

Przykład:

Niech dany będzie zestaw danych Zamówienia (poniżej) przedstawiający informacje o klientach oraz składanych/realizowanych przez nich zamówieniach. Należy wykonać normalizację danych do postaci:

- 1NF (ang. First Normal Form)
- 2NF
- 3NF
- BCNF (ang. Boyce-Codd Normal Form)

Własności postaci normalnych:

1NF – Relacja posiada **klucz** oraz wszystkie atrybuty są **atomowe**.

2NF – Każdy atrybut niekluczowy zależy od **całego** klucza.

3NF – Każdy atrybut niekluczowy zależy **bezpośrednio** od klucza.

BCNF – Dla **każdej** nietrywialnej zależności funkcyjnej po lewej stronie znajduje się **nadklucz**.

ZAMÓWIENIA:

NIP KLIENTA	NAZWISKO & IMIĘ	KOD TOWARU	NAZWA	CENA NETTO	GRUPA VAT	VAT
993-20-33-40	Kowal Anna	Z20	Cegła	10	I	23%
		P10	Piasek	100	II	7%
777-22-44-66	Nowak Ewa	Z20	Cegła	10	I	23%
311-11-11-11	Polak Kinga	K10	Cement	50	III	0%
		Z20	Cegła	10	I	23%

Powyższy zestaw danych opiszemy jako relację $Z(K, N, T, W, E, G, V)$.

Na podstawie przedstawionego powyżej zestawu danych, biorąc pod uwagę przedmiot modelowania (znaczenie poszczególnych danych/kolumn i powiązania między nimi) możemy wyznaczyć następujące zależności funkcyjne:

$K \rightarrow N$ (NIP KLIENTA określa jednoznacznie/identyfikuje NAZWISKO & IMIĘ albo inaczej pisząc NAZWISKO & IMIĘ zależy funkcyjnie od NIP-u KLIENTA)

$T \rightarrow WEG$ (KOD TOWARU określa jednoznacznie/identyfikuje jego NAZWĘ, CENĘ NETTO oraz GRUPĘ VAT do której towar należy).

$G \rightarrow V$ (GRUPA VAT określa jednoznacznie/identyfikuje stawkę VAT dla grupy)

Sprawdzenie kryteriów 1NF

Aby zweryfikować czy **spełniona jest 1NF** musimy ocenić czy atrybuty są atomowe (niepodzielne) oraz wyznaczyć klucz podstawowy relacji.

Atrybuty zadanego zestawu danych nie są atomowe, zatem dokonujemy podziału (pionowego) kolumny NAZWISKO & IMIĘ na dwie niezależne. Ponadto dokonujemy podziału (poziomego) zamówień Kowal Anny oraz Polak Kingi. W rezultacie otrzymujemy poniższy zestaw danych.

NIP KLIENTA	NAZWISKO	IMIĘ	KOD TOWARU	NAZWA	CENA NETTO	GRUPA VAT	VAT
993-20-33-40	Kowal	Anna	Z20	Cegła	10	I	23%
993-20-33-40	Kowal	Anna	P10	Piasek	100	II	7%
777-22-44-66	Nowak	Ewa	Z20	Cegła	10	I	23%
311-11-11-11	Polak	Kinga	K10	Cement	50	III	0%
311-11-11-11	Polak	Kinga	Z20	Cegła	10	I	23%

Aktualnie powyższy zestaw danych opiszemy jako relację Z(K, N, I, T, W, E, G, V) dla której zależności funkcyjne przedstawiają się następująco:

$K \rightarrow NI$

$T \rightarrow WEG$

$G \rightarrow V$

Wszystkie atrybuty są atomowe.

Wyznaczamy klucz relacji (patrz dokument „Wyjaśnienie zależności funkcyjnych oraz kluczy”).

KT	G	NIWEV
----	---	-------

Potencjalne klucze to KT albo KTG.

$KT^+ = KT$

$KT^+ = KNIT$

$KT^+ = KNITWEG$

$KT^+ = KNITWEGV$

KT jest kluczem podstawowym relacji.

Aktualnie relacja ZAMÓWIENIA spełnia kryteria 1NF.

Sprawdzenie kryteriów 2NF

Aby zweryfikować czy **spełniona jest 2NF** musimy ocenić czy każdy atrybut niekluczowy zależy od **całego** klucza.

Ponieważ kluczem jest KT oraz ponieważ NI występuje tylko w jednej zależności funkcyjnej $K \rightarrow NI$ tj. NAZWISKO oraz IMIĘ zależą tylko od K a nie zależą od T, możemy stwierdzić, że relacja nie spełnia kryteriów 2NF.

Dokonujemy zatem podziału relacji Z na dwie, wyodrębniając do jednej atrybuty niekluczowe naruszające kryterium 2NF (tj. NAZWISKO i IMIĘ) wraz z atrybutem kluczowym od którego zależą (tj. NIP KLIENTA). Natomiast w drugiej pozostawiamy wszystkie atrybuty kluczowe (tj. NIP KLIENTA oraz KOD TOWARU) oraz pozostałe atrybuty niekluczowe. W wyniku podziału otrzymamy dwie relacje KLIENCI oraz TOWARY.

