

Математические основы Баз данных

Лекция 0: Обзор курса и введение

Контакты

Фомченко Александр Валерьевич

- Email: alexander.mirea@mail.ru
- <https://github.com/alexander-mirea/DB1-2025>

Просьба старостам отправить мне контактные данные.

Описание раздела

1. Мотивация по изучению БД
2. Административные вопросы
3. Обзор курса
4. Обзор технологий, которые будут использоваться в курсе

Почему стоит изучить базы данных

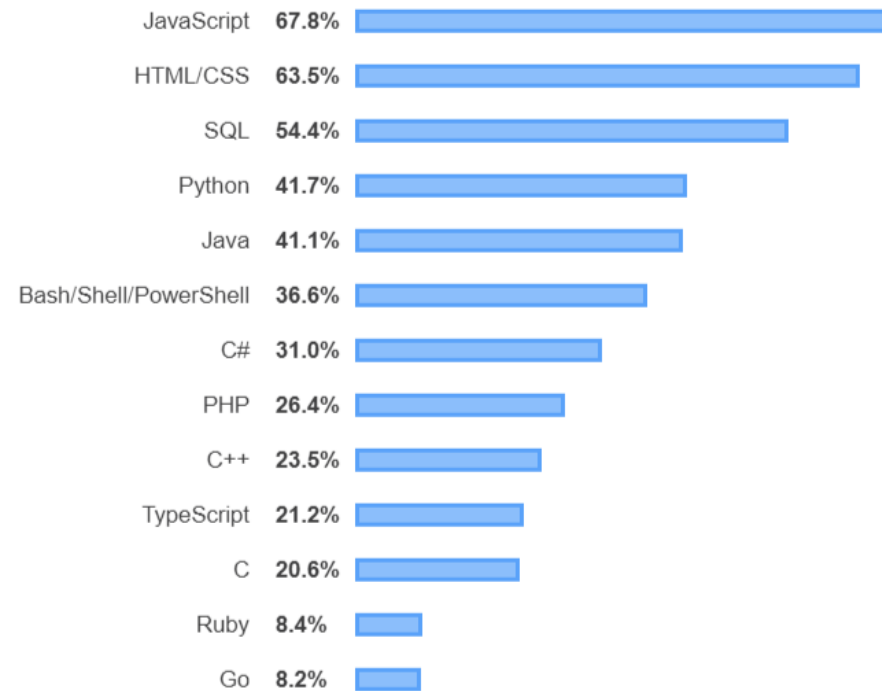


Most Popular Technologies

Programming, Scripting, and Markup Languages

All Respondents

Professional Developers



Где нужны знания SQL и СУБД

1. SQL/BI Developer
2. Junior Database Administrator
3. Аналитик данных
4. Data Engineer
5. Data Scientist
6. Software Engineer

Почему стоит изучить базы данных

- Сейчас количество данных резко увеличивается, поэтому важно разбираться в структуре обработки данных, а также наборе инфраструктуры.
- Тема баз данных содержит очень много фундаментальных идей из Computer Science

О чем этот курс

- Рассмотрим основы управления данными (data management)
- Как разрабатывать базы данных, осуществлять запросы к ним, строить приложения над ними
- Рассмотрим как работают системы управления базами данных
- Также затронем математические основы
- В основном будем говорить о проблемах на высоком уровне

Формат

- В этом семестре зачет, в следующем экзамен
- В этом семестре 5 практических заданий, при сдаче всех (с учетом посещаемости) – зачет автоматом, при сдаче первых двух – допуск к зачету
- Сдача практических заданий осуществляется в компьютерном классе, я могу дать дополнительное задание или задать вопрос по коду

Обзор технологий курса

- Python 3
- Sqlite
- Jupyter Notebook (Google Colab)

Перед практическим занятием рекомендую завести почту gmail и посмотреть как выглядит colab.

