

Нормализация

Аномалия добавления

- Невозможно добавить данные об одном объекте, не добавив данные о другом.

Имя	Фамилия	Курс	Предмет	Оценка	Дата Экзамена
Илон	Маск	2	Базы данных	5	22.06.2023
Илон	Маск	2	Программирование	4	22.06.2023
Линус	Торвальд	2	Базы данных	4	25.06.2023
Линус	Торвальд	2	Программирование	5	25.06.2023
?	?	?	Математика	?	30.06.2023

σ_{предикт}(R)

Результатом выборки является новое отношение, которое содержит только те записи из исходного отношения, которые

удовлетворяют заданному предикту

В SQL выборка реализована через инструкцию `SELECT`

3. Объединение таблиц

RUS

Соединением двух наборов строк можно определить новое отношение, которое включает все записи из первого

отношения с использованием второго для фильтрации

Отношение совместимо, если они состоят из одинаковых атрибутов и каждая пары атрибутов имеет одинаковый диапазон

в SQL возможно сделать объединение так:

```
SELECT * FROM таблица1  
UNION  
SELECT * FROM таблица2;
```

Синтаксически такие запросы в SQL разрешены, но синтаксисом они не являются. Приведен пример объединения из таблицы 3.1

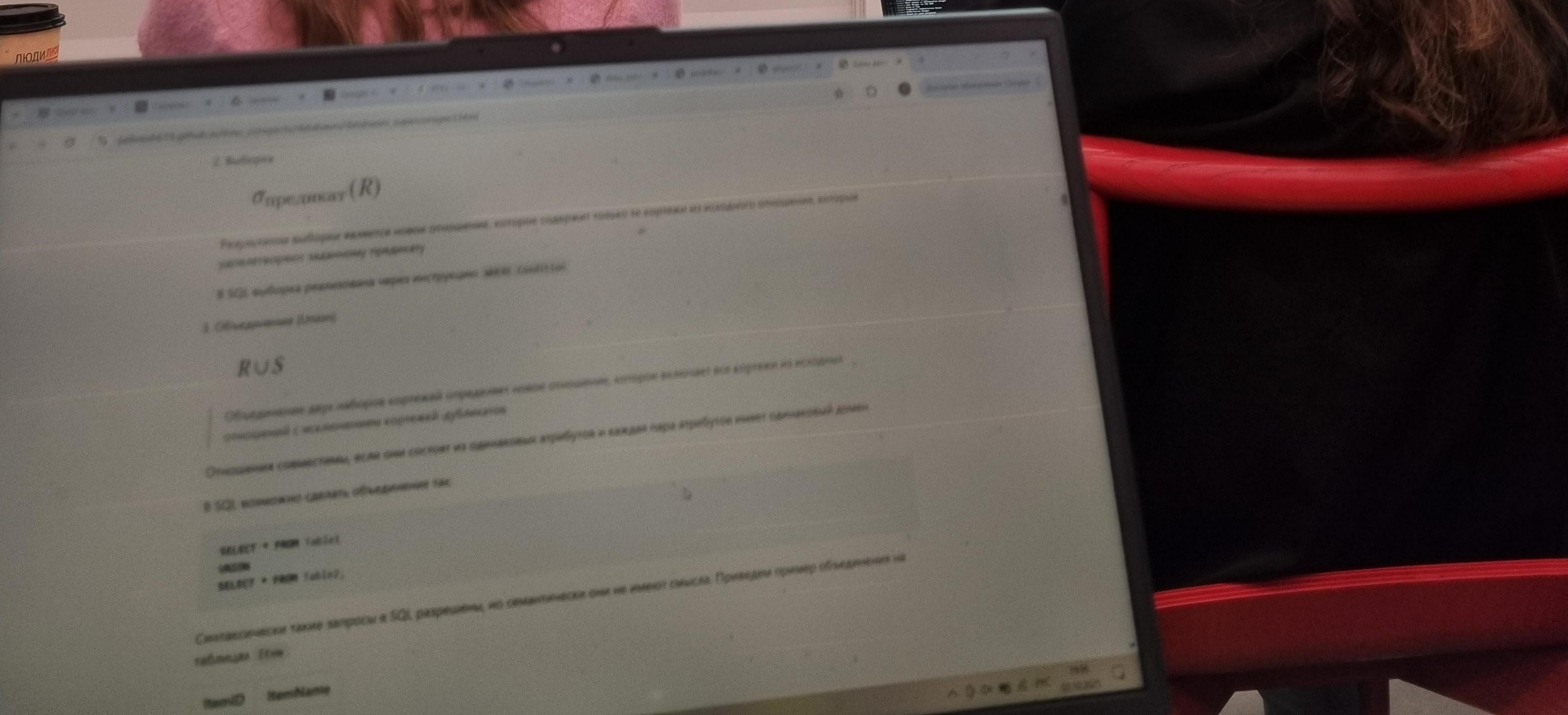
itemID	itemName	Base
1	Банан	Банан
2	Киви	Киви
3	Лимон	Лимон
4	Груша	Груша
5	Яблоко	Яблоко
6	Апельсин	Апельсин
7	Персик	Персик
8	Дыня	Дыня
9	Киви	Киви
10	Лимон	Лимон
11	Груша	Груша
12	Яблоко	Яблоко
13	Апельсин	Апельсин
14	Персик	Персик
15	Дыня	Дыня