KLIENCI:

NIP KLIENTA	NAZWISKO	IMIĘ
993-20-33-40	Kowal	Anna
993-20-33-40	Kowal	Anna
777-22-44-66	Nowak	Ewa
311-11-11-11	Polak	Kinga
311-11-11-11	Polak	Kinga

$L(K, N, I)$

$K \rightarrow NI$

$K^+ = K$

$K^+ = KNI$

Zatem K jest kluczem podstawowym relacji. Wszystkie atrybuty są atomowe. Ponieważ kluczem jest K oraz istnieje jedna zależność funkcyjna w postaci $K \rightarrow NI$ zatem możemy stwierdzić, że każdy atrybut niekluczowy zależy od całego klucza.

Aktualnie relacja KLIENCI spełnia kryteria 2NF.

Kontynuujemy weryfikację 2NF dla powstałej z podziału drugiej relacji (poniżej) Y(K, T, W, E, G).

TOWARY:

NIP KLIENTA	KOD TOWARU	NAZWA	CENA NETTO	GRUPA VAT	VAT
993-20-33-40	Z20	Cegła	10	I	23%
993-20-33-40	P10	Piasek	100	II	7%
777-22-44-66	Z20	Cegła	10	I	23%
311-11-11-11	K10	Cement	50	III	0%
311-11-11-11	Z20	Cegła	10	I	23%

$T \rightarrow WEG$

$G \rightarrow V$

Wyznaczamy klucz relacji:

KT	G	WEV
----	---	-----

Potencjalne klucze to KT albo KTG.

$KT^+ = KT$

$KT^+ = KTWEG$

$KT^+ = KTWEGV$

KT jest kluczem podstawowym relacji. Wszystkie atrybuty są atomowe. Ponieważ kluczem jest KT oraz ponieważ WE występuje tylko w jednej zależności funkcyjnej $T \rightarrow WEG$ tj. NAZWA oraz CENA NETTO zależą tylko od T a nie zależą od K, możemy stwierdzić, że relacja nie spełnia kryteriów 2NF. Podobnie V występujące w dwóch zależnościach funkcyjnych $T \rightarrow WEG$ oraz $G \rightarrow V$ zależy tranzytywnie (przechodnio) tylko od T a nie zależy od K. Dokonujemy zatem podziału relacji Y na dwie.

ZAMÓWIENIA:

NIP KLIENTA	KOD TOWARU
993-20-33-40	Z20
993-20-33-40	P10
777-22-44-66	Z20
311-11-11-11	K10
311-11-11-11	Z20

$Z(K, T)$

Brak zależności funkcyjnych dla w/w relacji, zatem kluczem podstawowym jest KT. Wszystkie atrybuty są atomowe. Brak atrybutów niekluczowych, zatem **relacja KLIENCI na pewno spełnia kryteria 2NF.**

TOWARY:

KOD TOWARU	NAZWA	CENA NETTO	GRUPA VAT	VAT
Z20	Cegła	10	I	23%
P10	Piasek	100	II	7%
Z20	Cegła	10	I	23%
K10	Cement	50	III	0%
Z20	Cegła	10	I	23%

$Y(T, W, E, G)$

$T \rightarrow WEG$

$G \rightarrow V$

Wyznaczamy klucz relacji:

T	G	WEV
---	---	-----

Potencjalne klucze to T albo TG.

$T^+ = T$

$T^+ = TWEV$

$T^+ = TWEV$

Zatem T jest kluczem podstawowym relacji. Wszystkie atrybuty są atomowe. Ponieważ kluczem jest T oraz istnieją zależności funkcyjne w postaci $T \rightarrow WEV$, $G \rightarrow V$ zatem możemy stwierdzić, że każdy atrybut niekluczowy zależy od całego klucza (V zależy tranzytywnie przez G).

Relacja TOWARY spełnia kryteria 2NF.

Sprawdzenie kryteriów 3NF

Aby zweryfikować czy **spełniona jest postać 3NF** musimy ocenić czy każdy atrybut niekluczowy zależy od **bezpośrednio** od klucza.

W przypadku niniejszego przykładu relacje KLIENCI oraz ZAMÓWIENIA spełniają 3NF. Natomiast kryterium 3NF nie spełnia relacja TOWARY.

TOWARY:

KOD TOWARU	NAZWA	CENA NETTO	GRUPA VAT	VAT
Z20	Cegła	10	I	23%
P10	Piasek	100	II	7%
Z20	Cegła	10	I	23%
K10	Cement	50	III	0%
Z20	Cegła	10	I	23%

$Y(T, W, E, G)$

$T \rightarrow WEV$

$G \rightarrow V$

Kluczem podstawowym relacji jest T.

