

# Введение в базы данных. Основы реляционной модели.

# Определение

- **База данных (БД, DB)** — совокупность самостоятельных данных любого типа, систематизированных таким образом, чтобы эти данные могли быть найдены и обработаны с помощью электронной вычислительной машины (ЭВМ, компьютер);
- Если номера телефонов Ваших друзей записаны в телефонной книжке мобильного телефона — это база данных телефонов, если разбросаны по разным носителям: часть на бумажных носителях, часть на визитных карточках, часть в сообщениях социальной сети, часть в Вашей памяти — то это не база данных!
- Если телефоны записаны на бумажных носителях или в памяти человека — они не могут быть обработаны ЭВМ;
- Если в папке находятся два файла (один текстовый с номером телефона и второй — фотография номера телефона) — они не систематизированы;

# СУБД

- **Система управления базами данных (СУБД)** — специальное программное обеспечение, которое управляет созданием и использованием баз данных с помощью ЭВМ. Благодаря систематизации данных, хранимых в базе СУБД может автоматически обрабатывать и находить данные по запросу.
- Например Вы хотите хранить все ваши фотографии в папке. Для этого вы создаете отдельную папку в ней подпапки с категориями, в подпапках уже фотографии, каждая фотография имеет уникальное названия в пределах подпапки. Таким образом корневая папка с подпапками – это ваша база данных (систематизированные данные и могут быть обработаны ЭВМ).
- Упрощенной СУБД в данном варианте будет выступать программа Windows Explorer, которая позволит вам находить фотографии по имени, перегруппировывать их, удалять старые и добавлять новые.

# Функции СУБД

Почему упрощенной?

- Непосредственное управление данными во внешней памяти (на жестких дисках);
- Управление буферами оперативной памяти для быстрого доступа к данным;
- Журнализация изменений, резервное копирование и восстановление базы данных после сбоев;
- Управление транзакциями (несколькими изменениями данных, воспринимаемых как одно целое);
- Поддержка языков БД (язык определения данных, язык манипулирования данными);
- Обеспечение безопасности доступа к данным;

# Классификация СУБД

- **По модели данных:**

Иерархические;

Сетевые;

Реляционные;

Объектно-ориентированные;

Объектно-реляционные.

- **По степени распределённости:**

Локальные (все части локальной СУБД размещаются на одном компьютере);

Распределённые (части СУБД могут размещаться на двух и более компьютерах).

- **По способу доступа к БД:**

Файл-серверные;

Клиент-серверные;

Встраиваемые.

# Реляционная СУБД

- **Реляционная база данных** — база данных, основанная на реляционной модели данных. **Отношение** — фундаментальное понятие реляционной модели данных. По этой причине модель и называется реляционной (от лат. relatio — «отношение», «зависимость», «связь»).
- Реляционная модель — математическая модель, поэтому отношение может быть описано следующим образом: Пусть даны  $n$  множеств  $D_1, D_2, D_3, \dots, D_n$ , тогда **отношение  $R$**  есть множество упорядоченных кортежей  $\langle d_1, d_2, \dots, d_n \rangle$ , где  $d_i \in D_k$ ,  $d_i$  — атрибут, а  $D_k$  — домен отношения  $R$ .
- **Отношение** обычно имеет простую графическую интерпретацию в виде **таблицы**, столбцы которой соответствуют атрибутам, а строки — кортежам, а в «ячейках» находятся значения атрибутов в кортежах.
- Термины «таблица», «строка», «столбец» могут использоваться только в неформальном контексте, при условии полного понимания, что эти более «дружественные» термины являются всего лишь приближением и не дают точного представления о сути обозначаемых понятий.
- **В дальнейшем под таблицей будет подразумеваться отношения реляционной модели**

# Понятие таблицы

- **Таблица реляционной базы данных** — объект базы данных, который хранит структурированную информацию в виде набора строк и столбцов.
- Реляционная база данных представляет собой множество взаимосвязанных **таблиц**, каждая из которых содержит информацию об объектах определенного вида.
- Каждая строка (запись / кортеж) таблицы (отношения) содержит данные об одном объекте (например, контакте из телефонной книжки), а столбцы (поля / атрибуты) таблицы содержат различные характеристики этих объектов - атрибуты (например, фамилия, имя, номер телефона).