О чем поговорим сегодня

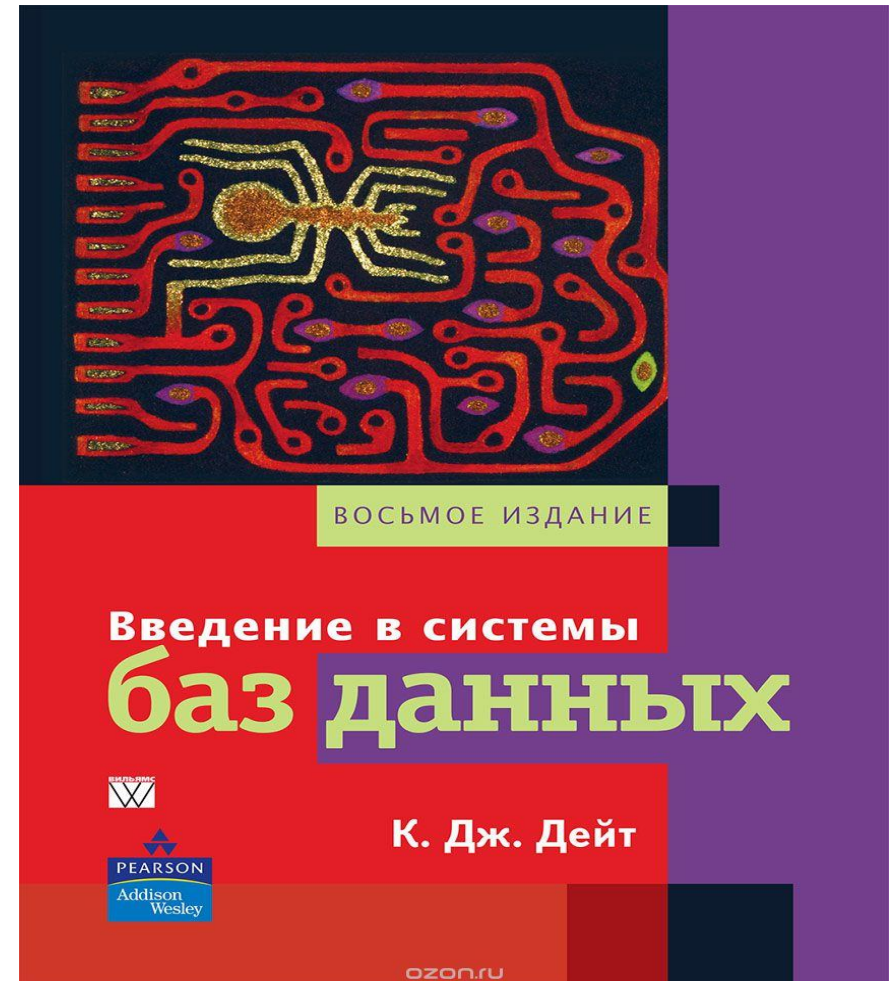
- Обзор литературы
- Природа данных, их польза и как управлять ими
- Способы хранения данных
- Развитие СУБД и современные решения (без углубления в историю)
- Реляционная модель как стандарт работы с данными

Лекции и литература

- Лекционные слайды покрывают большинство материала, который мы будем проходить
- Старайтесь повторять материал: через лекции, домашнее задание, рабочие материалы в классе

Введение в системы баз данных

- К. Дж. Дэйт
- Классика CS
- Много по теории, немного устарел по SQL



Quick Start SQL

- Можно заменить на любой аналог
- Синтаксис, примеры



+ ОДИН курс

- <https://teach-in.ru/course/database> (ВМК МГУ)

Введение

База Данных

- База данных — совокупность данных, организованных в соответствии с концептуальной структурой, описывающей характеристики этих данных и взаимоотношения между ними, которая поддерживает одну или более областей применения.
[\(ISO/IEC 2382:2015 - Information technology -- Vocabulary\)](#)
- В широком смысле база данных — это любой набор взаимосвязанной информации.

База Данных

- База данных (БД) – коллекция данных, содержащих информацию, релевантную для предприятия.
- Система Управления Базами Данных (СУБД) – коллекция связанных между собой данных и набор программ для работы с этими данными.

База Данных

- Основной целью СУБД является предоставление способа для хранения и забора информации о базе данных, который является как подходящим для пользователя, так и эффективным.
- Управление данных включает в себя как определение структур для хранения информации, так и предоставления механизмов для обработки информации.

Примеры приложений баз данных

- Информация о предприятии
 - Продажи
 - Бухгалтерия
 - HR
- Производство
- Банковское Дело
- Университеты
- Авиакомпания
- Телекоммуникации
- Электронная коммерция
- Документные базы данных

Требования к СУБД

- Разделение программ и данных
- Высокоуровневый язык запросов
- Целостность и согласованность
- Отказоустойчивость
- Защита и разграничение доступа



Требования к СУБД

- Целостность базы данных означает соответствие имеющейся в базе данных информации её внутренней логике, структуре и всем явно заданным правилам.

Вопрос

Является ли Excel СУБД?

Цели систем баз данных

Недостатки хранения информации в системе работы с файлами:

- Избыточность и несогласованность данных
- Сложность в заборе данных
- Изоляция данных
- Проблемы атомарности
- Аномалии конкурентного доступа
- Проблемы безопасности

Модель данных

– набор родовых понятий и признаков, которыми должны обладать все конкретные СУБД и управляемые ими БД, если они основываются на этой модели.

В основе структуры баз данных лежит модель данных – набор концептуальных инструментов для описания данных, отношений данных, семантики данных и ограничений целостности.

