

Программа учебной дисциплины «Управление данными»

Утверждена

Академическим советом ООП

Протокол № 10 от «26»_06__2018 г.

Автор	С.М. Диго, к.э.н., проф., sdigo@hse.ru А.А. Дружбаев, к.т.н., доц., druzhaev@hse.ru А.С.Акопов, д.т.н., проф., aakopov@hse.ru Т.К.Кравченко, д.э.н., профессор, tkravchenko@hse.ru
Число кредитов	10
Контактная работа (час.)	152
Самостоятельная работа (час.)	228
Курс	2 курс бакалавриата, Направление 38.03.05 «Бизнес-информатика»
Формат изучения дисциплины	Без использования онлайн курса

I. ЦЕЛЬ, РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ И ПРЕРЕКВИЗИТЫ

Целями освоения дисциплины «Управление данными» является формирование знаний, умений и навыков проектирования баз данных, администрирования баз данных и разработки приложений (внешних интерфейсов) баз данных.

Дисциплина «Управление данными» состоит из двух частей: «Базы данных» и «Хранилища данных».

Задачи дисциплины «Управление данными»:

- Изучить теорию реляционных баз данных;
- Получить представление о реализации промышленных баз данных;
- Овладеть навыками проектирования баз данных с использованием CASE-инструментов.
- Выработать компетенции в области выполнения запросов к базам данных и проектирования бизнес-логики баз данных.
- Хранилища данных (ХД) - причины возникновения.
- Архитектура ХД.
- Основные принципы проектирования ХД.
- Проектирование и разработка процесса наполнения ХД.
- Физическая модель ХД.
- Использование ХД в процессе анализа данных.
- Разработка WEB-приложений к ХД.

Данная программа раскрывает содержание дисциплины «Управление данными».

В результате изучения дисциплины студент должен:

знать:

- основные понятия теории баз данных и хранилищ данных;
- подходы к построению БД и ХД и сферы их применимости;
- особенности реляционной модели БД и многомерной модели ХД;
- понятие предметной области и способы ее описания;
- теорию реляционных БД: основные понятия, реляционная алгебра и реляционное исчисление, теория нормализации;
- теорию многомерных ХД: схемы типа «звезда», «снежинка», метод Data Vault Modeling, ETL-процедуры, ROLAP и др.
- методологии ER-моделирования, проектирование БД и ХД на основе ER-моделей;
- классификацию и способы задания ограничений целостности;
- языки описания и манипулирования данными разных классов;
- особенности работы в распределенной многопользовательской среде;
- методы организации доступа к данным;
- архитектуру программных средств, работающих с БД и ХД.

уметь:

- описывать предметную область, выявлять сущности и связи между ними;
- проектировать структуру БД и ХД вручную и с использованием выбранного CASE-средства;
- использовать инструментальные средства СУБД при реализации модели БД и ХД;
- разрабатывать бизнес-логику работы с БД и ХД;
- строить запросы, используя различные языковые средства;
- использовать БД и ХД при построении отчетов и разработке приложений.
- разрабатывать WEB-приложения к БД и ХД.

владеть:

- навыками проектирования структуры БД и ХД,
- навыками формирования SQL-запросов к БД и MDX-запросов к ХД,
- навыками анализа данных.

II. СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Часть 1. Базы данных

Тема 1. Введение в курс «Управление данными»

Структура курса, контрольные мероприятия, оценка за работу в семестре

Понятия данные, информация, управление данными. Аспекты изучения информации.

Понятие банка данных (БнД) Предпосылки создания БнД. Преимущества и недостатки БнД.

Компоненты БнД. База данных (БД) - ядро БнД. Программные средства БнД.

Языковые средства современных СУБД. Классификация языковых средств. Языки четвертого поколения и их свойства.

Технические средства БнД. Организационно-методические средства.

Понятие СУБД. Функции СУБД. Выбор СУБД.

Классификация БнД. Иерархические, сетевые БД, БД на инвертированных списках.

OLAP и OLTP-системы. Хранилища данных.

Смешанные БД. XML-базы данных.

Тенденции развития БнД.
Пользователи БнД.

Тема 2. Основы теории реляционных баз данных

Отличительные особенности реляционных моделей. Основные понятия: отношение, кортеж, атрибут, ключ (простой, составной; первичный, альтернативный). Домен.

Связывание таблиц. Внешний ключ.

Функциональные зависимости. Многозначные зависимости.

Операции реляционной алгебры. Трёхзначная логика (3VL).

Теория нормализации. Понятие 1НФ, 2НФ, 3НФ, 4НФ, 5НФ.

Алгоритм нормализации. Достоинства и недостатки подхода.

Тема 3. Инфологическое (концептуальное) моделирование предметной области

Определение предметной области. Требования, предъявляемые к ИЛМ. Способы описания предметной области.

Состав инфологической модели (ИЛМ). Требования, предъявляемые к ИЛМ.

ER-модели. Объекты (сущности) и классы объектов. Атрибуты объектов. Типы объектов.

Виды связей. Классы членства объектов в связи.

Базовая ER-модель.

Сравнение методик инфологического моделирования. Методология IDEF1X. Методология IE.

Модели описания предметной области. Компоненты UML, связанные с проектированием БД.

Тема 4. Проектирование баз данных – логическое и физическое моделирование

Общие сведения о даталогическом моделировании. Особенности даталогических моделей. Влияние особенностей модели на проектные решения. Факторы, влияющие на проектирование БД. Критерии оценки проекта БД.

