

Федеральное государственное автономное образовательное учреждение
высшего образования
"Национальный исследовательский университет
"Высшая школа экономики"

Московский институт электроники и математики им. А.Н Тихонова

Департамент компьютерной инженерии

ИЗУЧЕНИЕ ОСНОВ ЯЗЫКА SQL

**Методические указания к лабораторным работам № 1-4
по курсу "Базы данных"**

Москва

2022

Составитель: доцент, канд. техн. наук И.П. Карпова
аспирант А.С. Вендин

УДК 681.3

Изучение основ языка SQL: Методические указания к лабораторным работам 1–4 по курсу "Базы данных" / Московский институт электроники и математики НИУ ВШЭ; Сост.: И.П. Карпова, А.С. Вендин. – М., 2022. – 33 с.

Лабораторные работы посвящены изучению SQL – языка запросов к базам данных. Работы включают изучение операций реляционной алгебры, операторов определения данных, простых и сложных запросов к данным, представлений.

Для студентов II-IV курсов технических факультетов, изучающих автоматизированные информационные системы и системы баз данных.

Табл. 5. Ил. 6. Библиогр.: 5 назв.

Содержание

| | |
|--|----|
| ЦЕЛЬ ВЫПОЛНЕНИЯ РАБОТ | 4 |
| 1. ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ | 4 |
| 1.1. Общие положения..... | 4 |
| 1.2. Терминология..... | 4 |
| 1.3. Операции реляционной алгебры..... | 5 |
| 1.3.1. Основные операции реляционной алгебры..... | 5 |
| 1.3.2. Вспомогательные операции реляционной алгебры..... | 7 |
| 1.4. Введение в язык SQL..... | 8 |
| 1.4.1. Создание таблиц..... | 9 |
| 1.4.2. Команды модификации данных | 13 |
| 1.4.3. Извлечение данных из таблиц | 15 |
| 1.4.4. Операторы и предикаты | 18 |
| 1.4.5. Функции агрегирования | 21 |
| 1.4.6. Запрос SELECT на нескольких таблицах..... | 23 |
| 1.4.7. Подзапросы..... | 24 |
| 1.4.8. Самосоединение | 27 |
| 1.4.9. Замечания по использованию NULL-значений | 27 |
| 1.4.10. Оператор CASE | 28 |
| 1.4.11. Работа с представлениями..... | 30 |
| 1.4.12. Удаление объектов базы данных..... | 32 |
| 2. ВЫПОЛНЕНИЕ ЛАБОРАТОРНЫХ РАБОТ | 32 |
| Библиографический список..... | 33 |

ЦЕЛЬ ВЫПОЛНЕНИЯ РАБОТ

Цель выполнения лабораторных работ – изучение основ языка SQL и получение практических навыков работы с реляционными базами данных. Выполнение работ включает создание отношений (таблиц) базы данных, заполнение их данными, формирование запросов к данным и создание представлений.

1. ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

1.1. Общие положения

SQL (Structured Query Language) – это структурированный язык запросов к реляционным базам данных (БД). Базы данных, основанные на реляционной модели данных (РМД), являются в настоящее время наиболее широко распространёнными вследствие своей простоты и универсальности методов обработки данных. SQL является декларативным (не процедурным) языком, основанным на операциях реляционной алгебры.

Существуют три стандарта SQL [1], принятые ISO (International Standards Organization): SQL-89 (SQL-1), SQL-92 (SQL-2) и SQL-99/2003/2016 (SQL-3).

Большинство реляционных систем управления базами данных (СУБД) поддерживают стандарт SQL-92, который принят ISO в качестве основного стандарта. Многие версии (диалекты) имеют свои отличия, которые касаются, в основном, типов данных, синтаксиса отдельных команд и некоторых дополнительных возможностей.

1.2. Терминология

В основе РМД лежит понятие **отношения**, представляющего собой подмножество декартова произведения доменов [2, 3]. Отношение обозначается как R_{A_1, A_2, \dots, A_n} , где R – имя отношения, а A_i – атрибут, определенный на i -м домене. Элементы отношения $(a_1, a_2, \dots, a_n)_k$ называют **кортежами**. Длина кортежа (количество атрибутов, n) определяет *арность отношения*, количество кортежей – *мощность отношения*.

Обращение к отношению происходит по имени. Каждый атрибут отношения также имеет имя, принадлежит к определённому типу данных и характеризуется размером памяти, выделяемой под его хранение. Перечень атрибутов отношения с их типами и размерами называется **схемой отношения**.

Два отношения R_{A_1, A_2, \dots, A_n} и Q_{B_1, B_2, \dots, B_m} называют **односхемными**, если:

- 1) $n=m$,
- 2) $A_i \in S$ и $B_i \in S, \forall i=1, \dots, n$,

где S – семейство типов данных (например, все числовые типы образуют одно семейство, все символьные типы – другое семейство, все календарные типы – третье). Если хотя бы одно из условий 1)-2) не выполняется, то отношения являются **разносхемными**.

На атрибут (группу атрибутов) могут накладываться **ограничения целостности** – правила, которым должно соответствовать значение атрибута (или соотношение значений атрибутов).

Атрибут (группа атрибутов), значения которого являются уникальными в рамках отношения, идентифицируют кортеж и называются *потенциальными ключами*. Если ключ состоит из нескольких атрибутов, он называется *составным*. Ключей может быть несколько; основным является **первичный ключ** (*primary key*), его значения обязательны для каждой строки. Таблица может иметь только один первичный ключ (ПК).

Отношение обладает двумя важными свойствами:

1. В отношении не должно быть одинаковых кортежей, т.к. это множество.
2. Порядок кортежей в отношении несущественен.

Связи между отношениями реализуются с помощью внешнего ключа. **Внешний ключ** (*foreign key*) – это атрибут подчинённого (дочернего) отношения, который является копией первичного (*primary key*) или уникального (*unique*) ключа родительского отношения.

В стандартах SQL используются другие термины: отношение принято называть *таблицей*, кортеж – *строкой*, а атрибут – *столбцом* таблицы.

Каждая таблица хранит данные об одном типе объекта (сущности) предметной области (ПрО), причём одна строка таблицы содержит данные об одном **экземпляре** объекта данного типа. Например, таблица СТУДЕНТЫ может хранить данные обо всех студентах института, а отдельная строка этой таблицы – данные о конкретном студенте.

Данные в отношениях обрабатываются с помощью операций реляционной алгебры.

1.3. Операции реляционной алгебры

Операции реляционной алгебры (РА) применимы к реляционным отношениям (таблицам). Результатом выполнения операции реляционной алгебры также является отношение, построенное на основе одного или более исходных отношений. Различают унарные и бинарные операции: унарные выполняются над одним отношением, бинарные – над двумя.

Существует пять основных операций РА и три вспомогательных, которые могут быть выражены через основные. Их можно разбить на две группы [4]:

- 1) теоретико-множественные операции: объединение, пересечение, вычитание и декартово произведение;
- 2) специальные реляционные операции: выборка (селекция), проекция, соединение и деление.

1.3.1. Основные операции реляционной алгебры

1. Проекция (project).

Проекцией π отношения R_{A_1, A_2, \dots, A_n} на атрибуты A_i, A_j, \dots, A_k называют отношение R_{A_i, A_j, \dots, A_k} , где $\{A_i, A_j, \dots, A_k\} \subset \{A_1, A_2, \dots, A_n\}$.

Таким образом, это унарная операция, служащая для выбора подмножества атрибутов из отношения R . Она уменьшает арность отношения и может уменьшить мощность отношения за счёт исключения одинаковых кортежей.

Пример 1. Пусть имеется отношение $R_{A,B,C}$ (рис.1,а).

Тогда проекция $\pi_{A,C}(R)$ будет такой, как показано на рис.1,б.

| Отношение R | | |
|-------------|---|---|
| A | B | C |
| a | b | c |
| c | a | d |
| c | b | d |

| Проекция $\pi_{A,C}(R)$ | |
|-------------------------|---|
| A | C |
| a | c |
| c | d |

а) б)
Рис.1. Пример проекции отношения

2. Селекция (select).

Селекцией σ отношения R_{A_1, A_2, \dots, A_n} по условию F называют отношение R_{A_1, A_2, \dots, A_n} , содержащее кортежи исходного отношения, в которых значения атрибутов удовлетворяют условию F .

Таким образом, это унарная операция, результатом которой является подмножество кортежей исходного отношения, удовлетворяющих условиям на значения одного или нескольких атрибутов. Арность отношения в результате селекции не меняется.

Пример 2. Для отношения $R_{A,B,C}$ (рис. 2,а) селекция $\sigma_{C=d}(R)$ (при условии "значение атрибута C равно d") будет такой (рис. 2,б):

| Отношение R | | |
|-------------|---|---|
| A | B | C |
| a | b | c |
| c | a | d |
| c | b | d |

| Селекция $\sigma_{C=d}(R)$ | | |
|----------------------------|---|---|
| A | B | C |
| c | a | d |
| c | b | d |

а) б)
Рис.2. Пример селекции отношения

3. Декартово произведение (cartesian product).

Декартовым произведением отношений R_{A_1, \dots, A_n} и S_{B_1, \dots, B_k} называют отношение $T_{A_1, \dots, A_n, B_1, \dots, B_k}$, содержащее все комбинации кортежей исходных отношений.

Таким образом, это бинарная операция над разносхемными отношениями, соответствующая определению декартова произведения для РМД [3]. Кортежи результирующего отношения состоят из всех атрибутов исходных отношений.

