



# Bazy Danych

---

## Projektowanie. Normalizacja.

Krzysztof Regulski  
WIMiP, KISiM,  
regulski@agh.edu.pl  
B5, pok. 408

---

## Etapy projektowania baz danych:

---

1. **Specyfikacja wymagań użytkownika** - określenie zjawisk, dostępności i użyteczności danych, ich formatu i sposobów obliczeń, cele, zakres i kontekst systemu
2. **Projektowanie koncepcyjne** - projektowanie schematu E–R bazy. Użycie modelu E–R wpływa również na realizację pozostałych faz.
3. **Specyfikacja wymagań funkcjonalnych** - dokładny opis wymagań klienta i wszystkich przyszłych użytkowników systemu
4. **Projektowanie logiczne i fizyczne**
5. **Implementacja**

## – Projektowanie logicznej struktury bazy:

### » **Etap I:** określenie encji i zdefiniowanie atrybutów opisujących encje

- przyporządkowanie encji do zjawisk
- standaryzacja nazw i formatów
- identyfikacja źródeł danych

### » **Etap II:** określenie związków między encjami

- identyfikacja typu związków (relacji) (1-1, 1-M, N-M)

### » **Etap III:** normalizacja relacji

- obniżenie redundancji i wyeliminowanie anomalii (usuwania, wstawiania i aktualizacji)

## Wybór schematu relacji:

- Projekt bazy danych polega na znalezieniu właściwych schematów relacji tworzących bazę danych
- **Niewłaściwy projekt** (niewłaściwe schematy) mogą prowadzić do
  - » Redundancji (powtarzania informacji).
  - » Niemożności reprezentowania pewnych informacji.
  - » Anomalii związanych z operowaniem danymi (głównie modyfikacją danych)
- **Cele:**
  - » Unikanie redundancji danych
  - » Zapewnienie reprezentowania związków między danymi
  - » Zachowanie warunków integralności (umożliwienie kontroli warunków integralności podczas modyfikacji danych).

## **Model pojęciowy (konceptualny, podstawowy):**

- Zapis wymagań w postaci sformalizowanej, abstrahujący od problemów implementacyjnych („co” a nie „jak”)
- Zawiera
  - » modele graficzne
  - » tekstowe (ale też sformalizowane) uzupełnienia modeli graficznych
- Jest ważny, ponieważ:
  - » reprezentuje istotę wymagań
  - » nie zależy od zmiennych możliwości implementacji
  - » jest dobrą podstawą do projektowania systemów mających funkcjonować przez wiele lat

## Wymagania:

---

- Spełnienie wymagań użytkownika jest naszym celem działania
- Wymagania muszą więc być dokładnie znane
- Poprawne i kompletne sformułowanie wymagań na ogół nie jest proste
- Należy odróżniać *funkcje* systemu od *mechanizmów*, czyli sposobów implementacji funkcji
- Modelować dane niezależnie od przyszłej implementacji - *modelowanie E/R*
- Tworząc model pojęciowy nie można formułować zastrzeżeń co do realizowalności wymagań
- Zaprojektować dobry schemat - *normalizacja*

## Cel modelowania E-R:

---

- Model **E-R** służy do nieformalnego przedstawienia projektu bazy danych.
- Projekt ma postać graficzną zwaną diagramem **E-R** (entity-relationship diagram), diagramem **jednostka-związek** lub diagramem **związków encji**
- Istnieje procedura (pół)automatycznej transformacji diagramu **E-R** do konkretnej implementacji, na przykład do relacyjnej bazy danych.

# Encja (zbiory encji):

- **Encja (*entity*)** - model rzeczy, osób, zjawisk, pojęć itp., o których chcemy przechowywać informacje, które mają tożsamość (są rozróżnialne)

Np. konkretna osoba, firma, zdarzenie

- **Zbiór encji (*entity set*)** – zbiór (klasa) obiektów, które są tego samego typu (mają wspólne własności), inaczej: zbiór wystąpień (instancji) tej klasy

Np. zbiór osób, zbiór firm, zbiór określonych zdarzeń

zbiór encji

student		
12345	Jan	Pies
12346	Ewa	Lis
12347	Adam	Słoń

krotka

W diagramach E-R zbiory encji oznaczane są prostokątami:

student



# Atrybuty:

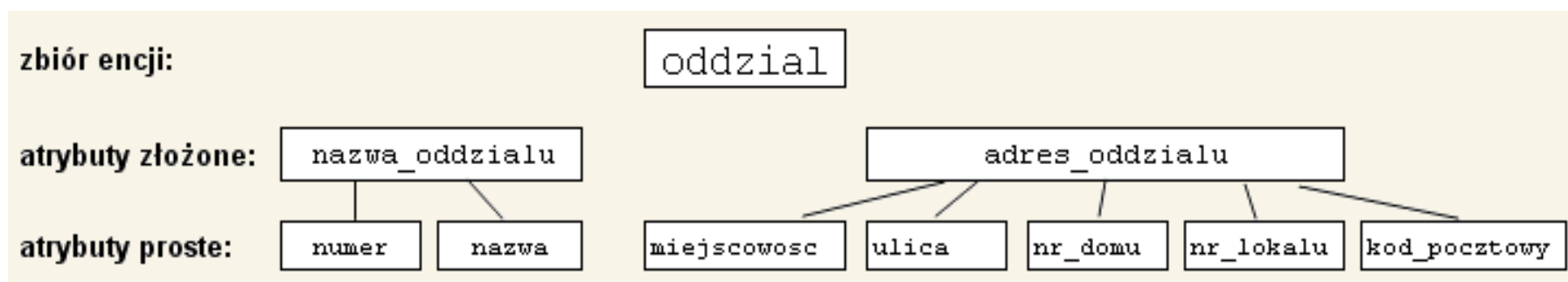
---

- Zbiór encji (i każda z nich) jest opisywany przez **zestaw atrybutów**, które odpowiadają własnościom posiadanym przez wszystkie obiekty reprezentowane przez ten zbiór. Np.:

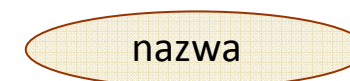
**student = (nr\_albumu, imie, nazwisko)**

- **Dziedzina atrybutu** – zbiór dopuszczalnych wartości, jakie może przyjmować dany atrybut na danym zbiorze encji. Dziedzina uzupełniona jest o wartość **NULL** oznaczającą, że encja nie posiada wartości atrybutu lub jest on nieznany.
- **Atrybuty:**
  - » proste vs. pochodne (mogą być obliczone na podstawie innych informacji w bazie),
  - » proste vs. złożone.

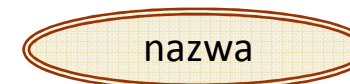
# Atrybuty proste vs. złożone (przykład):



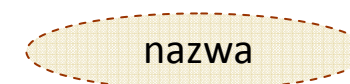
W diagramach atrybuty oznaczane są **elipsami**:



Atrybuty złożone **elipsami z podwójną linią**:

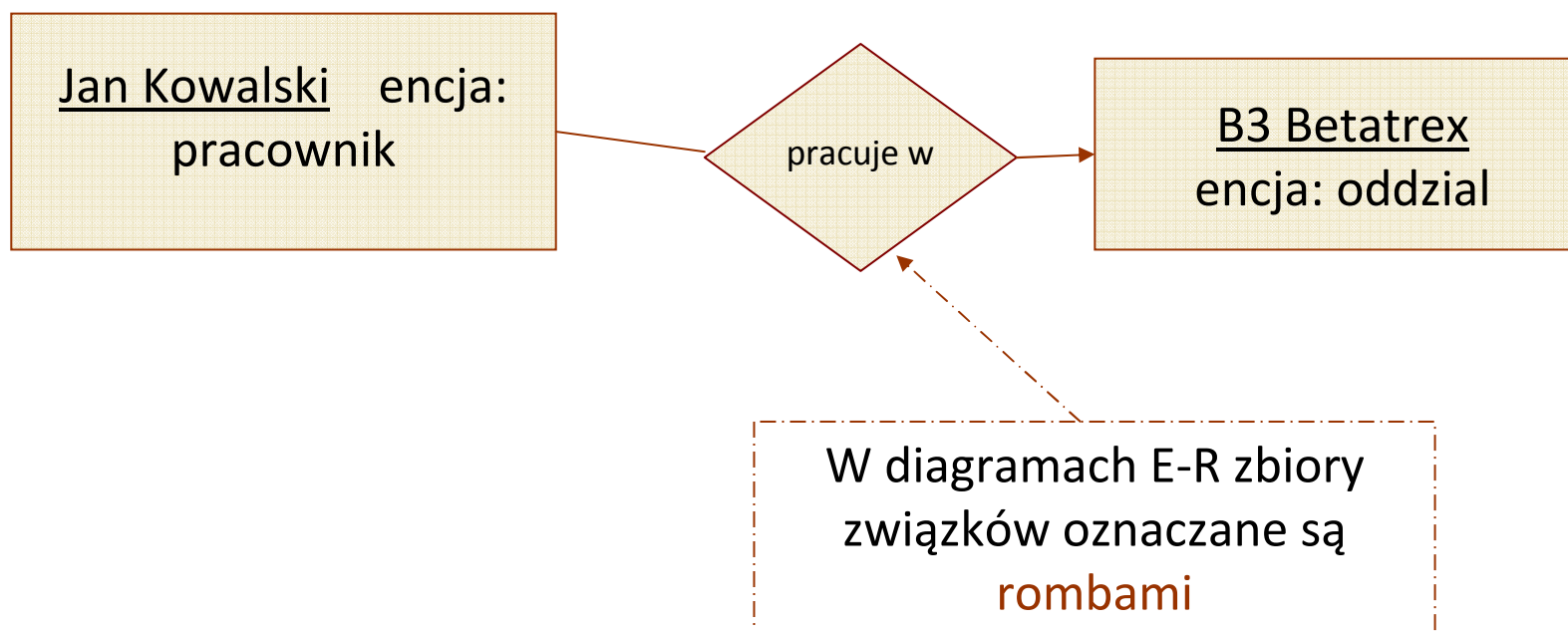


Atrybuty pochodne **elipsami z przerywanej linii**:

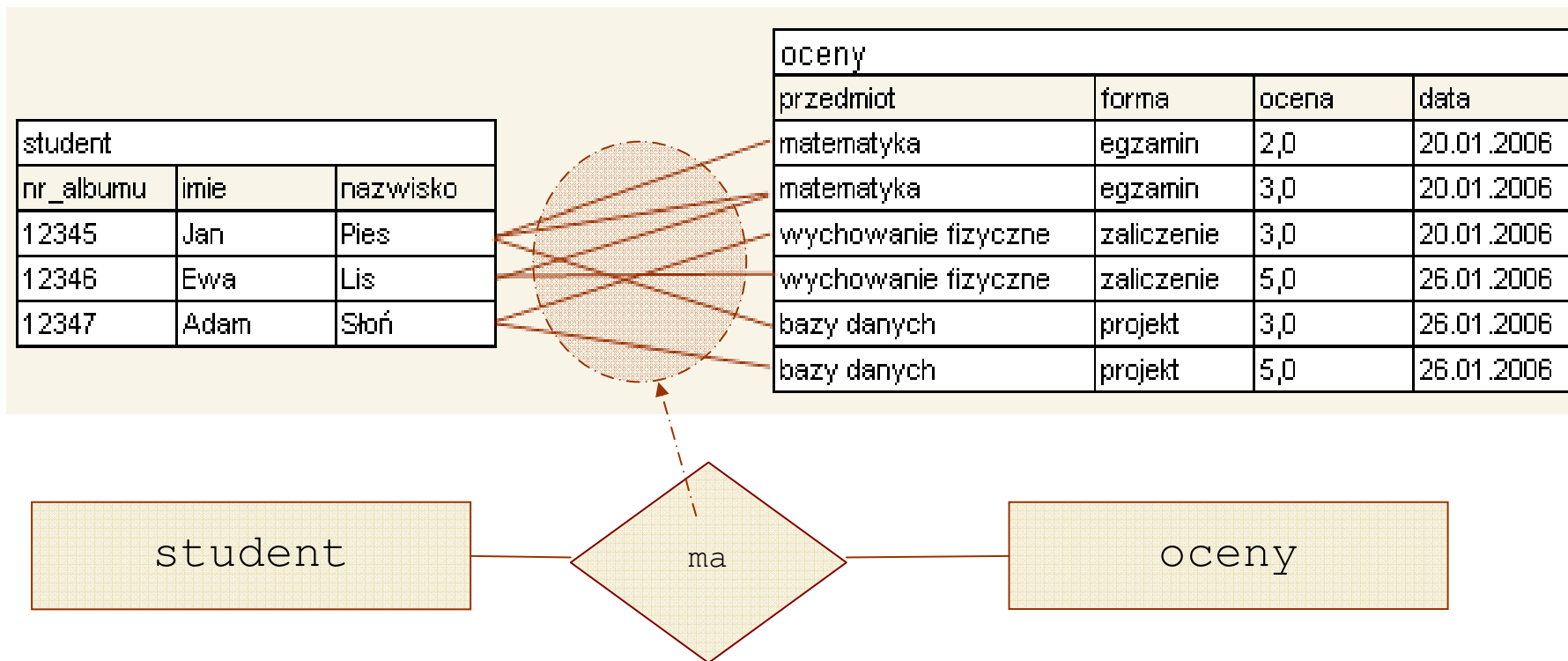


## Zbiory związków:

- **Związek** (*relationship*) reprezentuje relację pomiędzy pewnymi encjami (obiektami). Np.:



# Zbiór związków (przykład):



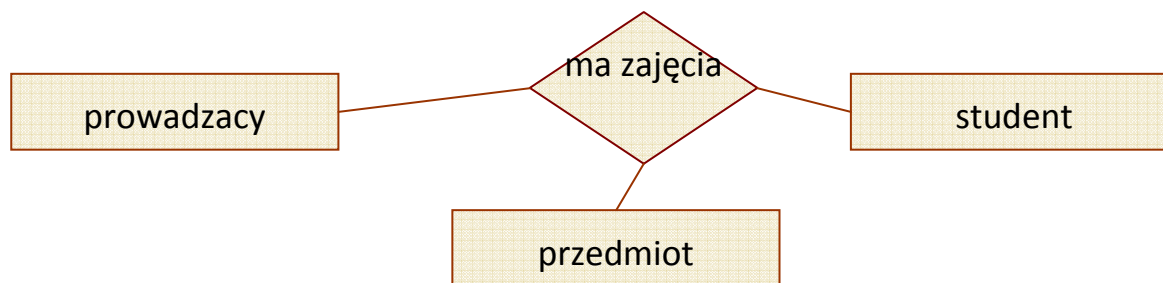
## Diagramy E-R – symbole podstawowe:

---

- **Prostokąty** – zbiory encji
- **Elipsy** – atrybuty
- **Elipsy – linia podwójna** – atrybuty złożone
- **Elipsy – linia przerywana** – atrybuty pochodne
- **Romby** – zbiory związków
- **Linie** – łączą zbiory encji ze zbiorami związków oraz atrybuty ze zbiorami encji i związków

## Stopień zbioru związków:

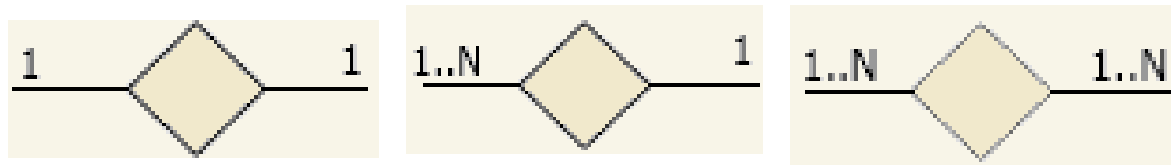
- **Stopień zbioru związków** (*degree of a relationship set*) określony jest przez liczbę zbiorów encji, które pozostają w danym związku.
- Związek dwóch zbiorów encji nazywany jest **binarnym** (stopień = 2).
- **Większość związków obserwowanych w świecie rzeczywistym to związki binarne.**
- W związku może uczestniczyć pewna, większa od 2 ilość zbiorów encji – związek ***n*-arny** (stopień = *n*) – mówimy wtedy o związku wieloargumentowym lub wielocłonowym.
  - » Np. zbiory encji: **przedmiot**, **przewodzący** oraz **student** mogą być połączone związkiem o stopniu = 3.



## Stopień przyporządkowania (krotność związków):

- **Stopień przyporządkowania** zbiorów encji do związku (relacji) określony jest przez ilość encji z każdego zbioru, które są ze sobą w tym związku powiązane.
- **Rodzaje związków binarnych** ze względu na stopień przyporządkowania:
  - *jeden do jeden*
  - *jeden do wielu*
  - *wiele do wielu*
- **W diagramie E-R** rodzaje związków są oznaczane w następujący sposób:
  - (grot strzałki) oznacza stopień przyporządkowania *jeden*
  - (brak grotu) – *wielu*

lub:



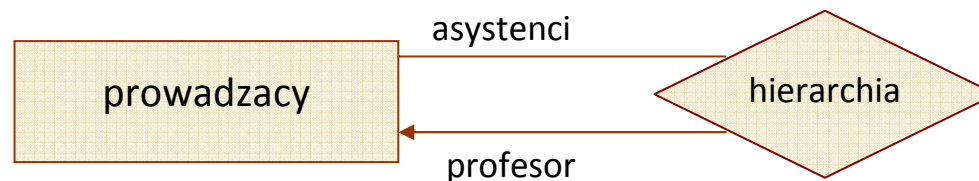
## Stopień uczestnictwa:

- Jeżeli każda encja zbioru **musi** brać udział w związku, wówczas jest on całkowity (uczestnictwo obowiązkowe); w przeciwnym razie – częściowy (uczestnictwo opcjonalne).
- **W diagramie E-R linie podwójne** łączą zbiory encji ze zbiorami związków w przypadku całkowitości.
- Występowanie encji pewnego zbioru (istnienie obiektu w rzeczywistości) może zależeć od występowania encji (obektu) innego zbioru.
- Jeżeli występowanie encji  $x$  zależy od  $y$  (*is existence dependent on*), to:
  - »  $y$  jest encją nadrzędną (dominującą) (*dominant entity*)
  - »  $x$  jest encją podrzędną (*subordinate entity*)
- Np. jeśli encja dominująca **student** zostanie usunięta z bazy, to wszystkie wystąpienia związanych z nią encji **oceny** także muszą zostać usunięte.