Проектирование структуры базы на основе использования ER-моделей. Влияние нотации ER-модели на сферу ее использования и алгоритм проектирования.

Алгоритм проектирования (алгоритм перехода от базовой ER-модели к структуре реляционной базы).

AllFusion ERwin Data Modeler. Уточнение терминологии. Интерфейс системы. Создание логических и физических моделей. Методология проектирования в среде ERwin Data Modeler.

Методология построения физических моделей. Типы данных. Физическое хранение данных (файлы, файловые группы, страницы, экстенды, индексы, секции). Влияние нормализации/денормализации.

Тема 5. Целостность данных

Понятие целостности и ограничения целостности. Классификация ограничений целостности. Причины, вызывающие нарушение ограничений целостности.

Описание ограничений целостности в CASE-средствах.

Способы задания ограничений целостности в современных СУБД: процедурный и декларативный способы задания ограничений целостности.

Ссылочная целостность (целосность связи).

Тема 6. Языки запросов. Введение

Понятие «язык запросов». Классификация языков запросов. Особенности языков запросов разных классов.

Классификация запросов. Особенности запросов разных классов.

Тема 7. Табличные языки запросов

Общая характеристика табличных языков запросов. Особенности реализации табличных языков запросов в современных СУБД.

Задание простых и сложных запросов. Влияние типа поля на формулирование запроса. Возможности совместной обработки нескольких файлов, связывание файлов. Вычисляемые поля. Упорядочение данных в ответе. Состав и порядок следования полей в ответе. Возможности группировки данных, получение подитогов. Использование агрегатных функций. Корректирующие запросы.

Запоминание и корректировка запросов.

Запросы на создание таблиц.

Табличные языки как «построители» SQL-запросов.

Тема 8. Язык SQL

Введение в SQL. Стандарты. Диалекты. Составные части языка SQL (DDL, DML, DCL). Расширения языка. Визуальные построители запросов. Структура языка и синтаксис основных операторов SQL. Запросы и подзапросы. Управление выполнением сценария.

Использование SQL для выборки данных из таблиц. Использование булевых и реляционных операторов создания более сложных предикатов. Использование специальных операторов в «условиях». Суммирование данных с помощью функций агрегирования. Форматирование результатов запросов. Использование множества таблиц в одном запросе. Вложенные запросы. Связанные подзапросы. Использование операторов EXIST, ANY, ALL, SOME. Использование предложения UNION. Ввод, удаление, изменение значений полей. Использование подзапросов с командами обновления. Создание таблиц, команда CREATE.

Табличные и скалярные функции. Представления. Работа с метаданными на основе системных представлений. Динамическое формирование запросов.

Планы выполнения запросов.

Тема 9. Триггеры и хранимые процедуры

Разработка триггеров БД. Условия выполнения триггера (триггерные события). Тело триггера. Использование триггеров для проверки корректности введенных данных и выполнения сложных ограничений целостности данных. Использование триггеров для обновления данных в смежных таблицах. Разработка хранимых процедур БД. Общее представление о языках разработки хранимых процедур (TSQL и PL/SQL).

Тема 10. Роль и функции администратора БД

Установка. Планы обслуживания. Резервное копирование и восстановление данных. Вопросы безопасности и управления доступом. Мониторинг и аудит. Перенос объектов БД и данных между серверами. Выделение дискового пространства. Загрузка и обновление данных. Управление пользователями (команды GRANT и REVOKE).

Тема 11. Распределенные БД

Понятие распределенных БД. Классификация БД.
Дополнительные критерии оценки СУБД.
Централизованные и распределенные системы.
Работа с базами данных в режимах "файл-сервер" и "клиент-сервер". Двух- и трехуровневые системы клиент-сервер.
Особенности работы с базами данных в многопользовательском режиме.
Обеспечение целостности БД в распределенных БД. Уровни изолированности в ANSI SQL. Механизм блокировок
Технологии тиражирования.
Проектирование распределенных баз данных.
Сетевые возможности современных СУБД.
Особенности создания баз данных, функционирующих в локальных сетях. Особенности программирования обработки баз данных при многопользовательском режиме доступа.
Проектирование приложений типа "клиент-сервер". Клиентское и серверное ПО.
Особенности работы в гетерогенной среде. Стандарты интерфейсов. Проблемы, возникающие при параллельном доступе, и пути их решения.

Тема 12. Вывод информации из баз данных.

Возможности генераторов отчетов современных СУБД. Задание формы и состава документа. Введение вычисляемых полей. Получение документов, включающих несколько степеней итогов. Получение документов на основе нескольких связанных файлов.

Графическое оформление документа.

Вывод документа на печать, на дисплей и в файл.

Использование генераторов форм для получения выходных документов.

Тема 13. Особенности разработки приложений баз данных под WEB на PHP и ASP.NET

Процедура инсталляции PHP и ASP.NET. Базовый синтаксис языков программирования PHP, VB.NET и C#. Особенности разработки приложений баз данных. Рекомендации по проектированию веб-интерфейсов баз данных. Обработка данных HTML-формы методом GET и POST. Проектирование физической схемы и объектов СУБД MySQL с использованием EMS SQL Manager for MySQL.

Тема 14. Организация доступа к данным из PHP и ASP.NET

Организация программного доступа к СУБД из PHP и ASP.NET. Изучение интерфейсов ADO, ADO.NET, ODBC и др. Выполнение запросов на чтение, запись и обновление данных из PHP-программы. Доступ к метаданным.