Виды (не все)

- Сетевая модель
- Иерархическая модель
- Модель данных инвертированных таблиц
- **Реляционная модель** (Relational Model)
- ER модель (Entity-Relationship Model)
- Документная модель (document-oriented)
- Модели данных, основанные на объектах (Object-based Data Models)

Модель данных SQL

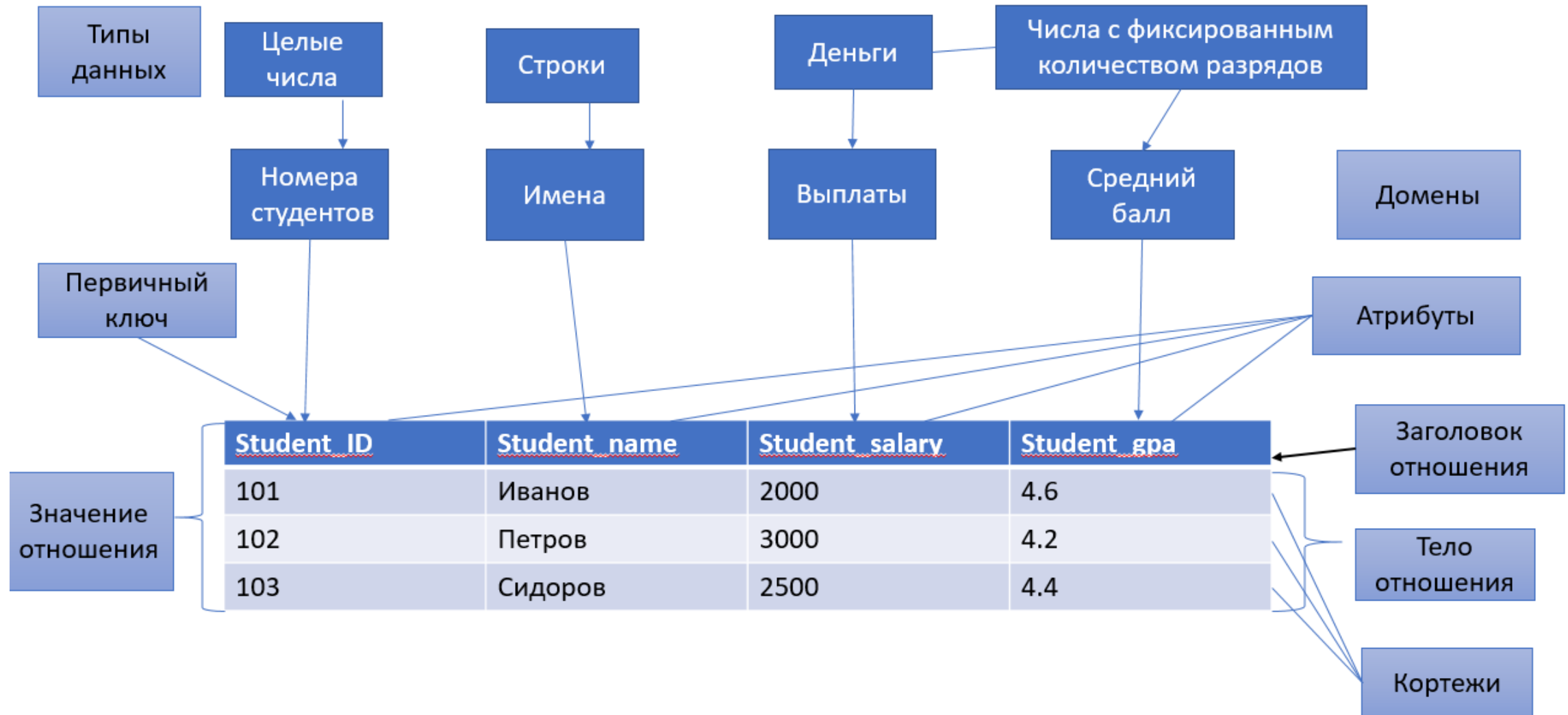
SQL-ориентированная БД представляет собой набор таблиц, каждая из которых в любой момент времени содержит некоторое *мультимножество* строк, соответствующих заголовку таблицы.

Модель данных SQL

Языки баз данных:

- DQL – Data Query Language
- DDL – Data Definition Language
 - Язык хранения и определения данных
- DML – Data Manipulation Language
- DCL – Data Control Language
- TCL – Transaction Control Language

Структура и отношения в модели данных SQL



Императивность и декларативность

На примере:

```
public static List<String> getStudentsFromIT() {  
    List<String> result = new ArrayList<>();  
    FileReader rdr = new FileReader("students.txt");  
    BufferedReader br = new BufferedReader(rdr);  
    String line = br.readLine();  
    while (line != null) {  
        String[] vals = line.split("\t");  
        String faculty = vals[2];  
        if (faculty.equals("IT"))  
            result.add(vals[1]);  
        line = br.readLine();  
    }  
    return result;  
}
```

```
SELECT name  
FROM Students  
WHERE faculty = 'IT'
```