Базы данных

Лекция 3. СУБД. Троичная логика. SQL: продолжение.





СИСТЕМА УПРАВЛЕНИЯ БАЗАМИ ДАННЫХ

• совокупность программных и лингвистических средств, обеспечивающих управление созданием и использованием баз данных

ПРЕИМУЩЕСТВА

- Совместный доступ
 - Экономия на обслуживании
- Непротиворечивость
 - Единая версия истины
- Контроль целостности
 - Соответствие данных внутренней логике системы
- Транзакционность
 - Надежный механизм обновления
- Независимость от данных
 - Упрощение сопровождения

ОСНОВНЫЕ СУЩЕСТВУЮЩИЕ РСУБД









- PostgreSQL
 - Open source and free
 - Наиболее соответствует стандарту SQL
- Oracle
 - Первыми создали коммерческое решение (1979 г.)
 - Много дополнительных плюшек
- MS SQL
 - Оригинальная идея получена от Sysbase
 - Хорошая интеграция с Microsoft решениями
 - Разнообразие процедурных расширений
- MySQL
 - Тоже open source
 - Тоже производится Oracle

БИНАРНАЯ ЛОГИКА





TRUE

FALSE

БИНАРНАЯ ЛОГИКА

AND	TRUE	FALSE
TRUE	TRUE	FALSE
FALSE	FALSE	FALSE
<u>OR</u>	TRUE	FALSE
OR TRUE	TRUE TRUE	TRUE

	<u>NOT</u>
TRUE	FALSE
FALSE	TRUE

БИНАРНАЯ ЛОГИКА



АБСТРАГИРУЕМСЯ ОТ ПРИВЫЧНОГО ИЛИ ОЧЕНЬ СИЛЬНО ГРУСТИМ

ТЕРНАРНАЯ (ТРОИЧНАЯ) ЛОГИКА

- Что делать, если значение в ячейке неизвестно?
- Как проверить, что значение действительно неизвестно?
- Как учитывать в расчетах такие ячейки?
- Как правильно их агрегировать?

ТЕРНАРНАЯ (ТРОИЧНАЯ) ЛОГИКА

- Допустимые значения логического выражения:
 - TRUE
 - FALSE
 - UNKNOWN
- Как оперировать со значениями только из булевой логики
- **Как оперировать со значениями типа UNKNOWN**
- В SQL используется обозначение NULL



CPABHEHИE C NULL

- NULL = 1
- NULL <> 1
- NULL > 1
- NULL = NULL



CPABHEHИE C NULL

\bullet NULL = 1	NULL
--------------------	------

• NULL = NULL NULL



- (NULL = 1) OR (1 = 0)
- (NULL = 1) OR (1 = 1)
- (NULL = 1) AND (1 = 0)
- (NULL = 1) AND (1 = 1)
- NOT (NULL = NULL)



- (NULL = 1) OR (1 = 0)
 - NULL OR FALSE
- (NULL = 1) OR (1 = 1)
 - NULL OR TRUE
- (NULL = 1) AND (1 = 0)
 - NULL AND FALSE
- (NULL = 1) AND (1 = 1)
 - NULL AND TRUE
- NOT (NULL = NULL)
 - NOT NULL



- X AND TRUE
- X AND FALSE
- X OR TRUE
- X OR FALSE



• X AND TRUE

• X AND FALSE

• X OR TRUE

• X OR FALSE

X

FALSE

TRUE

X



• NULL OR FALSE NULL

• NULL OR TRUE TRUE

• NULL AND FALSE FALSE

• NULL AND TRUE NULL

• NOT NULL NULL



Т.е. если NULL как-то влияет на значение логического выражения, результат не определен

- Ключевые слова
 - WHERE
 - HAVING

строго требуют значение TRUE, поэтому условие NOT FALSE не является достаточным

```
SELECT col

FROM T

WHERE col = NULL;
```

```
SELECT col

FROM T

WHERE col = NULL;
```

По такому запросу не будет отобрана ни одна запись

```
SELECT col

FROM T

WHERE col = NULL

OR NOT (col = NULL);
```

```
SELECT col

FROM T

WHERE col = NULL

OR NOT (col = NULL);
```