	Фамилия	Имя	Номер
➤	Сергеев	Сергей	+38 (050) 123-12-45
	Петров	Петр	+38 (050) 123-12-34
	Иванов	Иван	+38 (050) 123-12-78

# Строки и столбцы таблицы

☰	Фамилия ▼	Имя ▼	Номер ▼
➤	Сергеев	Сергей	+38 (050) 123-12-45
	Петров	Петр	+38 (050) 123-12-34
	Иванов	Иван	+38 (050) 123-12-78

☰	Фамилия ▼	Имя ▼	Номер ▼
	Сергеев	Сергей	+38 (050) 123-12-45
	Петров	Петр	+38 (050) 123-12-34
	Иванов	Иван	+38 (050) 123-12-78
➤	Давыдов	Давид	+38 (067) 122-11-69

☰	Фамилия ▼	Имя ▼	Номер ▼	В Черном списке ▼
	Сергеев	Сергей	+38 (050) 123-12-45	Нет
	Петров	Петр	+38 (050) 123-12-34	Нет
➤	Иванов	Иван	+38 (050) 123-12-78	Да
	Давыдов	Давид	+38 (067) 122-11-69	Нет



# БД - Мои фотографии

Категория	Название фотографии	Дата путешествия	Дата фотографии	Размер файла	Фотография
Свадебное путешествие	Поход в ресторан	10/7/2015 4:51:44 PM	10/7/2015 4:51:44 PM	3Мб	
Свадебное путешествие	Красивый закат	10/7/2015 4:51:44 PM	10/8/2015 4:51:44 PM	4Мб	
Свадебное путешествие	Прозрачная водичка	10/7/2015 4:51:44 PM	10/9/2015 4:51:44 PM	3,4Мб	
Отпуск в Египте	Верхом на верблюде	10/25/2016 4:51:44 PM	10/26/2016 4:51:44 PM	5Мб	
Отпуск в Египте	Путешествие на квадрациклах	10/25/2016 4:51:44 PM	10/28/2016 4:51:44 PM	4Мб	
Отпуск в Египте	Подводный мир	10/25/2016 4:51:44 PM	10/25/2016 4:51:44 PM	5Мб	
Отпуск в Египте	Поход в ресторан	10/25/2016 4:51:44 PM	10/25/2016 4:53:10 PM	8Мб	
Отпуск в Египте	Пирамиды	10/25/2016 4:51:44 PM	10/25/2016 4:51:44 PM	4,5Мб	

В поле таблицы можно хранить любой тип информации:

- строка,
- целое число,
- число с плавающей точкой,
- логическое значение,
- дату,
- время,
- дату и время,
- текст,
- массив байт.

# Операции над таблицами

- **Операции над таблицами (отношениями)** (операции реляционной алгебры) – совокупность операций, которые могут быть выполнены над таблицами (отношениями). Результатом любой операции является таблица (отношение).
- Используемые на практике операции реляционной алгебры:
  1. **Переименование атрибутов;**
  2. **Объединение;**
  3. **Пересечение;**
  4. **Вычитание;**
  5. **Декартово произведение;**
  6. **Выборка (ограничение);**
  7. **Проекция;**
  8. **Соединение (комбинация декартового произведения, выборки и проекции).**

# Переименование атрибутов

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23

Таблица А

id	surname	phone_number
1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23
2	Петров	+38(099) 677-34-00

Таблица А

# Объединение

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23

Таблица А

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23
6	Дмитриев	+38(050) 111-22-33
7	Семенов	+38(066) 000-34-78

Таблица В

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23
6	Дмитриев	+38(050) 111-22-33
7	Семенов	+38(066) 000-34-78

Таблица (A UNION B)

# Объединение с дубликатами\*

Отсутствует в реляционной алгебре, так как нарушает свойство отношения не содержать дубликатов кортежей, но на практике возможно использование

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23

Таблица А

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23
6	Дмитриев	+38(050) 111-22-33
7	Семенов	+38(066) 000-34-78

Таблица В

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23
2	Петров	+38(099) 677-34-00
1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23
6	Дмитриев	+38(050) 111-22-33
7	Семенов	+38(066) 000-34-78

Таблица (A UNION ALL B)

# Пересечение

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23

Таблица А

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23
6	Дмитриев	+38(050) 111-22-33
7	Семенов	+38(066) 000-34-78

Таблица В

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23

Таблица (A INTERSECT B)

# Вычитание

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23

Таблица A

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23
6	Дмитриев	+38(050) 111-22-33
7	Семенов	+38(066) 000-34-78

Таблица B

ID	Фамилия	Номер телефона
2	Петров	+38(099) 677-34-00
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23

Таблица (A MINUS B)

# Декартово произведение

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23

Таблица А

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23
6	Дмитриев	+38(050) 111-22-33
7	Семенов	+38(066) 000-34-78