Нормализация

Аномалия удаления

- При удалении одних данных мы невольно теряем другие.

Имя	Фамилия	Курс	Предмет	Оценка	Дата Экзамена
Илон	Маск	2	Базы данных	5	22.06.2023
Илон	Маск	2	Программирование	4	22.06.2023
Линус	Торвальд	2	Базы данных	4	25.06.2023
Линус	Торвальд	2	Программирование	5	25.06.2023
Линус	Торвальд	2	Математика	5	30.06.2023



Нормализация

Аномалия обновления

- Чтобы изменить какое-то значение, нам приходится обновлять его в нескольких местах.

Имя	Фамилия	Курс	Предмет	Оценка	Дата Экзамена
Илона	Маск	2	Базы данных	5	22.06.2023
Илон	Маск	2	Программирование	4	22.06.2023
Линус	Торвальд	2	Базы данных	4	25.06.2023
Линус	Торвальд	2	Программирование	5	25.06.2023

2. Выборка
 $\sigma_{\text{условие}}(R)$

Результатом выборки являются новые отношения, которые содержат только те записи из исходного отношения, которые удовлетворяют заданному условию.

В SQL выборка реализуется через инструкцию `SELECT` с子запросом.

3. Объединение таблиц

RUS

Объединение двух наборов записей определяет новое отношение, которое включает все записи из каждого, соединенного с исключениями коррекции дубликатов.

Объединение симметрическое, если оно состоит из одинаковых атрибутов и каждая пара записей имеет одинаковый результат.

В SQL возможно сделать объединение так:

```
SELECT * FROM Table1
UNION
SELECT * FROM Table2;
```

Синтаксис таких запросов в SQL различается, но симметрический результат не имеет смысла. Приведены примеры объединения из таблицы `Employees`:

