



Основы хранилищ и витрин данных

Программа электронного курса

Модулей: 4

Основы хранилищ и витрин данных

Тип: лекции

Лекция 1. Концепция хранилища данных

Определение и свойства хранилища данных. Детализированные и агрегированные данные. Метаданные. Виды метаданных. Краткий обзор моделей хранилищ данных.

Лекция 2. Многомерная модель данных

Многомерное представление данных. Базовые понятия многомерной модели. Структура многомерного куба. Работа с измерениями.

Лекция 3. Многомерные и реляционные хранилища данных

Агрегаты и нерегламентированные запросы. Многомерные хранилища (MOLAP). Реляционные хранилища (ROLAP). Схемы «звезда», «созвездие» и «снежинка». Сравнение MOLAP и ROLAP. Гибридная (HOLAP) модель.

Лекция 4. Витрины данных

Что такое витрины данных? Пример централизованного ХД с витринами данных. Проектирование витрины данных.

Лекция 5. Загрузка данных в хранилище

Поддержка историчности. Медленно меняющиеся измерения и их типы. Неполная загрузка данных. Многопоточная организация процесса загрузки данных. Постзагрузочные операции.

Лекция 6. Метод Data Vault

История появления. Основы концепции и определение Data Vault. Основные компоненты модели Data Vault. Обобщенная схема хранилища данных. Трехуровневая архитектура. Преимущества и недостатки Data Vault.

Самопроверочный тренажер

25 вопросов с выбором вариантов ответа.

Работа с хранилищами в Vertica и Loginom. Часть 1

Тип: пошаговая симуляция

Занятие 1. Знакомство с Vertica

Обзор Vertica. Что такое Vertica. Строковые и колончатые БД. Индексы и проекции. Взаимодействие с другим инструментами. Виртуальная машина. Порядок скачивания. Порядок установки. Импорт конфигурации. Запуск виртуальной машины. Vertica Management Console.

Занятие 2. Средство администрирования базы данных DBeaver

Обзор DBeaver. Проектирование БД. Что такое DBeaver. Преимущества. Поддерживаемые БД. Порядок скачивания. Порядок установки.

Занятие 3. Создание подключения к Vertica в Loginom

Подключение к Vertica из других приложений. Порядок скачивания драйверов. Установка драйвера ODBC. Установка драйвера JDBC. Создание подключения из DBeaver. Создание подключения из Loginom.

Работа с хранилищами в Vertica и Loginom. Часть 2

Тип: пошаговая симуляция

Занятие 1. Проектирование хранилища - схема "Снежинка"

Схема "снежинка". Подготовка к проектированию. Наборы данных. Описание полей. Выбор таблицы фактов. Выбор таблиц измерений. Итоговая схема хранилища. Создание БД в DBeaver. Добавление пустой базы в Vertica Management Console. Создание подключения к базе. Создание таблиц. Добавление полей. Типы данных. Выбор ключей.

Занятие 2. Наполнение хранилища

Импорт данных. Настройка полей. Демонстрация наборов. Регистрация драйвера. Настройка подключения в Loginom. Добавление узла подключения. Настройка экспорта. Узел экспорта в БД. Необходимые настройки. Прочие настройки. Порядок выполнения.

Занятие 3. Создание семантического слоя. Получение срезов

Определение семантического слоя. Семантический слой для аналитика. Способы хранения метаданных. Получение срезов. Компонент "Метаданные". Компонент "Все продажи". Компонент "Каталог товаров". Импорт без использования фильтра. Доступные фильтры. Импорт с использованием фильтра. Проектирование компонентов.

Работа с хранилищами в Vertica и Loginom. Часть 3

Тип: пошаговая симуляция

Занятие 1. Проектирование хранилища - схема "Data Vault"

Схема "Data Vault". Поддержка историчности. Обзор данных. Выбор бизнес-ключей. Таблицы-концентраторы. Ссылки. Спутники. Итоговая логическая схема. Создание БД в DBeaver.

Занятие 2. Наполнение хранилища

Импорт данных. Добавление технических метаданных. Настройка подключения. Экспорт: концентраторы. Экспорт: ссылки. Экспорт: спутники. Загрузка новых данных. Обновление данных. Добавление новой сущности.

Занятие 3. Создание семантического слоя. Получение срезов

Компонент "Метаданные". Компонент "Каталог товаров". Компонент "Все продажи". Импорт без использования фильтра. Доступные фильтры. Импорт с использованием фильтра. Проектирование компонентов.

Контрольное задание

2 задачи.

Темы: получение срезов из хранилища.

Самопроверочный тренажер

5 вопросов (вписать вариант ответа в поле).