Пример 3. Пусть имеются отношение $R(A,B)$ и отношение $S(C,D,E)$ (рис. 3,а).

Тогда декартово произведение $R \times S$ будет таким (рис. 3,б):

| Отношение R | |
|-------------|---|
| A | B |
| 1 | 4 |
| 2 | 5 |
| 3 | 6 |

| Отношение S | | |
|-------------|---|---|
| C | D | E |
| g | h | a |
| a | b | c |

| Декартово произведение R×S | | | | |
|----------------------------|---|---|---|---|
| A | B | C | D | E |
| 1 | 4 | g | h | a |
| 1 | 4 | a | b | c |
| 2 | 5 | g | h | a |
| 2 | 5 | a | b | c |
| 3 | 6 | g | h | a |
| 3 | 6 | a | b | c |

а) б)
Рис.3. Пример декартова произведения отношений

4. Объединение (union).

Объединение – бинарная операция над односхемными отношениями. Объединением отношений R_{A_1, \dots, A_n} и S_{A_1, \dots, A_n} называют отношение T_{A_1, \dots, A_n} , при этом кортеж t_i входит в T , если он принадлежит R или S :

$$t_i \in T: t_i \in R \sqcup t_i \in S.$$

То есть объединение включает в себя все кортежи обоих исходных отношений без повторов.

Примечание: имена полей односхемных отношений могут быть разными, достаточно, чтобы совпадало количество полей и типы данных соответствующих полей.

5. Разность (except).

Разность – бинарная операция над односхемными отношениями. Разностью $T = R - S$ отношений R_{A_1, \dots, A_n} и S_{A_1, \dots, A_n} называется отношение T , включающее множество кортежей отношения R , не входящих в S :

$$t_i \in T: t_i \in R \sqcap \text{not}(t_i \in S).$$

Пример 4. Пусть имеются отношение $R(A,B,C)$ и отношение $S(A,B,C)$ (рис. 4,а). Тогда разность $R-S$ будет такой (рис. 4,б):

| Отношение R | | |
|-------------|---|---|
| A | B | C |
| a | b | c |
| c | a | d |
| c | h | c |

| Отношение S | | |
|-------------|---|---|
| A | B | C |
| g | h | a |
| a | b | c |
| h | d | d |
| d | a | x |

| Разность R-S | | |
|--------------|---|---|
| A | B | C |
| c | a | d |
| c | h | c |

а) б)
Рис.4. Пример разности отношений

1.3.2. Вспомогательные операции реляционной алгебры

6. Пересечение (intersect).

Пересечение – бинарная операция над односхемными отношениями. Пересечением $T = R \cap S$ отношений R_{A_1, \dots, A_n} и S_{A_1, \dots, A_n} называют отношение T_{A_1, \dots, A_n} , при этом кортеж t_i входит в T , если он принадлежит обоим исходным отношениям:

$$t_i \in T: t_i \in R \sqcap t_i \in S.$$

Пересечение можно выразить через разность:

$$R \cap S = R - (R - S).$$

7. Соединение (join).

Соединение – бинарная операция над разносхемными отношениями. Эта операция определяет подмножество декартова произведения двух разносхемных отношений. Кортеж декартова произведения входит в результирующее отношение, если выполняется условие соединения F , которое задаёт соотношение значений атрибутов разных таблиц. Соединение может быть выражено так:

$$R \bowtie_F S = \sigma_F (R \times S)$$

Если условием является равенство значений двух атрибутов исходных отношений, такая операция называется **эквисоединением**. *Естественным* называется эквисоединение по одинаковым атрибутам исходных отношений. (Одинаковые атрибуты отношений – это, например, первичный ключ родительского отношения и внешний ключ дочернего отношения).

Пример 5. Пусть имеются отношения $R(A,B,C)$ и $S(A,D,E)$ (рис. 5,а). Тогда естественное соединение $R \bowtie S$ будет таким, как показано на рис. 5,б.

| Отношение R | | |
|-------------|---|---|
| A | B | C |
| a | b | c |
| c | a | d |
| c | h | c |
| g | b | d |

| Отношение S | | |
|-------------|---|---|
| A | D | E |
| g | h | a |
| c | b | c |
| h | d | d |

| Соединение $R \bowtie S$ | | | | |
|--------------------------|---|---|---|---|
| A | B | C | D | E |
| c | a | d | b | c |
| c | h | c | b | c |
| g | b | d | h | a |

Рис.5. Пример естественного соединения отношений

8. Деление (division).

Пусть отношение R содержит атрибуты $\{r_1, r_2, \dots, r_k, r_{k+1}, \dots, r_n\}$, а отношение S – атрибуты $\{r_{k+1}, \dots, r_n\}$. Тогда результирующее отношение T содержит атрибуты $\{r_1, r_2, \dots, r_k\}$. Кортеж t_i входит в T , если его декартово произведение с отношением S входит в R :

$$t_i \in T : t_i \times S \subset R.$$

Деление может быть выражено так:

$$R / S = \pi_{r_1, \dots, r_k}(R) - \pi_{r_1, \dots, r_k}((\pi_{r_1, \dots, r_k}(R) \times S) - R).$$

Пример 6. Пусть имеются отношения $R(A,B,C,D)$ и $S(C,D)$ (рис. 6,а). Тогда частное R/S будет таким, как показано на рис. 6,б.

| Отношение R | | | |
|-------------|---|---|---|
| A | B | C | D |
| a | b | c | b |
| a | b | g | h |
| c | f | g | h |
| c | f | c | b |
| a | v | c | b |
| c | v | g | h |

| Отношение S | |
|-------------|---|
| C | D |
| c | b |
| g | h |

| Частное R/S | |
|-------------|---|
| A | B |
| a | b |
| c | f |

Рис.6. Пример операции деления

1.4. Введение в язык SQL

Язык работы с базами данных должен предоставлять пользователям следующие возможности:

- создавать базу данных, включая таблицы с полным описанием их структуры;
- выполнять основные операции манипулирования данными (добавление, изменение, удаление данных);
- выполнять запросы, осуществляющие преобразование данных в необходимую информацию.

Создание базы данных обычно выполняется автоматически при установке самой СУБД, поэтому в данном пособии не рассматривается.

Для реализации этих функций SQL включает три группы средств:

- DDL (Data Definition Language) – язык определения данных;
- DML (Data Manipulation Language) – язык манипулирования данными;
- DCL (Data Control Language) – язык управления данными.

По стандарту ANSI подмножество команд DCL является частью DDL.

В командах SQL не различаются прописные и строчные буквы (за исключением строчных литералов). Символы и строки символов заключаются в **одинарные кавычки**, например, 'N', 'учебник'. Однострочные комментарии начинаются с двух минусов (--), многострочные заключаются в символы /* и */.

Каждая команда заканчивается символом ';'. Значения параметров команд, принятые по умолчанию, выделены подчеркиванием, например, ALL.

Примечание. Примем следующие обозначения для описания синтаксиса:

- { } – содержимое скобок рассматривается как единое целое для остальных символов;
- | – заменяет слово ИЛИ;
- [] – содержимое этих скобок является необязательным;
- < > – содержимое этих скобок заменяется соответствующими ключевыми словами, литералами, идентификаторами или выражениями (в зависимости от контекста);
- ... – всё, что предшествует этим символам, может повторяться произвольное число раз;
- ..., – всё, что предшествует этим символам, может повторяться произвольное число раз, каждое вхождение отделяется запятой.

В соответствии со стандартов ISO идентификатор определяется как последовательность символов длиной не более 128, начинающаяся с буквы латинского алфавита и содержащая буквы латинского алфавита, цифры и знак подчеркивания '_'. В большинстве СУБД накладываются более жёсткие ограничения на длину идентификатора, но можно использовать знаки _, # и \$ как буквы.

Синтаксис команд и примеры, приведённые в данном пособии, соответствуют синтаксису СУБД PostgreSQL.

Описание процесса установки PostgreSQL 11 можно посмотреть по адресу: <https://info-comp.ru/sisadminst/684-install-postgresql-11-on-windows.html>

1.4.1. Создание таблиц

Создание нового пустого отношения (таблицы) выполняется с помощью команды DDL **CREATE TABLE**. Упрощённый синтаксис этой команды:

```
CREATE TABLE [ IF NOT EXISTS ] table_name ( [  
  { column_name data_type [ column_constraint [ ... ] ]  
  | table_constraint }  
  [, ... ]  
  ] )
```

Ограничения целостности столбца имеет следующий синтаксис:

column_constraint:

```
[ CONSTRAINT constraint_name ]  
{ NOT NULL |  
  NULL |
```

```
CHECK ( expression ) |
DEFAULT default_expr |
UNIQUE index_parameters |
PRIMARY KEY index_parameters |
REFERENCES reftable [ ( refcolumn ) ]
[ ON DELETE action ] [ ON UPDATE action ] }
```

table_constraint :

```
[ CONSTRAINT constraint_name ]
{ CHECK ( expression ) |
  UNIQUE ( column_name [, ... ] ) index_parameters |
  PRIMARY KEY ( column_name [, ... ] ) index_parameters |
  FOREIGN KEY ( column_name [, ... ] )
    REFERENCES reftable [ ( refcolumn [, ... ] ) ] ]
[ ON DELETE action ] [ ON UPDATE action ] }
```

При указании IF NOT EXISTS таблица будет создана только в том случае, если такая таблица не существует. Если таблица (или другой объект с таким именем) существует, будет выдано сообщение об ошибке.