EmployeeID	Lastname	Firstname	Address
1	Schafer	Ernest	92187 Maingate 9
2	Koch	Patricia	92187 Maingate 9
3	Hardy	Mark	92187 Maingate 9
4	Colmenar	Sergio	92187 Maingate 9
5	Spiegelburg	Hannover	92187 Maingate 9
6	Legendre	Helga	92187 Maingate 9
7	Frerichs	Yuri	92187 Maingate 9
8	Janzen	Yuri	92187 Maingate 9
9	Adler	Yuri	92187 Maingate 9
10	Hoogstra	Yuri	92187 Maingate 9
11	Remus	Yuri	92187 Maingate 9
12	Frederiksen	Yuri	92187 Maingate 9
13	Gallichan	Yuri	92187 Maingate 9
14	Montgomery	Yuri	92187 Maingate 9
15	Montgomery	Yuri	92187 Maingate 9
16	Montgomery	Yuri	92187 Maingate 9
17	Montgomery	Yuri	92187 Maingate 9
18	Montgomery	Yuri	92187 Maingate 9
19	Montgomery	Yuri	92187 Maingate 9
20	Montgomery	Yuri	92187 Maingate 9
21	Montgomery	Yuri	92187 Maingate 9
22	Montgomery	Yuri	92187 Maingate 9
23	Montgomery	Yuri	92187 Maingate 9
24	Montgomery	Yuri	92187 Maingate 9
25	Montgomery	Yuri	92187 Maingate 9
26	Montgomery	Yuri	92187 Maingate 9
27	Montgomery	Yuri	92187 Maingate 9
28	Montgomery	Yuri	92187 Maingate 9
29	Montgomery	Yuri	92187 Maingate 9
30	Montgomery	Yuri	92187 Maingate 9
31	Montgomery	Yuri	92187 Maingate 9
32	Montgomery	Yuri	92187 Maingate 9
33	Montgomery	Yuri	92187 Maingate 9
34	Montgomery	Yuri	92187 Maingate 9
35	Montgomery	Yuri	92187 Maingate 9
36	Montgomery	Yuri	92187 Maingate 9
37	Montgomery	Yuri	92187 Maingate 9
38	Montgomery	Yuri	92187 Maingate 9
39	Montgomery	Yuri	92187 Maingate 9
40	Montgomery	Yuri	92187 Maingate 9
41	Montgomery	Yuri	92187 Maingate 9
42	Montgomery	Yuri	92187 Maingate 9
43	Montgomery	Yuri	92187 Maingate 9
44	Montgomery	Yuri	92187 Maingate 9
45	Montgomery	Yuri	92187 Maingate 9
46	Montgomery	Yuri	92187 Maingate 9
47	Montgomery	Yuri	92187 Maingate 9
48	Montgomery	Yuri	92187 Maingate 9
49	Montgomery	Yuri	92187 Maingate 9
50	Montgomery	Yuri	92187 Maingate 9
51	Montgomery	Yuri	92187 Maingate 9
52	Montgomery	Yuri	92187 Maingate 9
53	Montgomery	Yuri	92187 Maingate 9
54	Montgomery	Yuri	92187 Maingate 9
55	Montgomery	Yuri	92187 Maingate 9
56	Montgomery	Yuri	92187 Maingate 9
57	Montgomery	Yuri	92187 Maingate 9
58	Montgomery	Yuri	92187 Maingate 9
59	Montgomery	Yuri	92187 Maingate 9
60	Montgomery	Yuri	92187 Maingate 9
61	Montgomery	Yuri	92187 Maingate 9
62	Montgomery	Yuri	92187 Maingate 9
63	Montgomery	Yuri	92187 Maingate 9
64	Montgomery	Yuri	92187 Maingate 9
65	Montgomery	Yuri	92187 Maingate 9
66	Montgomery	Yuri	92187 Maingate 9
67	Montgomery	Yuri	92187 Maingate 9
68	Montgomery	Yuri	92187 Maingate 9
69	Montgomery	Yuri	92187 Maingate 9
70	Montgomery	Yuri	92187 Maingate 9
71	Montgomery	Yuri	92187 Maingate 9
72	Montgomery	Yuri	92187 Maingate 9
73	Montgomery	Yuri	92187 Maingate 9
74	Montgomery	Yuri	92187 Maingate 9
75	Montgomery	Yuri	92187 Maingate 9
76	Montgomery	Yuri	92187 Maingate 9
77	Montgomery	Yuri	92187 Maingate 9
78	Montgomery	Yuri	92187 Maingate 9
79	Montgomery	Yuri	92187 Maingate 9
80	Montgomery	Yuri	92187 Maingate 9
81	Montgomery	Yuri	92187 Maingate 9
82	Montgomery	Yuri	92187 Maingate 9
83	Montgomery	Yuri	92187 Maingate 9
84	Montgomery	Yuri	92187 Maingate 9
85	Montgomery	Yuri	92187 Maingate 9
