцелевыми накопленными данными о деятельности корпорации в прошлом



Сферы применения

~~OLAP~~

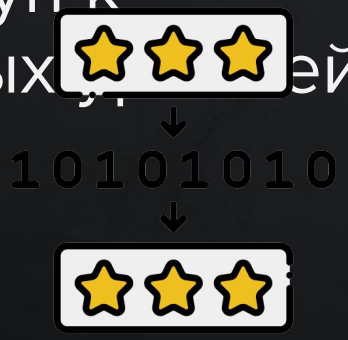
- Превращения набора количественных показателей в качественные
- Помощи аналитикам, менеджерам и управляющим в формировании своего собственного видения данных

Ключевые характеристики OLAP

- **Быстрота -** означает, что система должна обеспечивать выдачу большинства ответов пользователям в пределах приблизительно пяти секунд

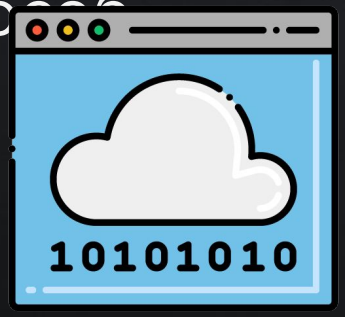
Ключевые характеристики OLAP

- **Разделимость** - означает, что система дает возможность выполнять все требования защиты данных и реализовывать распределенный и одновременный доступ к данным для различных уровней пользователей



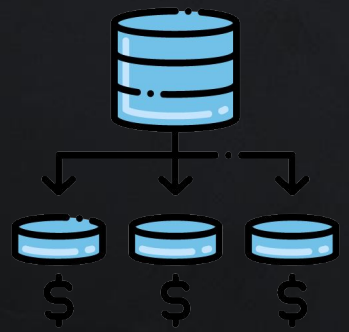
Ключевые характеристики OLAP

- **Многомерность** - ключевое требование. Система должна обеспечить многомерное концептуальное представление данных, включая полную поддержку для иерархий и множественных иерархий, поскольку это определяет наиболее логичный спо
анализировать бизнес



Ключевые характеристики OLAP

- **Доступность** - необходимая информация должна быть получена там, где она необходима, независимо от ее объема и места хранения



Ключевые характеристики OLAP

- **Анализ** - означает, что система может справляться с любым логическим и статистическим анализом, характерным для данного приложения, и обеспечивает его сохранение в виде, доступном для конечного пользователя

Этапы построения OLAP-системы

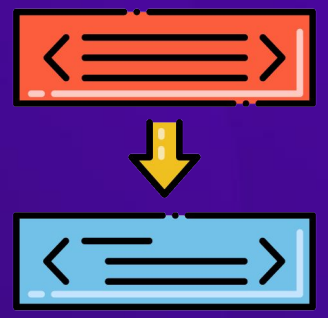
- Хранилище данных
- Многомерные OLAP-кубы
- Система ключевых показателей (KPI)
- Интеллектуальный анализ данных
- Доступ через web-интерфейс