W relacji TOWARY występuje tranzytywność tj. atrybut niekluczowy V zależy od G ($G \rightarrow V$), które z kolei zależy od klucza tj. T. Zatem V nie zależy bezpośrednio od T lecz poprzez G. Mamy naruszenie kryterium 3NF. Dokonujemy zatem podziału relacji Y na dwie, wyodrębniając do jednej atrybut niekluczowy naruszający kryterium 3NF (tj. VAT) wraz z atrybutem kluczowym od którego zależą (tj. GRUPA VAT). Natomiast w drugiej pozostawiamy wszystkie atrybuty kluczowe (tj. KOD

TOWARU) oraz pozostałe atrybuty niekluczowe (NAZWA, CENA NETTO oraz GRUPA VAT). W wyniku podziału otrzymujemy.

TOWARY:

KOD TOWARU	NAZWA	CENA NETTO	GRUPA VAT
Z20	Cegła	10	I
P10	Piasek	100	II
Z20	Cegła	10	I
K10	Cement	50	III
Z20	Cegła	10	I

$Y(T, W, E, G)$

$T \rightarrow WEG$

Kluczem podstawowym relacji jest T. Wszystkie atrybuty są atomowe. Ponieważ kluczem jest T oraz istnieje zależność funkcyjna w postaci $T \rightarrow WE$ zatem możemy stwierdzić, że każdy atrybut niekluczowy zależy od całego klucza. Ponadto każdy atrybut niekluczowy tj. W, E, G zależy bezpośrednio od klucza.

Relacja TOWARY spełnia kryteria 3NF.

Ponadto w wyniku podziału otrzymujemy:

GRUPY:

GRUPA VAT	VAT
I	23%
II	7%
I	23%
III	0%
I	23%

$U(G, V)$

$G \rightarrow V$

Kluczem podstawowym relacji jest G. Wszystkie atrybuty są atomowe. Ponieważ kluczem jest G oraz istnieje zależność funkcyjna w postaci $G \rightarrow V$ zatem możemy stwierdzić, że każdy atrybut niekluczowy zależy od całego klucza (w tym przypadku istnieje jeden atrybut niekluczowy V). Ponadto każdy atrybut niekluczowy tj. V zależy bezpośrednio od klucza.

Relacja GRUPY spełnia kryteria 3NF.

Sprawdzenie kryteriów BCNF

Aby zweryfikować czy spełniona jest postać **BCNF** musimy sprawdzić czy dla **każdej** nietrywialnej zależności funkcyjnej po lewej stronie znajduje się **nadklucz**.

Przykład (typy zależności funkcyjnych):

Niech dana będzie relacja **PODATNICY**(**NIP**, **PESEL**, **IMIE**, **NAZWISKO**)

Zależność trywialna: $NS \rightarrow N$ albo $S \rightarrow S$

Zależność nietrywialna: $NP \rightarrow NIW$ albo $S \rightarrow SIW$

Zależność całkowicie nietrywialna: $N \rightarrow IW$ albo $S \rightarrow IW$ albo $S \rightarrow N$ albo $N \rightarrow S$

W analizowanym zadaniu wszystkie otrzymane relacje tj. **KLIENCI**, **ZAMÓWIENIA**, **TOWARY**, **GRUPY** spełniają kryteria **BCNF**.

KLIENCI:

NIP KLIENTA	NAZWISKO	IMIE
993-20-33-40	Kowal	Anna
993-20-33-40	Kowal	Anna
777-22-44-66	Nowak	Ewa
311-11-11-11	Polak	Kinga
311-11-11-11	Polak	Kinga

Relacja: $L(K, N, I)$

Klucz podstawowy: **K**

Zależności funkcyjne: $K \rightarrow NI$ (zależność całkowicie nietrywialna)

Nadklucz: **K**

Relacja KLIENCI spełnia kryteria BCNF.

ZAMÓWIENIA:

NIP KLIENTA	KOD TOWARU
993-20-33-40	Z20
993-20-33-40	P10
777-22-44-66	Z20
311-11-11-11	K10
311-11-11-11	Z20

Relacja: $Z(K, T)$

Klucz podstawowy: **K, T**

Zależności funkcyjne: brak

Nadklucz: **K, T**

Relacja ZAMÓWIENIA spełnia kryteria BCNF.

TOWARY:

KOD TOWARU	NAZWA	CENA NETTO	GRUPA VAT
Z20	Cegła	10	I
P10	Piasek	100	II
Z20	Cegła	10	I
K10	Cement	50	III
Z20	Cegła	10	I

Relacja: Y(T, W, E, G)

Klucz podstawowy: T

Zależności funkcyjne: $T \rightarrow WEG$ (zależność całkowicie nietrywialna)

Nadklucz: T

Relacja TOWARY spełnia kryteria BCNF.

GRUPY:

GRUPA VAT	VAT
I	23%
II	7%
I	23%
III	0%
I	23%

Relacja: U(G, V)

Klucz podstawowy: G

Zależności funkcyjne: $G \rightarrow V$ (zależność całkowicie nietrywialna)

Nadklucz: G

Relacja GRUPY spełnia kryteria BCNF.