По такому запросу тоже не будет отобрана ни одна запись

• Предикат IS NULL позволяет отбирать строки со значениями NULL

```
SELECT col

FROM T

WHERE col IS NULL;
```

• Предикат IS NULL позволяет отбирать строки со значениями NULL

```
SELECT col

FROM T

WHERE col IS NULL;
```

- 1 NOT IN (0)
- 1 NOT IN (1)

- 1 NOT IN (0)
- 1 NOT IN (1)
- 1 NOT IN (NULL)
- 1 NOT IN (2, NULL)
- 1 NOT IN (NULL, 1)

```
• 1 NOT IN (0)
                           TRUE
• 1 NOT IN (1)
                           FALSE
• 1 NOT IN (NULL)
                           NULL
• 1 NOT IN (2, NULL)
                           NULL
• 1 NOT IN (NULL, 1)
```

FALSE

- IS DISTINCT FROM
 - Аналогично операции '<>'
 - Относится к NULL как к известному значению
 - Oracle и MS SQL не умеют 🕾
 - Зато PostgreSQL умеет ©

NULL B 3AΠPOCAX: IS DISTINCT FROM

attr_1	attr_2
1	1
1	2
1	NULL
2	1
2	2
2	NULL
NULL	1
NULL	2
NULL	NULL

```
SELECT attr_1
    , attr_2
FROM table_1
WHERE attr_1 IS DISCTINCT FROM attr_2;
```

NULL B 3AΠPOCAX: IS DISTINCT FROM

attr_1	attr_2	attr_1 IS DISTINCT FROM attr_2
1	1	FALSE
1	2	TRUE
1	NULL	TRUE
2	1	TRUE
2	2	FALSE
2	NULL	TRUE
NULL	1	TRUE
NULL	2	TRUE
NULL	NULL	FALSE

NULL B 3AΠPOCAX: IS DISTINCT FROM

attr_1	attr_2	attr_1 IS DISTINCT FROM attr_2
1	1	FALSE
1	2	TRUE
1	NULL	TRUE
2	1	TRUE
2	2	FALSE
2	NULL	TRUE
NULL	1	TRUE
NULL	2	TRUE
NULL	NULL	FALSE

NULL B 3AПPOCAX: COALESCE

- COALESCE (attr1, attr2, ..., attrn)
 - Функция возвращает первое отличное от NULL значение
 - Число аргументов ограничивается только вашей фантазией

NULL B 3AПPOCAX: COALESCE

attr_1	attr_2
1	1
1	2
1	NULL
2	1
2	2
2	NULL
NULL	1
NULL	2
NULL	NULL

SELECT coalesce(attr_1, attr_2) AS attr
FROM table
WHERE attr 1 IS DISCTINCT FROM attr_2;

NULL B 3AПPOCAX: COALESCE

attr_1	attr_2
1	1
1	2
1	NULL
2	1
2	2
2	NULL
NULL	1
NULL	2
NULL	NULL

attr			
1			
1			
1			
2			
2			
2			
1			
2			
NULL			

NULL В АГРЕГИРУЮЩИХ ФУНКЦИЯХ

• При подсчетах в агрегирующих функциях NULL не учитывается

```
SELECT sum(attr)
, min(attr)
, max(attr)
, avg(attr)
, count(attr)
FROM table;
```

NULL В АГРЕГИРУЮЩИХ ФУНКЦИЯХ

```
attr
NULL
3
10
2
15
NULL
28
NULL
11
80
```

```
SELECT sum(attr)
, min(attr)
, max(attr)
, avg(attr)
, count(attr)
FROM table;
```

NULL В АГРЕГИРУЮЩИХ ФУНКЦИЯХ

attr

NULL



sum(attr)	min(attr)	max(attr)	avg(attr)	count(attr)
150	1	80	18.75	8

NULL

NULL

F571: BOOLEAN TEST

• Расширяет оператор IS на все значения логического типа данных: TRUE, FALSE, NULL;