Тема 15. Языки программирования PHP и VB.NET

Особенности языков программирования PHP и VB.NET. Объектно-ориентированные возможности PHP. Логические операторы и функции PHP. Подключение внешних библиотек. Организация доступа к различным СУБД. Использование графической библиотеки. Глобальные массивы. Работа с cookies и сессиями. Работа с файловой системой.

Тема 16. Разработка приложений баз данных под WEB

Расширенные возможности PHP. Разработка сложных веб-интерфейсов (на примере системы авторизации пользователей на сайте и форума).

Часть 2. Хранилища данных

Тема 17. Хранилища данных (ХД) - причины возникновения.

Проблема хранения данных при принятии решений. Причины появления Хранилищ данных. Применение хранилищ данных. Ключевые области применения - отчетность, интеллектуальный анализ данных, операционное реагирование (КРІ). Отличия Хранилищ от операционных и прочих баз данных. Хранилище = "единый прозрачный источник данных о бизнесе компании для ее руководства".

Место хранилища в IT-инфраструктуре предприятия. Компоненты хранилища и их связь с прочими элементами IT-инфраструктуры. Понятие основных данных (НСИ). Системы управления НСИ.

Тема 18. Архитектура ХД.

Типовые архитектуры хранилищ данных, рассматривается глобальное хранилище данных, централизованное хранилище данных, распределенное хранилище данных, киоски (витрины) данных, взаимосвязанные киоски данных, независимые киоски данных, корпоративная информационная фабрика, хранилище данных с архитектурой шины данных, федеративное хранилище данных.

Тема 19. Основные принципы проектирования ХД

Инструменты (CASE - средства) для проектирования ХД. Методы проектирования Центрального хранилища данных. Метод многомерного моделирования: схема «Звезда», схема «Снежинка», преимущества и недостатки.

Подход к построению ХД -Data Vault.

Моделирование временных данных.

Жизненный цикл разработки хранилища данных.

Тема 20. Проектирование и разработка процесса наполнения ХД

Основные этапы задачи наполнения хранилища. Подходы ETL/ELT, различия между ними, преимущества и недостатки. Основные ETL/ELT инструменты. Задача отслеживания изменений в контексте хранилища. Медленно меняющиеся измерения (SCD).

Тема 21. Физическая модель ХД

Денормализация данных, оптимизация запросов, использование индексов, партиций. Использование специальных конструкций языка SQL.

Тема 22. Использование ХД в процессе анализа данных.

История возникновения систем бизнес-анализа (BI), основные игроки и тенденции на рынке BI. Место ХД в системах BI. Системы принятия решений, OLAP. Системы отчетности. Язык запросов MDX.

6.2. Содержание практических занятий.

Часть 1. Базы данных

Тема 2. Основы теории реляционных баз данных

Студенты на практическом занятии должны освоить основы работы с СУБД MS Access и на примере простой готовой базы данных (в качестве примера предлагается БД Orders={«Продавцы-Покупатели-Заказы»}) познакомиться с важнейшими характеристиками, такими как реляционная таблица, запись, поле, формат поле, ключевое поле, сложные ключи, индексы, схема базы данных. Типы данных в полях. Форматы данных. Значения по умолчанию. Связывание таблиц. Виды соединений.

Далее, студенты должны выбрать некоторую предметную область, например, база данных по обработке заказов, база данных по регистрации людей в аэропорту, база данных интернет-магазина, складская база данных и т.д. и для выбранной базы данных таблицы, заполнив их произвольными данными.

Домашнее задание: доработать выбранный проект базы данных до 5 – 7 реляционных таблиц, заполненных данными.

Тема 3. Инфологическое (концептуальное) моделирование предметной области

На примере исходных документов (например, отчет о продажах по филиалам и клиентам фирмы) требуется изучить процедуру построения нормализованной 3-й формы (3-НФ) для некоторой базы данных, например, Orders (Продажи).

При этом необходимо осуществить построение нормализованных форм (Первая нормальная форма) → (Вторая нормальная форма) → (Третья нормальная форма) путем исключения частных и транзитивных (прямых) зависимостей неключевых атрибутов от ключевых в формате MS Excel в виде плоских (нормализованных) таблиц.

Далее, после нормализации, требуется создать из отчета по продажам реляционную базу данных, состоящую из трех таблиц в MS Access.

Примеры исходного (ненормализованного) и нормализованного документов прилагаются в качестве образца.

Домашнее задание: Взять некоторый исходный документ (или документы), например, отчет о продажах, построить для него 1-ую, 2-ую и 3-ую нормальные формы с последующей реализацией соответствующей базы данных в MS Access.

Тема 4. Логическое и физическое проектирование баз данных

Реализация различных уровней проектирования СУБД с использованием ErWin для предметно-ориентированной («домашней») базы данных:

Концептуальное проектирование — построение семантической модели предметной области, то есть информационной модели наиболее высокого уровня абстракции.

Логическое (дatalogическое) проектирование — создание схемы базы данных на основе конкретной модели данных, например, реляционной модели данных в нотации IDEF 1X. Для реляционной модели данных даталогическая модель — набор схем отношений, обычно с указанием первичных ключей, а также «связей» между отношениями, представляющих собой внешние ключи.

Физическое проектирование — создание схемы базы данных для конкретной СУБД. Специфика конкретной СУБД может включать в себя ограничения на именованное объектов базы данных, ограничения на поддерживаемые типы данных и т.п.

