

Normalizacja: 1NF, 2NF i 3NF (ściągą)

Normalizacja: przed i po arkusz > 3NF

PRZED: jedna tabela z arkusza

nr	klient	miasto	telefony	produkty
10	Anna	Kraków	600100100, 600100101	Monitor 24
11	Anna	Kraków	600100100, 600100101	Klawiatura, Mysz
12	Piotr	Gdańsk	500200200	Kabel HDMI

☐ to samo zapisane dwa razy (redundancja)
☒ kilka wartości w jednym polu (łamać 1NF)
 zmiana miasta Anny = poprawka w każdym jej wierszu
 nie da się zapisać klienta bez zamówienia

PO (3NF): cztery tabele, każda informacja raz

klient

id	imie	miasto
1	Anna	Kraków
2	Piotr	Gdańsk

produkt

id	nazwa	cena
1	Klawiatura	120
2	Mysz	45
3	Monitor 24	650

zamowienie

id	klient_id
10	1
11	1
12	2

pozycja

zamowienie_id	produkt_id	ilosc
10	3	1
11	1	1
11	2	1

telefon(klient_id, numer) osobno: jeden numer w jednym wierszu

Normalizacja to proces organizowania danych w bazie, polegający na wyeliminowaniu nadmiarowych danych oraz niespójnych zależności.

Trzy postacie

Postać	Warunek	Sygnal w tabeli	Naprawa
1NF	pole ma jedną wartość, brak powtarzających się grup kolumn, jest klucz główny	600100100, 600100101 w jednej komórce; kolumny telefon1, telefon2	lista na osobne wiersze albo do osobnej tabeli
2NF	1NF plus kolumna niekluczowa zależy od CAŁEGO klucza	klucz złożony (nr, produkt), a cena zależy tylko od produkt	kolumnę do tabeli, w której ta część klucza jest kluczem głównym
3NF	2NF plus brak zależności przechodnich	miasto zależy od kod, a nie od klucza tabeli	obie kolumny do osobnej tabeli, w macierzystej zostaje klucz obcy

Tabela z kluczem jednokolumnowym, która jest w 1NF, jest już w 2NF: jednoczęściowy klucz nie ma części. Zależność funkcyjna: znając X, znasz Y.

Anomalie

Anomalia	Przykład na arkuszu sklepu
aktualizacji	Anna zmienia miasto: poprawka w każdym jej wierszu, jeden pominięty i baza jest niespójna
wstawiania	nowego klienta bez zamówienia nie ma gdzie wpisać
usuwania	skasowanie zamówienia 12 kasuje jedyny ślad po Piotrze

Metoda na pytanie z rysunkiem

- Lista wartości po przecinku w polu albo powtarzająca się grupa kolumn? Nie jest w 1NF.
- Klucz złożony, a kolumna zależy tylko od jego części? Jest w 1NF, nie w 2NF.
- Kolumna niekluczowa zależy od innej niekluczowej? Jest w 2NF, nie w 3NF.
- Żadne „tak” nie padło? Tabela jest w 3NF.

Pierwsze „tak” kończy sprawdzanie i wskazuje odpowiedź.

Wynik na przykładzie sklepu

przed: zamówienia_arkusz(nr, klient, miasto, kod, telefony, produkty, ceny)

po 3NF:

klient(id PK, imie, miasto, kod)

telefon(id PK, klient_id FK, numer)

produkt(id PK, nazwa, cena, kategoria)

zamowienie(id PK, klient_id FK)

pozycja(zamowienie_id FK, produkt_id FK, ilosc)

Cechy relacyjnej bazy danych, które z tego wychodzą: dane w tabelach, klucz główny unikalny i różny od NULL w każdej tabeli, powiązania przez klucze obce, każda informacja zapisana raz.

Pułapki

- Postaci sprawdza się po kolei: nie ma 3NF bez 2NF ani 2NF bez 1NF.
- 2NF dotyczy wyłącznie tabel z kluczem złożonym; przy kluczu z jednej kolumny nie ma czego łamać.
- Zależność przechodnia to zależność między dwiema kolumnami NIEKLUCZOWYMI, a nie od klucza.
- Dzielenie tabeli zawsze zostawia klucz obcy w miejscu, z którego kolumny wyszły.
- Denormalizacja (cena skopiowana do pozycji zamówienia, żeby została ceną historyczną) jest świadomym wyjątkiem, nie odpowiedzią na egzaminie.
- Podział na tabele mnoży złączenia w kwerendach: to nie błąd projektu, tylko jego koszt.

Egzamin praktyczny: baza z arkusza jest już znormalizowana, dlatego kwerenda „wykorzystująca relację” musi łączyć tabele po kluczu obcym ze schematu podanego w zadaniu.