

**ТЕМА: НОРМАЛИЗАЦИЯ**

ДИСЦИПЛИНА: ВВЕДЕНИЕ В БАЗЫ ДАННЫХ

ДЛЯ СТУДЕНТОВ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ  
6B06105 «DATA SCIENCE»

СТАРШИЙ ПРЕПОДАВАТЕЛЬ КЛЮЕВА Е.Г.

# ПЛАН ЛЕКЦИИ

- Введение в нормализацию;
- Первая нормальная форма;
- Вторая нормальная форма;
- Третья нормальная форма;
- Нормальная форма Бойса-Кодда;
- Четвертая нормальная форма;
- Пятая нормальная форма;
- Бизнес-модель процесса нормализации.

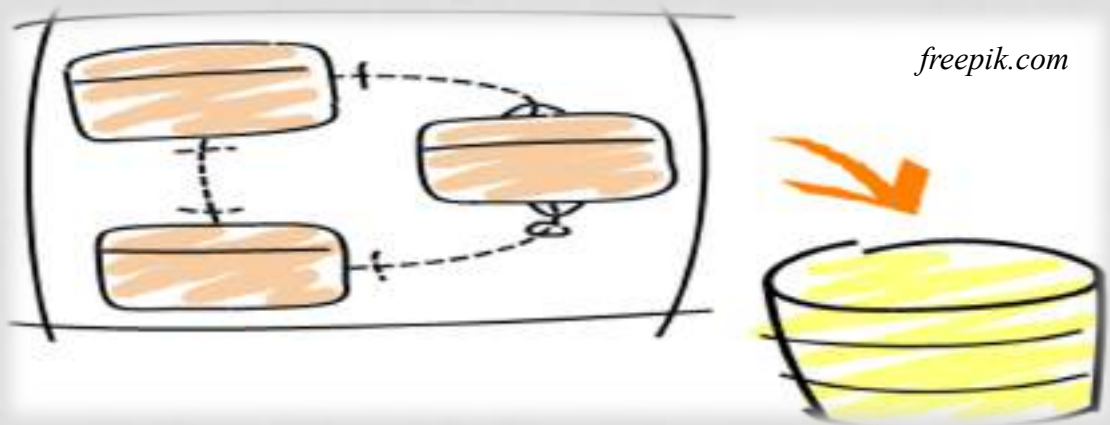


*freepik.com*

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

Рассмотрим пример:

ПОСТАВКИ (Номер поставки, Название товара, Цена товара, Количество, Дата поставки, Название поставщика, Адрес поставщика)



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

*Аномалии* - это противоречие между моделью предметной области и физической моделью данных, поддерживаемой средствами конкретной СУБД.

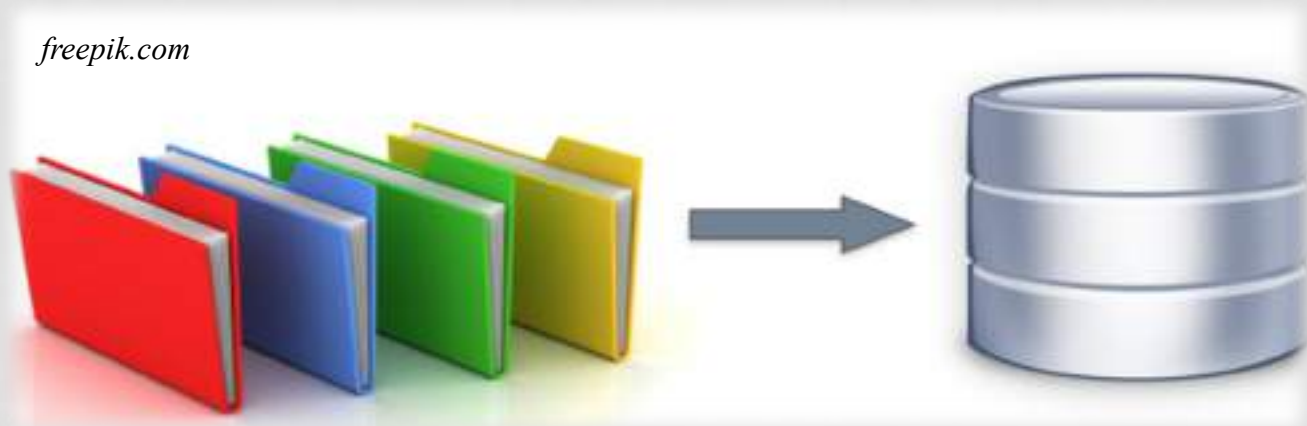


*freepik.com*

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

Различают следующие виды аномалий:

- аномалии вставки (INSERT);
- аномалии обновления (UPDATE);
- аномалии удаления (DELETE).



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

1. Аномалия обновления может возникнуть, если у какого-либо поставщика изменился адрес. Изменения должны быть внесены во все кортежи, соответствующие поставкам этого поставщика; в противном случае данные будут противоречивы.

2. Аномалия удаления возникнет при удалении записей обо всех поставках одного поставщика все данные о поставщике будут утеряны.

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

3. Аномалия добавления возникнет, если с поставщиком заключен договор, но поставок от него еще не было. Информация о таком поставщике не может быть внесена в отношение ПОСТАВКИ, т.к. для него не определён ключ (номер поставки и название товара) и другие обязательные поля.