Изучение интерфейса ErWin и создание учебного проекта в ErWin для ранее выбранной (домашней) базы данных. К примеру, для базы данных Orders («Продавцы-Покупатели-Заказы», см. рисунок ниже) строится ER-модель.

Домашнее задание: Построить в ER-диаграмму к домашней базе данных (не менее 6 реляционных таблиц) с использованием ErWin.

Тема 5. Целостность данных

Реализаций ссылочной целостности между таблицами в ErWin. Настройка раз Описание ограничений целостности в ErWin. Каскадное обновление таблиц. Каскадное удаление таблиц. Изучение конструкций FOREIGN KEY и REFERENCES.

Домашнее задание: обеспечить целостность данных к проектируемой БД.

Тема 7. Табличные языки запросов

Введение в язык формирования запросов на языке SQL. Изучаются прежде всего, так называемые простые табличные запросы класса SELECT * FROM TABLE. При этом для освоения языка SQL используется СУБД Access. Запросы строятся на языке SQL для извлечения данных их базы данных. На примере домашней базы данных освоить построение запросов к одной таблице с помощью мастера запросов MS Access, а затем провести анализ запросов в редакторе SQL.

Запросы следует формировать с использованием ключевых слов: SELECT, FROM, AS, WHERE. Также следует рассмотреть построение запросов на выборку данных с условиями (использование WHERE). При этом желательно использовать логические операторы OR, AND, NOT в условиях, а также выполнить сортировку результатов (оператор ORDER BY).

Далее предлагается выполнить простое упражнение по реинжинирингу домашней базы данных (MS Access), в которой реализованы реляционные связи между таблицами, а также запросы в систему ErWin.

Домашнее задание: построить несколько запросов к домашней базе данных как с помощью мастера запросов в MS Access, так и напрямую в редакторе запросов вида:

SELECT * FROM TABLE, SELECT * FROM TABLE WHERE и другие простые запросы на извлечение данных из таблицы.

Выполнить реинжиниринг домашней базы данных в ErWin.

Тема 8. Язык SQL

Запросы к нескольким таблицам (пояснить формирование оператора JOIN). Запросы с условиями, которые включают специальные операторы IN, BETWEEN, ALL, LIKE, IS NULL. Запросы с функциями агрегирования: SUM, MAX, MIN и AVG. Использование оператора GROUP BY. Оператор HAVING и особенности его применения с оператором GROUP BY.

Использование функций для получения вычисляемых столбцов. Вложенные запросы с оператором EXISTS. Команды DML (INSERT, DELETE и UPDATE). Команда создания таблиц CREATE TABLE. Контроль вводимых данных: проверка данных на NULL (NOT NULL), проверка на уникальность (UNIQUE), проверка данных по условию (CHECK и CHECK <с условиями>), присвоение значения по умолчанию (DEFAULT)

Домашнее задание: построить 5 сложных запросов к домашней базе данных с использованием освоенных ключевых функций.

Тема 9. Триггеры и хранимые процедуры

Введение в язык SQL-процедур. Параметрические SQL-процедуры. Системные переменные и способы их вывода. Определение и назначение триггеров. Примеры триггеров, обеспечивающих целостность и контроль вводимых данных.

Домашнее задание: разработать два триггера и одну Хранимую процедуру к своей БД.

Тема 10. Роль и функции администратора БД

Управление пользователями. Создание резервных копий БД. Миграция данных между различными системами (используя DTS). Экстракция БД в виде SQL-скрипта.

Распределение привилегий и защита от несанкционированного доступа: «логин» и пароль» к базе данных, команда GRANT, аргументы ALL и PUBLIC, передача привилегий с использованием GRANT OPTION. Лишение привилегий пользователя (REVOKE)

Домашнее задание: создать нескольких пользователей в БД MS SQL Server и наделить их различными правами доступа на объекты БД (например, только чтение, только добавление, полный доступ и др.), используя команды: GRANT, REVOKE, ALL и PUBLIC и др. Создать резервную копию БД.

Тема 11. Распределенные БД

Изучение процедуры инсталляции и настройки сетевого доступа для клиент-серверных СУБД, например, MS SQL Server или Oracle. Организация переноса данных между различными серверами БД с использованием DTS (Data Transformation Services).

Домашняя работа: Обеспечить перенос данных с одного сервера СУБД на другой сервер (или с одного instance (экземпляра БД) на другой instance), используя DTS. Организовать пакетную загрузку данных в отдельные таблицы с помощью команды BULK INSERT (T-SQL).

Тема 12. Вывод информации из баз данных.

Изучение генераторов отчетов, обеспечивающих возможность вывода информации из СУБД.

Введение в ODBC и ADO. Создание DNS для учебной базы данных.

Разработка внешних приложений с помощью офисных пакетов (например, Excel).

Формирование макроса доступа к базе данных. Модификация макроса и разработка модуля для чтения данных из БД.

Концепция диалогового интерфейса. Разработка схемы граф-схемы диалога решения прикладной задачи. Введение в VBA (операции чтения/записи данных с листа, чтения/записи данных с формы, обработка данных формы и т.п.)

Разработка в Excel диалоговой формы вывода списка запросов, хранящихся в Excel-таблице.. Разработка модуля коннекции к базе данных, модуля чтения данных АО SQL-запросу и модуля вывода отобранных данных на рабочий лист Excel-книги.

Домашняя работа: Вывод данных из СУБД с помощью генератора отчетов. Разработка простого приложения баз данных под MS Excel на VBA и ADO для вывода информации из СУБД.