## Role:

- Zbiory encji pozostające w związku niekoniecznie muszą być różne, wówczas encje takie występują w różnych **rolach**, a związek taki nazywamy **związkiem rekurencyjnym**.
- **W diagramie E-R** role oznaczane są przez etykiety przy liniach łączących symbol encji (prostokąt) i związku (romb). Oznaczanie ról nie jest obowiązkowe, stosuje się je dla podkreślenia semantyki związku.



# Decyzje projektowe:

- **Zbiory encji czy atrybuty**

Wybór jest zależny od rodzaju modelowanego przedsięwzięcia (fragmentu rzeczywistości) oraz od znaczenia poszczególnych atrybutów.

- **Zbiory encji czy zbiory związków**

Możliwa wskazówka: związek wyznacza pewną akcję (zdarzenie), która zachodzi (zaszła) pomiędzy encjami (obiektami).

- **Związki binarne czy związki o wyższym stopniu**

Związek o stopniu większym niż 2 zawsze może być zastąpiony przez pewną liczbę związków binarnych. Jednak związki n-arne mogą lepiej oddawać semantykę relacji (jawne określenie, że chodzi o związek więcej niż 2 obiektów).

- **Przykład:**

- 1. Adres jako atrybut**

Adres może być traktowany jako atrybut (na przykład encji **student**), wówczas jego wartością jest tekst, np: "Poznań, ul. Kwiatowa 5 m. 12".

- 2. Adres jako jednostka**

Adres może być traktowany jako niezależna jednostka o atrybutach: **Miejscowość, Ulica, Nr Domu, Nr Lokalu.**

## Ograniczenia - warunki integralności

---

- W ramach projektowania baz danych specyfikuje się pewne ograniczenia ilościowe i zależności pomiędzy encjami lub związkami, które mają zachodzić **zawsze**.
- Ograniczenia te i zależności określają **warunki integralności** rozważnej bazy i są następnie sprawdzane i dotrzymywane.
- Do najistotniejszych warunków ogólnych, specyfikowalnych w modelu E-R należą:
  - » stopień przyporządkowania zbiorów encji do związku,
  - » zależność występowania encji z różnych zbiorów,
  - » całkowitość vs. częściowość udziału zbioru encji w zbiorze związków.

## Klucze w zbiorze encji:

---

- **Nadkluczem** (*super key*) jest każdy zbiór atrybutów, którego wartości jednoznacznie identyfikują encję w zbiorze encji (częściowa zależność funkcjonalna).
- **Kluczem kandydującym** (*candidate key*) nazywany jest nadklucz o minimalnej liczbie atrybutów w zbiorze encji (pełna zależność funkcjonalna).
- W danym zbiorze może występować kilka kluczy kandydujących, jeden z nich (dowolnie wybrany) nazywany jest **kluczem głównym** (*primary key*).
- **W diagramie E-R** klucze główne są oznaczane przez podkreślenie etykiety atrybutu.

## Klucze w zbiorze związków:

---

- Jeżeli zbiór związków posiada atrybuty o odpowiednich własnościach, na podstawie tych atrybutów można zdefiniować dla niego (jednoznacznie identyfikujące) **klucze główne** tak, jak dla zbiorów encji.
- *Złożenie* kluczy głównych zbiorów encji biorących udział w związku tworzy **nadklucz** dla zbioru związków.
- Sposób zdefiniowania **klucza głównego** zależy od stopnia związku oraz stopnia przyporządkowania poszczególnych zbiorów encji.
  - » Dla binarnego związku *jeden do jeden* każdy z kluczy głównych (encji) może być kluczem głównym.
  - » Dla binarnego związku *wielu do wielu* para kluczy głównych encji jest kluczem głównym.

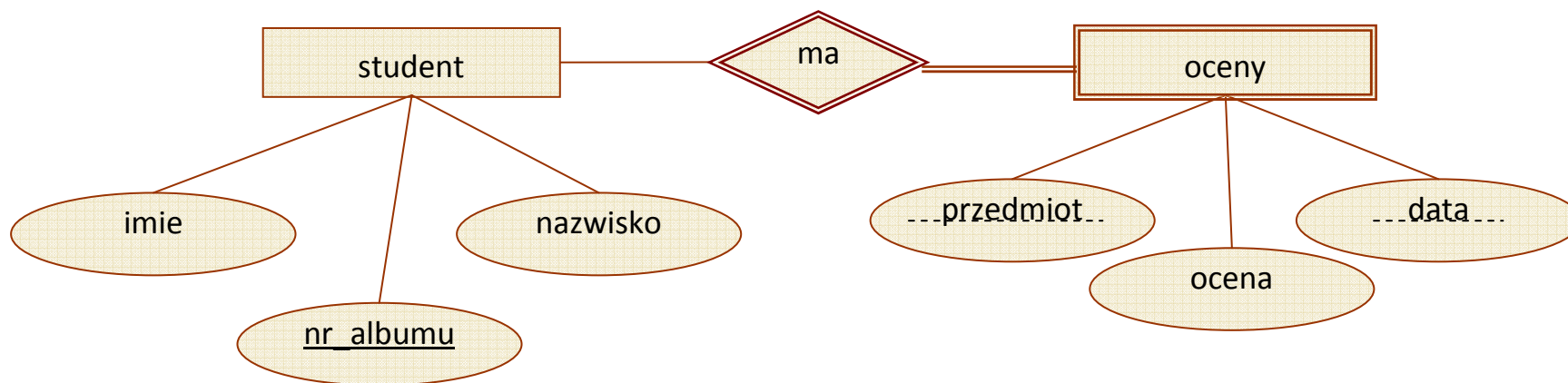
## Zbiory słabe encji:

---

- Zbiór encji, który nie posiada klucza głównego (nie ma odpowiednich atrybutów), jest nazywany **zbiorem słabym encji** (*weak entity set*) w przeciwieństwie do zbioru silnego - (*strong entity set*).
- **Informacyjność** zbioru słabych encji jest uwarunkowana istnieniem innego silnego zbioru encji, który pozostaje ze słabym w związku *jeden do wielu*. Encje tego zbioru silnego dominują również nad encjami słabego w sensie zależności występowania.
- W zbiorze słabych encji występować musi wówczas **klucz częściowy** (*partial key*), który pozwala na rozróżnienie (częściowe) jego encji względem dominującego zbioru silnego.
- **Kluczem głównym** zbioru słabego encji jest kompozycja klucza głównego dominującego zbioru silnego encji i klucza częściowego zbioru słabego.

## Zbiory słabe encji:

- W diagramie E-R **klucze częściowe** oznaczane są przez **podkreślenie** etykiety atrybutu **linią przerywaną**.
- Związki między słabą jednostką a jednostkami jej właścicieli nazywamy **związkami identyfikującymi**.
- W notacji graficznej zbiory słabe encji oraz zbiory związków identyfikujących rysujemy **podwójną linią**.