• SQL: 1999

- Поддерживается не всеми СУБД, есть в:
 - PostgreSQL
 - MySQL

РЕЗЮМЕ

- При написании запросов нужно помнить:
 - В какой СУБД вы работаете и на каком диалекте SQL пишете
 - В каких полях вашей таблицы допустимо присутствие NULL
 - Как именно работает троичная логика в разных операциях
 - Корректно ли отработает ваш запрос, если где-то встретится NULL
- Как написать хороший запрос:
 - Сначала подумать
 - Потом написать

DML: SELECT



```
SELECT [DISTINCT] select_item_comma_list
  FROM table_reference_comma_list
  [WHERE conditional_expression]
  [GROUP BY column_name_comma_list]
[HAVING conditional_expression]
  [ORDER BY order item comma list];
```

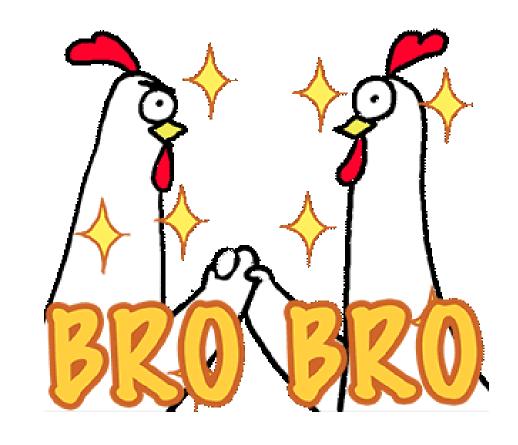
НЕСКОЛЬКО ИСТОЧНИКОВ

соединить

объединить

пересечь

вычесть



СОЕДИНЕНИЕ ТАБЛИЦ

- CROSS JOIN (полное декартово произведение таблиц)
- INNER JOIN (исключает несовпадающие строки)
- OUTER JOIN (содержит несовпадающие строки):
 - LEFT [OUTER] JOIN
 - RIGHT [OUTER] JOIN
 - FULL [OUTER] JOIN

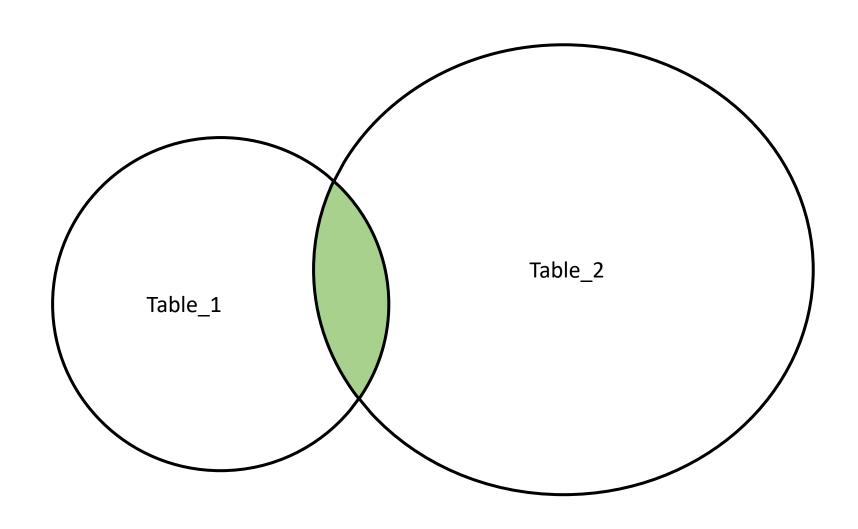
CROSS JOIN

На выходе: полное декартово произведение 2 таблиц

«Сцепление» строк 2 таблиц по заданном условию

```
SELECT column_name_comma_list
   FROM table_1
INNER JOIN table_2
ON table_1.column_name = table_2.column_name;
```

Только те строки, для которых условие соединения истинно



PROD_ID	PROD_NAME
123	Computer
124	Laptop
125	Scanner
126	Printer
127	Camera
128	Mobile phone

MANUFACTURER_NAME	PROD_ID
Lenovo	123
Lenovo	124
HP	123
HP	124
HP	125
HP	126
Samsung	128
Lenovo	125
Samsung	123
Samsung	129
LG	129

