

Избыточность баз данных означает наличие дублирующейся информации в реляционной базе данных или системе управления данными, где идентичные или похожие данные существуют в нескольких местах. Избыточность информации ведет не только к тому, что требуется увеличение объема накопителей, но и приводит к аномалиям в базе данных.

Примеры неверного логического проектирования, которые могут привести к избыточности данных:

- Дублирование информации в разных таблицах без необходимости;
- Некорректное использование связей между таблицами, что приводит к избыточности данных;
- Создание лишних атрибутов в таблицах, содержащих информацию, которая уже хранится в других таблицах.

Для устранения избыточности баз данных необходимо провести следующие **действия**:

1. Проанализировать структуру базы данных и выявить дублирующиеся данные;
2. Определить основные ключи и установить правильные связи между таблицами;
3. Оптимизировать структуру базы данных, удалив лишние атрибуты и таблицы;
4. Разработать нормализованную структуру базы данных, соблюдая нормальные формы;
5. Пересоздать базу данных с устраненными недостатками логической структуры.

Пример: Допустим, у нас есть таблица, в которой хранятся данные списка преподавателей и списка предметов, которые они ведут.

Код преподавателя	Преподаватель	Предметы	Код предмета
1	Иванов	Математический анализ	1
2	Петров	Химия	2
3	Сидоров	Физика	3
4	Куприянова	Дискретная математика	4
5	Петров	Физическая культура	5
6	Рожкова	Аналитическая геометрия	6

Избыточность данных в этой таблице заключается в том, что любой преподаватель может вести несколько предметов, как преподаватель Петров и для каждого нового предмета приходится добавлять новые записи в таблицу. Один преподаватель может вести разные предметы, а разные предметы могут вести разные преподаватели. Но это еще не всё. Допустим, появился новый предмет и мы хотим его добавить в существующую таблицу базы данных, но мы еще не нашли преподавателя для этого предмета. А вписать в таблицу предмет нужно уже сейчас. В этом случае мы должны присвоить значение NULL каждому атрибуту преподавателя, но делать это никак нельзя, так как атрибут «Код преподавателя» является первичным ключом отношения. При попытке это сделать мы получим ошибку. Это является аномалией включения. Чтобы избавиться от аномалии включения нужно разбить таблицу на две: таблица преподавателей и таблица предметов.

Код преподавателя	Преподаватель
1	Иванов

2	Петров
3	Сидоров
4	Куприянова
5	Рожкова

Код предмета	Предмет
1	Математический анализ
2	Химия
3	Физика
4	Дискретная математика
5	Физическая культура
6	Аналитическая геометрия