Расшифровка элементов описания приведена в табл. 1.

Обратите внимание: в качестве имен таблиц, полей и т.д. нельзя использовать ключевые слова языка SQL!

Таблица 1. Описание элементов команды CREATE TABLE

| Элемент | Описание |
|--------------------------------|--|
| <i>table_name</i> | Имя таблицы, обычный идентификатор. Должно быть уникальным среди имен объектов (таблиц и т.д.) в рамках базы данных. |
| <i>column_name</i> | Имя поля (столбца) таблицы, обычный идентификатор. |
| <i>data_type</i> | Тип данных поля. Можно использовать один из следующих типов: – INT, SMALLINT, BIGINT – целые числа; – NUMERIC [(длина [, точность])] – числа с фиксированной запятой; – FLOAT4, FLOAT8, REAL – вещественные числа; – CHAR[(длина)] – символьные строки фиксированной длины – VARCHAR(длина) – символьные строки переменной длины; – DATE, TIMESTAMP – дата, дата/время; – TIME – время. (Более подробно о типах данных можно прочитать в документации: https://www.postgresql.org/docs/9.6/datatype.html). |
| DEFAULT <i>default_expr</i> | Выражение, значение которого присваивается полю по умолчанию, если при добавлении записи значение этого поля не вводится явно. |
| <i>column_constraint</i> | Ограничения целостности поля: – NULL – поле может не иметь значения (быть пустым); – NOT NULL – данный столбец не принимает значения NULL; – PRIMARY KEY – первичный ключ (обязательный, уникальный и |

| | |
|---|---|
| | <p>единственный на таблицу);</p> <ul style="list-style-type: none"> – UNIQUE – уникальность значения поля в пределах таблицы; – CHECK (<i>expression</i>) – условие, которому должно удовлетворять значение поля; – REFERENCES <i>reftable</i> [(<i>refcolumn</i>)] – внешний ключ, который ссылается на столбец <i>refcolumn</i> родительской таблицы <i>reftable</i>. |
| <i>table_constraint</i> | <p>Ограничения целостности таблицы:</p> <ul style="list-style-type: none"> – CHECK (<i>expression</i>) – условие, которому должны удовлетворять значения нескольких полей; – FOREIGN KEY [(<i>field_list1</i>)] REFERENCES <i>reftable</i> [(<i>field_list2</i>)] – составной внешний ключ: <i>field_list1</i> – список полей, принадлежащих текущей таблице, <i>field_list2</i> – поля таблицы <i>reftable</i>; – PRIMARY KEY (<i>field_list</i>) – составной первичный ключ; – UNIQUE (<i>field_list</i>) – уникальность комбинации значений полей в пределах таблицы. |
| ON DELETE <i>action</i> (ON UPDATE <i>action</i>) | <p><i>action</i> – действие, производимое при удалении строки (обновлении значения уникального поля) в родительской таблице:</p> <p>NO ACTION – в случае нарушения ограничения целостности внешнего ключа (ВнК) система выдаст сообщение об ошибке. Для отложенных ограничений ошибка произойдёт в момент их проверки, если строки, ссылающиеся на удалённую запись, по-прежнему будут существовать. Этот вариант действия подразумевается по умолчанию.</p> <p>RESTRICT – в случае нарушения ограничения целостности ВнК система выдаст сообщение об ошибке, как только обнаружит нарушение (аналогично NO ACTION), но эта проверка будет неоткладываемой.</p> <p>CASCADE – каскадное удаление строк в подчиненной таблице.</p> <p>SET NULL – установка значения NULL в поле (поля) внешнего ключа.</p> <p>SET DEFAULT – установка значения по умолчанию в поле (поля) ВнК.</p> <p>Обратите внимание: если значение по умолчанию не удовлетворяет ограничению внешнего ключа, операция завершится ошибкой!</p> |

Если размер поля не указан, то принимается значение, принятое в данной СУБД по умолчанию для данного типа. Для всех СУБД по умолчанию точность для числовых типов NUMERIC (или NUMBER) равна 0, размер поля типа CHAR равен 1, для типа VARCHAR размер указывать обязательно, а для календарных типов размер указывать нельзя.

Для типов DATE и TIME поддерживается арифметика дат, например:

(<текущая дата> + 1) – завтра;

(<дата1> – <дата2>) – количество дней, прошедших между двумя датами;

(<текущее время> + '1 hour') – через час;

(<текущее время> – '20 minute') – 20 минут назад.

Получить текущую дату можно с помощью специальной функции, имя которой зависит от СУБД: *sysdate* – для Oracle; *now()* – для Access; *getdate()* – для MS SQL Server и т.д. В PostgreSQL для получения текущей даты/времени соответствующего типа можно использовать следующие SQL-совместимые функции: *current_date*, *current_time*, *current_timestamp*, *localtime* и *localtimestamp*.

Стандарт SQL включает понятие неопределённого или неизвестного значения – NULL-значения. Для обязательных полей устанавливается ограничение NOT NULL. Это означает, что при изменении значения этого поля или при до-

бавлении новых записей таблицы значение этого поля не может быть неопределенным (NULL). Ограничение NOT NULL можно наложить на поле только один раз, иначе возникает ошибка.

Для любого поля с помощью конструкции DEFAULT default_expr может быть задано значение по умолчанию. Оно используется в тех случаях, когда при добавлении данных в таблицу значение этого поля не указывается.

Для ограничений целостности (ОЦ) можно задавать имена:

CONSTRAINT name constraint_definition

Если имена ограничениям целостности не задавать, то система сама назначит эти имена. Например, в Oracle сообщение о нарушении такого ограничения целостности может выглядеть так: ORA-02290: check constraint (SYSTEM.SYS_C0018280) violated (где SYS_C0018280 – название ОЦ).

Примеры создания таблиц:

1. Таблица "Отделы" с полями "Номер отдела" (ПК), "Название отдела":

```
CREATE TABLE depart
(   depno   NUMERIC(2)   CONSTRAINT pk_dep PRIMARY KEY,
    name    VARCHAR(80)  NOT NULL);
```

2. Таблица "Сотрудники" с полями "Номер отдела" (внешний ключ), "Табельный номер сотрудника" (ПК), "ФИО сотрудника", "Должность", "Оклад", "Дата рождения", "Пол", "Телефон", "Дата поступления на работу":

```
CREATE TABLE emp
(   depno   NUMERIC(2)   CONSTRAINT ref_dep REFERENCES depart,
    tabno   CHAR(5)      CONSTRAINT pk_emp PRIMARY KEY,
    name    VARCHAR(50)  NOT NULL,
    post    VARCHAR(35)  NOT NULL,
    salary  NUMERIC(7,2) NOT NULL DEFAULT 15000
        CONSTRAINT check_salary CHECK (salary > 12000),
    born    DATE         NOT NULL,
    sex     CHAR         NOT NULL,
    phone   CHAR(11),
    edate   DATE         DEFAULT current_date);
```

3. Таблица "Дети сотрудников" с полями "Табельный номер родителя" (внешний ключ), "Имя ребенка", "Пол", "Дата рождения":

```
CREATE TABLE children
(   tabno   CHAR(5)   CONSTRAINT ref_emp REFERENCES emp(tabno),
    name    VARCHAR(50) NOT NULL,
    sex     CHAR,
    born    DATE,
    CONSTRAINT pk_child PRIMARY KEY (tabno, name),
        /* составной первичный ключ */
    CONSTRAINT check_sex CHECK (sex IN ('м', 'ж')));
```

Обратите внимание:

- общие ограничения целостности (ограничения на таблицу) и составные ключи указываются через запятую после последнего поля;

- если внешний ключ ссылается на первичный ключ (ПК) другого отношения, имена полей ПК можно не указывать (см. создание таблицы *etp*);
- типы полей внешнего ключа **должны совпадать** с типами полей первичного (или уникального) ключа, на который он ссылается;
- если внешний ключ составной, список полей входящий в ключ, указывается после перечисления всех полей таблицы с ключевым словом **FOREIGN KEY**:

```
create table exam    -- "Расписание зачетов", основная таблица
(   eclass  NUMERIC(3),      -- аудитория
    edate    DATE,           -- дата
    edisc    VARCHAR(60),    -- дисциплина
    UNIQUE (eclass, edate));

create table tab     -- "Зачеты", подчинённая таблица
(   tid     NUMERIC(6) PRIMARY KEY,
    tclass  NUMERIC(3),      -- аудитория
    tdate   DATE,           -- дата
    tgroup  CHAR(6),        -- группа
    FOREIGN KEY (tclass, tdate) REFERENCES exam(eclass, edate));
```

После создания таблицы ее можно заполнять данными.

1.4.2. Команды модификации данных

К командам модификации данных (DML) относятся добавление, удаление и изменение (обновление) записи (строки таблицы). При выполнении этих команд проверяются все установленные для таблицы ограничения целостности. В случае нарушения любого ограничения целостности или возникновения других проблем (переполнение памяти, например) команда DML не выполняется и выдаётся сообщение об ошибке. Если же команда выполнилась успешно, выдается информация о количестве обработанных строк.

INSERT – добавление записи в таблицу. Синтаксис команды:

```
INSERT INTO <имя таблицы> [(<список имен полей>)]
VALUES (<список выражений>) | <запрос>;
```

Под <запросом> подразумевается команда **SELECT** (см. ниже), результаты работы которой добавляются в указанную таблицу.