Тема 13. Особенности разработки приложений баз данных под WEB на PHP и ASP.NET

Инсталляция PHP, Apache, IIS и EMS SQL Manager for MySQL. Инсталляция MySQL (пакет Denwer). Перенос проекта домашней СУБД (взять из курсового проекта) на платформу MySQL с помощью EMS SQL Manager for MySQL

Домашняя работа: Перенести проект домашней СУБД на платформу MySQL с помощью EMS SQL Manager for MySQL.

Тема 14. Организация доступа к данным из PHP и ASP.NET

Организация программного доступа из PHP к СУБД MySQL. Организация программного доступа из PHP к СУБД MS SQL Server с использованием php_mssql.dll. Организация программного доступа из ASP.NET к СУБД MS SQL Server с использованием ADO.

Домашняя работа: Разработать простую процедуру на PHP, обеспечивающую успешный коннект к СУБД MySQL. Выполнить простые запросы (SELECT) из PHP к СУБД с выгрузкой результатов на WEB-страницу.

Тема 15. Языки программирования PHP и VB.NET

Разработка сценария сайта и диалоговых форм. Программирование диалоговых форм на HTML и PHP. Разработка веб-интерфейса для ввода данных в таблицу БД, веб-интерфейса для корректировки данных. Разработка «главной» формы с диалогом типа «меню».

Домашняя работа: Разработать простой веб-интерфейс к домашней СУБД (MySQL) с использованием PHP.

Тема 16. Разработка приложений баз данных под WEB

Разработка форм вывода данных из СУБД в HTML- форму типа «таблица». Отладка основных режимов работы сайта (диалог и реакция на ошибки ввода). Отработка формы для сброса данных в Excel файл. Разработка веб-интерфеса для аутентификации пользователей на сайте.

Домашняя работа: Разработать веб-интерфес для аутентификации пользователей на своем сайте с использованием MySQL и PHP.

Часть 2. Хранилища данных

Тема 17. Хранилища данных (ХД) - причины возникновения.

Проектирование хранилища данных - метод многомерного моделирования
Отработка принципов моделирования
Разработка ER-модели ХД по заданной предметной области

Домашняя работа: Разработать ER- модель ХД по заданной предметной области с использованием схемы типа «звезда»

Тема 18. Архитектура ХД.

Проектирование хранилища данных - метод DataVault
Отработка принципов моделирования
Разработка ER-модели ХД по заданной предметной области

Домашняя работа: Доработать ER- модель ХД по заданной предметной области с использованием метода DataVault. Развить схему ER-модели до уровня «снежинка»

Тема 19. Основные принципы проектирования ХД

Проектирование хранилища данных - проектирование временных данных
Отработка принципов моделирования
Разработка ER-модели по заданной предметной области
Мини-контрольная работа №1 (Проектирование ХД)
Разработка ER-модели по заданной предметной области

Домашняя работа: Расширить ER- модель ХД по заданной предметной области за счет включения данных зависящих от времени, как в таблицу фактов, так и в одну из существующих таблиц измерений (реализовать временно-зависимый справочник).

Тема 20. Проектирование и разработка процесса наполнения ХД

Разработка требований к процессу загрузки данных
Отработка принципов формирования требований к источникам данных, описания трансформаций, маппингов.
Оформление требований

Домашняя работа: Разработать хранимую процедуру, обеспечивающую последовательную загрузку данных из текстовых файлов в ХД.

Тема 21. Физическая модель ХД

Проектирование аналитического хранилища данных в среде MS SQL Server (SSAS) – домашнее задание

- Разработка структуры источника данных для аналитического хранилища
- Проектирование измерений хранилища
- Проектирование куба хранилища
- Создание вычисляемых показателей

Домашняя работа: Создать проект SSAS в среде MS SQL Server, включающий как создание кубов, так и создание измерений, а также вычисляемых показателей.

Тема 22. Использование ХД в процессе анализа данных.

- Использование языков запросов к многомерным данным (MDX)
- Анализ данных при помощи MDX запросов

Домашняя работа: Используя средство построения MDX-запросов (Power Pivot) построить OLAP-отчет к кубу принадлежащему ранее сохраненному SSAS-проекту.

III. ОЦЕНИВАНИЕ

В соответствии с Рабочим учебным планом применяются следующие формы текущего и итогового контроля:

- (1) контрольная работа по проверке навыков построения ER-модели - М₁
- (2) контрольная работа по проверке знания языка SQL - М₂
- (3) домашнее задание по проектированию базы данных – М₃
- (4) контрольная работа по проектированию хранилища данных- М₄
- (5) контрольная работа по разработке Web-приложений баз данных и хранилищ данных- М₅
- (6) реферат (проектирование и описание домашней БД и прототипа ХД)- М₆

Реферат должен включать следующие разделы:

- краткое описание предметной области;
- нормализация таблиц;
- построение диаграммы (схемы) СУБД в СА ErWin;
- реализация физической модели СУБД;
- заполнение СУБД тестовыми данными;
- написание различных запросов к СУБД (не менее 10 запросов);
- разработка хранимых процедур и триггеров (не менее 1 хранимой процедуры и 2-х триггеров);
- вывод информации из СУБД (с использованием генераторов отчетов и программирования).
- разработка прототипа ХД на основе БД.
- разработка хранимой процедуры для загрузки данных в ХД.
- создание проекта SSAS в MS SQL Server.

- разработка MDX-запроса к ХД.
- разработка WEB-интерфейса с ХД.