Хранилище данных

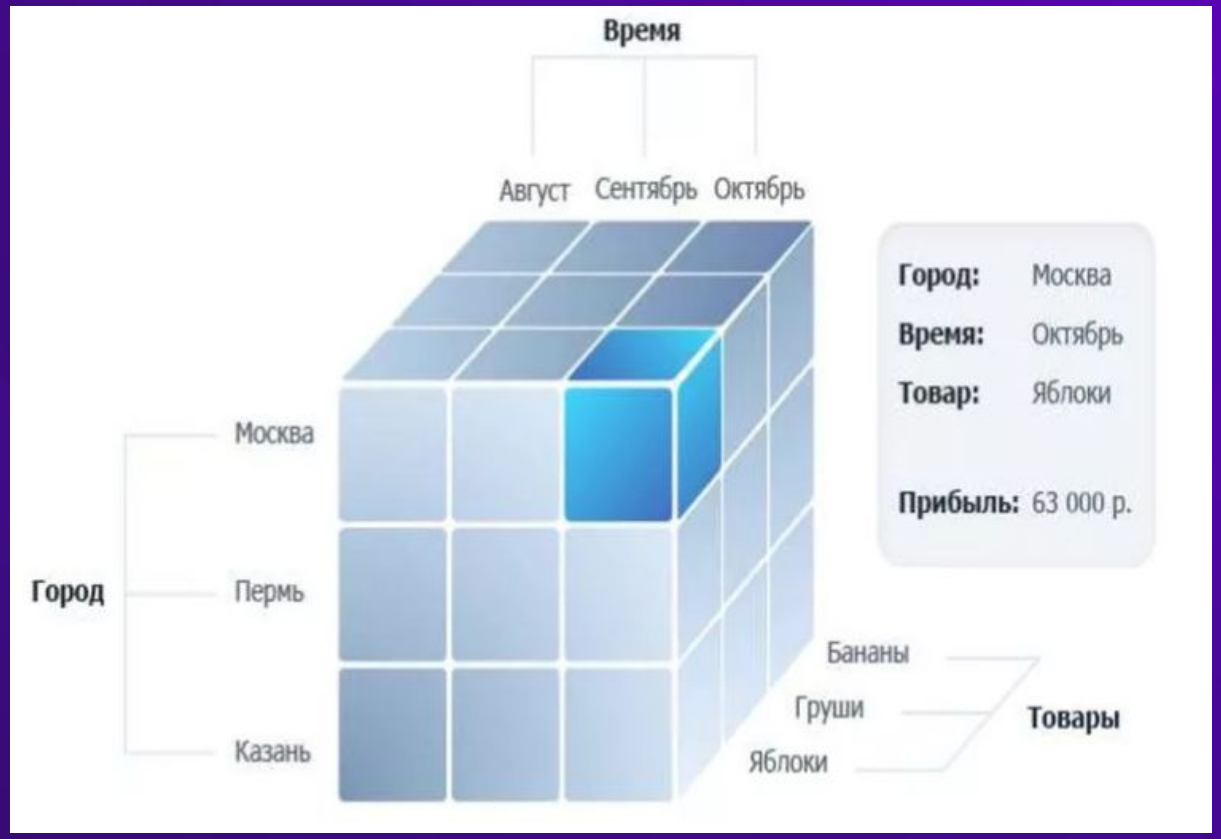
Хранилище данных является основой OLAP-системы. Процедуры загрузки с определённой периодичностью пополняют хранилище из различных источников , выполняя проверку и предварительную обработку.

Многомерные OLAP-кубы

На основе хранилища данных строятся многомерные OLAP-кубы, позволяющие в реальном времени осуществлять анализ данных и формировать отчёты в различных разрезах и произвольной глубиной детализации.



Пример OLAP куба



Система ключевых показателей (KPI)

Показатели деятельности предприятия, которые помогают организации в достижении стратегических и операционных целей, создаётся на базе OLAP-кубов для мониторинга и оценки бизнес-процессов, а также для оповещения ответственных сотрудников о фактах отклонения.

Интеллектуальный анализ данных

Модели интеллектуального анализа данных проектируются для автоматического прогнозирования наиболее важных показателей бизнеса, а также извлечения скрытых закономерностей из накопленной информации.

Доступ через web-интерфейс

Чаще всего доступ осуществляется через специализированный web-портал, позволяющий работать с OLAP-кубами и отчётами, обладающий административным интерфейсом.