86	Montgomery	Yuri	92187 Maingate 9
87	Montgomery	Yuri	92187 Maingate 9
88	Montgomery	Yuri	92187 Maingate 9
89	Montgomery	Yuri	92187 Maingate 9
90	Montgomery	Yuri	92187 Maingate 9
91	Montgomery	Yuri	92187 Maingate 9
92	Montgomery	Yuri	92187 Maingate 9
93	Montgomery	Yuri	92187 Maingate 9
94	Montgomery	Yuri	92187 Maingate 9
95	Montgomery	Yuri	92187 Maingate 9
96	Montgomery	Yuri	92187 Maingate 9
97	Montgomery	Yuri	92187 Maingate 9
98	Montgomery	Yuri	92187 Maingate 9
99	Montgomery	Yuri	92187 Maingate 9
100	Montgomery	Yuri	92187 Maingate 9
101	Montgomery	Yuri	92187 Maingate 9
102	Montgomery	Yuri	92187 Maingate 9
103	Montgomery	Yuri	92187 Maingate 9
104	Montgomery	Yuri	92187 Maingate 9
105	Montgomery	Yuri	92187 Maingate 9
106	Montgomery	Yuri	92187 Maingate 9
107	Montgomery	Yuri	92187 Maingate 9
108	Montgomery	Yuri	92187 Maingate 9
109	Montgomery	Yuri	92187 Maingate 9
110	Montgomery	Yuri	92187 Maingate 9
111	Montgomery	Yuri	92187 Maingate 9
112	Montgomery	Yuri	92187 Maingate 9
113	Montgomery	Yuri	92187 Maingate 9
114	Montgomery	Yuri	92187 Maingate 9
115	Montgomery	Yuri	92187 Maingate 9
116	Montgomery	Yuri	92187 Maingate 9
117	Montgomery	Yuri	92187 Maingate 9
118	Montgomery	Yuri	92187 Maingate 9
119	Montgomery	Yuri	92187 Maingate 9
120	Montgomery	Yuri	92187 Maingate 9
121	Montgomery	Yuri	92187 Maingate 9
122	Montgomery	Yuri	92187 Maingate 9
123	Montgomery	Yuri	92187 Maingate 9
124	Montgomery	Yuri	92187 Maingate 9
125	Montgomery	Yuri	92187 Maingate 9
126	Montgomery	Yuri	92187 Maingate 9
127	Montgomery	Yuri	92187 Maingate 9
128	Montgomery	Yuri	92187 Maingate 9
129	Montgomery	Yuri	92187 Maingate 9
130	Montgomery	Yuri	92187 Maingate 9
131	Montgomery	Yuri	92187 Maingate 9
132	Montgomery	Yuri	92187 Maingate 9
133	Montgomery	Yuri	92187 Maingate 9
134	Montgomery	Yuri	92187 Maingate 9
135	Montgomery	Yuri	92187 Maingate 9
136	Montgomery	Yuri	92187 Maingate 9
137	Montgomery	Yuri	92187 Maingate 9
138	Montgomery	Yuri	92187 Maingate 9
139	Montgomery	Yuri	92187 Maingate 9
140	Montgomery	Yuri	92187 Maingate 9
141	Montgomery	Yuri	92187 Maingate 9
142	Montgomery	Yuri	92187 Maingate 9
143	Montgomery	Yuri	92187 Maingate 9
144	Montgomery	Yuri	92187 Maingate 9
145	Montgomery	Yuri	92187 Maingate 9
146	Montgomery	Yuri	92187 Maingate 9
147	Montgomery	Yuri	92187 Maingate 9
148	Montgomery	Yuri	92187 Maingate 9
149	Montgomery	Yuri	92187 Maingate 9
150	Montgomery	Yuri	92187 Maingate 9
151	Montgomery	Yuri	92187 Maingate 9
152	Montgomery	Yuri	92187 Maingate 9
153	Montgomery	Yuri	92187 Maingate 9
154	Montgomery	Yuri	92187 Maingate 9
155	Montgomery	Yuri	92187 Maingate 9
156	Montgomery	Yuri	92187 Maingate 9
157	Montgomery	Yuri	92187 Maingate 9
158	Montgomery	Yuri	92187 Maingate 9
159	Montgomery	Yuri	92187 Maingate 9
160	Montgomery	Yuri	92187 Maingate 9
161	Montgomery	Yuri	92187 Maingate 9
162	Montgomery	Yuri	92187 Maingate 9
163	Montgomery	Yuri	92187 Maingate 9
164	Montgomery	Yuri	92187 Maingate 9
165	Montgomery	Yuri	92187 Maingate 9
166	Montgomery	Yuri	92187 Maingate 9
167	Montgomery	Yuri	92187 Maingate 9
168	Montgomery	Yuri	92187 Maingate 9