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

Процесс проектирования базы данных представляет собой *процесс нормализации* схем отношений, причем каждая следующая нормальная форма обладает свойствами лучшими, чем предыдущая.





# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

*Нормализация* — пошаговый обратимый процесс композиций или декомпозиций исходных отношений, обладающих лучшими свойствами при включении, изменении, удалении данных, назначении им ключей и выявлении всех функциональных зависимостей.

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

В теории реляционных баз данных обычно определяется следующая последовательность нормальных форм:

- первая нормальная форма (1NF);
- вторая нормальная форма (2NF);
- третья нормальная форма (3NF);
- нормальная форма Бойса-Кодда (BCNF);
- четвертая нормальная форма (4NF);
- пятая нормальная форма, или нормальная форма проекции-соединения (5NF или PJ/NF).

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

Основные свойства нормальных форм:

- каждая следующая нормальная форма в некотором смысле лучше предыдущей;
- при переходе к следующей нормальной форме свойства предыдущих нормальных свойств сохраняются.

*freepik.com*



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

Модель некоторой предметной области опишем следующим неформальным текстом:

1. сотрудники выполняют проекты;
2. проекты состоят из нескольких заданий;
3. каждый сотрудник может участвовать в одном или нескольких проектах, или временно не участвовать ни в каких проектах;
4. над каждым проектом может работать несколько сотрудников, или временно проект может быть приостановлен, тогда над ним не работает ни один сотрудник;

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

Модель предметной области опишем следующим неформальным текстом:

5. Над каждым заданием в проекте работает ровно один сотрудник;

6. Каждый сотрудник числится в одном отделе;

7. Каждый сотрудник имеет телефон, находящийся в отделе сотрудника.

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

1. О каждом сотруднике необходимо хранить табельный номер и его фамилию. Табельный номер является уникальным для каждого сотрудника.

2. Каждый отдел имеет уникальный номер.

3. каждый проект имеет номер и наименование. Номер проекта является уникальным.

4. Каждая работа из проекта имеет номер, уникальный в пределах проекта. Работы в разных проектах могут иметь одинаковые номера.

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

## СОТРУДНИКИ\_ОТДЕЛЫ\_ПРОЕКТЫ

(*H\_Comp*, *ФАМ*, *H\_Отд*, *Тел*, *H\_ПР*, *Назв\_ПР*, *H\_Зад*)

Потенциальный ключ: {*H\_Comp*, *H\_ПР*}.



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

<i><b>Н_ Сотр</b></i>	<i><b>ФАМ</b></i>	<i><b>Н_Отд</b></i>	<i><b>Тел</b></i>	<i><b>Н_ПР</b></i>	<i><b>Наз_ПР</b></i>	<i><b>Н_ Зад</b></i>
<i>1</i>	Иванов	1	11-22-33	<i>1</i>	Космос	1
<i>1</i>	Иванов	1	11-22-33	<i>2</i>	Климат	1
<i>2</i>	Петров	1	11-22-33	<i>1</i>	Космос	2
<i>3</i>	Сидоров	2	22-33-11	<i>1</i>	Космос	3
<i>3</i>	Сидоров	2	22-33-11	<i>2</i>	Климат	2



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

## Аномалии вставки (INSERT)

В отношении **СОТРУДНИКИ\_ОТДЕЛЫ\_ПРОЕКТЫ** нельзя вставить данные

- 1) о сотруднике, который пока не участвует ни в одном проекте;
- 2) о проекте, над которым пока не работает ни один сотрудник.

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

## Аномалии обновления (UPDATE)

Фамилии сотрудников, наименования проектов, номера телефонов повторяются во многих кортежах отношения. Поэтому если сотрудник меняет фамилию, или проект меняет наименование, или меняется номер телефона, то такие изменения необходимо одновременно выполнить во всех местах, где эта фамилия, наименование или номер телефона встречаются, иначе отношение станет некорректным.

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

## Аномалии удаления (DELETE)

При удалении некоторых данных может произойти потеря другой информации. Например, если закрыть проект "Космос" и удалить все строки, в которых он встречается, то будут потеряны все данные о сотрудниках проекта. Если по проекту временно прекращены работы, то при удалении данных о работах по этому проекту будут удалены и данные о самом проекте.

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

В отношении  $R$  атрибут  $Y$  функционально зависит от атрибута  $X$  ( $X$  и  $Y$  могут быть составными) в том и только в том случае, если каждому значению  $X$  соответствует в точности одно значение  $Y$ :  $R.X \rightarrow R.Y$ .

*Функциональная зависимость реквизитов* - зависимость, при которой в каждом экземпляре отношения значению ключевого реквизита соответствует только одно значение описательного реквизита.

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

**СОТРУДНИКИ\_ОТДЕЛЫ\_ПРОЕКТЫ**  
(Н\_Сотр, ФИО, Н\_Отд, Тел, Н\_ПР, Назв\_ ПР,  
Н\_Зад)

Возможные ключи:

**Н\_Сотр**

**Н\_ПР**

**Н\_Сотр , Н\_ПР**



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

## СОТРУДНИКИ\_ОТДЕЛЫ\_ПРОЕКТЫ

(Н\_Сотр, ФАМ, Н\_Отд, Тел, Н\_ПР, Назв\_ПР, Н\_Зад)

Н\_Сотр → ФАМ

Н\_Сотр → Н\_Отд

Н\_Сотр → ТЕЛ



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

**СОТРУДНИКИ\_ОТДЕЛЫ\_ПРОЕКТЫ**  
(Н\_Сотр, ФАМ, Н\_Отд, Тел, Н\_ПР, Назв\_ПР,  
Н\_Зад)

**Н\_ПР → Назв\_ПР**

**Н\_Сотр, Н\_ПР → Н\_Зад**



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

Функциональная зависимость  $R.X \rightarrow R.Y$  называется полной, если атрибут  $Y$  не зависит функционально от любого точного подмножества  $X$ .