Таблица В

ID	Фамилия	Номер телефона	ID_1	Фамилия_1	Номер телефона_1
1	Иванов	+38(050) 123-45-67	1	Иванов	+38(050) 123-45-67
1	Иванов	+38(050) 123-45-67	3	Сергеев	+38(050) 987-65-23
1	Иванов	+38(050) 123-45-67	6	Дмитриев	+38(050) 111-22-33
1	Иванов	+38(050) 123-45-67	7	Семенов	+38(066) 000-34-78
3	Сергеев	+38(050) 987-65-23	1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23	3	Сергеев	+38(050) 987-65-23
3	Сергеев	+38(050) 987-65-23	6	Дмитриев	+38(050) 111-22-33
3	Сергеев	+38(050) 987-65-23	7	Семенов	+38(066) 000-34-78
4	Николаев	+38(050) 987-65-23	1	Иванов	+38(050) 123-45-67
4	Николаев	+38(050) 987-65-23	3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23	6	Дмитриев	+38(050) 111-22-33
4	Николаев	+38(050) 987-65-23	7	Семенов	+38(066) 000-34-78
5	Васильев	+38(066) 987-11-23	1	Иванов	+38(050) 123-45-67
5	Васильев	+38(066) 987-11-23	3	Сергеев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23	6	Дмитриев	+38(050) 111-22-33
5	Васильев	+38(066) 987-11-23	7	Семенов	+38(066) 000-34-78
2	Петров	+38(099) 677-34-00	1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00	3	Сергеев	+38(050) 987-65-23
2	Петров	+38(099) 677-34-00	6	Дмитриев	+38(050) 111-22-33
2	Петров	+38(099) 677-34-00	7	Семенов	+38(066) 000-34-78

Таблица (A TIMES B)



# Выборка (ограничение)

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23

Таблица **A**

ID	Фамилия	Номер телефона
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23

Таблица **B WHERE ID > 3**

# Проекция

id	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23
9	Васильев	+38(050) 123-45-67

Таблица А

Фамилия
Николаев
Иванов
Петров
Васильев
Сергеев

Таблица А[Фамилия]

# Соединение (Комбинирование операций)

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
2	Петров	+38(099) 677-34-00
3	Сергеев	+38(050) 987-65-23
4	Николаев	+38(050) 987-65-23
5	Васильев	+38(066) 987-11-23

Таблица А

ID	Фамилия	Номер телефона
1	Иванов	+38(050) 123-45-67
3	Сергеев	+38(050) 987-65-23
6	Дмитриев	+38(050) 111-22-33
7	Семенов	+38(066) 000-34-78

Таблица В

Фамилия	Фамилия_1
Иванов	Иванов
Сергеев	Сергеев

**$C = (A \text{ TIMES } B) \text{ WHERE } A.ID = B.ID$**   
Таблица C[A.Фамилия, B.Фамилия]

# Выводы

1. **База данных (БД, DB)** — совокупность самостоятельных данных любого типа, систематизированных таким образом, чтобы эти данные могли быть найдены и обработаны с помощью электронной вычислительной машины (ЭВМ, компьютер);
2. **Система управления базами данных (СУБД)** — специальное программное обеспечение, которое управляет созданием и использованием баз данных с помощью ЭВМ. Благодаря систематизации данных, хранимых в базе СУБД может автоматически обрабатывать и находить данные по запросу;
3. В рамках данного курса будут рассмотрена реляционная локальная клиент-серверная СУБД PostgreSQL (Реляционная и локальная???)
4. **Реляционная база данных** — база данных, основанная на реляционной модели данных;
5. Реляционная база данных представляет собой множество взаимосвязанных **таблиц**, каждая из которых содержит информацию об объектах определенного вида;
6. **Таблица реляционной базы данных** — объект базы данных, который хранит структурированную информацию в виде набора строк и столбцов.
7. Каждая строка (запись) таблицы содержит данные об одном объекте (например, контакте из телефонной книжки), а столбцы (поля) таблицы содержат различные характеристики этих объектов - атрибуты (например, фамилия, имя, номер телефона);
8. В поле таблицы можно хранить любой тип информации: строка, целое число, число с плавающей точкой, логическое значение, дату, время, дату и время, текст, массив байт.
9. Реляционная алгебра определяет следующие операции над отношениями (**таблицами**): Переименование атрибутов, Объединение, Пересечение, Вычитание, Декартово произведение, Выборка (ограничение), Проекция, Соединение (комбинация декартового произведения, выборки и проекции).