Нормализация

Функциональная зависимость

- Функциональная зависимость $A \rightarrow B$ означает, что каждому значению атрибута A соответствует ровно одно значение атрибута B. Говоря проще, зная A, мы однозначно определяем B.
- НомерПаспорта \rightarrow ФИО (Зная номер паспорта, мы найдем только одного человека).
- НомерЗаказа \rightarrow ДатаЗаказа, Сумма (Зная номер заказа, мы узнаем его дату и сумму).

Нормализация

Первая нормальная форма

- Переменная отношения находится в первой нормальной форме тогда и только тогда, когда в любом допустимом значении отношения каждый его кортеж содержит только одно значение каждого из атрибутов.

Повторяющиеся группы

ИСУ	Фамилия	Номер 1	Номер 2
1234	Маск	111-11-11	
2452	Торвальд	222-22-22	444-44-44
2132	Гейтс		333-33-33

Неатомарные атрибуты

ИСУ	Фамилии	Номер 1
1234	Маск	111-11-11
2452	Торвальд	222-22-22,444-44-44
2132	Гейтс	

1-ая нормальная форма

ИСУ	Фамилия	Номер
1234	Маск	111-11-11
2452	Торвальд	222-22-22
2452	Торвальд	444-44-44
2132	Гейтс	333-33-33

Отношение находится в первой нормальной форме (1НФ) тогда и только тогда, когда

- в отношении нет повторяющихся групп (атрибутов с одинаковым смыслом)
- все атрибуты атомарны
- у отношения есть ключ

Нормализация

Вторая нормальная форма

- Переменная отношения находится во второй нормальной форме тогда и только тогда, когда она находится в первой нормальной форме и каждый неключевой атрибут неприводимо зависит от каждого её потенциального ключа.

Потенциальный ключ — {ИСУ, Предмет}

ИСУ	Имя	Фамилия	Курс	Предмет	Оценка
1234	Илон	Маск	2	Базы данных	4
1234	Илон	Маск	2	Программирование	5
2452	Линус	Торвальд	2	Базы данных	5
2452	Линус	Торвальд	2	Программирование	5

ИСУ	Имя	Фамилия	Курс
1234	Илон	Маск	2
2452	Линус	Торвальд	2

ИСУ	Предмет	Оценка
1234	Базы данных	5
1234	Программирование	4
2452	Базы данных	5
2452	Программирование	5

Нормализация

Третья нормальная форма

- Переменная отношения находится в третьей нормальной форме, когда она находится во второй нормальной форме, и отсутствуют транзитивные функциональные зависимости неключевых атрибутов от ключевых.

Преподаватель	Кафедра	Номер
Иванов	ИС	111-11-11
Петров	ИС	111-11-11
Сидоров	ИКТ	222-22-22

Преподаватель	Кафедра
Иванов	ИС
Петров	ИС
Сидоров	ИКТ

Кафедра	Номер
ИС	111-11-11
ИКТ	222-22-22

Нормализация

Нормальная форма Бойса-Кодда

- Переменная отношения находится в НФБК, когда она находится в третьей нормальной форме и ключевые атрибуты составного ключа не зависят от неключевых атрибутов
- Таблицы с простым первичным ключом, находящиеся в ЗНФ, автоматически соответствуют НФБК

Потенциальные ключи: {Номер слота, Время начала}, {Номер слота, Время окончания}, {Лаба, Время начала}, {Лаба, Время окончания}

Номер слота	Время начала	Время окончания	Лаба
1	09:30	10:30	1
1	11:00	12:00	1
1	14:00	15:30	1
2	10:00	11:30	1

Проблема: существует функциональная зависимость Тариф \rightarrow Номер карта. Детерминант (Тариф) не является потенциальным ключом

Нормализация

Высшие нормальные формы

- Четвертая нормальная форма (4НФ): Борется с многозначными зависимостями. Это когда один атрибут определяет множество независимых значений других атрибутов.
- Пятая нормальная форма (5НФ): Борется с зависимостями соединения без потерь, когда отношение можно без искажений восстановить только из трех и более проекций.