В предложении **VALUES** указываются выражения, порождающие значения атрибутов новой записи таблицы. **Выражение** может включать вызовы функций, определенных в данной СУБД, константы, знак операций конкатенации строк (|| – две вертикальных черты подряд) или знаки арифметических операций: –, +, *, /. Типы значений выражений должны соответствовать типам полей таблицы. Если это требование не соблюдается, система пытается применить неявное преобразование данных к типу поля, в которое вносится значение. Если ей это не удастся, она выводит сообщение об ошибке и не добавляет запись в таблицу. Строки и даты должны заключаться в одинарные кавычки. Формат даты в СУБД PostgreSQL желательно указывать как 'YYYY-MM-DD' или 'YYYY/MM/DD' (например, '2022-10-20').

Если значения устанавливаются не для всех полей или порядок значений не соответствует тому порядку полей, который был установлен при создании таблицы, то после имени таблицы в скобках приводится список полей в соответствии со списком значений.

В тех случаях, когда при добавлении записи значение какого-либо поля неизвестно, его можно не устанавливать, пропустив это поле в списке полей или указав для него значение **NULL** (но только для тех полей, на которые не наложено ограничение целостности **NOT NULL**).

Если в списке полей отсутствует какое-либо поле таблицы, то ему будет присвоено значение **NULL** или значение по умолчанию (**DEFAULT**), если оно определено в командах **CREATE TABLE** или **ALTER TABLE**.

Для ясности можно также явно указать значения по умолчанию для отдельных столбцов или всей строки:

```
INSERT INTO <имя таблицы> VALUES (DEFAULT, DEFAULT);
```

```
INSERT INTO <имя таблицы> VALUES DEFAULT VALUES;
```

Одна команда может вставить сразу несколько строк:

```
INSERT INTO <имя таблицы> VALUES (), (), ();
```

Пример: Добавить в таблицу "Сотрудники" новую запись:

```
INSERT INTO emp (depno, tabno, name, post, salary, born, phone)
VALUES(3, '00112', 'Попов В.Г.', 'экономист', 45400, '1979-12-23', '115-34-11');
```

1 строка создана.

В данном случае дата рождения вводится как строка '1979-12-23'. А в качестве даты поступления сотрудника на работу будет установлена текущая дата, т.к. для поля **edate** определено значение по умолчанию **DEFAULT** и в команде **INSERT** значение не вводится.

Пример: Добавить в таблицу "Дети" 2 записи:

```
INSERT INTO children VALUES
(819, 'Иван', 'м', '2020-01-10'), (819, 'Мария', 'ж', '2020-01-10');
```

UPDATE – обновление данных в таблице. Синтаксис:

```
UPDATE <имя таблицы>
SET {<имя поля> = <выражение>}...
[WHERE <условие>];
```

Команда обновления изменяет в указанной таблице значения указанных полей тех записей, которые удовлетворяют заданному условию отбора (**WHERE <условие>**). Если условие не указано, обновления применяются ко всем записям таблицы.

Пример: Изменить должность и зарплату сотрудника Попова В.Г., табельный номер '00112':

```
UPDATE emp
SET post = 'ст. экономист', salary = salary*1.1
WHERE tabno = '00112';
```

1 строка обновлена.

DELETE – удаление записей из таблицы. Синтаксис этой команды:

```
DELETE FROM <имя таблицы>  
[ WHERE <условие> ];
```

Эта команда удаляет из указанной таблицы те записи, которые удовлетворяют заданному условию отбора (**WHERE** <условие>).

Внимание! Если не указывать условие выбора записей, то все записи таблицы будут удалены без предупреждения и без запроса на подтверждение!

Пример: Удалить запись о сотруднике Попове В.Г., табельный номер '00112':

```
DELETE FROM emp  
WHERE tabno = '00112';  
1 строка удалена.
```

1.4.3. Извлечение данных из таблиц

Извлечение данных из таблиц БД выполняется с помощью команды **SELECT** (селекция). Эта команда не изменяет данные в БД.

Результатом выполнения команды **SELECT** является временная таблица, которая помещается в курсор (специальную область памяти СУБД) и обычно сразу выводится на экран. Упрощённый синтаксис этой команды:

```
SELECT [ ALL | DISTINCT ] <список вывода>  
FROM <имя таблицы 1> [ <алиас имени таблицы 1> ]  
[ , ... [ , имя таблицы n> [ <алиас имени таблицы n> ] ] ]  
[ WHERE <условие отбора записей> ]  
[ GROUP BY <имя поля 1> [ , ... [ , <имя поля k> ] ] ]  
[ HAVING <условие отбора групп> ]  
{ UNION [ ALL ] | INTERSECT | EXCEPT }  
SELECT ... – второй запрос ]  
[ ORDER BY <имя поля 1> [ ASC | DESC ]  
[ , ... [ , <имя поля m> [ ASC | DESC ] ] ] ] ];
```

<список вывода>:

```
{ * | <выражение 1> [ AS <алиас поля 1> ]  
[ , ... [ , <выражение i> [ AS <алиас поля i> ] ] ] }
```

Порядок конструкций в команде **SELECT** не может быть изменён.

Расшифровка элементов описания приведена в табл. 2.

Таблица 2. Элементы команды SELECT

| Элемент | Описание |
|-----------------|--|
| <список вывода> | Все, что находится между словами SELECT и FROM, называется списком вывода (или списком выбора). Он может состоять из спис- |

| | |
|-------------------------|---|
| | ка элементов, разделённых запятыми, или символа *. Выражение может включать имена полей таблиц, знаки операций, вызовы функций и константы. |
| <имя таблицы> | Имя таблицы или представления (представления – см. раздел 1.4.11). |
| <алиас> | Временный синоним имени таблицы, определённый только внутри данного запроса . |
| <условие выбора строки> | Условие, которое может быть истинным или ложным для каждой записи из таблицы (таблиц), определённых предложением FROM. |
| <условие выбора группы> | Условие, которое может быть истинным или ложным для каждой группы записей, полученных после вычисления GROUP BY. |
| <имя поля> | Имя поля (столбца) таблицы. |
| <позиция> | Номер элемента в <списке выбора> (целое число). |

Список выбора (или список вывода) определяет схему результата (временной таблицы). Список выбора может быть модифицирован одним из двух ключевых слов:

DISTINCT – предикат удаления из результата повторяющихся записей.

ALL – предикат, обратный к **DISTINCT**. Это значение используется по умолчанию, его можно не указывать.

Рассмотрим основные предложения (фразы) команды **SELECT**:

SELECT – после этого ключевого слова указывается **список выбора** – список выражений, которые будут образовывать результат. Каждому выражению можно сопоставить временный синоним (алиас столбца), который будет *названием поля* результирующей таблицы, например:

```
(sal*0.87+bonus) AS salary
```

Частица **AS** здесь является необязательной. Если название столбца результата содержит пробелы или другие запрещенные для идентификаторов символы, его нужно заключать в кавычки (например, 'зарплата, руб.').

Алиас можно использовать также в **group by** и **order by**, например:

```
select depno dep, name from emp order by dep;
```

Если надо вывести все поля из тех таблиц, к которым обращается данный запрос, можно указать символ *. В этом случае сначала будут выведены поля таблицы, стоящей первой в предложении **FROM**, затем – второй и т.д. Поля, относящиеся к одной таблице, будут выводиться в том порядке, в каком они были указаны при создании таблицы.

FROM – в этой части указывается имя таблицы (имена таблиц или представлений, см. далее), из которой будут извлекаться данные. СУБД PostgreSQL, как и некоторые другие системы, позволяет опускать эту часть команды, тогда **SELECT** можно использовать, например, для вызова функций:

```
SELECT current_date;
```

WHERE – эта часть содержит условия выбора отдельных записей.

GROUP BY – объединяет в одну группу все записи с одинаковым значением указанного поля (или комбинации полей). Каждой такой группе в результирующей таблице соответствует **одна запись**.

HAVING – позволяет указать условия выбора для групп записей.

ORDER BY – упорядочивает результирующие строки по значению одного или нескольких полей: **ASC** – по возрастанию, **DESC** – по убыванию.

Порядок выполнения операции **SELECT** такой:

1. Выбор из указанной во фразе **FROM** таблицы тех записей, которые удовлетворяют условию отбора (**WHERE**).
2. Группировка полученных записей (**GROUP BY**).
3. Выбор тех групп, которые удовлетворяют условию отбора групп (**HAVING**).
4. Сортировка записей в указанном порядке (**ORDER BY**).
5. Извлечение из записей тех полей, которые заданы в списке выбора, и формирование результирующей таблицы.

Если во фразе **FROM** указаны две и более таблицы, то эта последовательность действий выполняется для **декартова произведения указанных таблиц**.

Отношения для примеров приведены в таблицах 3-5.