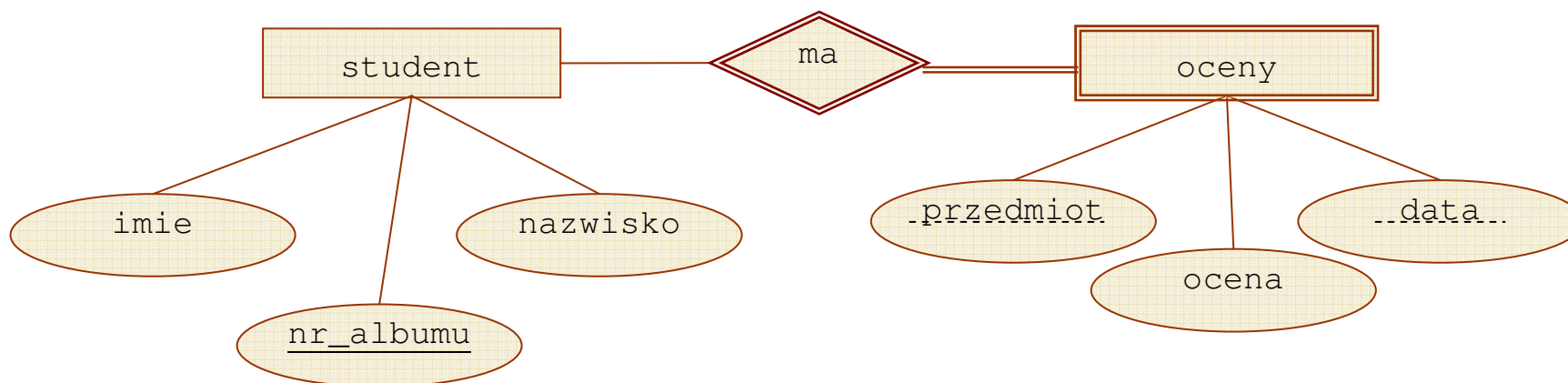


## Diagramy E-R → tabele

- Opis bazy danych za pomocą diagramu E-R może być przekształcony w opis w postaci kolekcji tabel.
- Istnienie kluczy głównych pozwala na jednoznaczne wyrażenie zbiorów encji i zbiorów związków w formie tabel.
- **Zasady ogólne**
  - » Każdemu zbiorowi encji odpowiada tabela (tabela przyjmuje nazwę zbioru encji).
  - » Każdemu zbiorowi związków (jeżeli ten posiada atrybuty) odpowiada tabela (tabela przyjmuje nazwę zbioru związków).
  - » Kolumny tabel reprezentują atrybuty opisujące odpowiednio encje lub związki.
  - » Wiersze tabel odpowiadają encjom lub związkom
- **Zbiór silny encji** jest reprezentowany bezpośrednio
- Dla **zbioru słabego encji** tabela musi zawierać dodatkowo atrybuty stanowiące klucz główny tabeli nadrzędnej.



# Diagramy E-R → tabele (przykład)

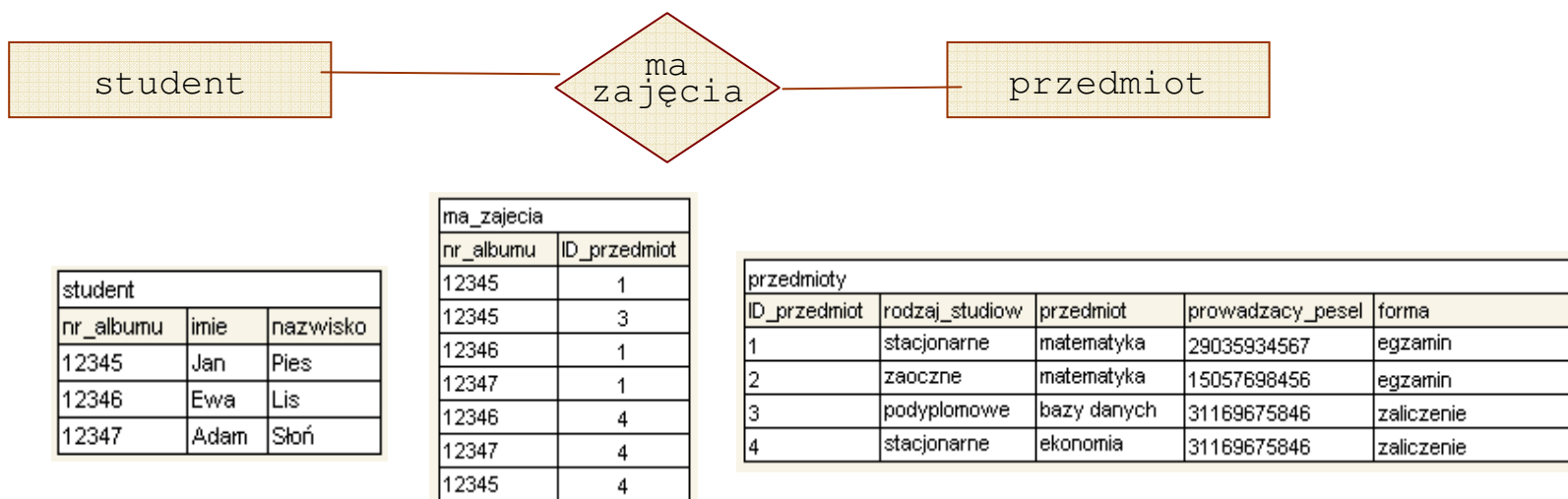


student		
nr_albumu	imie	nazwisko
12345	Jan	Pies
12346	Ewa	Lis
12347	Adam	Stoń

oceny			
nr_albumu	przedmiot	ocena	data
12345	matematyka	2,0	20.01.2006
12346	matematyka	3,0	20.01.2006
12347	wychowanie fizyczne	3,0	20.01.2006
12346	bazy danych	3,0	26.01.2006
12347	bazy danych	5,0	26.01.2006

## Diagramy E-R → tabele

- Jeżeli **zbiór związków** nie ma własnego klucza głównego, wówczas tabela musi dodatkowo zawierać odpowiednią do własności kombinację kluczy głównych swoich zbiorów encji.
  - » Np. Związek *wielu do wielu* jest reprezentowany przy pomocy tabeli o dodatkowych kolumnach odpowiadających kluczom głównym zbiorów encji biorących udział w związku.



# „Student lubi piwo...”

L.p.	Imię	Nazwisko	Marka piwa	Nazwa piwa	Gatunek	Alkohol [%]
1.	Stefan	Żeromski	Okocim	Palone	lager	7
2.	Stefan	Żeromski	Żywiec	Porter	porter bałtycki	9,5
3.	Stefan	Żeromski	Cornelius	ALE ALE	ale	5,8
4.	Eliza	Orzeszkowa	Ciechan	Miodowe	pilzner	5,7
5.	Eliza	Orzeszkowa	Redd's	Cranberry	pilzner	4,5
6.	Adam	Mickiewicz	Okocim	Porter	porter bałtycki	8,3
7.	Friedrich	Nietzsche	Paulaner	Hefe-Weissbier Naturtrüb	pszeniczne	5,5
8.	Friedrich	Nietzsche	Ciechan	Pszeniczne	pszeniczne	5,6
9.	Tadeusz	Konwicky	Żywiec	Żywiec Beer	pilzner	5,6
10.	Rafał	Wojaczek	Żywiec	Żywiec Beer	pilzner	5,6
11.	Wisława	Szyborska	Żywiec	Żywiec Beer	pilzner	5,6
12.	Johann	Goethe	Żywiec	Żywiec Beer	pilzner	5,6
13.	James	Joyce	Carlsberg	Carlsberg	lager	5
14.	Francis	Fitzgerald	Haineken	Haineken Lager Beer	lager	5
15.	Maria	Skłodowska-Curie	Żywiec	Desperados	pilzner	6
16.	Zofia	Nałkowska	Łomża	Miodowe	pilzner	5,7
17.	Fiodor	Dostojewski	Tyskie	Gronie	pilzner	5,6
18.	Fryderyk	Chopin	Tyskie	Gronie	pilzner	6,6
19.	Stefan	Batory	Tyskie	Gronie	pilzner	7,6

# „Student lubi piwo...”

STUDENT			PIWO			
L.p.	Imię	Nazwisko	Marka piwa	Nazwa piwa	Gatunek	Alkohol [%]
1.	Stefan	Żeromski	Okocim	Palone	lager	7
2.	Stefan	Żeromski	Żywiec	Porter	porter bałtycki	9,5
3.	Stefan	Żeromski	Cornelius	ALE ALE	ale	5,8
4.	Eliza	Orzeszkowa	Ciechan	Miodowe	pilzner	5,7
5.	Eliza	Orzeszkowa	Redd's	Cranberry	pilzner	4,5
6.	Adam	Mickiewicz	Okocim	Porter	porter bałtycki	8,3
7.	Friedrich	Nietzsche	Paulaner	Hefe-Weissbier Naturtrüb	pszeniczne	5,5
8.	Friedrich	Nietzsche	Ciechan	Pszeniczne	pszeniczne	5,6
9.	Tadeusz	Konwicki	Żywiec	Żywiec Beer	pilzner	5,6
10.	Rafał	Wojaczek	Żywiec	Żywiec Beer	pilzner	5,6
11.	Wisława	Szyborska	Żywiec	Żywiec Beer	pilzner	5,6
12.	Johann	Goethe	Żywiec	Żywiec Beer	pilzner	5,6
13.	James	Joyce	Carlsberg	Carlsberg	lager	5
14.	Francis	Fitzgerald	Haineken	Haineken Lager Beer	lager	5
15.	Maria	Skłodowska-Curie	Żywiec	Desperados	pilzner	6
16.	Zofia	Nałkowska	Łomża	Miodowe	pilzner	5,7
17.	Fiodor	Dostojewski	Tyskie	Gronie	pilzner	5,6
18.	Fryderyk	Chopin	Tyskie	Gronie	pilzner	6,6
19.	Stefan	Batory	Tyskie	Gronie	pilzner	7,6