Нормализация

Многозначные зависимости

- X многозначно определяет Y тогда и только тогда, когда для каждого значения X существует набор значений Y , и этот набор не зависит от значений других атрибутов Z .
- Если мы зафиксируем значение X , то получим независимые друг от друга множества соответствующих значений Y и Z .
- Многозначная зависимость возникает, когда у одной сущности есть две или более независимых характеристики “многие-ко-многим”.

Нормализация Пример 4НФ

Преподаватель	Студент	Предмет
Иванов	Петров	Математика
Иванов	Петров	Физика
Иванов	Сидоров	Математика
Иванов	Сидоров	Физика
Петров	Иванов	История

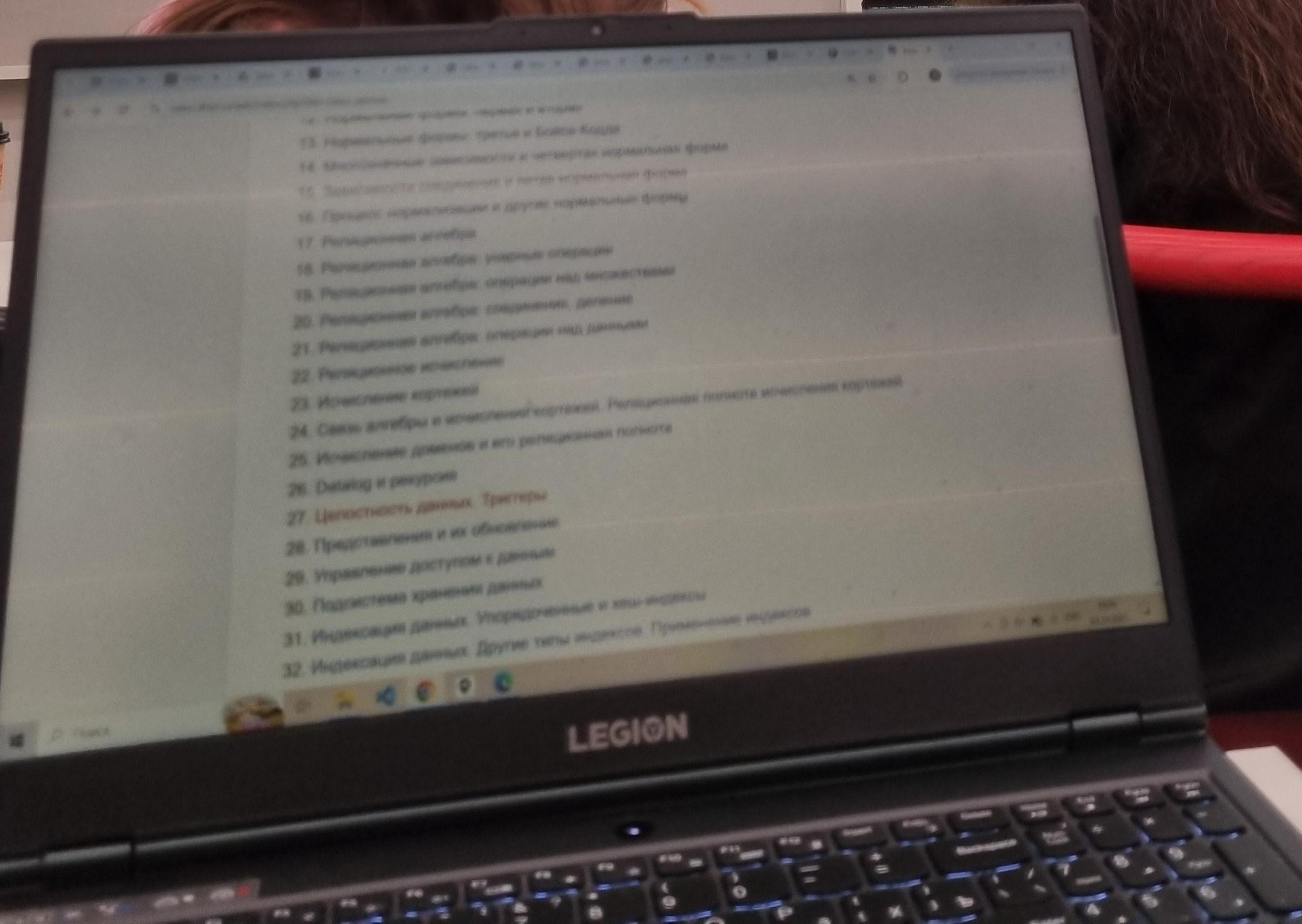
- Преподаватель Иванов ведет два предмета (Математика и Физика) и у него два студента (Петров и Сидоров). Чтобы отразить все возможные парные комбинации, нам пришлось создать 4 строки.
- Аномалия Вставки: Чтобы добавить преподавателю Иванову третьего студента, нам нужно будет вставить две новых строки: одну для Математики, другую для Физики.
- Удаление: Если студент Сидоров перестал изучать Физику у Иванова, мы должны удалить только одну строку? Но какую? Удаление одной строки нарушит целостность данных о том, что Иванов ведет и Физику, и что Сидоров у него учится.