- Имеем 2 таблицы:
 - PRODUCT (ID продукта, Наименование продукта)
 - PROD_MAN (Наименование производителя, ID продукта)
- Хотим получить связку «продукт-производитель»:
 - 1. Соединяем таблицы PRODUCT и PROD_MAN
 - 2. Оставляем интересующие нас колонки

```
SELECT prod_name
    , manufacturer_name
FROM product p
INNER JOIN prod_man pm
ON p.prod id = pm.prod_id;
```

PROD_ID	PROD_NAME
123	Computer
124	Laptop
125	Scanner
126	Printer
127	Camera
128	Mobile phone

MANUFACTURER_NAME	PROD_ID
Lenovo	123
Lenovo	124
HP	123
HP	124
HP	125
HP	126
Samsung	128
Lenovo	125
Samsung	123
Samsung	129
LG	129

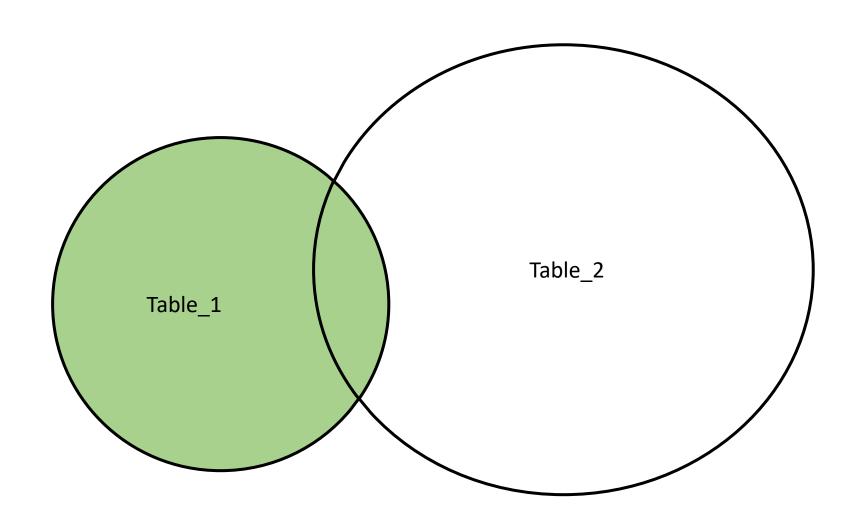
INNER JOIN: РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ

PROD_NAME	MANUFACTURER_NAME
Computer	Lenovo
Computer	HP
Computer	Samsung
Laptop	Lenovo
Laptop	HP
Scanner	HP
Scanner	Lenovo
Printer	HP
Mobile phone	Samsung

«Сцепление» строк 2 таблиц по заданном условию

```
SELECT column_name_comma_list
   FROM table_1
   LEFT JOIN table_2
   ON table_1.column_name = table_2.column_name;
```

- В результате присутствуют **все** строки «левой» таблицы
- Те строки, которые не соединяются с «правой» таблицей, все равно попадают в результат
- Поля от «правой» таблицы в таких строках заполняются специальным значением NULL



- Задача:
 - Выяснить производителей всех продуктов, имеющихся в магазине
- Решение:
 - Использовать LEFT JOIN вместо INNER JOIN в предыдущей задаче

```
SELECT PROD_NAME
    , MANUFACTURER_NAME
FROM PRODUCT P
LEFT JOIN PROD_MAN PM
ON P.PROD_ID = PM.PROD_ID;
```

PROD_ID	PROD_NAME
123	Computer
124	Laptop
125	Scanner
126	Printer
127	Camera
128	Mobile phone

MANUFACTURER_NAME	PROD_ID
Lenovo	123
Lenovo	124
HP	123
HP	124
HP	125
HP	126
Samsung	128
Lenovo	125
Samsung	123
Samsung	129
LG	129