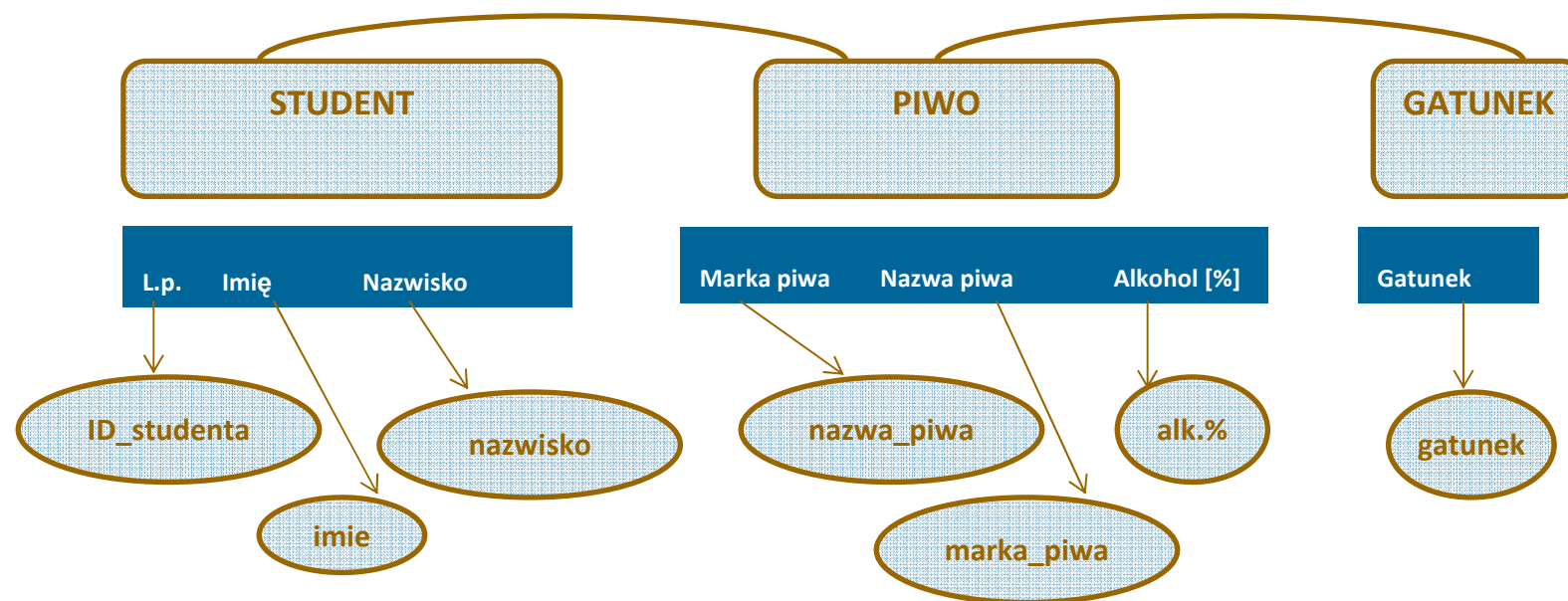
# „Student lubi piwo...”

STUDENT			PIWO		GATUNEK	PIWO
L.p.	Imię	Nazwisko	Marka piwa	Nazwa piwa	Gatunek	Alkohol [%]
1.	Stefan	Żeromski	Okocim	Palone	lager	7
2.	Stefan	Żeromski	Żywiec	Porter	porter bałtycki	9,5
3.	Stefan	Żeromski	Cornelius	ALE ALE	ale	5,8
4.	Eliza	Orzeszkowa	Ciechan	Miodowe	pilzner	5,7
5.	Eliza	Orzeszkowa	Redd's	Cranberry	pilzner	4,5
6.	Adam	Mickiewicz	Okocim	Porter	porter bałtycki	8,3
7.	Friedrich	Nietzsche	Paulaner	Hefe-Weissbier Naturtrüb	pszeniczne	5,5
8.	Friedrich	Nietzsche	Ciechan	Pszeniczne	pszeniczne	5,6
9.	Tadeusz	Konwicky	Żywiec	Żywiec Beer	pilzner	5,6
10.	Rafał	Wojaczek	Żywiec	Żywiec Beer	pilzner	5,6
11.	Wisława	Szyborska	Żywiec	Żywiec Beer	pilzner	5,6
12.	Johann	Goethe	Żywiec	Żywiec Beer	pilzner	5,6
13.	James	Joyce	Carlsberg	Carlsberg	lager	5
14.	Francis	Fitzgerald	Haineken	Haineken Lager Beer	lager	5
15.	Maria	Skłodowska-Curie	Żywiec	Desperados	pilzner	6
16.	Zofia	Nałkowska	Łomża	Miodowe	pilzner	5,7
17.	Fiodor	Dostojewski	Tyskie	Gronie	pilzner	5,6
18.	Fryderyk	Chopin	Tyskie	Gronie	pilzner	6,6
19.	Stefan	Batory	Tyskie	Gronie	pilzner	7,6

# „Student lubi piwo...”

STUDENT			PIWO		GATUNEK	PIWO
L.p.	Imię	Nazwisko	Marka piwa	Nazwa piwa	Gatunek	Alkohol [%]
1.	Stefan	Żeromski	Okocim	Palone	lager	7
2.	Stefan	Żeromski	Żywiec	Porter	porter bałtycki	9,5
3.	Stefan	Żeromski	Cornelius	ALE ALE	ale	5,8
4.	Eliza	Orzeszkowa	Ciechan	Miodowe	pilzner	5,7
5.	Eliza	Orzeszkowa	Redd's	Cranberry	pilzner	4,5
6.	Adam	Mickiewicz	Okocim	Porter	porter bałtycki	8,3
7.	Friedrich	Nietzsche	Paulaner	Hefe-Weissbier Naturtrüb	pszeniczne	5,5
8.	Friedrich	Nietzsche	Ciechan	Pszenciczne	pszeniczne	5,6
9.	Tadeusz	Konwicki	Żywiec	Żywiec Beer	pilzner	5,6
10.	Rafał	Wojaczek	Żywiec	Żywiec Beer	pilzner	5,6
11.	Wisława	Szyborska	Żywiec	Żywiec Beer	pilzner	5,6
12.	Johann	Goethe	Żywiec	Żywiec Beer	pilzner	5,6
13.	James	Joyce	Carlsberg	Carlsberg	lager	5
14.	Francis	Fitzgerald	Haineken	Haineken Lager Beer	lager	5
15.	Maria	Skłodowska-Curie	Żywiec	Desperados	pilzner	6
16.	Zofia	Nałkowska	Łomża	Miodowe	pilzner	5,7
17.	Fiodor	Dostojewski	Tyskie	Gronie	pilzner	5,6
18.	Fryderyk	Chopin	Tyskie	Gronie	pilzner	6,6
19.	Stefan	Batory	Tyskie	Gronie	pilzner	7,6

# „Student lubi piwo...”



STUDENT	
ID_studenta	Char(20)
imie	Char(20)
nazwisko	Char(20)

PIWO	
marka_piwa	Char(20)
nazwa_piwa	Char(20)
alk.	Float

GATUNEK	
gatunek	Char(20)


# „Student lubi piwo...”


STUDENT	
ID_studenta	Char(20)
imie	Char(20)
nazwisko	Char(20)

PIWO	
marka_piwa	Char(20)
nazwa_piwa	Char(20)
alk.	Float

GATUNEK	
gatunek	Char(20)



STUDENT	
 ID_studenta	Char(20) NN
imie	Char(20)
nazwisko	Char(20)

PIWO	
 ID_piwa	Char(20) NN
marka_piwa	Char(20)
nazwa_piwa	Char(20)
alk.	Float


GATUNEK	
 ID_gatunku	Char(20) NN (PK)
gatunek	Char(20)