Каждая из форм текущего контроля оценивается по 10-балльной шкале.

Общая оценка за текущий контроль (по 10-балльной шкале) рассчитывается по формуле:

$$O_{\text{текущий}} = 0,1 \times M_1 + 0,1 \times M_2 + 0,2 \times M_3 + 0,1 \times M_4 + 0,1 \times M_5 + 0,4 \times M_6$$

- где M_1 – оценка за первую контрольную работу;
 M_2 – оценка за вторую контрольную работу;
 M_3 – оценка за домашнюю работу;
 M_4 – оценка за третью контрольную работу;
 M_5 – оценка за четвертую контрольную работу;
 M_6 – оценка за реферат.

Экзамен проводится в виде теста. Критерии выставления оценок за экзамен:

- 100% теста – 10 баллов
- 90% теста – 9 баллов
- 80% теста – 8 баллов
- 70% теста – 7 баллов
- 60% теста – 6 баллов
- 50% теста – 5 баллов
- 40% теста – 4 балла
- 30% теста – 3 балла
- 20% теста – 2 балла
- 10% теста – 1 балл

При определении накопленной оценки (по 10-балльной шкале) аудиторная работа и самостоятельная внеаудиторная работа не оцениваются. Поэтому накопленная оценка совпадает с оценкой за текущий контроль и рассчитывается по формуле:

$$O_{\text{накопленная}} = 0,6 \cdot O_{\text{текущий}} + 0,4 \cdot O_{\text{ауд}} + 0,0 \cdot O_{\text{сам.работа}},$$

- где $O_{\text{текущий}}$ – оценка за текущий контроль;
 $O_{\text{ауд}}$ – оценка за аудиторную работу;
 $O_{\text{сам.работа}}$ – оценка за самостоятельную работу.
 $O_{\text{ауд}} = 0,5$ посещение лекций + $0,5$ посещение практических занятий
Критерии выставления оценки за посещение лекций:

- 100% посещаемости – 10 баллов
- 90% посещаемости – 9 баллов
- 80% посещаемости – 8 баллов
- 70% посещаемости – 7 баллов
- 60% посещаемости – 6 баллов
- 50% посещаемости – 5 баллов

- 40% посещаемости – 4 балла
- 30% посещаемости – 3 балла
- 20% посещаемости – 2 балла
- 10% посещаемости – 1 балл

Критерии выставления оценки за посещение практических занятий:

- 100% посещаемости – 10 баллов
- 90% посещаемости – 9 баллов
- 80% посещаемости – 8 баллов
- 70% посещаемости – 7 баллов
- 60% посещаемости – 6 баллов
- 50% посещаемости – 5 баллов
- 40% посещаемости – 4 балла
- 30% посещаемости – 3 балла
- 20% посещаемости – 2 балла
- 10% посещаемости – 1 балл

Результирующая оценка $O_{результ}$ формируется на основе итоговой оценки за экзамен (по 10-балльной шкале) и накопленной оценки. Результирующая оценка рассчитывается по формуле:

$$O_{результ} = 0,4 \cdot O_{экзамен} + 0,6 \cdot O_{накопленная} ,$$

где $O_{экзамен}$ – оценка за итоговый контроль (экзамен);
 $O_{накопленная}$ – накопленная оценка.

При формировании результирующей оценки на основе весовых коэффициентов в ведомость выставляется целая часть полученного числа.

IV. ПРИМЕРЫ ОЦЕНОЧНЫХ СРЕДСТВ

4.1 Вопросы для оценки качества освоения дисциплины

1. Дайте определение реляционной База Данных? Роль системы управления базами данными (СУБД) в организации.
2. Архитектура современных СУБД. В чем суть логического и инфологического моделирования?
3. Этапы проектирования СУБД. Каковы принципы построения ER-модели?
4. Нормализация БД. В чем особенности первой, второй и третьей нормальной формы?
5. Какие задачи по проектированию БД могут решаться с помощью инструментальных средств ErWin/BrWin.
6. Каково назначение языков DDL и DML. Как создать таблицу БД используя DDL.
7. Какие существуют операторы языка SQL для чтения и записи данных?
8. Какие существуют операторы языка SQL для обновления (редактирования) и удаления данных?

9. Какие существуют операторы языка SQL предоставления прав доступа к отдельным объектам БД различным пользователям?
10. В чем отличия архитектуры клиент-сервер от архитектуры файл-сервер? Каковы особенности работы БД в многопользовательском режиме?
11. Роль интерфейсов ODBC и ADO в проектах разработки приложений баз данных под Windows?
12. Какие основные функции языка программирования PHP используются для организации доступа к объектам СУБД MySQL Server в WEB-приложениях.
13. Как осуществить доступ к БД из приложений под Windows, используя VBA и ADO.
14. Опишите основные этапы проектирования приложений БД под WEB. Как осуществить передачу данных с WEB-страницы серверу БД?
15. Перечислите функции администратора БД. Как осуществить физический перенос объектов БД с данными между различными серверам (например, в MS SQL Server).
16. Перечислите основные инструментальные средства инструментальные средства СУБД MS SQL Server?
17. Каково предназначение инструментального средства инструментального средства служб Reporting Services?
18. Каково назначение оператора BULK INSERT в Transact SQL?
19. Каково назначение функций – указателей в Transact SQL?
20. Перечислите основные инструментальные средства инструментальные средства СУБД Oracle?
21. Какие функции управления данными могут быть реализованы с помощью Oracle enterprise manager?
22. Из каких ключевых блоков состоит программа PL/SQL?
23. Как осуществить запрос нескольких строк из базы данных, используя курсоры PL/SQL?
24. Какие задачи, связанные с проектирование СУБД MySQL Server могут быть выполнены с помощью программы EMS SQL Manager for MySQL?
25. Каковы основные причины внедрения ХД в организации? Чем ХД отличается от транзакционной БД?
26. Опишите основные особенности ROLAP, HOLAP и MOLAP.
27. Опишите схему типа «звезда» и «снежинка».
28. Особенности метода DataVault.
29. Какие типы инфо-кубов поддерживаются в современных ХД.
30. Зачем используются агрегаты в ХД.
31. Расскажите об основных функциональных возможностях языка MDX.
32. Опишите принципы формирования OLAP-отчетности.
33. Опишите процедуру создания аналитического хранилища данных в среде MS SQL Server (SSAS).
34. Что такое ETL-процедура? Опишите возможную схему загрузки данных в ХД из плоских файлов.