Таблица 3. Отношение "Сотрудники" (Emp)

| TabNo | DepNo | Name | Post | Salary | Born | Phone |
|-------|-------|--------------|-------------------|---------|------------|-----------|
| 988 | 1 | Рюмин В.П. | начальник отдела | 48600.0 | 01.02.1970 | 115-26-12 |
| 909 | 1 | Серова Т.В. | вед. программист | 48600.0 | 20.10.1981 | 115-91-19 |
| 829 | 1 | Саргсян А.В. | экономист | 43500.0 | 03.10.1988 | null |
| 819 | 1 | Тамм Л.В. | экономист | 43500.0 | 13.11.1985 | 115-91-19 |
| 100 | 2 | Волков Л.Д. | программист | 46500.0 | 16.10.1982 | null |
| 110 | 2 | Буров Г.О. | бухгалтер | 42880.0 | 22.05.1975 | 115-46-32 |
| 023 | 2 | Малова Л.А. | гл. бухгалтер | 59240.0 | 24.11.1974 | 114-24-55 |
| 130 | 2 | Лукина Н.Н. | бухгалтер | 42880.0 | 12.07.1979 | 115-46-32 |
| 034 | 3 | Перова К.В. | делопроизводитель | 33000.0 | 24.04.1988 | 115-12-70 |
| 002 | 3 | Сухова К.А. | начальник отдела | 48600.0 | 08.06.1967 | 115-12-69 |
| 056 | 5 | Павлов А.А. | директор | 80000.0 | 05.05.1969 | 115-33-44 |
| 087 | 5 | Котова И.М. | секретарь | 35000.0 | 16.09.1990 | 115-33-65 |
| 088 | 5 | Кроль А.П. | зам. директора | 70000.0 | 18.04.1974 | 115-33-01 |
| 454 | 3 | Беленко П.А. | делопроизводитель | 33000.0 | 20.05.1996 | 115-12-70 |

Таблица 4. Отношение "Отделы" (Depart)

| DepNo | Name |
|-------|-----------------------------|
| 2 | Бухгалтерия |
| 3 | Отдел кадров |
| 4 | Отдел технического контроля |
| 1 | Плановый отдел |
| 5 | Администрация |

Таблица 5. Отношение "Дети"(Children)

| TabNo | Name | Born | Sex |
|-------|-------|------------|-----|
| 988 | Вадим | 03.05.2005 | м |
| 110 | Ольга | 18.07.2010 | ж |
| 087 | Илья | 19.02.2012 | м |
| 087 | Анна | 26.12.2014 | ж |
| 909 | Инна | 25.01.2008 | ж |
| 909 | Роман | 21.11.2006 | м |
| 909 | Антон | 06.03.2009 | м |

Примеры:

1. Вывести все записи (строки) из таблицы "Отделы":

```
SELECT *
FROM depart;
```

| DEPNO | NAME |
|-------|-----------------------------|
| 2 | Бухгалтерия |
| 3 | Отдел кадров |
| 4 | Отдел технического контроля |
| 1 | Плановый отдел |
| 5 | Администрация |

2. Вывести список сотрудников с указанием должности и зарплаты за вычетом подоходного налога, упорядочить по отделам и фамилиям:

```
SELECT depno, name, post, salary*0.87 AS sal  
FROM emp  
ORDER BY depno, name;
```

| DEPNO | NAME | POST | SAL |
|-------|--------------|-------------------|---------|
| 1 | Рюмин В.П. | начальник отдела | 42195.0 |
| 1 | Саргсян А.В. | экономист | 37845.0 |
| 1 | Серова Т.В. | вед. программист | 42195.0 |
| 1 | Тамм Л.В. | экономист | 37845.0 |
| 2 | Буров Г.О. | бухгалтер | 37305.6 |
| 2 | Волков Л.Д. | программист | 40455.0 |
| 3 | Беленко П.А. | делопроизводитель | 27840.0 |
| ... | ... | ... | ... |

3. Вывести список должностей с окладом в порядке убывания оклада:

```
SELECT DISTINCT post, salary  
FROM emp  
ORDER BY 2 DESC; -- 2 – номер элемента списка выбора
```

| POST | SALARY |
|-------------------|--------|
| директор | 80000 |
| зам.директора | 70000 |
| гл. бухгалтер | 59240 |
| вед. программист | 48600 |
| начальник отдела | 48600 |
| ... | ... |
| секретарь | 35000 |
| делопроизводитель | 33000 |

1.4.4. Операторы и предикаты

Расширение возможностей команд языка SQL достигается за счёт применения различных **операторов, предикатов и функций**.

Операторы:

- символные: || – конкатенация строк;
- арифметические: +, -, *, /;
- сравнения: =, >, <, >=, <=, <>;
- логические: AND, OR, NOT.

Примеры:

4. Составить список сотрудников второго и третьего отдела, имеющих зарплату выше 40000 рублей после уплаты подоходного налога 13%:

```
SELECT depno, name, salary*0.87 AS sal
FROM emp
WHERE salary*0.87>40000 AND (depno=2 OR depno=3)
ORDER BY name;
```

| DEPNO | NAME | SALARY |
|-------|-------------|---------|
| 2 | Волков Л.Д. | 40455.0 |
| 2 | Малова Л.А. | 51538.8 |
| 3 | Сухов К.А. | 42195.0 |

5. Составить список сотрудников первого отдела с указанием должности:

```
SELECT post || ' ' || name AS ename
FROM emp
WHERE depno=1
ORDER BY 1 DESC;
```

| ENAME |
|------------------------------|
| экономист Тамм Л.В. |
| экономист Саргсян А.В. |
| начальник отдела Рюмин В.П. |
| вед. программист Серова Т.В. |

Обратите внимание на добавление пробела между полями: || ' ' ||.

Предикаты, используемые в командах:

- **IN** – предикат вхождения в список:
 <выражение> IN (<список значений>)
– определяет множество значений, с которыми будет сравниваться значение <выражения>. Предикат считается истинным, если значение выражения равно хотя бы одному из элементов множества.
- **BETWEEN** – предикат нахождения в диапазоне:
 <выражение> BETWEEN <значение1> AND <значение2>
– определяет, входит ли значение <выражения> в указанные границы. Если значение выражения меньше, чем <значение1>, или больше, чем <значение2>, предикат возвращает "ложь".
- **LIKE** – предикат подобия:
 <выражение> LIKE 'образец'
– используется для сравнения строк, применяется только к полям типа CHAR, VARCHAR. Возможно использование шаблонов (символов расширения):
 '_' – один любой символ,
 '%' – произвольное количество символов (в т.ч., ни одного).
Экранировать специальное значение символов '_' и '%' можно с помощью ключевого слова ESCAPE и символа экранирования, например:
 where <строка> LIKE '_#%%' ESCAPE '#';

Символ экранирования (escape) может быть любым. В примере первый символ % будет искажаться как символ, а второй имеет специальное значение. По умолчанию в качестве специального символа экранирования выбрана обратная косая черта, например:

<имя поля> like '_%' -- найдет все строки типа '1%', '2%' и т.д.

С помощью слова ESCAPE можно выбрать другой, а чтобы включить символ экранирования в шаблон поиска, его надо продублировать.

- **IS NULL** – предикат неопределённого значения:

<выражение> IS NULL

– определяет, установлено ли значение поля; возвращает истину, если не установлено. Другие предикаты и операторы сравнения возвращают неопределённый результат (NULL), если хотя бы один из операндов имеет значение NULL. Значение NULL интерпретируется как "ложь". Например, условие a=null никогда не выполняется (оно тождественно ложно).

Все эти предикаты могут комбинироваться с операцией "не": NOT IN, NOT LIKE, NOT BETWEEN, IS NOT NULL.

Примеры:

6. Вывести список программистов и ведущих программистов:

```
SELECT depno, name, post
FROM emp
WHERE post like '%программист%';
```

| DEPNO | NAME | POST |
|-------|-------------|------------------|
| 1 | Серова Т.В. | вед. программист |
| 2 | Волков Л.Д. | программист |

7. Увеличить на 10% оклады начальникам отделов и программистам:

```
UPDATE emp
SET salary=salary*1.1
WHERE post LIKE 'нач%отдел%' OR post LIKE '%программист%';
```

4 строки обновлено.

8. Вывести список сотрудников старше 40 лет из 1-го и 3-го отделов:

```
SELECT depno, name, extract (years from age(current_date, born)) as age
FROM emp
WHERE depno IN (1, 3) AND
extract (years from age(current_date, born)) > 40;
```

| DEPNO | NAME | AGE |
|-------|-------------|-----|
| 1 | Рюмин В.П. | 49 |
| 3 | Сухова К.А. | 71 |

Примечание: функция age возвращает тип, в котором хранится информация о числе полных лет, месяцев и дней между двумя датами, функция extract(years from age) выбирает из полученного значения только год.

9. Список сотрудников, не имеющих телефонов:

```
SELECT tabno, name, post
FROM emp
WHERE phone IS NULL;
```

| TABNO | NAME | POST |
|-------|--------------|-------------|
| 100 | Волков Л.Д. | программист |
| 829 | Саргсян А.В. | экономист |

10. Список сотрудников, родившихся в 80-е годы XX века:

```
SELECT tabno, name, born, post
FROM emp
WHERE born BETWEEN '1980-01-01' AND '1989-12-31';
```

| TABNO | NAME | BORN | POST |
|-------|--------------|------------|-------------------|
| 909 | Серова Т.В. | 1981-10-20 | вед. программист |
| 829 | Саргсян А.В. | 1988-10-03 | экономист |
| 100 | Волков Л.Д. | 1982-10-16 | программист |
| 034 | Перова К.В. | 1988-04-24 | делопроизводитель |
| 819 | Тамм Л.В. | 1985-11-13 | экономист |

1.4.5. Функции агрегирования

Для подсчёта различных агрегированных значений (по группе записей) стандарт SQL включает т.н. функции агрегирования:

- COUNT(*) – определяет количество строк (записей) в результате.
- MAX(<поле>), MIN(<поле>) – определяет максимальное (минимальное) значение указанного поля в результирующем множестве.
- SUM(<поле>) – определяет арифметическую сумму значений указанного числового поля в результирующем множестве записей.
- AVG(<поле>) – определяет среднее арифметическое значений указанного числового поля в результирующем множестве записей.