20. Реляционная алгебра: соединение, деление
21. Реляционная алгебра: операции над данными
22. Реляционное исчисление
23. Исчисление кортежей
24. Связь алгебры и исчисления кортежей. Реляционная полнота исчисления кортежей
25. Исчисление доменов и его реляционная полнота
26. Datalog и рекурренты
27. Целостность данных. Триплеты
28. Представления и их обновление
29. Управление доступом к данным
30. Подсистема хранения данных
31. Индексация данных. Упорядоченные и хеш-индексы
32. Индексация данных. Другие типы индексов. Применение индексов
33. Императивное подмножество SQL
34. Data Control Language
35. Хранимые процедуры
36. Хранимые процедуры и функции: операторы
37. Курсоры
38. Хранимые функции
39. Транзакции. Восстановление. Классический алгоритм

Нормализация

Пятая нормальная форма

- Отношение находится в 5НФ, если каждая нетривиальная зависимость соединения в нем следует из его потенциальных ключей.
- Отношение в 5НФ не может быть без потерь декомпозировано на несколько меньших отношений (на две, три и более), а потом корректно восстановлено. Любая такая попытка приведет либо к потере данных, либо к появлению "phantomных" данных.



Нормализация

Пример 5НФ

Поставщик	Деталь	Проект
Smith	Болт	A
Smith	Гайка	B
Jones	Болт	A

- Бизнес-правило: "Если поставщик S поставляет деталь P, и деталь P используется в проекте J, и поставщик S работает над проектом J, то поставщик S поставляет деталь P для проекта J".
- Это правило является циклическим ограничением. Оно означает, что наша таблица должна быть закрыта относительно этого правила. Если существуют три пары (S,P), (P,J) и (S,J), то должна существовать и тройка (S,P,J).
- Аномалия: Если мы попробуем декомпозировать эту таблицу на любые две проекции (например, (Поставщик, Деталь) и (Деталь, Проект)), а потом обратно их соединить, мы получим фиктивные (ложные) строки, которых не было в исходном отношении.

Нормализация

Пример staff_property_inspection

- Отчет о осмотрах объектов недвижимости сотрудниками.
- Проблема: Повторяющиеся группы данных, адрес объекта и имя сотрудника дублируются для каждого осмотра.

property_no	property_address	inspection_date	inspection_time	comments	staff_no	staff_name	car_registration
PG4	6 Lawrence St...	18-Oct-00	10.00	Need to replace	SG37	Ann Beech	M231JGR
PG4	6 Lawrence St...	22-Apr-01	09.00	In good order	SG14	David Ford	M533HDR
PG4	6 Lawrence St...	1-Oct-01	12.00	Damp rot in bathroom	SG14	David Ford	N721HFR

Нормализация

1 НФ

- Устраниены повторяющиеся группы. Первичный ключ: (property_no, inspection_date).
- Атрибут property_address частично зависит только от property_no, что нарушает 2НФ.

staff_property_inspection

property_no	inspection_date	inspection_time	property_address	comments	staff_no	staff_name	car_registration
PG4	18-Oct-00	10.00	6 Lawrence	Need to replace	SG37	Ann Beech	M231JGR
PG4	22-Apr-01	09.00	6 Lawrence	In good order	SG14	David Ford	M533HDR
PG4	1-Oct-01	12.00	6 Lawrence	Damp rot in	SG14	David Ford	N721HFR