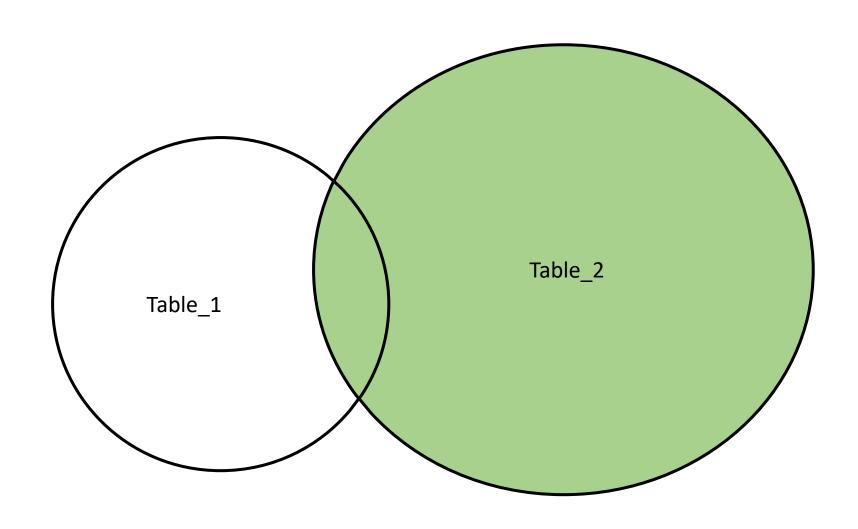
LEFT JOIN: РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ

PROD_NAME	MANUFACTURER_NAME
Computer	Lenovo
Computer	HP
Computer	Samsung
Laptop	Lenovo
Laptop	HP
Scanner	HP
Scanner	Lenovo
Printer	HP
Camera	NULL
Mobile phone	Samsung

«Сцепление» строк 2 таблиц по заданном условию

```
SELECT column_name_comma_list
  FROM table_1
RIGHT JOIN table_2
  ON Table1.column_name = Table2.column_name;
```

- В результате присутствуют все строки «правой» таблицы
- Те строки, которые не соединяются с «левой» таблицей, все равно попадают в результат
- Поля от «левой» таблицы в таких строках заполняются специальным значением NULL



- Задача:
 - Выяснить продукты всех производителей, имеющихся на складе
- Решение:
 - Использовать RIGHT JOIN вместо LEFT JOIN в предыдущей задаче

```
SELECT PROD_NAME
    , MANUFACTURER_NAME
FROM PRODUCT P
RIGHT JOIN PROD_MAN PM
ON P.PROD ID = PM.PROD ID;
```

PROD_ID	PROD_NAME
123	Computer
124	Laptop
125	Scanner
126	Printer
127	Camera
128	Mobile phone

MANUFACTURER_NAME	PROD_ID
Lenovo	123
Lenovo	124
HP	123
HP	124
HP	125
HP	126
Samsung	128
Lenovo	125
Samsung	123
Samsung	129
LG	129

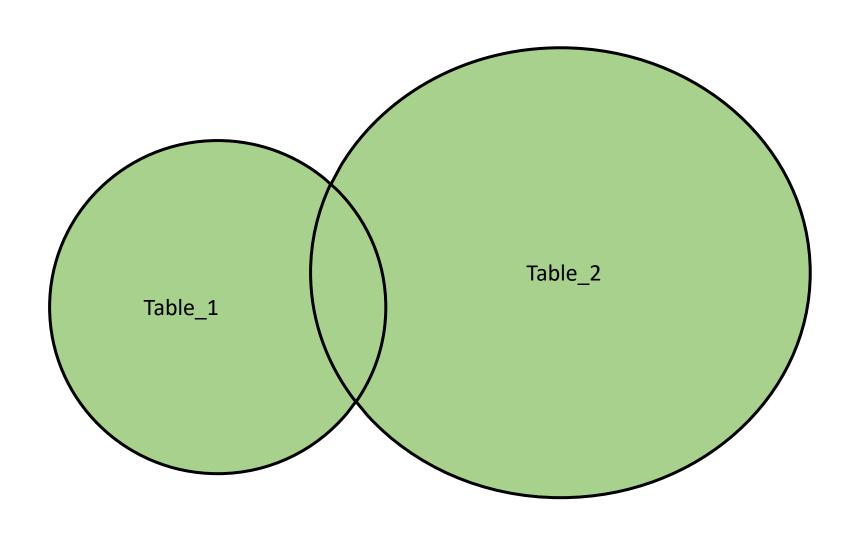
RIGHT JOIN: РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ

PROD_NAME	MANUFACTURER_NAME
Computer	Lenovo
Computer	HP
Computer	Samsung
Laptop	Lenovo
Laptop	HP
Scanner	HP
Scanner	Lenovo
Printer	HP
Mobile phone	Samsung
NULL	Samsung
NULL	LG

«Сцепление» строк 2 таблиц по заданном условию

```
SELECT column_name_comma_list
  FROM table_1
  FULL JOIN table_2
  ON table_1.column_name = table_2.column_name;
```

- В результирующей таблице присутствуют все строки «левой» и «правой» таблиц
- Иными словами, является комбинацией левого и правого соединения



```
SELECT PROD_NAME
    , MANUFACTURER_NAME
FROM PRODUCT P
FULL JOIN PROD_MAN PM
ON P.PROD ID = PM.PROD ID;
```

PROD_ID	PROD_NAME
123	Computer
124	Laptop
125	Scanner
126	Printer
127	Camera
128	Mobile phone

MANUFACTURER_NAME	PROD_ID
Lenovo	123
Lenovo	124
HP	123
HP	124
HP	125
HP	126
Samsung	128
Lenovo	125
Samsung	123
Samsung	129
LG	129

FULL JOIN: результат выполнения

PROD_NAME	MANUFACTURER_NAME
Computer	Lenovo
Computer	HP
Computer	Samsung
Laptop	Lenovo
Laptop	HP
Scanner	HP
Scanner	Lenovo
Printer	HP
Camera	NULL
Mobile phone	Samsung
NULL	Samsung
NULL	LG

- Не является отдельным видом соединения
- Бывает полезен для некоторого типа задач
- Задача:
 - Дана таблица RELATIONS (PERSON_ID, PERSON_NM, FATHER_ID)
 - Необходимо сформировать таблицу с именами отцов и детей

PERSON_ID	PERSON_NM	FATHER_ID
1	Иванов Иван Иванович	10
2	Иванов Николай Иванович	1
3	Петров Дмитрий Сергеевич	12
4	Петров Анатолий Дмитриевич	3
5	Петров Петр Дмитриевич	3
6	Степанов Сергей Иванович	13
7	Степанов Матвей Степанович	6
8	Яковлев Андрей Степанович	6

```
SELECT t2.person_nm AS father_nm
, t1.person_nm AS son_nm
FROM relations t1 -- дети
INNER JOIN relations t2 -- отцы
ON t1.father id = t2.person id;
```

PERSON_ID	PERSON_NM	FATHER_ID
1	Иванов Иван Иванович	10
2	Иванов Николай Иванович	1
3	Петров Дмитрий Сергеевич	12
4	Петров Анатолий Дмитриевич	3
5	Петров Петр Дмитриевич	3
6	Степанов Сергей Иванович	13
7	Степанов Матвей Степанович	6
8	Яковлев Андрей Степанович	6

SELF-JOIN: РЕЗУЛЬТАТ ВЫПОЛНЕНИЯ

FATHER_NM	SON_NM
Иванов Иван Иванович	Иванов Николай Иванович
Петров Дмитрий Сергеевич	Петров Анатолий Дмитриевич
Петров Дмитрий Сергеевич	Петров Петр Дмитриевич
Степанов Сергей Иванович	Степанов Матвей Степанович
Степанов Сергей Иванович	Яковлев Андрей Степанович

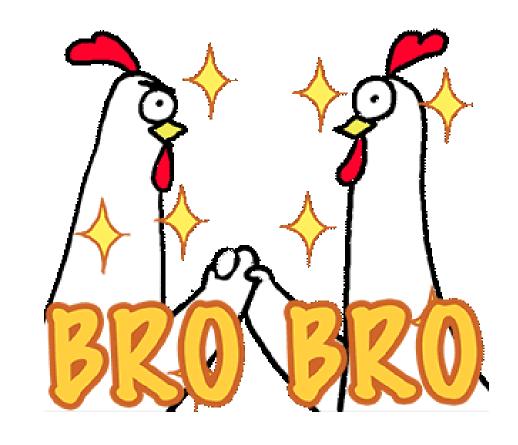
НЕСКОЛЬКО ИСТОЧНИКОВ

соединить

объединить

пересечь

вычесть



КОМБИНИРОВАНИЕ НЕСКОЛЬКИХ ИСТОЧНИКОВ

- UNION
- INTERSECT
- EXCEPT

```
Query_A
{UNION | INTERSECT | EXCEPT}
Query_B
```

UNION (ALL)