# „Student lubi piwo...”

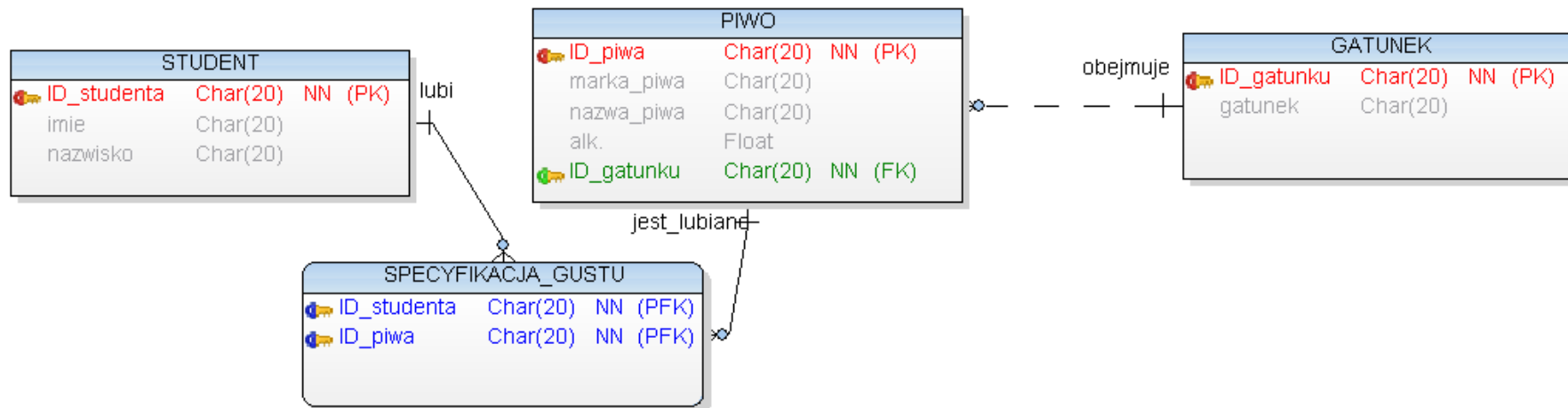
STUDENT			PIWO		GATUNEK	
L.p.	Imię	Nazwisko	Marka piwa	Nazwa piwa	Gatunek	Alkohol [%]
1.	Stefan	Żeromski	Okocim	Palone	lager	7
2.	Stefan	Żeromski	Żywiec	Porter	porter bałtycki	9,5
3.	Stefan	Żeromski	Cornelius	ALE ALE	ale	5,8
4.	Eliza	Orzeszkowa	Ciechan	Miodowe	pilzner	5,7
5.	Eliza	Orzeszkowa	Redd's	Cranberry	pilzner	4,5
6.	Adam	Mickiewicz	Okocim	Porter	porter bałtycki	8,3
7.	Friedrich	Nietzsche	Paulaner	Hefe-Weissbier Naturtrüb	pszeniczne	5,5
8.	Friedrich	Nietzsche	Ciechan	Pszeniczne	pszeniczne	5,6
9.	Tadeusz	Konwicky	Żywiec	Żywiec Beer	pilzner	5,6
10.	Rafał	Wojaczek	Żywiec	Żywiec Beer	pilzner	5,6
11.	Wisława	Szyborska	Żywiec	Żywiec Beer	pilzner	5,6
12.	Johann	Goethe	Żywiec	Żywiec Beer	pilzner	5,6
13.	James	Joyce	Carlsberg	Carlsberg	lager	5
14.	Francis	Fitzgerald	Haineken	Haineken Lager Beer	lager	5
15.	Maria	Skłodowska-Curie	Żywiec	Desperados	pilzner	6
16.	Zofia	Nałkowska	Łomża	Miodowe	pilzner	5,7
17.	Fiodor	Dostojewski	Tyskie	Gronie	pilzner	5,6
18.	Fryderyk	Chopin	Tyskie	Gronie	pilzner	6,6
19.	Stefan	Batory	Tyskie	Gronie	pilzner	7,6

# „Student lubi piwo...”

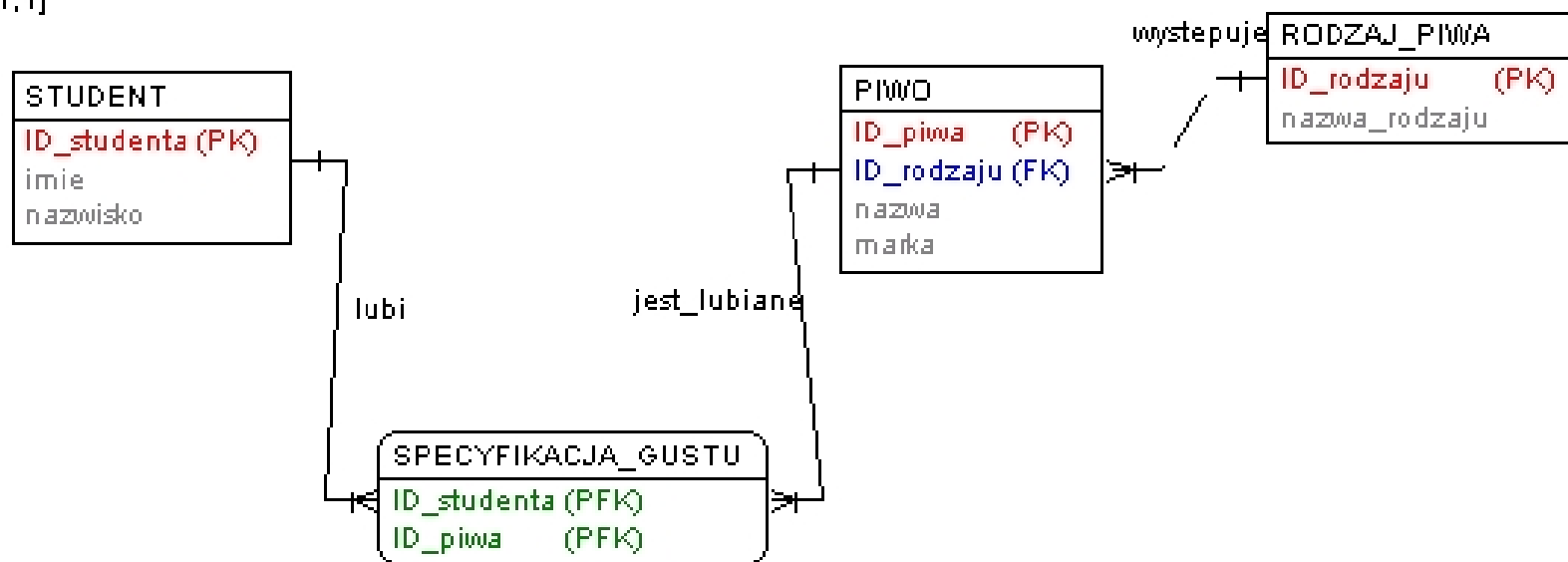
STUDENT	
 ID_studenta	Char(20) NN
imie	Char(20)
nazwisko	Char(20)

PIWO	
 ID_piwa	Char(20) NN
marka_piwa	Char(20)
nazwa_piwa	Char(20)
alk.	Float

GATUNEK	
 ID_gatunku	Char(20) NN (PK)
gatunek	Char(20)



[1.1]



# Normalizacja.

---

## Przykład wad bazy danych:

### – Rozważmy schemat relacji:

» *pracownik\_oddzialu* = *<nazwa\_oddzialu, numer\_oddzialu, adres\_oddzialu, pesel, imie, nazwisko, stanowisko, wynagrodzenie>*

### – Redundancja:

» Dane *nazwa\_oddzialu, numer\_oddzialu, adres\_oddzialu* są pamiętane dla każdego pracownika

» Marnowanie miejsca (przestrzeni potrzebnej dla przechowywania danych)

» Komplikacje (aktualizacja, błąd, usuwanie...)

### – Brak możliwości reprezentowania pewnych informacji

» Nie można reprezentować informacji o oddziałach, których powołanie dopiero jest planowane i nie zatrudniły jeszcze pracowników

» Rozwiązaniem mogło by być zastosowanie wartości NULL, ale takie rozwiązanie także stwarza pewne problemy

» Nie możemy zaprezentować w prosty sposób wszystkich pracowników z danego miasta, ponieważ miejscowość nie jest wyodrębnionym atrybutem

## Rodzaje anomalii:

---

- **Anomalia dołączania** - wadą jest to, że musimy wpisywać wszystko albo nic,
  - » np. planuje się powołanie nowego oddziału, nie możemy jednak zapisać w bazie miejscowości, adresu i nazwy zakupionego budynku, dopóki nowy oddział nie zatrudni pracowników
- **Anomalia aktualizacji** - np. oddział firmy został przeniesiony, przez co musimy aktualizować jego adres dla każdego pracownika, a przez przypadek omijamy Jana Kowalskiego
- **Anomalia usuwania** – zamknięto oddział firmy, w związku z czym usunięto z bazy wszystkie rekordy zawierające jego nazwę, tym samym usunięto wszystkie dane dotyczące pracowników

## Dekompozycja:

- Można **zdekomponować** schemat relacji – *pracownik\_oddzialu* :
  - » *oddzial* =  $\langle nazwa\_oddzialu, numer\_oddzialu, miejscowosc\_oddzialu, adres\_oddzialu \rangle$
  - » *pracownik* =  $\langle numer\_oddzialu, pesel, imie, nazwisko, stanowisko, wynagrodzenie \rangle$

- **Wszystkie atrybuty** oryginalnego schematu ( $R$ ) muszą się pojawić w dekompozycji ( $R_1, R_2$ ):

$$R = R_1 \cup R_2$$

- W przypadku gdy relacja nie posiada właściwej postaci należy dokonać dekompozycji relacji  $R$  na  $\langle R_1, R_2, \dots, R_n \rangle$
- Proces „dochodzenia” do „właściwej” postaci określa się mianem **normalizacji**

## Postacie normalne:

---

- Pierwsza (1PN)
- Druga (2PN)
- Trzecia (3PN)
- Boyce'a-Codd'a (PNBC)
- Czwarta (4PN)
- Piąta (5PN)



## Pierwsza postać normalna

---

- Relacja jest w pierwszej postaci normalnej, jeśli każda wartość atrybutu w każdej krotce tej relacji jest **wartością elementarną**, czyli nierozkładalną.
- Relacja jest w pierwszej postaci normalnej, jeśli nie ma **powtarzających się grup**.