Нормализация 2 НФ

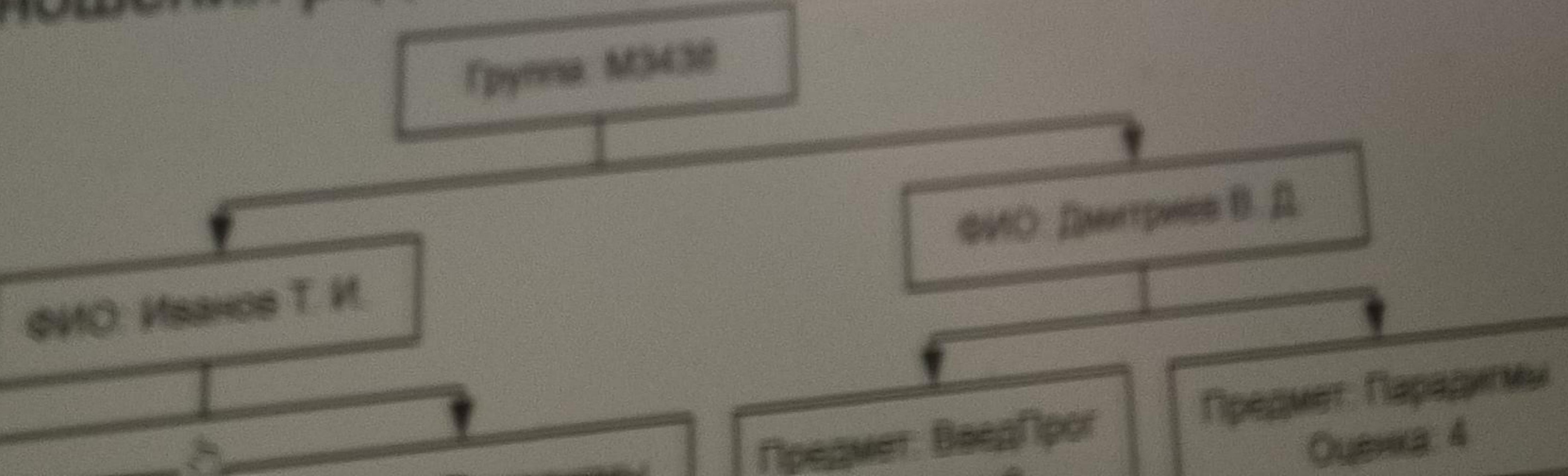
- Устранена частичная зависимость. Неключевые атрибуты зависят от всего первичного ключа.
- Атрибут `staff_name` транзитивно зависит от первичного ключа через `staff_no`, что нарушает ЗНФ.

property_inspection						
property_no	inspection_date	inspection_time	comments	staff_no	staff_name	car_registration
PG4	18-Oct-00	10.00	Need to	SG37	Ann Beech	M231JGR
PG4	22-Apr-01	09.00	In good	SG14	David Ford	M533HDR
PG4	1-Oct-01	12.00	Damp rot	SG14	David Ford	N721HFR

property	
property_no	property_address
PG4	6 Lawrence St...
PG16	5 Novar Dr...

Иерархическая модель данных

- Предложена и реализована IBM в 1966
- Представление данных
 - Дерево записей
 - Отношения родитель – ребенок



Нормализация НФБК

inspection

property_no	inspection_date	inspection_time	comments	staff_no
PG4	18-Oct-00	10.00	Need to	SG37
PG4	22-Apr-01	09.00	In good	SG14
PG4	1-Oct-01	12.00	Damp rot in	SG14

staff_car

staff_no	inspection_date	car_registration
SG37	18-Oct-00	M231JGR
SG14	22-Apr-01	M533HDR
SG14	1-Oct-01	N721HFR

SQL DELETE

```
[ WITH [ RECURSIVE ] with_query [, ...] ]
DELETE FROM [ ONLY ] table_name [*] [ [ AS ] alias ]
[ USING from_item [, ...] ]
[ WHERE condition | WHERE CURRENT OF cursor_name ]
[ RETURNING [ WITH ( { OLD | NEW } AS output_alias [, ...] ) ]
  { * | output_expression [ [ AS ] output_name ] } [, ...] ]
```

SQL ON DELETE

- RESTRICT
- NO ACTION
- SET NULL
- SET DEFAULT