- Позволяет объединять результаты двух и более запросов:
 - С одинаковым числом столбцов
 - Порядок столбцов (семантически) в каждом запросе одинаков
 - Типы данных соответствующих столбцов совместимы
- Результатом является объединение результатов запросов
- Если в Query_A есть строка X и в Query_B есть идентичная строка X, то в результате объединения останется только 1 строка X
- Если нужно, чтобы при объединении строка X была записана дважды (или более раз), можно использовать UNION ALL

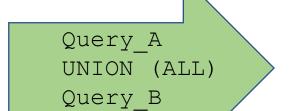
UNION (ALL)

Результат выполнения Query_A

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Мавлютов	Максим
Лукьянов	Денис
Митюрин	Максим

Результат выполнения Query_B

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Тюрюмина	Элла
Роздухова	Нина





UNION

Результат выполнения Query_A

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Мавлютов	Максим
Лукьянов	Денис
Митюрин	Максим

Query_A

Query_B

UNION

Результат выполнения Query_B

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Тюрюмина	Элла
Роздухова	Нина

	Фамилия	Имя
	Меркурьева	Надежда
	Халяпов	Александр
	Тюрюмина	Элла
•	Мавлютов	Максим
	Лукьянов	Денис
	Митюрин	Максим
	Роздухова	Нина

UNION ALL

Результат выполнения Query_A

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Мавлютов	Максим
Лукьянов	Денис
Митюрин	Максим

Результат выполнения Query_B

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Тюрюмина	Элла
Роздухова	Нина

	Фамилия	Имя
A	Меркурьева	Надежда
	Халяпов	Александр
Query_A	Мавлютов	Максим
UNION ALL Query B	Лукьянов	Денис
20.017_7	Митюрин	Максим
	Меркурьева	Надежда
	Халяпов	Александр
	Тюрюмина	Элла
	Роздухова	Нина

UNION (ALL)

Фамилия

Иванов

Петров

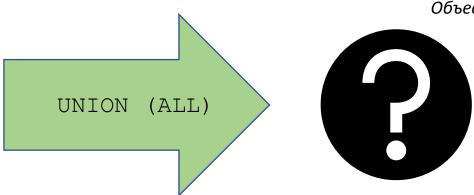
Иванов

Сидоров

Фамилия

Иванов

Петров



UNION ALL

Фамилия

Иванов

Петров

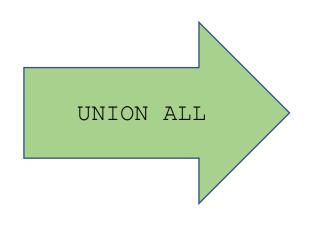
Иванов

Сидоров

Фамилия

Иванов

Петров



Фамилия

Иванов

Петров

Иванов

Сидоров

Иванов

Петров

UNION

Фамилия

Иванов

Петров

Иванов

Сидоров

Фамилия

Иванов

Петров



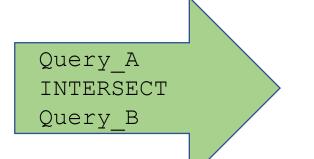
INTERSECT

Результат выполнения Query_A

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Мавлютов	Максим
Лукьянов	Денис
Митюрин	Максим

Результат выполнения Query_B

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Тюрюмина	Элла
Роздухова	Нина





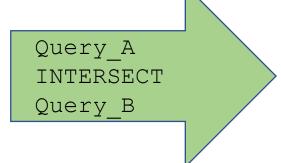
INTERSECT

Результат выполнения Query_A

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Мавлютов	Максим
Лукьянов	Денис
Митюрин	Максим

Результат выполнения Query_B

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Тюрюмина	Элла
Роздухова	Нина