*Функционально полная зависимость* неключевых атрибутов заключается в том, что каждый неключевой атрибут функционально зависит от составного ключа, но не находится в функциональной зависимости ни от какой части составного ключа.



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

## **СОТРУДНИКИ\_ОТДЕЛЫ\_ПРОЕКТЫ**

(Н\_Сотр, ФАМ, Н\_Отд, Тел, Н\_ПР, Назв\_ПР, Н\_Зад)

Возможный составной ключ:

**Н\_Сотр, Н\_ПР**

*freepik.com*



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

## СОТРУДНИКИ\_ОТДЕЛЫ\_ПРОЕКТЫ

(Н\_Сотр, ФАМ, Н\_Отд, Тел, Н\_ПР, Назв\_ПР, Н\_Зад)

Функционально частичная зависимость:

$\text{Н\_Сотр, Н\_ПР} \rightarrow \text{ФАМ}$

$\text{Н\_Сотр, Н\_ПР} \rightarrow \text{Н\_Отд}$

$\text{Н\_Сотр, Н\_ПР} \rightarrow \text{ТЕЛ}$

$\text{Н\_Сотр, Н\_ПР} \rightarrow \text{Назв\_ПР}$

Функционально полная зависимость:

$\text{Н\_Сотр, Н\_ПР} \rightarrow \text{Н\_Зад}$

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

Функциональная зависимость  $R.X \rightarrow R.Y$  называется *транзитивной*, если существует такой атрибут  $Z$ , что имеются функциональные зависимости  $R.X \rightarrow R.Z$  и  $R.Z \rightarrow R.Y$  и отсутствует функциональная зависимость  $R.Z \rightarrow R.X$ .



# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

## **СОТРУДНИКИ\_ОТДЕЛЫ\_ПРОЕКТЫ**

(Н\_Сотр, ФАМ, Н\_Отд, Тел, Н\_ПР, Назв\_ПР, Н\_Зад)

Функциональная зависимость:

**Н\_Сотр  $\rightarrow$  Н\_Отд**

Транзитивная зависимость:

**Н\_Отд  $\rightarrow$  ТЕЛ**

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

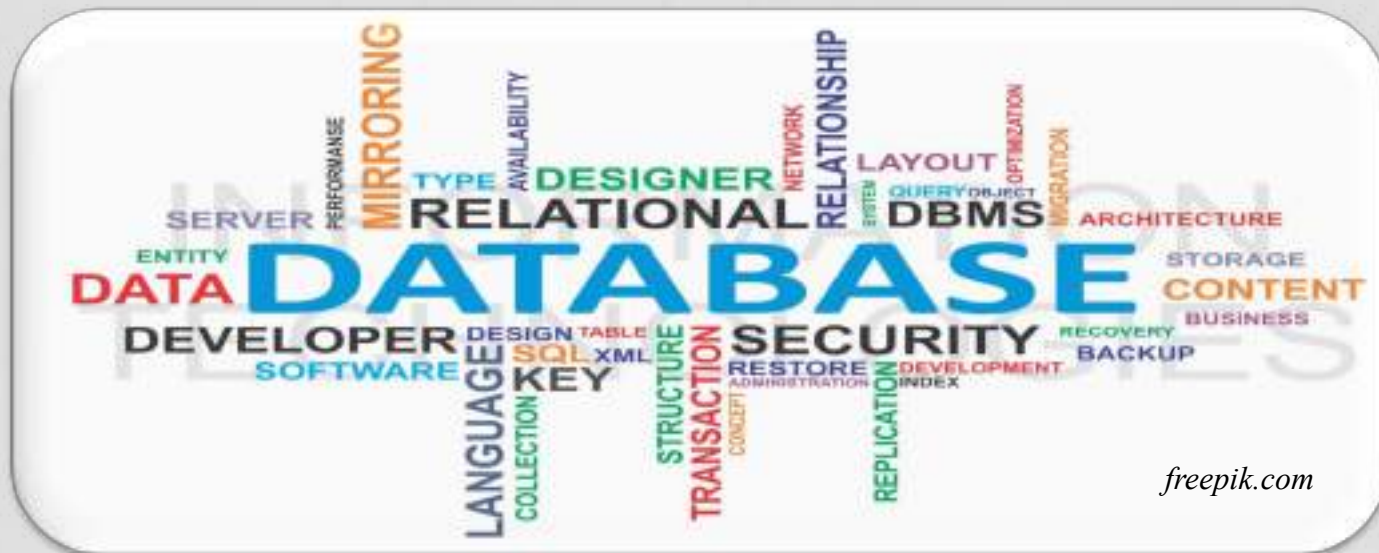
*Неключевым* (описательным) атрибутом называется любой атрибут отношения, не входящий в состав первичного ключа.