# Przykład:

<b>pracownik oddzialu</b>								
numer_ oddzialu	nazwa_o ddzialu	miestowosc _oddzialu	adres_ oddzialu	pesel	imie	nazwisko	stanowisko	wynagrodze nie
A1	Alfaton	Kraków	Słoneczna 2/5	21108106713	Jan	Głuchomolski	Sprzedawca	2 700,00 zł
A1	Alfaton	Kraków	Słoneczna 2/5	15057698456	Wiktor	Sakowski	Kierowca	1 800,00 zł
A2	Alfabras	Kraków	Fabryczna 13/7	13015778123	Magdalena	Waligórska	Sprzedawca	2 700,00 zł
B1	Betatrix	Hel	Leśna 10	30127623435	Ignacy	Kuchta	Kierownik sekcji	3 500,00 zł
B1	Betatrix	Hel	Leśna 10	29035934567	Marzena	Sępieł	Kierowca	1 800,00 zł
B1	Betatrix	Hel	Leśna 10	31169675846	Jadwiga	Makusz	Sprzedawca	2 700,00 zł
B2	Betanor	Gdynia	Widokowa 12/26	21016922987	Stanisław	Marelkowska	Kierowca	1 800,00 zł
B3	Betarex	Gdynia	Portowa 23	17037487678	Zuzanna	Wał	Sprzedawca	2 700,00 zł

<b>oddzial</b>			
numer_ oddzialu	nazwa_ oddzialu	miestowosc_ oddzialu	adres_ oddzialu
A1	Alfaton	Kraków	Słoneczna 2/5
A2	Alfabras	Kraków	Fabryczna 13/7
B1	Betatrix	Hel	Leśna 10
B2	Betanor	Gdynia	Widokowa 12/26
B3	Betarex	Gdynia	Portowa 23

<b>pracownik</b>					
pesel	imie	nazwisko	stanowisko	wynagrodzenie	numer_ oddzialu
21108106713	Jan	Głuchomolski	Sprzedawca	2 700,00 zł	A1
15057698456	Wiktor	Sakowski	Kierowca	1 800,00 zł	A1
13015778123	Magdalena	Waligórska	Sprzedawca	2 700,00 zł	A2
30127623435	Ignacy	Kuchta	Kierownik sekcji	3 500,00 zł	B1
29035934567	Marzena	Sępieł	Kierowca	1 800,00 zł	B1
31169675846	Jadwiga	Makusz	Sprzedawca	2 700,00 zł	B1
21016922987	Stanisław	Marelkowska	Kierowca	1 800,00 zł	B2
17037487678	Zuzanna	Wał	Sprzedawca	2 700,00 zł	B3

## Zależności funkcyjne i wielowartościowe:

- **Zależność funkcyjna** oznacza, że znając wartość jednego atrybutu, zawsze możemy określić wartość innego. Symbolem stosowanym w teorii relacji jest strzałka umieszczona pomiędzy dwoma atrybutami, na przykład:

$$X \longrightarrow Y \quad (X \text{ określa } Y)$$

**przykład:** gdy znamy numer PESEL naszego pracownika, możemy określić jego nazwisko

- **Zależność wielowartościowa** oznacza, że znając wartość jednego atrybutu, możemy zawsze określić wartości zbioru innego atrybutu. W teorii relacji używa się symbolu zależności wielowartościowej w postaci podwójnej strzałki, na przykład:

$$X \longrightarrow\longrightarrow Y \quad (X \text{ określa wiele } Y)$$

**przykład:** znając numer oddziału możemy określić nazwiska wszystkich zatrudnionych pracowników

## Teoria postaci normalnych relacji:

Formalnie **zależność danych** można zdefiniować następująco. Schemat relacji oznaczamy przez  $R\langle A_1 \dots A_N \rangle$ , gdzie  $A_1 \dots A_N$  są atrybutami relacji. Niech  $X$  i  $Y$  będą podzbiorami zbioru atrybutów.

$$X \subset \{A_1 \dots A_N\}, Y \subset \{A_1 \dots A_N\}.$$

Zależność danych zapisujemy w postaci:

$$X \rightarrow Y$$

i mówimy, że podzbiór atrybutów  $Y$  **zależy funkcyjnie** od podzbioru atrybutów  $X$ , jeżeli nie jest możliwe, by relacja  $R$  zawierała dwie krotki mające składowe zgodne (tzn. identyczne dla wszystkich atrybutów ze zbioru  $X$ ) i jednocześnie co najmniej jedną niezgodną składową dla atrybutów ze zbioru  $Y$ .

## Klucz:

---

- **Kluczem relacji** nazywamy taki zbiór atrybutów tej relacji, których kombinacje wartości jednoznacznie identyfikują każdą krotkę tej relacji a żaden podzbiór tego zbioru nie posiada tej własności. W kluczu nie może zawierać się wartość NULL.
- Klucz jest **kluczem prostym**, jeżeli powyżej opisany zbiór jest jednoelementowy - w przeciwnym razie mówimy o **kluczu złożonym**.
- W ogólności, w relacji można wyróżnić wiele kluczy, które nazywamy **kluczami potencjalnymi**. Wybrany klucz spośród kluczy potencjalnych nazywamy **kluczem głównym**.

## Druga i trzecia postać normalna:

---

- Relacja jest w **drugiej** postaci normalnej, jeśli jest w 1PN oraz każdy atrybut tej relacji **nie wchodzący** w skład żadnego klucza potencjalnego jest w **pełni funkcyjnie zależny** od wszystkich kluczy potencjalnych tej relacji.
- Relacja jest w **2PN** jeżeli każdy atrybut nie wchodzący w skład klucza **zależy od klucza a nie od jego części**.
- Relacja będąca w pierwszej postaci normalnej, jest równocześnie w **drugiej** postaci normalnej, jeśli wszystkie jej klucze potencjalne są **kluczami prostymi**.
- Dana relacja jest w **trzeciej** postaci normalnej, jeśli jest ona w drugiej postaci normalnej i każdy jej atrybut nie wchodzący w skład żadnego klucza potencjalnego **nie jest przechodnio funkcyjnie zależny** od żadnego klucza potencjalnego tej relacji.

## Druga i trzecia postać normalna:

---

Inaczej mówiąc, **wszystkie niekluczowe kolumny są określane**

**kluczem, całym kluczem i tylko kluczem,  
*„...tak nam dopomóż Codd”***

# Przykład:

Protokoły zaliczeń i egzaminów

Rodzaj studiów	Rok studiów	Rok akademicki	Przedmiot	Prowadzący	Forma	Termin	Student		Ocena	Data
				Imię i nazwisko			nr albumu	Nazwisko i imię		
stacjonarne	I	2005/2006	matematyka	prof. Jan Kot	egzamin	I	12345	Pies Jan	2,0	20.01.2006
stacjonarne	I	2005/2006	matematyka	prof. Jan Kot	egzamin	I	12346	Lis Ewa	3,0	20.01.2006
zaoczne	I	2005/2006	matematyka	prof. Jan Kot	egzamin	II	12347	Stoń Adam	3,0	20.01.2006
stacjonarne	I	2005/2006	wychowanie fizyczne	mgr Janina Mysz	zaliczenie	I	12345	Pies Jan	5,0	26.01.2006
stacjonarne	I	2005/2006	wychowanie fizyczne	mgr Janina Mysz	zaliczenie	I	12346	Lis Ewa	3,0	26.01.2006
zaoczne	I	2005/2006	wychowanie fizyczne	mgr Janina Mysz	zaliczenie	II	12347	Stoń Adam	3,0	26.01.2006

Wartości atrybutów nie są elementarne



# Przykład (c.d.):

**Klucz potencjalny:** <rodzaj\_studiow, rok\_studiow, rok\_akademicki, przedmiot, forma, termin, student\_nr\_albumu>

## Zależności funkcyjne

Protokoły zaliczeń i egzaminów

rodzaj_studiow	rok_studiow	rok_akademicki	przedmiot	przewadzacy_tytul	przewadzacy_imie	przewadzacy_nazwisko	forma	termin	student_nr_albumu	student_imie	student_nazwi	ocena	data
stacjonarne	I	2005/2006	matematyka	prof.	Jan	Kot	egzamin	I	12345	Jan	Pies	2,0	20.01.2006
stacjonarne	I	2005/2006	matematyka	prof.	Jan	Kot	egzamin	I	12346	Ewa	Lis	3,0	20.01.2006
zaoczne	I	2005/2006	matematyka	prof.	Jan	Kot	egzamin	II	12347	Adam	Słoń	3,0	20.01.2006
stacjonarne	I	2005/2006	wychowanie fizyczne	mgr	Janina	Mysz	zaliczenie	I	12345	Jan	Pies	5,0	26.01.2006
stacjonarne	I	2005/2006	wychowanie fizyczne	mgr	Janina	Mysz	zaliczenie	I	12346	Ewa	Lis	3,0	26.01.2006
zaoczne	I	2005/2006	wychowanie fizyczne	mgr	Janina	Mysz	zaliczenie	II	12347	Adam	Słoń	3,0	26.01.2006