Правила уточнения использования агрегирующих функций:

COUNT (<поле>) – подсчёт количества ненулевых значений поля;

COUNT (distinct <поле>) – подсчёт количества разных значений поля;

SUM (distinct <поле>) – суммирование разных значений поля;

AVG (distinct <поле>) – среднее арифметическое разных значений поля.

В некоторых СУБД (Oracle, PostgreSQL) есть еще возможность использовать

COUNT (<условие>):

Примеры:

11. Посчитать количество сотрудников предприятия:

```
SELECT count(*), ' человек(a)'
FROM emp;
```

| | |
|----------|------------|
| count(*) | человек(a) |
| 14 | человек(a) |

12. Определить минимальную и максимальную зарплату сотрудников:

```
SELECT min(salary) AS minsal, max(salary) AS maxsal  
FROM emp;
```

| minsal | maxsal |
|--------|--------|
| 33000 | 80000 |

13. Определить среднюю зарплату сотрудников 3-го отдела:

```
SELECT avg(salary) AS avg3  
FROM emp  
WHERE depno=3;
```

| avg3 |
|-------|
| 38200 |

Агрегирующие функции можно комбинировать с фразой **GROUP BY**: в этом случае подсчёт будет производиться для каждой группы записей с одинаковым значением комбинации полей, указанных в **GROUP BY**.

Примеры:

14. Посчитать количество сотрудников по отделам:

```
SELECT depno, count(*), 'сотрудник(а)'  
FROM emp  
GROUP BY depno;
```

| DEPNO | count(*) | сотрудник(а) |
|-------|----------|--------------|
| 1 | 4 | сотрудник(а) |
| 2 | 4 | сотрудник(а) |
| 3 | 3 | сотрудник(а) |
| 5 | 3 | сотрудник(а) |

15. Сумма зарплат по отделам:

```
SELECT depno, sum(salary) AS sumsal  
FROM emp  
GROUP BY depno;
```

| DEPNO | sumsal |
|-------|--------|
| 1 | 193700 |
| 2 | 193000 |
| 3 | 118350 |
| 5 | 185000 |

16. Количество сотрудников, получающих зарплату более 30000:

```
SELECT count(salary > 30000.00) from emp;
```

При использовании фразы **GROUP BY** существует правило, которого надо строго придерживаться: в списке выбора могут быть указаны **только функции агрегирования, константы и поля, перечисленные в GROUP BY**. Нарушение этого правила является синтаксической ошибкой.

Если при использовании фразы **GROUP BY** необходимо вывести не все группы, а только те, которые удовлетворяют некоторому условию, то условие

для результатов агрегирующих функций можно указать только в части **HAVING**. Это ограничение определяется порядком вычисления команды **SELECT**: если указать условие на агрегирующую функцию в части **WHERE**, то на момент проверки этого условия значения агрегирующих функций ещё не подсчитано, поэтому его невозможно проверить. Ниже приведён пример использования фразы **HAVING**.

17. Вывести список отделов, в которых количество сотрудников больше трёх:

```
SELECT depno as "Отдел", count(*) as "Количество сотрудников"
FROM emp
GROUP BY depno
HAVING count(*)>3;
```

| Отдел | Количество сотрудников |
|-------|------------------------|
| 1 | 4 |
| 2 | 4 |

1.4.6. Запрос **SELECT** на нескольких таблицах

Запрос **SELECT** на нескольких таблицах основан на декартовом произведении исходных таблиц. Если указать условие соответствия значений полей разных таблиц, то получится соединение таблиц. Для полей с одинаковыми названиями нужно указывать имя таблицы (или алиас) перед именем поля, разделяя их точкой (т.н. квалифицированная ссылка).

Примеры:

18. Запрос по двум таблицам. Вывести список сотрудников с детьми:

```
SELECT e.name, c.name AS child, c.born
FROM emp e, children c /* e, c – алиасы таблиц */
WHERE e.tabno = c.tabno /* условие соединения */
ORDER BY e.name, c.born;
```

| NAME | CHILD | BORN |
|-------------|-------|------------|
| Буров Г.О. | Ольга | 18.07.2010 |
| Котова И.М. | Илья | 19.02.2012 |
| Котова И.М. | Анна | 26.12.2014 |
| Рюмин В.П. | Вадим | 03.05.2005 |
| Серова Т.В. | Роман | 21.11.2006 |
| Серова Т.В. | Инна | 25.01.2008 |
| Серова Т.В. | Антон | 06.03.2009 |

19. Посчитать количество сотрудников по отделам:

```
SELECT d.name, count(*), 'сотрудник(а)'
FROM emp e, depart d
WHERE e.depno=d.depno
GROUP BY d.depno, d.name;
```

| NAME | COUNT(*) | СОТРУДНИК(А) |
|----------------|----------|--------------|
| Плановый отдел | 4 | сотрудник(а) |

| | | |
|---------------|---|--------------|
| Бухгалтерия | 4 | сотрудник(а) |
| Отдел кадров | 2 | сотрудник(а) |
| Администрация | 3 | сотрудник(а) |

Данные об отделе 4, в котором нет сотрудников, выводиться не будут, т.к. для записи о 4-м отделе не выполняется условие соединения (e.depno=d.depno).

1.4.7. Подзапросы

Подзапросом называется запрос **SELECT**, который находится внутри другой команды **SQL**. Подзапросы можно разделить на следующие группы в зависимости от возвращаемых результатов:

- Скалярные – это подзапросы, возвращающие единственное значение.
- Векторные – подзапросы, возвращающие от 0 до нескольких однотипных элементов (список элементов).
- Табличные – это подзапросы, возвращающие таблицу.

Подзапросы бывают коррелированные и некоррелированные. **Коррелированные** подзапросы содержат ссылки на значения полей в запросе верхнего уровня, а некоррелированные – не содержат. **Некоррелированный** подзапрос вычисляется один раз для запроса верхнего уровня, а коррелированный – для каждой строки запроса верхнего уровня.

Сначала рассмотрим использование подзапросов в команде **SELECT**. Подзапросы могут располагаться в разных частях этой команды:

- в части **FROM** – табличные некоррелированные;
- в части **WHERE** – любые;
- в части **HAVING** – любые;
- в части **SELECT** – только скалярные.

Подзапрос всегда стоит справа от оператора сравнения или предиката. Следующие операторы используются для модификации операторов сравнения:

- **ALL** – оператор, эквивалентный понятию "все". Например:
> **ALL** (< **ALL**) – больше (меньше) каждого значения элементов результирующего множества.
- **ANY (SOME)** – оператор, эквивалентный понятию "любой".
= **ANY** – равно одному из значений элементов результирующего множества (эквивалентно использованию предиката **IN**).
> **ANY** (< **ANY**) – больше (меньше) любого значения элементов результирующего множества.
- **EXISTS** – оператор, эквивалентный понятию "существует". Может использоваться с логическим оператором **NOT**.

Если список, модифицированный оператором **ALL**, содержит **NULL**-значение, то результирующий запрос будет пуст, т.к. нельзя сравнить никакое значение с **NULL**-значением.

Выражение <>**ANY**(...) не эквивалентно **NOT IN**: оно выполняется всегда, кроме случаев сравнения со списком **NULL**-значений.

Примеры использования подзапросов в части **WHERE**:

20. Выдать список сотрудников, имеющих детей:

а) с помощью операции соединения таблиц (без подзапроса):

```
SELECT DISTINCT e.*  
FROM emp e, children c  
WHERE e.tabno = c.tabno;
```

б) с помощью некоррелированного векторного подзапроса:

```
SELECT *  
FROM emp  
WHERE tabno IN (SELECT tabno FROM children);
```

в) с помощью коррелированного табличного подзапроса:

```
SELECT *  
FROM emp e  
WHERE EXISTS (SELECT * FROM children c WHERE e.tabno = c.tabno);
```

Оператор EXISTS возвращает "истину", если подзапрос выбирает хотя бы одну строку, и "ложь", если результат подзапроса пуст.

| TABNO | DEPNO | NAME | POST | SALARY | BORN | PHONE |
|-------|-------|-------------|------------------|---------|------------|-----------|
| 988 | 1 | Рюмин В.П. | начальник отдела | 48600.0 | 01.02.1970 | 115-26-12 |
| 087 | 5 | Котова И.М. | секретарь | 35000.0 | 16.09.1990 | 115-33-65 |
| 110 | 2 | Буров Г.О. | бухгалтер | 42880.0 | 22.05.1975 | 115-46-32 |
| 023 | 2 | Малова Л.А. | гл. бухгалтер | 59240.0 | 24.11.1954 | 114-24-55 |

21. Выдать список сотрудников, оклад которых выше среднего на предприятии (некоррелированный скалярный подзапрос):

```
SELECT depno, name, post, salary  
FROM emp  
WHERE salary > (SELECT avg(salary) FROM emp);
```

| DEPNO | NAME | POST | SALARY |
|-------|-------------|----------------|--------|
| 2 | Малова Л.А. | гл. бухгалтер | 59240 |
| 5 | Павлов А.А. | директор | 80000 |
| 5 | Кроль А.П. | зам. директора | 70000 |

Примеры использования подзапросов в части FROM:

22. Выдать список сотрудников, имеющих оклады выше среднего по отделу:

а) с помощью подзапроса в части WHERE:

```
SELECT depno, name, post, salary  
FROM emp e  
WHERE salary > (SELECT avg(salary) FROM emp m  
WHERE m.depno = e.depno);
```

б) с помощью подзапроса в части FROM:

```
SELECT e.depno, name, post, salary  
FROM emp e,  
(SELECT depno, avg(salary) AS avgsal FROM emp GROUP BY depno) m  
WHERE e.depno = m.depno AND e.salary > avgsal;
```

Второй вариант будет работать быстрее, т.к. подзапрос в части FROM вычисляется один раз, а коррелированный подзапрос в части WHERE вычисляется

ся для каждой строки запроса верхнего уровня (в нашем случае – для каждого сотрудника).

| DEPNO | NAME | POST | SALARY |
|-------|-------------|------------------|---------|
| 1 | Рюмин В.П. | начальник отдела | 48600.0 |
| 1 | Серова Т.В. | вед. программист | 48600.0 |
| 2 | Малова Л.А. | гл. бухгалтер | 59240.0 |
| 3 | Сухова К.А. | начальник отдела | 48600.0 |
| 5 | Павлов А.А. | директор | 80000.0 |
| 5 | Кроль А.П. | зам.директора | 70000.0 |

Пример использования подзапросов в части HAVING:

23. Выдать список отделов, в которых средние оклады ниже среднего оклада по предприятию:

```
SELECT depno, avg(salary)
FROM emp
GROUP BY depno
HAVING avg(salary) < (SELECT avg(salary) FROM emp);
```

| DEPNO | avg(salary) |
|-------|-------------|
| 1 | 46000 |
| 2 | 47875 |
| 3 | 38200 |

Предложение **UNION** позволяет объединять результаты нескольких запросов **SELECT** для реализации соответствующей операции реляционной алгебры. Результаты этих запросов должны быть построены по одной схеме. Предложение **ORDER BY** может встречаться в таком запросе один раз – в конце последнего предложения **SELECT**.

Пример использования операции UNION:

24. Посчитать количество сотрудников по всем отделам (включая те отделы, в которых нет сотрудников):

```
SELECT depno, count(*), 'человек(a)'
FROM emp
GROUP BY depno
UNION
SELECT depno, 0, 'человек'
FROM depart
WHERE depno NOT IN (SELECT depno FROM emp)
ORDER BY 1; /* упорядочение по первому столбцу */
```

| DEPNO | count(*) | человек(a) |
|-------|----------|------------|
| 1 | 4 | человек(a) |
| 2 | 4 | человек(a) |
| 3 | 2 | человек(a) |
| 4 | 0 | человек |
| 5 | 3 | человек(a) |

А с помощью подзапроса в части **SELECT** можно запрос из примера 23 написать гораздо короче. (К сожалению, использование подзапроса в части **SELECT** поддерживается не всеми СУБД, но Oracle и PostgreSQL поддерживают такие подзапросы).

Пример использования подзапросов в части **SELECT**:

25. Подсчёт количества сотрудников по всем отделам (включая те отделы, в которых нет сотрудников):

```
SELECT depno, (SELECT count(*) FROM emp e
                WHERE e.depno = d.depno) AS cnt
FROM depart d;
```

1.4.8. Самосоединение

В команде **SELECT** можно обратиться к одной и той же таблице несколько раз. При этом для каждой таблицы необходимо задать свой алиас, чтобы можно было обращаться к полям этих таблиц. Система будет выполнять такой запрос на основе декартова произведения таблиц, поэтому необходимо указывать условие соединения. А для того чтобы исключить соединение записи таблицы с самой собой в запросе на самосоединение необходимо также указывать условие типа "не равно" (<>, >, <).

Пример использования самосоединения:

26. Вывести список детей, у которых есть младшие братья или сёстры:

```
SELECT e.name, c1.name AS child1, c1.born AS born1,
       c2.name AS child2, c2.born AS born2
FROM children c1, children c2, emp e
WHERE c1.tabno=e.tabno           -- первое условие соединения
      AND c1.tabno=c2.tabno     -- второе условие соединения
      AND c1.born<c2.born       -- условие исключения
ORDER BY 1, 3;
```

| NAME | CHILD1 | BORN1 | CHILD2 | BORN2 |
|-------------|--------|------------|--------|------------|
| Котова И.М. | Илья | 19.02.1987 | Анна | 26.12.1989 |
| Серова Т.В. | Роман | 21.11.2006 | Инна | 25.01.2008 |
| Серова Т.В. | Роман | 21.11.2006 | Антон | 06.03.2009 |
| Серова Т.В. | Инна | 25.01.2008 | Антон | 06.03.2009 |

1.4.9. Замечания по использованию NULL-значений

Понятие неопределённого значения (**NULL**-значения) включено в стандарт SQL. Это значение не зависит от типа данных и может быть присвоено полю любого типа.

Значение **NULL** несравнимо ни с каким другим значением, даже со значением **NULL**. То есть при сравнении двух неопределённых значений А и В (**A IS NULL** и **B IS NULL**) условие (**A=B**) даст **NULL**, и это будет интерпретировано как "ложь". Т.о., в SQL применяется трёхзначная логика ("да", "нет", "не знаю"), о чём надо помнить при формировании условий сравнения.

Пример влияния NULL-значений на результат:

27. Вывести список сотрудников, причём сначала – тех, у которых номер телефона начинается с '114', а затем – всех остальных:

```
SELECT e.*
  FROM emp e
     WHERE phone LIKE '114%'
UNION ALL
SELECT e.*
  FROM emp e
     WHERE phone not LIKE '114%';
```

В результате выполнения данного запроса те сотрудники, у которых нет телефона (phone IS NULL), выведены не будут, хотя с точки зрения привычной двузначной логики условие (phone LIKE '114%' UNION ALL phone NOT LIKE '114%') даёт полное множество событий.

Тем не менее, из правила о несравнимости NULL-значений ни с какими другими значениями есть исключения:

- предложение GROUP BY объединяет все NULL-значения в одну группу,
- предикат DISTINCT <имя поля> оставляет только одно значение NULL,
- функция AVG не учитывает NULL-значения, и сумма значений поля делится на количество ненулевых (IS NOT NULL) значений.

1.4.10. Оператор CASE

Многие СУБД поддерживают оператор CASE. Этот оператор может быть использован в одной из двух синтаксических форм записи:

1-я форма:

```
CASE <проверяемое выражение>
  WHEN <сравниваемое выражение 1> THEN <возвращаемое значение 1>
  [...]
  WHEN <сравниваемое выражение N> THEN <возвращаемое значение N>]
[ELSE <возвращаемое значение>]
END
```

2-я форма:

```
CASE
  WHEN <предикат 1> THEN <возвращаемое значение 1>
  [...]
  WHEN <предикат N> THEN <возвращаемое значение N>]
[ELSE <возвращаемое значение>]
END
```

Все предложения WHEN должны иметь одинаковую синтаксическую форму, то есть нельзя смешивать первую и вторую формы. При использовании первой синтаксической формы условие WHEN удовлетворяется, как только значение проверяемого выражения станет равным значению выражения, указанного в предложении WHEN. При использовании второй синтаксической формы условие WHEN удовлетворяется, как только предикат принимает значе-

ние TRUE. При удовлетворении условия оператор CASE возвращает значение, указанное в соответствующем предложении THEN. Если ни одно из условий WHEN не выполнилось, то будет использовано значение, указанное в предложении ELSE. При отсутствии ELSE, будет возвращено NULL-значение. Если удовлетворены несколько условий, то будет возвращено значение предложения THEN первого из них, так как остальные просто не будут проверяться.

Примеры использования операции CASE:

28. Вывести список сотрудников с указанием количества сыновей и дочерей:

```
SELECT e.name,  
       count(case when sex='м' then 1 else null end) sons,  
       count(case when sex='ж' then 1 else null end) daughters  
FROM   emp e, children c  
WHERE  e.tabno = c.tabno  
GROUP BY e.name;
```

| NAME | SONS | DAUGHTERS |
|-------------|------|-----------|
| Буров Г.О. | 0 | 1 |
| Котова И.М. | 1 | 1 |
| Рюмин В.П. | 1 | 0 |
| Серова Т.В. | 2 | 1 |

Обратите внимание: бездетные сотрудники в список не вошли.

29. Вывести список сотрудников с указанием тех, у кого самая высокая и самая низкая зарплата на предприятии:

```
SELECT e.name, salary,  
       (case salary  
         when (select max(salary) from emp) then 'самая высокая'  
         when (select min(salary) from emp) then 'самая низкая'  
         else 'средняя'  
        end) as note  
FROM   emp e;
```

| NAME | SALARY | NOTE |
|--------------|--------|---------------|
| Рюмин В.П. | 48600 | средняя |
| Серова Т.В. | 48600 | средняя |
| Волков Л.Д. | 46500 | средняя |
| Перова К.В. | 33000 | самая низкая |
| Буров Г.О. | 42880 | средняя |
| Малова Л.А. | 59240 | средняя |
| Сухова К.А. | 48600 | средняя |
| Лукина Н.Н. | 42880 | средняя |
| Саргсян А.В. | 43500 | средняя |
| Тамм Л.В. | 43500 | средняя |
| Павлов А.А. | 80000 | самая высокая |
| Котова И.М. | 35000 | средняя |
| Кроль А.П. | 70000 | средняя |
| Беленко П.А. | 33000 | самая низкая |

1.4.11. Работа с представлениями

Представление (view, обзор) – это хранимый запрос, создаваемый на основе команды *SELECT*. Представление реально не содержит данных. Запрос, определяющий представление, выполняется тогда, когда к представлению происходит обращение с другим запросом, например, *SELECT*, *UPDATE* и т.д.