*freepik.com*

# ВВЕДЕНИЕ В НОРМАЛИЗАЦИЮ

Два или более атрибута *взаимно независимы*, если ни один из этих атрибутов не является функционально зависимым от других.



# ПЕРВАЯ НОРМАЛЬНАЯ ФОРМА

<i>Н_ Сотр</i>	<i>ФАМ</i>	<i>Н_ Отд</i>	<i>Тел</i>	<i>Н_ПР</i>	<i>Наз_ПР</i>	<i>Н_ Зад</i>
1	Иванов	1	11-22-33	1,2	Космос,Климат	1, 1
2	Петров	1	11-22-33	1	Космос	2
3	Сидоров	2	22-33-11	1,2	Космос,Климат	3,2

# ПЕРВАЯ НОРМАЛЬНАЯ ФОРМА

Отношение находится в 1NF когда все входящие в него атрибуты являются атомарными (неделимыми).

Таблица находится в 1NF когда ни одна из ее строк не содержит в любом своем поле более одного значения и ни одно из ее ключевых полей не пусто.





# ПЕРВАЯ НОРМАЛЬНАЯ ФОРМА

<i>Н_ Сотр</i>	<i>ФАМ</i>	<i>Н_ Отд</i>	<i>Тел</i>	<i>Н_ ПР</i>	<i>Наз_ ПР</i>	<i>Н_ Зад</i>
<i>1</i>	Иванов	<i>1</i>	11-22-33	<i>1</i>	Космос	<i>1</i>
<i>1</i>	Иванов	<i>1</i>	11-22-33	<i>2</i>	Климат	<i>1</i>
<i>2</i>	Петров	<i>1</i>	11-22-33	<i>1</i>	Космос	<i>2</i>
<i>3</i>	Сидоров	<i>2</i>	22-33-11	<i>1</i>	Космос	<i>3</i>
<i>3</i>	Сидоров	<i>2</i>	22-33-11	<i>2</i>	Климат	<i>2</i>

# ВТОРАЯ НОРМАЛЬНАЯ ФОРМА

Отношение находится во 2NF когда оно уже в 1NF и нет неключевых атрибутов, зависящих от части сложного ключа.

Отношение находится во 2NF в том случае, когда оно находится в 1NF, и каждый неключевой атрибут функционально полно зависит от первичного составного ключа.



# ВТОРАЯ НОРМАЛЬНАЯ ФОРМА

## **СОТРУДНИКИ\_ОТДЕЛЫ\_ПРОЕКТЫ**

(Н\_Сотр, ФАМ, Н\_Отд, Тел, Н\_ПР, Назв\_ПР, Н\_Зад)

Ключ: {Н\_Сотр, Н\_ПР}

Н\_Сотр  $\rightarrow$  ФАМ

Н\_Сотр  $\rightarrow$  Н\_Отд

Н\_Сотр  $\rightarrow$  ТЕЛ

Н\_ПР  $\rightarrow$  Назв\_ПР

Н\_Сотр, Н\_ПР  $\rightarrow$  Н\_Зад

# ВТОРАЯ НОРМАЛЬНАЯ ФОРМА

Исходное отношение:

**СОТРУДНИКИ\_ОТДЕЛЫ\_ПРОЕКТЫ**

Отношения после декомпозиции:

**СОТРУДНИКИ\_ОТДЕЛЫ**

**ПРОЕКТЫ**

**ЗАДАНИЯ**

*freepik.com*



# ВТОРАЯ НОРМАЛЬНАЯ ФОРМА

Отношение СОТРУДНИКИ\_ОТДЕЛЫ

<i><b>Н_Сотр</b></i>	<b>ФАМ</b>	<b>Н_Отд</b>	<b>Тел</b>
<i>1</i>	Иванов	1	11-22-33
<i>2</i>	Петров	1	11-22-33
<i>3</i>	Сидоров	2	22-33-11

# ВТОРАЯ НОРМАЛЬНАЯ ФОРМА

## Отношение ПРОЕКТЫ

<i><b>Н_ПР</b></i>	<i><b>Назв_ПР</b></i>
<i><b>1</b></i>	Космос
<i><b>2</b></i>	Климат

# ВТОРАЯ НОРМАЛЬНАЯ ФОРМА

## Отношение ЗАДАНИЯ

<i><b>Н_Сотр</b></i>	<i><b>Н_ПР</b></i>	<i><b>Н_Зад</b></i>
<i>1</i>	<i>1</i>	<i>1</i>
<i>1</i>	<i>2</i>	<i>1</i>
<i>2</i>	<i>1</i>	<i>2</i>
<i>3</i>	<i>1</i>	<i>3</i>
<i>3</i>	<i>2</i>	<i>2</i>

## ВТОРАЯ НОРМАЛЬНАЯ ФОРМА

### Оставшиеся аномалии вставки (INSERT)

В отношении **СОТРУДНИКИ\_ОТДЕЛЫ** нельзя вставить кортеж (4, Пушников, 1, 22-33-11), т.к. при этом получится, что два сотрудника из 1-го отдела (Иванов и Пушников) имеют разные номера телефонов, а это противоречит модели предметной области.



## ВТОРАЯ НОРМАЛЬНАЯ ФОРМА

### Оставшиеся аномалии обновления (UPDATE)

Одни и те же номера телефонов повторяются во многих кортежах отношения. Поэтому если в отделе меняется номер телефона, то такие изменения необходимо одновременно выполнить во всех кортежах, где этот номер телефона встречаются.

