

Temat: Normalizacja bazy danych - ćwiczenia.

Normalizacja bazy danych to proces polegający na usuwaniu z niej błędów i nieprawidłowości dotyczących struktury tabel.

Jak należy przeprowadzić proces normalizacji bazy danych? W tabeli zostały zgromadzone dane przedstawiające sprzedaż w pewnej firmie:

Tabela - Dane

Kod towaru	Nazwa produktu	Kod klienta	Nazwa klienta	Ilość
TV32	Telewizor 32 cale	JKO	Jan Kowalski	1
OCD	Odtwarzacz CD	TIK	Tomek Iksiński	2
OCD	Odtwarzacz CD	KKO	Kuba Kowalski	3
LOD	Lodówka	KKO	Kuba Kowalski	2

Widać wyraźnie, że część danych w tabeli powtarza się – **redundancja** (*redundancja dla baz danych oznacza, że te same dane pojawiają się w wielu miejscach bazy danych*). Zjawisko to dotyczy głównie baz zbudowanych z jednej tabeli. Powstająca w wyniku redundancji nadmiarowość jest zjawiskiem niekorzystnym, gdyż powoduje niepotrzebny wzrost objętości bazy danych. Nadmiarowość dotyczy pól: kod towaru, nazwa produktu, kod i nazwa klienta. Sprawa skomplikuje się, gdy Kuba Kowalski zmienia nazwę na przykład na Kuba Tobiasz Kowalski. W wyniku takiej zmiany trzeba będzie zmieniać wszystkie pola – wszystkie wiersze, w których występuje Kuba Kowalski. Kolejnym problemem będzie jak zapisać produkt, który nie został jeszcze zakupiony. Taka postać tabeli oznacza, że towar można dopisać tylko wtedy, gdy ktoś go kupi. Takie podejście do gromadzenia danych jest oczywiście błędne i komplikuje obsługę bazy danych. W celu uniknięcia opisanych problemów musimy znormalizować tabelę. Należy podzielić ją na tyle tabel, ile obiektów i działań potrzeba opisać. W tym wypadku trzy: Klienci, Towary, Sprzedaż.

Podsumowując: w wyniku normalizacji otrzymamy trzy tabele.

Tabela – „Klienci”:

Kod klienta	Nazwa klienta
JKO	Jan Kowalski
TIK	Tomek Iksiński
KKO	Kuba Kowalski

Tabela – „Towary”:

Kod towaru	Nazwa produktu
TV32	Telewizor 32 cale
OCD	Odtwarzacz CD
LOD	Lodówka

Tabela „Sprzedaż”

Kod towaru	Kod klienta	Ilość
TV32	JKO	1
OCD	TIK	2
OCD	KKO	3
LOD	KKO	2

Normalizacja bazy danych polega na sprawdzeniu zgodności jej wykonania (projektu) z pierwszą, drugą i trzecią postacią normalną.

Baza danych do prawidłowego (efektywnego) działania musi spełniać założenia normalności.

Przejdźmy do definicji postaci normalnych:

Pierwsza postać normalna

Relacja jest w pierwszej postaci normalnej, jeśli wartości atrybutów są elementarne (dalej niepodzielne, jak najmniejsze porcje informacji). Oznacza, że są to pojedyncze wartości określonego typu.

Druga postać normalna

Relacja pomiędzy tabelami jest w drugiej postaci normalnej, jeżeli jest w pierwszej postaci normalnej i nie zawiera redundancji.

Druga postać normalna powstaje w wyniku utworzenia oddzielnych tabel dla zestawów wartości, odnoszących się do wielu rekordów, a następnie powiązania tak powstałych tabel za pomocą klucza zewnętrznego.

Trzecia postać normalna

Relacja pomiędzy tabelami jest w trzeciej postaci normalnej, jeśli jest ona w drugiej postaci normalnej i wartości jej atrybutów zależą tylko od klucza podstawowego, a nie od siebie nawzajem.

Postać ta powstaje w wyniku wyeliminowania pól, które nie zależą od klucza podstawowego (pola te automatycznie wchodzą w skład kolejnej tabeli).

Po zakończeniu procesu normalizacji należy dokonać stosownych modyfikacji w schematach tabel. Na pewno trzeba będzie uzupełnić system powiązań relacyjnych. Prawdopodobnie pojawią się relacje jedno-do-jednego charakterystyczne dla tabel zawierających dodatkowe informacje o obiektach.

Ćwiczenie 1

Dokonaj normalizacji bazy danych przechowującej informacje o osobach.
Której postaci normalnej nie spełnia ta tabela? Dlaczego?

Tabela 1

id	Dane_osobowe
1	Jan Nowak, 80-870, Gdańsk, Reja 25
2	Paweł Kowalski, 80-125, Gdańsk, Kartuska 32
3	Anna Nowak, 81-150, Sopot, Morska 13

Ćwiczenie 2

Dokonaj normalizacji bazy danych przechowującej informacje o zwierzętach.
Której postaci normalnej nie spełnia ta tabela? Dlaczego?

Tabela 2

zwierzę	opiekun	adres
Iwica Saba	Jan Król	Lewin, Ruska 6
Ilew King	Jan Król	Lewin, Ruska 6
szympanś Kiko	Anna Las	Lewin, Kłodzka 5/6

Ćwiczenie 3

Dokonaj normalizacji bazy danych przechowującej informacje o zawodach sportowych.
Której postaci normalnej nie spełnia ta tabela? Dlaczego?

Tabela 3

dane zawodnika	dyscyplina sportowa	miasto	data	czy jest mistrzem
Lewis Hamilton	Formuła 1	Nurburgring	11-10-2020	Tak
Dawid Kubacki	Skoki narciarskie	Szczyrk	17-10-2020	Tak
Iga Świątek	Tenis	Paryż	10-10-2020	Tak