Назначение представлений:

1. Хранение сложных запросов.
2. Представление данных в виде, удобном пользователю.
3. Соккрытие конфиденциальной информации.
4. Предоставление дифференцированного доступа к данным.

Создание представления выполняется командой **CREATE VIEW**:

```
CREATE [ OR REPLACE ] VIEW <имя представления>  
  [ (<список имён столбцов> ) ]  
  AS <запрос> [ WITH CHECK OPTION ];
```

Запрос (команда *SELECT*), на основании которого создаётся представление, называется **определяющим запросом**, а таблицы, к которым происходит обращение в определяющем запросе – **базовыми таблицами**. Определяющий запрос по стандарту SQL не может включать предложение *ORDER BY*.

Если не указывать имена столбцов, то они получают названия по именам, перечисленным в списке выбора определяющего запроса. Указывать имена столбцов представления обязательно, если список выбора содержит агрегирующие функции или столбцы с одинаковыми именами из разных таблиц.

Примеры:

30. Создать представление "Сотрудники с детьми" (для удобного представления данных о детях сотрудников):

```
CREATE VIEW emp_child(depno, name, child, sex, born)  
  AS SELECT e.depno, e.name, c.name, c.sex, c.born  
  FROM emp e, children c  
  WHERE e.tabno = c.tabno;
```

```
SELECT * FROM emp_child;
```

| DEPNO | NAME | CHILD | SEX | BORN |
|-------|-------------|-------|-----|------------|
| 2 | Буров Г.О. | Ольга | ж | 18.07.2001 |
| 2 | Котова И.М. | Илья | м | 19.02.1987 |
| 2 | Котова И.М. | Анна | ж | 26.12.1989 |
| ... | ... | ... | ... | ... |
| 1 | Серова Т.В. | Антон | м | 06.03.2009 |

31. Создать представление "Сотрудники 2-го отдела" (для предоставления полного доступа к данным о сотрудниках 2-го отдела начальнику этого отдела):

```
CREATE VIEW emp2  
  AS SELECT *  
  FROM emp  
  WHERE depno = 2;
```

```
SELECT * FROM emp2;
```

| TABNO | DEPNO | NAME | POST | SALARY | BORN | PHONE |
|-------|-------|-------------|---------------|--------|----------|-----------|
| 110 | 2 | Буров Г.О. | бухгалтер | 42880 | 22.05.75 | 115-46-32 |
| 100 | 2 | Волков Л.Д. | программист | 46500 | 16.10.82 | |
| 130 | 2 | Лукина Н.Н. | бухгалтер | 42880 | 12.07.79 | 115-46-32 |
| 023 | 2 | Малова Л.А. | гл. бухгалтер | 59240 | 24.11.54 | 114-24-55 |

32. Создать представление "Сотрудники" (без данных о зарплате, для сокрытия конфиденциальной информации):

```
CREATE VIEW employees  
AS SELECT tabno, depno, name, post, born, phone  
FROM emp;
```

Представление может быть обновляемым и не обновляемым.

Обновляемым является представление, при обращении к которому можно обновить базовую таблицу.

Пример обновления базовой таблицы *emp* через представление *emp2*:

```
UPDATE emp2  
SET salary = 48000  
WHERE tabno = '100';
```

Изменения будут произведены в базовой таблице и отразятся в представлении.

```
SELECT * FROM emp2;
```

| TABNO | DEPNO | NAME | POST | SALARY | BORN | PHONE |
|-------|-------|-------------|---------------|--------------|----------|-----------|
| 110 | 2 | Буров Г.О. | бухгалтер | 42880 | 22.05.75 | 115-46-32 |
| 100 | 2 | Волков Л.Д. | программист | 48000 | 16.10.82 | |
| 130 | 2 | Лукина Н.Н. | бухгалтер | 42880 | 12.07.79 | 115-46-32 |
| 023 | 2 | Малова Л.А. | гл. бухгалтер | 59240 | 24.11.54 | 114-24-55 |

Вносимые изменения могут выйти за рамки определяющего запроса и поэтому не будут видны через представление. Если необходимо защитить данные от такого вмешательства, то нужно в команде создания представления указать ключевые слова **WITH CHECK OPTION**: тогда система отвергнет изменения, выходящие за рамки определяющего запроса.

По стандарту SQL-2 представление не является обновляемым, если определяющий запрос:

- 1) содержит ключевое слово *DISTINCT*;
- 2) содержит множественные операции (*UNION* и др.);
- 3) содержит предложение *GROUP BY*;
- 4) ссылается на другое необновляемое представление;
- 5) содержит вычисляемые выражения в списке выбора;
- 6) выбирает данные более чем из одной таблицы.

Во многих СУБД только первые 4 условия делают представление полностью необновляемым. Если список выбора содержит отдельные поля и вычисляемые выражения, то изменять нельзя только вычисляемые выражения. В Oracle если данные выбираются из нескольких таблиц, то можно изменить те поля

списка выбора, которые относятся к дочерней таблице: ее данные не дублируются в результате (Postgres последнюю возможность не поддерживает). Такие представления называются **частично обновляемыми**.

1.4.12. Удаление объектов базы данных

Удаление объектов БД выполняется с помощью команды **DROP**.

- **DROP TABLE** – удаление таблицы:

```
DROP TABLE [ IF EXISTS ] name [, ...] [ CASCADE | RESTRICT ]
```

Таблица будет удалена без дополнительного запроса на подтверждение вместе с данными и некоторыми другими объектами, существование которых зависит от наличия таблицы (индексы, триггеры и проч.). При указании *CASCADE* вместе с таблицей каскадно удаляются все зависящие от неё объекты БД (другие таблицы). Если указать *RESTRICT*, то при наличии зависящих от удаляемой таблицы объектов операция будет отменена.

- **DROP VIEW** – удаление представления:

```
DROP VIEW <имя представления>;
```

2. ВЫПОЛНЕНИЕ ЛАБОРАТОРНЫХ РАБОТ

Различают две версии SQL – интерактивную и встроенную. Интерактивный SQL подразумевает наличие командной строки или окна, в которое вводятся команды. Результаты выполнения этих команд выводятся на экран. Встроенный SQL включается в различные утилиты, входящие в состав СУБД (например, в утилиты для формирования отчётов или экранных форм).

Выполнение лабораторных работ заключается в создании таблиц (фрагмента базы данных) и запросов в соответствии с вариантом задания, который выдает преподаватель, в режиме работы с интерактивным SQL.

Лабораторная работа №1 посвящена созданию и заполнению таблиц базы данных. Каждая таблица должна иметь первичный ключ и обязательные поля (*not null*). Таблицы должны быть связаны по внешнему ключу.

Лабораторная работа №2 заключается в написании запросов к созданным отношениям. Необходимо предложить два варианта написания одного из запросов (по выбору) и проанализировать пути и эффективность их выполнения.

Третья работа посвящена созданию представлений. Для каждого представления необходимо проверить с помощью запросов модификации данных, является ли оно обновляемым, и объяснить полученный результат.

Лабораторная работа №4 заключается в реализации операций реляционной алгебры средствами SQL, возможно, с использованием отношений, созданных по заданию работы №1. Операции РА должны быть реализованы в точном соответствии с их определением. Для операции "разность" нельзя использовать EXCEPT, а для пересечения нельзя использовать INTERSECT (они реализованы не во всех СУБД и надо уметь обходиться без этих ключевых слов).

Перед началом работы необходимо получить у преподавателя имя и пароль и инструкции по запуску СУБД, под управлением которой будут проводиться работы.

Если СУБД не установлена на рабочей станции, то её можно установить самостоятельно, осуществляя установку под каждую конкретную ОС. После установки сервер запускается и слушает обращения на порту с номером 5432.

Для взаимодействия с сервером необходимо клиентское приложение для написания и выполнения запросов. Существует множество инструментов, которые можно использовать для работы с СУБД, но одним из клиентов, предлагаемых для работы с PostgreSQL является pgAdmin [5]. Если приложение не установлено, то оно скачивается с сайта, ссылка на который приведена, и устанавливается на рабочую станцию.

Через pgAdmin происходит подключение к серверу. Если базы нет, то она создается пользователем. Затем создается схема в рамках этой базы. По умолчанию можно использовать public схему. Затем уже в существующей схеме можно выполнять задания: создание таблиц и выполнение запросов.

Библиографический список

1. Грабер М. Введение в SQL. – М.: Лори, 2008. – 378 с.
2. Карпова И.П. Базы данных. Курс лекций и материалы для практических занятий. – Учебное пособие. – Издательство "Питер", 2013. – 240 с.
3. Коннолли Т., Бегг К. Базы данных: проектирование, реализация, сопровождение. Теория и практика, 3-е изд.: Пер. с англ.: Уч. пос. – М.: Изд. дом "Вильямс", 2003. – 1440 с.
4. Документация по PostgreSQL и Postgres Pro [Электронный ресурс]. – URL: <https://postgrespro.ru/docs> – (Дата обращения: 06.04.2022).
5. Open Source administration and development platform for PostgreSQL [Электронный ресурс]. – URL: <https://www.pgadmin.org/>. – pgAdmin. – (Дата обращения: 06.04.2022).