4.2 Тематика заданий по различным формам текущего контроля

Примерная тематика курсовых работ

Проектирование баз и хранилищ данных для конкретных предметных областей

1. Проектирование БД (ХД) работы склада производственного предприятия.

2. Проектирование БД (ХД) работы склада предприятия розничной торговли.
3. Проектирование БД (ХД) работы склада оптовой торговли.
4. Проектирование БД (ХД) интернет-магазина.
5. Проектирование БД (ХД) интернет-аукциона.
6. Проектирование БД (ХД) службы онлайн-объявлений.
7. Проектирование БД (ХД) сайта поиска работы.
8. Проектирование БД (ХД) социальной сети.
9. Проектирование БД (ХД) новостной интернет-службы.
10. Проектирование БД (ХД) регистрации обращений граждан в государственные органы.
11. Проектирование БД (ХД) для учета контингента студентов ВУЗа
12. Проектирование БД (ХД) для управления работой компьютерного класса
13. Проектирование БД (ХД) для начисления заработной платы
14. Проектирование БД (ХД) для контроля успеваемости студентов.
15. Проектирование БД (ХД) детского сада.
16. Проектирование БД (ХД) спортивной школы.
17. Проектирование БД (ХД) фитнес-центра.
18. Проектирование БД (ХД) центра детского творчества
19. Проектирование БД (ХД) для расчета заработной платы.
20. Проектирование БД (ХД) для учета домашних финансов.
21. Проектирование БД (ХД) для домашней библиотеки.
22. Проектирование БД (ХД) для районной библиотеки.
23. Проектирование БД (ХД) для домашней видеотеки.
24. Проектирование БД (ХД) для пункта проката видеофильмов.
25. Проектирование БД (ХД) кинотеатра.
26. Проектирование БД (ХД) риэлтерского агентства.
27. Проектирование БД (ХД) для учета услуг, оказываемых юридической консультационной фирмой.
28. Проектирование БД (ХД) автосервиса.
29. Проектирование БД (ХД) автозаправочной станции.
30. Проектирование БД (ХД) центра по продаже автомобилей.
31. Проектирование БД (ХД) службы такси.
32. Проектирование БД (ХД) для учета кадров коммерческого предприятия.
33. Проектирование БД (ХД) для ведения документооборота предприятия.
34. Проектирование БД (ХД) для ведения бухгалтерского учета в малом бизнесе.
35. Проектирование БД (ХД) службы знакомств.
36. Проектирование базы данных туристического агентства.
37. Проектирование базы данных туристического клуба
38. Проектирование БД (ХД) поликлиники.
39. Проектирование базы данных аптеки.
40. Проектирование базы данных гостиницы.
41. Проектирование базы данных издательства.
42. Проектирование базы данных для учета расчетов с клиентами в банке.
43. Проектирование базы данных для учета выдачи и погашения кредитов.
44. Проектирование базы данных ведения операций по пластиковым карточкам.
45. Проектирование базы данных строительной фирмы.
46. Проектирование базы данных учета телефонных звонков.
47. Проектирование базы данных торговой организации.
48. Проектирование базы данных учета транспортных средств в ГИБДД.
49. Проектирование базы данных пункта по ремонту бытовой техники.

50. Проектирование БД для пункта проката автомобилей.

V. РЕСУРСЫ

5.1 Основная литература

1. Стружкин Н.П. Базы данных: проектирование : практикум: учебник для вузов. Изд. ЮРАЙТ. 2016. (25 экз.).
2. Советов Б. Я., Цехановский В. В., Чертовской В. Д. БАЗЫ ДАННЫХ 3-е изд., пер. и доп. Учебник для прикладного бакалавриата. Изд. ЮРАЙТ. 2018. [Режим доступа: <https://biblio-online.ru/book/bazy-dannyh-415601>]
3. Парфенов Ю. П. ; под науч. ред. Папуловской Н.В. ПОСТРЕЛЯЦИОННЫЕ ХРАНИЛИЩА ДАННЫХ. Учебное пособие для вузов. Изд. ЮРАЙТ. 2018. [Режим доступа: <https://biblio-online.ru/book/postrelyacionnye-hranilischa-dannyh-415695>]
4. Тузовский А. Ф. ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА WEB-ПРИЛОЖЕНИЙ. Учебное пособие для академического бакалавриата. Изд. ЮРАЙТ. 2018. [Режим доступа: <https://www.biblio-online.ru/book/proektirovanie-i-razrabotka-web-prilozheniy-413954>]