Преимущества OLAP системы на предприятии

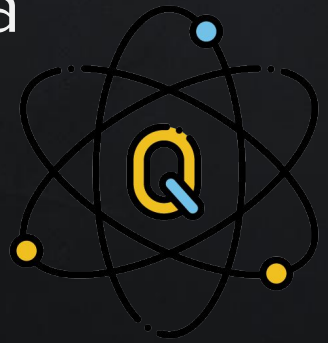
- Интегрировать данные различных информационных систем, создав единую версию правды
- Проектировать новые отчеты без участия программистов

Преимущества OLAP системы

- В реальном времени анализировать данные по любым категориям и показателям бизнеса на любом уровне детализации
- Производить мониторинг и прогнозирование ключевых показателей бизнеса

Недостатки OLAP системы

- Высокие требования к аппаратному обеспечению, подготовке и знаниям административного персонала и конечных пользователей
- Высокие затраты на реализацию проекта внедрения

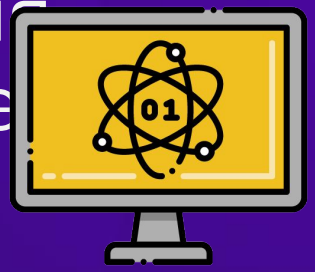


Вопрос 2

Хранилища данных

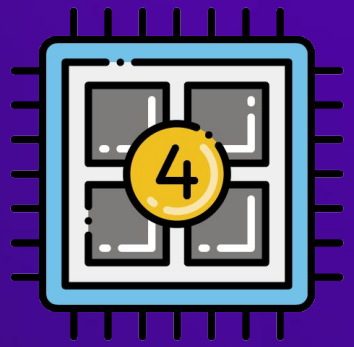
Хранилище данных

предметно-ориентированный,
интегрированный,
неизменяемый и
поддерживающий хронологию
набор данных,
предназначенный для
обеспечений принятия
управленческих реше



Концепция хранилищ данных

это концепция подготовки данных для последующего анализа



Выполнение положений

- Интеграция и согласование данных из различных источников
- Разделение наборов данных, используемых системами выполнения транзакций и системами поддержки принятия решений

Исходные данные

- Исходные данные для анализа формируются системами операционной обработки, поступают из электронных архивов и от поставщиков информации, например: онлайн информационных агентств (Lenta.ru).
- Эти источники слабо связаны между собой. Соответственно, данные, которые они предоставляют, имеют различную структуру и форматы представления.

Предметная область

- Хранилище должно разрабатываться с учетом специфики предметной области, а не приложений, оперирующих данными
- Структура хранилища должна отражать представления аналитика об информации, с которой ему приходится работать

Требования к хранилищам данных

1. Ориентация на предметную область
2. Интегрированность
3. Неизменность данных
4. Поддержка хронологии

Интегрированность

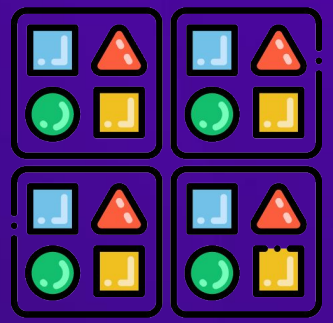
- Информация в хранилище загружается из приложений, созданных разными разработчиками
- Необходимо объединить данные этих приложений, приведя их к единому синтаксическому и семантическому виду

Неизменяемость данных

- Данные после загрузки в аналитических системах остаются неизменными
- Внесение каких-либо изменений, кроме добавления записей не предполагаются

Поддержка хронологии

- Учет хронологии достигается введением ключевых атрибутов «ДАТА» и/или «ВРЕМЯ» в структуры хранилища данных



Поддержка хронологии:

1. Маленькие
2. Средние
3. Большие
4. Сверхбольшие

Киоски данных

специализированное тематическое хранилище, обслуживающее одно из направлений деятельности организации

Выводы

- Эффект от правильной организации, стратегического и оперативного планирования развития бизнеса трудно заранее оценить в цифрах, но очевидно, что он в десятки и даже сотни раз может превзойти затраты на реализацию таких систем.

Выводы

- Однако не следует и заблуждаться. Эффект обеспечивает не сама система, а люди с ней работающие. Современные аналитические системы не являются системами искусственного интеллекта и они не могут ни помочь, ни помешать в принятии решения.
- Их цель своевременно обеспечить менеджера всей информацией необходимой для принятия решения в удобном виде.

Благодарю
за внимание!