## ВТОРАЯ НОРМАЛЬНАЯ ФОРМА

### Оставшиеся аномалии удаления (DELETE)

При удалении некоторых данных по-прежнему может произойти потеря другой информации. Например, если удалить сотрудника Сидорова, то будет потеряна информация о том, что в отделе номер 2 находится телефон 22-33-11.

## ВТОРАЯ НОРМАЛЬНАЯ ФОРМА

Отношение СОТРУДНИКИ\_ОТДЕЛЫ

<i><b>Н_Сотр</b></i>	<b>ФАМ</b>	<b>Н_Отд</b>	<b>Тел</b>
<i>1</i>	Иванов	1	11-22-33
<i>2</i>	Петров	1	11-22-33
<i>3</i>	Сидоров	2	22-33-11

# ТРЕТЬЯ НОРМАЛЬНАЯ ФОРМА

Отношение находится в 3NF тогда, когда отношение находится в 2NF и все неключевые атрибуты взаимно независимы.

Отношение находится в третьей нормальной форме (3NF) в том и только в том случае, если находится в 2NF и каждый неключевой атрибут нетранзитивно зависит от первичного ключа.

# ТРЕТЬЯ НОРМАЛЬНАЯ ФОРМА

## Отношение СОТРУДНИКИ

<i><b>Н_Сотр</b></i>	<i><b>ФАМ</b></i>	<i><b>Н_Отд</b></i>
<i>1</i>	Иванов	1
<i>2</i>	Петров	1
<i>3</i>	Сидоров	2

# ТРЕТЬЯ НОРМАЛЬНАЯ ФОРМА

## Отношение ОТДЕЛЫ

<i><b>Н_ОТД</b></i>	<i><b>Тел</b></i>
<i><b>1</b></i>	<b>11-22-33</b>
<i><b>2</b></i>	<b>22-33-11</b>

## ТРЕТЬЯ НОРМАЛЬНАЯ ФОРМА

Реляционная модель, состоящая из четырех отношений **СОТРУДНИКИ**, **ОТДЕЛЫ**, **ПРОЕКТЫ**, **ЗАДАНИЯ**, находящихся в третьей нормальной форме, является адекватной описанной модели предметной области, и требует наличия только тех триггеров, которые поддерживают ссылочную целостность.

# НОРМАЛЬНАЯ ФОРМА БОЙСА-КОДДА

Отношение "Поставки"

Номер поставщика PNUM	Наименование поставщика PNAME	Номер детали DNUM	Поставляемое количество VOLUME
1	Фирма 1	1	100
1	Фирма 1	2	200
1	Фирма 1	3	300
2	Фирма 2	1	150
2	Фирма 2	2	250
3	Фирма 3	1	1000



# НОРМАЛЬНАЯ ФОРМА БОЙСА-КОДДА

Потенциальные ключи:

**{PNUM, DNUM}** и **{PNAME, DNUM}**.

Функциональные зависимости:

**PNUM  $\rightarrow$  PNAME**

**PNAME  $\rightarrow$  PNUM**

**{PNUM, DNUM}  $\rightarrow$  VOLUME**

**{PNUM, DNUM}  $\rightarrow$  PNAME**

**{PNAME, DNUM}  $\rightarrow$  VOLUME**

**{PNAME, DNUM}  $\rightarrow$  PNUM**

# НОРМАЛЬНАЯ ФОРМА БОЙСА-КОДДА

*Детерминант* - любой атрибут, от которого полностью функционально зависит некоторый другой атрибут.

Отношение находится в нормальной форме Бойса-Кодда (НФБК), если оно находится во 3НФ и в нем отсутствуют зависимости атрибутов первичного ключа от неключевых атрибутов.

# НОРМАЛЬНАЯ ФОРМА БОЙСА-КОДДА

Отношение "Поставки" не находится в НФБК, т.к. имеются зависимости ( $PNUM \rightarrow PNAME$  и  $PNAME \rightarrow PNUM$ ), детерминанты которых не являются потенциальными ключами.

*freepik.com*



# НОРМАЛЬНАЯ ФОРМА БОЙСА-КОДДА

## Отношение Поставщики

<b>Номер поставщика PNUM</b>	<b>Наименование поставщика PNAME</b>
1	Фирма 1
2	Фирма 2
3	Фирма 3

# НОРМАЛЬНАЯ ФОРМА БОЙСА-КОДДА

## Отношение Поставки-2

<b>Номер поставщика PNUM</b>	<b>Номер детали DNUM</b>	<b>Поставляемое количество VOLUME</b>
1	1	100
1	2	200
1	3	300
2	1	150
2	2	250
3	1	1000

# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

Отношение "Абитуриенты-Факультеты-Предметы"

Абитуриент	Факультет	Предмет
Иванов	Математический	Математика
Иванов	Математический	Информатика
Иванов	Физический	Математика
Иванов	Физический	Физика
Петров	Математический	Математика
Петров	Математический	Информатика

# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

Отношение Абитуриенты-Факультеты-Предметы\_1

Номер Абитуриента	Номер Факультета	Номер Предмета
1	1	1
1	1	2
1	2	1
1	2	3
2	1	1
2	1	2

# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

Отношение "Абитуриенты"

**Номер  
Абитуриента**

**Абитуриент**

1

Иванов

2

Петров



# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

## Отношение "Факультеты"

<b>Номер Факультета</b>	<b>Факультет</b>
1	Математический
2	Физический

# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

## Отношение "Предметы"

Номер Предмета	Предмет
1	Математика
2	Информатика
3	Физика

# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

## Аномалия вставки (INSERT)

При попытке добавить в отношение "Абитуриенты-Факультеты-Предметы" новый кортеж, например (Сидоров, Математический, Математика), мы обязаны добавить также и кортеж (Сидоров, Математический, Информатика), т.к. все абитуриенты математического факультета обязаны иметь один и тот же список сдаваемых предметов.

# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

## Аномалия удаления (DELETE)

При попытке удалить кортеж (Иванов, Математический, Математика), мы обязаны удалить также и кортеж (Иванов, Математический, Информатика).

Если удалить кортеж (Иванов, Физический, Математика), а вместе с ним и кортеж (Иванов, Физический, Физика), то будет потеряна информация о предметах, которые должны сдаваться на физическом факультете.

## ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

Декомпозиция отношения "Абитуриенты-Факультеты-Предметы" для устранения аномалий не может быть выполнена на основе функциональных зависимостей, т.к. это отношение не содержит никаких функциональных зависимостей. Это отношение является полностью ключевым, т.е. ключом отношения является все множество атрибутов.

# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

В отношении  $R(A, B, C)$  существует *многозначная зависимость* (multi valid dependence, MVD)  $R.A \twoheadrightarrow R.B$  в том случае, если множество значений  $B$ , соответствующее паре значений  $A$  и  $C$ , зависит только от  $A$  и не зависит от  $C$ .



# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

В отношении "Абитуриенты-Факультеты-Предметы" имеется *многозначная зависимость*:

**Факультет → → Абитуриент|Предмет**



# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

*Тривиальной* называется такая многозначная зависимость  $A \twoheadrightarrow B$ , для которой атрибут  $B$  является подмножеством атрибута  $A$  и  $A \cup B = R$ , где  $R$  – рассматриваемое отношение.

*Многозначная зависимость* вида  $X \twoheadrightarrow\rightarrow Y|Z$  называется нетривиальной многозначной зависимостью, если не существует функциональных зависимостей  $X \rightarrow Y$  и  $X \rightarrow Z$ .



# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

В отношении "Абитуриенты-Факультеты-Предметы" имеется именно *нетривиальная многозначная зависимость*:

**Факультет  $\rightarrow\rightarrow$  Абитуриент|Предмет**



# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

Отношение находится в четвертой нормальной форме (4NF) тогда и только тогда, когда отношение находится в НФБК и не содержит нетривиальных многозначных зависимостей.



# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

## Теорема (Фейджина)

Отношение  $R(A, B, C)$  можно спроецировать без потерь в отношения  $R1(A, B)$  и  $R2(A, C)$  в том и только в том случае, когда существует  $MVDA \rightarrow \rightarrow B|C$  ( что равнозначно наличию двух зависимостей  $A \rightarrow B$  и  $A \rightarrow C$ ).

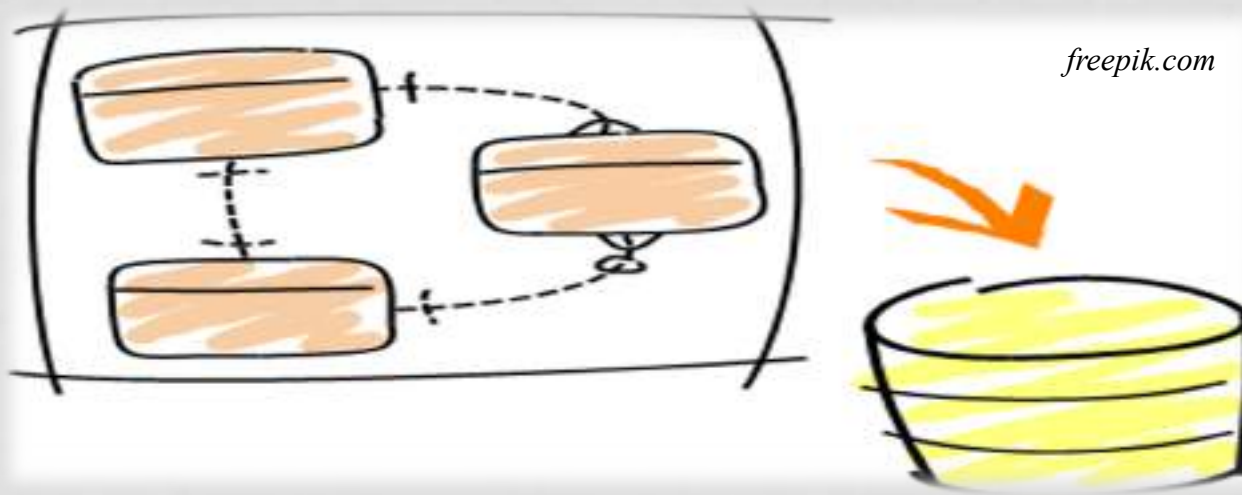
# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

Под *проецированием без потерь* понимается такой способ декомпозиции отношения путем применения операции проекции, при котором исходное отношение полностью и без избыточности восстанавливается путем естественного соединения полученных отношений.



# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

*Зависимость соединения без потерь* называется свойство декомпозиции, которое означает, что при комбинировании отношений с помощью операции естественного соединения не возникают фиктивные строки.



# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

Отношение "Факультеты-Абитуриенты"

Факультет	Абитуриент
Математический	Иванов
Физический	Иванов
Математический	Петров

# ЧЕТВЕРТАЯ НОРМАЛЬНАЯ ФОРМА

Отношение "Факультеты-Предметы"

Факультет	Предмет
Математический	Математика
Математический	Информатика
Физический	Математика
Физический	Физика

# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Отношение R

X	Y	Z
1	1	2
1	2	1
2	1	1
1	1	1



# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Проекция  $R_1 = R[X, Y]$

X	Y
1	1
1	2
2	1

# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Проекция  $R_2=R[X,Z]$

**X**

**Z**

1

2

1

1

2

1

# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Проекция  $R_3 = R[Y, Z]$

<b>Y</b>	<b>Z</b>
1	2
2	1
1	1

# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Отношение  $R$  не восстанавливается ни по одному из попарных соединений:

**R1 JOIN R2**

**R1 JOIN R3**

**R2 JOIN R3**



# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

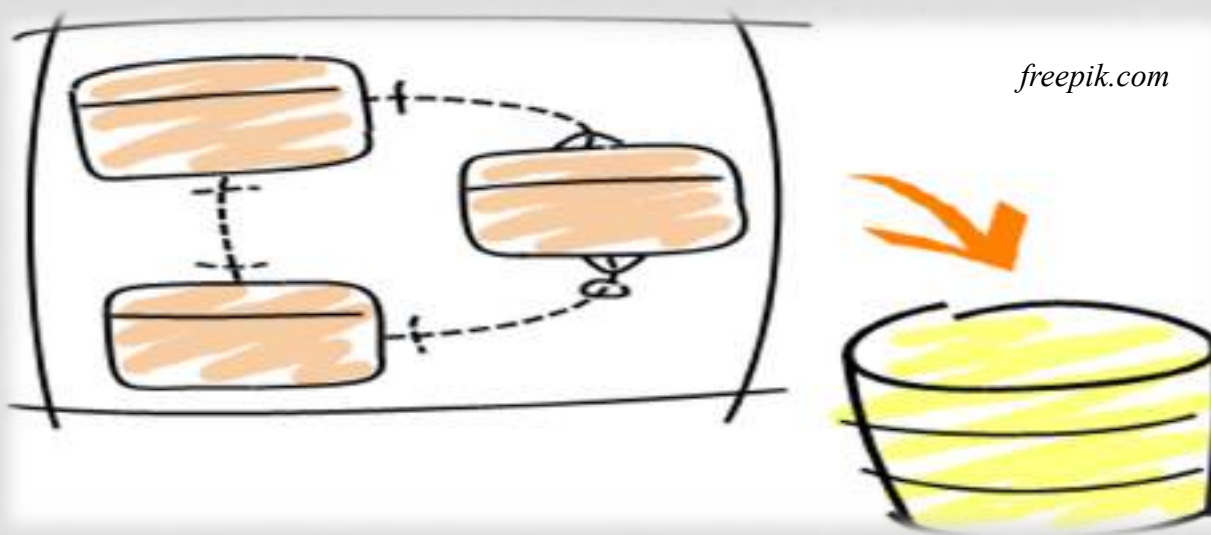
R1 JOIN R2

<b>X</b>	<b>Y</b>	<b>Z</b>
1	1	2
1	1	1
<b>1</b>	<b>2</b>	<b>2</b>
1	2	1
2	1	1

# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Однако отношение  $R$  восстанавливается соединением всех трех проекций:

$$R1 \text{ JOIN } R2 \text{ JOIN } R3 = R$$



# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Отношение  $R(X, Y, \dots, Z)$  удовлетворяет *зависимости соединения*  $*$   $(X, Y, \dots, Z)$  в том и только в том случае, когда  $R$  восстанавливается без потерь путем соединения своих проекций на  $X, Y, \dots, Z$ .



# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Можно предположить, что отношение R из примера удовлетворяет следующей зависимости соединения:

$$*(XY, XZ, YZ)$$





## ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Зависимость соединения  $*(A, B, \dots Z)$  называется **нетривиальной зависимостью соединения**, если выполняется два условия:

- одно из множеств атрибутов  $A, B, \dots Z$  не содержит потенциального ключа отношения  $R$ ;
- ни одно из множеств атрибутов не совпадает со всем множеством атрибутов отношения  $R$ .

# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Зависимость соединения  $*(A, B, \dots Z)$  называется *тривиальной зависимостью соединения*, если выполняется два условия:

- либо все множества атрибутов  $A, B, \dots Z$  содержат потенциальный ключ отношения  $R$ ;
- либо одно из множеств атрибутов совпадает по всем множествам атрибутов отношения  $R$ .

# ПЯТАЯ НОРМАЛЬНАЯ ФОРМА

Отношение находится в пятой нормальной форме (5NF) тогда и только тогда, когда любая имеющаяся зависимость соединения является тривиальной.



# БИЗНЕС-МОДЕЛЬ ПРОЦЕССА НОРМАЛИЗАЦИИ

## Шаг 1 (Приведение к 1НФ)

На первом шаге задается одно или несколько отношений, отображающих понятия предметной области. По модели предметной области выписываются обнаруженные функциональные зависимости. Все отношения автоматически находятся в 1НФ.