5.2 Дополнительная литература:

5. Акопов А.С. Учебно-методическое пособие по курсу «Базы данных» по теме Разработка web-приложений на стороне сервера. Утверждено Учебно-методическим советом факультета «Бизнес-информатика» ГУ-ВШЭ в качестве учебного пособия для студентов 20.04.2010.
6. Веллинг, Л. Разработка веб-приложений с помощью PHP и MySQL. Вильямс, 2010. - 847 с.
7. Джеффри Д. Ульман, Дженнифер Уид «Введение в системы баз данных», изд. Лори, 2006 г.
8. Линн Бейли, Майкл Моррисон. Изучаем PHP и MySQL. Издательство: Эксмо, 2010 г.
9. Герасимов Н.А. Разработка диалоговых систем в среде VBA,- Российская Экономическая Академия им. Г.В. Плеханова, М., 2004 г.
10. Герасимов Н.А. Проектирование диалоговых баз данных в среде VBA – учебно-методическое пособие (Практический курс), ГУ-ВШЭ, 2005, 171.
11. Герасимов Н.А. Практикум по разработке диалоговых систем и баз данных в Web-среде – М., РИО Российская Экономическая Академия им. Г.В. Плеханова , 2005, 108 с
12. Герасимов Н.А. Практикум по языку SQL в среде СУБД ACCESS – М., РИО Российская Экономическая Академия им. Г.В. Плеханова , 2007 г.
13. Герасимов Н.А. Практикум по разработке баз данных в СУБД IBM DB2 – М., РИО Российская Экономическая Академия им. Г.В. Плеханова , 2008, 89 с
14. Герасимов Н.А. Практикум по разработке Web-приложений на языке JavaScript - М., РИО Российская Экономическая Академия им. Г.В. Плеханова , 2010, 107 с

15. Дьюсен Р. SQL Server 2008 для начинающих разработчиков: Пер с англ.- СПб.:БХВ-Петербург, 2009.- 704с.
16. Кузнецов С.Д. Объектно-реляционные базы данных: прошедший этап или недооцененные возможности? Труды Института системного программирования, т. 13, часть 2, М., ИСП РАН, 2007, стр. 115-140.
17. Маклаков С. В. Создание информационных систем с ALLFusion Modeling Suite. М.: ДИАЛОГ-МИФИ, 2005 – 432 с.
18. «Репликация Microsoft SQL Server 2005/2008»-М.: Эком Паблишерз, 2009, 288 страниц
19. Туманов, В. Е. Проектирование хранилищ данных для систем бизнес-аналитики. М. БИНОМ. Лаборатория знаний, 2010. - 615 с.
20. <http://www.sql.ru>
21. [http://msdn.microsoft.com/ru-ru/library/bb418431\(v=SQL.10\).aspx](http://msdn.microsoft.com/ru-ru/library/bb418431(v=SQL.10).aspx)
22. <http://www.intuit.ru/catalog/database/>
23. <http://www.microsoft.com/technet/prodtechnol/sql/2005/default.mspx>
24. <http://www.microsoft.com/sql>

5.3 Программное обеспечение

№ п/п	Наименование	Условия доступа
1.	Microsoft Windows 7 Professional RUS Microsoft Windows 10 Microsoft Windows 8.1 Professional RUS	<i>Из внутренней сети университета (договор)</i>
2.	Microsoft Office Professional Plus 2010	<i>Из внутренней сети университета (договор)</i>
3	Microsoft SQL Server	<i>Из внутренней сети университета (договор)</i>
4	CA ErWin Data Modeler	<i>Бесплатная демо-версия доступна по адресу: https://erwin.com/software-trials/</i>
5	PHP	<i>Бесплатное ПО, доступное по адресу: http://www.denwer.ru/</i>
6	MySQL	<i>Бесплатное ПО, доступное по адресу: http://www.denwer.ru/</i>

5.4 Профессиональные базы данных, информационные справочные системы, интернет-ресурсы (электронные образовательные ресурсы)

№ п/п	Наименование	Условия доступа

	<i>Профессиональные базы данных, информационно-справочные системы</i>	
1.	Электронно-библиотечная система Юрайт	URL: https://biblio-online.ru/
	<i>Интернет-ресурсы (электронные образовательные ресурсы)</i>	
2.	Открытое образование	URL: https://openedu.ru/
3	ИНТУИТ	URL: http://www.intuit.ru/

5.5 Материально-техническое обеспечение дисциплины

Учебные аудитории для лекционных занятий по дисциплине обеспечивают использование и демонстрацию тематических иллюстраций, соответствующих программе дисциплины в составе:

- ПЭВМ с доступом в Интернет (операционная система, офисные программы, антивирусные программы);
- мультимедийный проектор с дистанционным управлением.

Учебные аудитории для лабораторных и самостоятельных занятий по дисциплине оснащены персональными компьютерами, с возможностью подключения к сети Интернет и доступом к электронной информационно-образовательной среде НИУ ВШЭ.

Авторы программы:

НИУ-ВШЭ профессор С.М. Диго
(место работы) (занимаемая должность) (инициалы, фамилия)
Автор программы:

НИУ-ВШЭ доцент А.А. Дружаев
(место работы) (занимаемая должность) (инициалы, фамилия)

НИУ-ВШЭ профессор А.С. Акопов
(место работы) (занимаемая должность) (инициалы, фамилия)

НИУ-ВШЭ профессор Т.К. Кравченко
(место работы) (занимаемая должность) (инициалы, фамилия)