Пересечение результатов запросов

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр

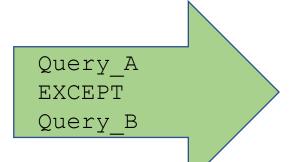
EXCEPT

Результат выполнения Query_A

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Мавлютов	Максим
Лукьянов	Денис
Митюрин	Максим

Результат выполнения Query_B

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Тюрюмина	Элла
Роздухова	Нина



Разность результатов запросов



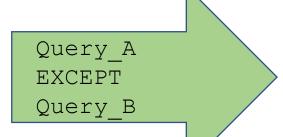
EXCEPT

Результат выполнения Query_A

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Мавлютов	Максим
Лукьянов	Денис
Митюрин	Максим

Результат выполнения Query_B

Фамилия	Имя
Меркурьева	Надежда
Халяпов	Александр
Тюрюмина	Элла
Роздухова	Нина



Разность результатов запросов

Фамилия	Имя
Мавлютов	Максим
Лукьянов	Денис
Митюрин	Максим

NULL B 3A IPOCAX

- Для теоретико-множественных операций
 - UNION
 - INTERSECT
 - EXCEPT

NULL = NULL

• HO! Для операции UNION ALL NULL <> NULL

SQL ЗАПРОС: БУДЕТ ЛИ РАБОТАТЬ?

```
SELECT t1.attr1 1, ..., t1.attr1 N, t2.attr2 1, ..., t2.attr2 M
  FROM t1
 INNER | LEFT | RIGHT | FULL JOIN t2
    ON t1.attr1 1 = t2.attr2 1 AND t1.attr1 N > t2.attr2 M
 WHERE tl.attrl k BETWEEN X AND Y
    OR t2.attr2 1 LIKE 'a%'
 GROUP BY t1.attr1 2, t2.attr2 2
HAVING SUM(t1.attr1 3) > 0
 UNION
SELECT t3.attr3 1, ..., t3.attr3 N, t4.attr4 1, ..., t4.attr4 M
  FROM t3
 INNER | LEFT | RIGHT | FULL JOIN t4
    ON t3.attr3 1 > t4.attr4 2;
```

СВЯЗЬ РЕЛЯЦИОННОЙ АЛГЕБРЫ С SQL

Реляционная алгебра	SQL	
Теоретико-множественные операции		
Объединение	UNION	
Пересечение	INTERSECT	
Разность	EXCEPT	
Реляционные операции		
Ограничение	WHERE	
Проекция	SELECT	
Соединение	JOIN	
Деление	-	

ПРИМЕРЫ ЗАПРОСОВ

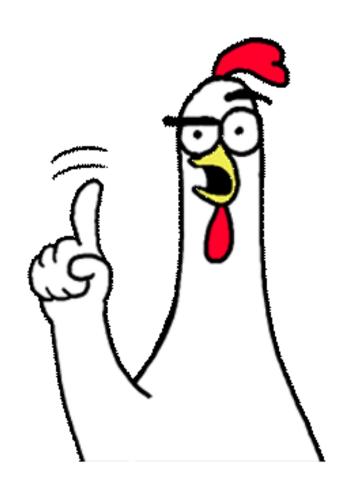
Реляционная алгебраSQLИз таблицы Product(product_id, product_name, product_type, product_price)
выделить все наименования продуктов с ценой не менее 150 рублейP150 = Product[product_price >= 150]
R = P150[product_name]SELECT DISTINCT product_name
FROM product
WHERE product_price >= 150;

Пусть даны таблицы Product(product_id, product_name, product_type, product_price) и Shop(shop_id, shop_name, product_id). Найти все магазины, в которых продается мыло с ценой менее 150 рублей

```
PS = Product[product_type =
'Mыло'][product_price < 150]

SS = PS[PS[product_id] =
Shop[product_id]]Shop
SSN = SS[shop_name]
OTBET: SSN.

SELECT DISTINCT shop_name
FROM product p
INNER JOIN shop sh
ON p.product_id = sh.product_id
WHERE product_type = 'Mыло'
AND product_price < 150;
```



Вопросы?