# БИЗНЕС-МОДЕЛЬ ПРОЦЕССА НОРМАЛИЗАЦИИ

## Шаг 2 (Приведение к 2НФ)

Если в некоторых отношениях обнаружена зависимость атрибутов от части сложного ключа, то проводим декомпозицию этих отношений на несколько отношений следующим образом: те атрибуты, которые зависят от части сложного ключа выносятся в отдельное отношение вместе с этой частью ключа. В исходном отношении остаются все ключевые атрибуты.

# БИЗНЕС-МОДЕЛЬ ПРОЦЕССА НОРМАЛИЗАЦИИ

## Шаг 3 (Приведение к 3НФ)

Если в некоторых отношениях обнаружена зависимость некоторых неключевых атрибутов других неключевых атрибутов, то проводим декомпозицию этих отношений следующим образом: те неключевые атрибуты, которые зависят других неключевых атрибутов выносятся в отдельное отношение.

# БИЗНЕС-МОДЕЛЬ ПРОЦЕССА НОРМАЛИЗАЦИИ

## Шаг 4 (Приведение к НФБК)

Если имеются отношения, содержащие несколько потенциальных ключей, то необходимо проверить, имеются ли функциональные зависимости, детерминанты которых не являются потенциальными ключами. Если такие функциональные зависимости имеются, то необходимо провести дальнейшую декомпозицию отношений. Те атрибуты, которые зависят от детерминантов, не являющихся потенциальными ключами выносятся в отдельное отношение вместе с детерминантами.

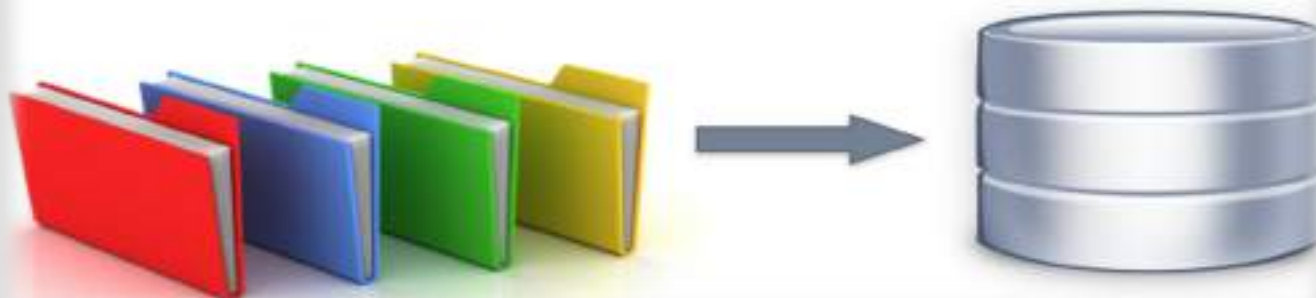


# БИЗНЕС-МОДЕЛЬ ПРОЦЕССА НОРМАЛИЗАЦИИ

## Шаг 5 (Приведение к 4НФ)

Если в отношениях обнаружены нетривиальные многозначные зависимости, то необходимо провести декомпозицию для исключения таких зависимостей.

*freepik.com*





# БИЗНЕС-МОДЕЛЬ ПРОЦЕССА НОРМАЛИЗАЦИИ

## Шаг 6 (Приведение к 5НФ)

Если в отношениях обнаружены нетривиальные зависимости соединения, то необходимо провести декомпозицию для исключения и таких зависимостей.



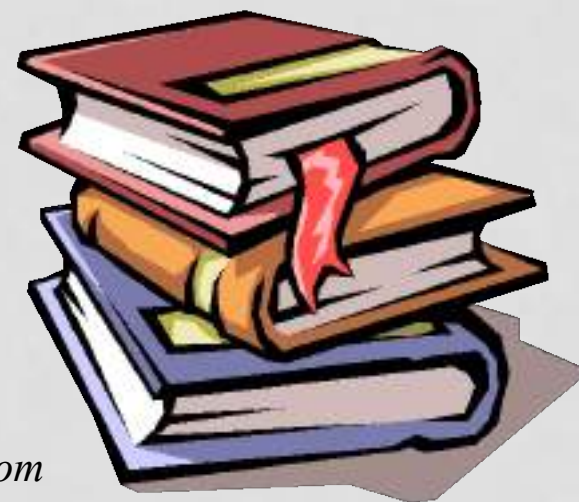
## ЗАДАНИЯ ДЛЯ СРС

1. Для чего предназначен механизм нормализации?
2. Перечислите основные свойства нормальных форм.
3. Какая последовательность нормальных форм обычно выделяется в теории реляционных баз данных?
4. В чем заключается теорема Фейджина?
5. К какой нормальной форме обычно заканчивается процесс проектирования реляционной базы данных на практике?

# РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Т. Конноли, К. Бегг. Базы данных: Проектирование, реализация и сопровождение. Теория и практика [Текст] : учебное пособие / 3-е изд. - М. ; СПб. ; Киев : Вильямс, 2018. - 1440 с.

2. К.Дж. Дейт. Введение в системы баз данных. М., 2018. - 1328 с.



***ЛЕКЦИЯ ОКОНЧЕНА.  
БЛАГОДАРЮ ЗА ВНИМАНИЕ!***