Powtarzające się grupy

prowadzacy			
pesel	tytul	imie	nazwisko
29035934567	prof.	Jan	Kot
31169675846	mgr	Janina	Mysz
15057698456	dr hab. inż.	Wojciech	Kuna

przedmioty						
id_przedmiot	rodzaj_studiow	rok_studiow	rok_akademicki	przedmiot	prowadzacy_pesel	forma
1	stacjonarne	I	2005/2006	matematyka	29035934567	egzamin
2	zaoczne	II	2005/2006	matematyka	15057698456	egzamin
3	stacjonarne	I	2005/2006	wychowanie fizyczne	31169675846	zaliczenie
4	zaoczne	II	2005/2006	wychowanie fizyczne	31169675846	zaliczenie

protokoly				
id_przedmiot	termin	nr_albumu	ocena	data
1	I	12345	2,0	20.01.2006
1	I	12346	3,0	20.01.2006
2	II	12347	3,0	20.01.2006
3	I	12345	5,0	26.01.2006
3	I	12346	3,0	26.01.2006
4	I	12347	3,0	26.01.2006

**Klucz główny relacji protokoly:**  
 $\langle id\_przedmiot, termin, nr\_albumu \rangle$

student		
nr_albumu	imie	nazwisko
12345	Jan	Pies
12346	Ewa	Lis
12347	Adam	Stoń

## Ta relacja nie jest w trzeciej postaci normalnej

Pracownik	PESEL	KodPocztowy	Miejscowość	Województwo
Jan Kowalski	12345678911	32-082	Bolechowice	małopolskie
Adam Kot	98977796666	30-150	Kraków	małopolskie
Ewa Lis	76281976372	32-082	Bolechowice	małopolskie



## Zależność funkcjonalna przechodnia

---

- Niech  $X$ ,  $Y$  i  $Z$  będą trzema rozłącznymi podzbiorami atrybutów danej relacji
- $Z$  jest przechodnio funkcjonalnie zależny od  $X$ , jeśli  $Z$  jest funkcjonalnie zależny od  $Y$  i  $Y$  jest funkcjonalnie zależny od  $X$  natomiast  $X$  nie jest zależny od  $Y$  i  $Y$  nie jest zależny od  $Z$

## Forma normalna Boyce-Codd'a

---

- Jest uzupełnieniem trzeciej postaci normalnej i jest niezbędna w przypadku gdy atrybuty będące kandydatami na klucze są:
  - » wielokrotne,
  - » złożone,
  - » nakładające się na siebie

- Relacja jest w postaci **Boyce-Codd'a (BCPN)** jeżeli dla każdej nietrywialnej zależności między podzbiarami relacji zbiór będący wyznacznikiem jest zbiorem identyfikującym tej relacji
- Zależność  $X \rightarrow Y$  jest trywialna jeżeli  $Y$  jest podzbiorem  $X$
- Definicja BCPN zastępuje definicje, pierwszej, drugiej i trzeciej formy normalnej dodatkowo je poszerzając

## Forma normalna Boyce-Codd'a

<b>IdStudenta</b>	<b>Seminarium</b>	<b>Opiekun</b>
1	metalurgia	Kowalski
1	inżynieria wiedzy	Kozłowski
2	inżynieria wiedzy	Janowski
3	metalurgia	Kowalski
4	informatyka	Zajac

- opiekun może mieć tylko jedno seminarium, więc kluczem w relacji może być: podzbiór ***IdStudenta + Seminarium*** lub ***IdStudenta + Opiekun***
- Przy tej drugiej opcji nie spełniona jest III postać normalna, ponieważ klucz ***IdStudenta + Opiekun*** nie determinują seminarium, która zależy tylko i wyłącznie od opiekuna. Z tego powodu jako kandydata na klucz główny powinno się wziąć pod uwagę ***IdStudenta + Seminarium***
- zależność ***Opiekun*** → ***Seminarium*** jest funkcjonalna i nietrywialna a ***Opiekun*** nie jest zbiorem identyfikującym

## Forma normalna Boyce-Codd'a

---

- Założmy, że chcemy przydzielić nazwiska prowadzących do odpowiednich seminariów przed rozpoczęciem zapisów studentów na konkretny przedmiot.
- Nie możemy tego zrobić, ponieważ nazwisko studenta jest częścią klucza głównego i z tego powodu, przy wstawianiu nowego wiersza tabeli trzeba wstawić nazwisko studenta (anomalia dołączania).
- Również może zaistnieć anomalia w przypadku, gdy wystąpi potrzeba zmiany nazwiska prowadzącego przedmiot. Wtedy będzie trzeba wstawić nowe nazwisko do wielu wierszy jednocześnie.
- Aby zaradzić tym problemom należy sprowadzić tabelę do postaci Boyce-Codda. W tym celu zostaną utworzone 3 tabele:



## Forma normalna Boyce-Codd'a

IdStudenta	Opiekun
1	Kowalski
1	Kozłowski
2	Janowski
3	Kowalski
4	Zajęc

Seminarium	Opiekun
metalurgia	Kowalski
inżynieria wiedzy	Kozłowski
inżynieria wiedzy	Janowski
informatyka	Zajęc

IdStudenta	Seminarium
1	metalurgia
1	inżynieria wiedzy
2	inżynieria wiedzy
3	metalurgia
4	informatyka

## Forma normalna Boyce-Codd'a

<b>IdPracownika</b>	<b>Zawód</b>	Wykształcenie	Stawka
1	ślusarz	podstawowe	5,20
1	tokarz	zawodowe	5,30
2	ślusarz	zawodowe	5,50
3	tokarz	zawodowe	5,30
4	ślusarz	podstawowe	5,20

kluczem w relacji jest podzbiór ***IdPracownika, Zawód*** lub ***IdPracownika, Wykształcenie***

zależność ***Zawód, Wykształcenie***  $\rightarrow$  ***Stawka*** jest funkcjonalna i nietrywialna

a ***Zawód, Wykształcenie*** nie jest zbiorem identyfikującym

## Podsumowanie:

---

- Normalizacja ma na celu takie przekształcenie relacji, by uniknąć **redundancji i anomalii**.
- Przekształcenie relacji do kolejnych postaci normalnych wiąże się najczęściej ze **zmniejszeniem ilości pamięci** potrzebnej do **przechowania informacji**.
- Unikanie powtórzeń pozwala na **łatwiejszą i szybszą aktualizację** danych.
- Doprowadzenie bazy do wysokiej postaci normalizacji może **spowolnić odczyt** w dużych bazach ze względu na skomplikowany schemat danych.
- W większości przypadków po znormalizowaniu bazy danych przychodzi kolej na rozważenie możliwości wykonania **odwrotnej operacji (denormalizacji)**, polegającej na połączeniu niektórych znormalizowanych tabel, a to z myślą o przyspieszeniu dostępu do pewnych danych.

## Czwarta forma normalna

---

- Relacja jest w czwartej formie normalnej (IV PN) wtedy i tylko wtedy, gdy jest w trzeciej postaci normalnej i nie zawiera nietrywialnej wielowartościowej zależności atrybutów

- Podzbiór atrybutów  $Y$  jest wielowartościowo funkcjonalnie zależny od podzbioru  $X$  w schemacie  $R$ , jeżeli dla dowolnej relacji  $r$  w schemacie  $R$  i dla dowolnej pary krotek  $t_1$  i  $t_2$  z relacji  $r$  istnieje taka para krotek że:

$$s_1[X]=s_2[X]=t_1[X]=t_2[X] \text{ i}$$

$$s_1[Y]=t_1[Y] \text{ i } s_1[R-X-Y]=t_2[R-X-Y] \text{ i}$$

$$s_2[Y]=t_2[Y] \text{ i } s_2[R-X-Y]=t_1[R-X-Y]$$

# Zależność wielowartościowa

X

Y

R-X-Y

krotka	Nazwisko	Imię dziecka	Znajomość języków
$t_1$	Kot	Ania	niemiecki
$t_2$	Kot	Jaś	angielski
$s_1$	Kot	Ania	angielski
$s_2$	Kot	Jaś	niemiecki
	Słoń	Ola	niemiecki
	Słoń	Ola	angielski

$$t_1[X]=t_2[X]=s_1[X]=s_2[X]=(\text{Kot})$$

$$s_1[Y]=t_1[Y]=(\text{Ania}) \text{ i}$$

$$s_1[R-X-Y]=t_2[R-X-Y]=(\text{angielski}) \text{ i}$$

$$s_2[Y]=t_2[Y]=(\text{Jaś}) \text{ i}$$

$$s_2[R-X-Y]=t_1[R-X-Y]=(\text{niemiecki})$$

## Piąta postać normalna

---

- Relacja jest w piątej formie normalnej (V PN) wtedy i tylko wtedy, gdy jest w czwartej postaci normalnej i nie istnieje jej rozkład odwracalny na zbiór mniejszych tabel.

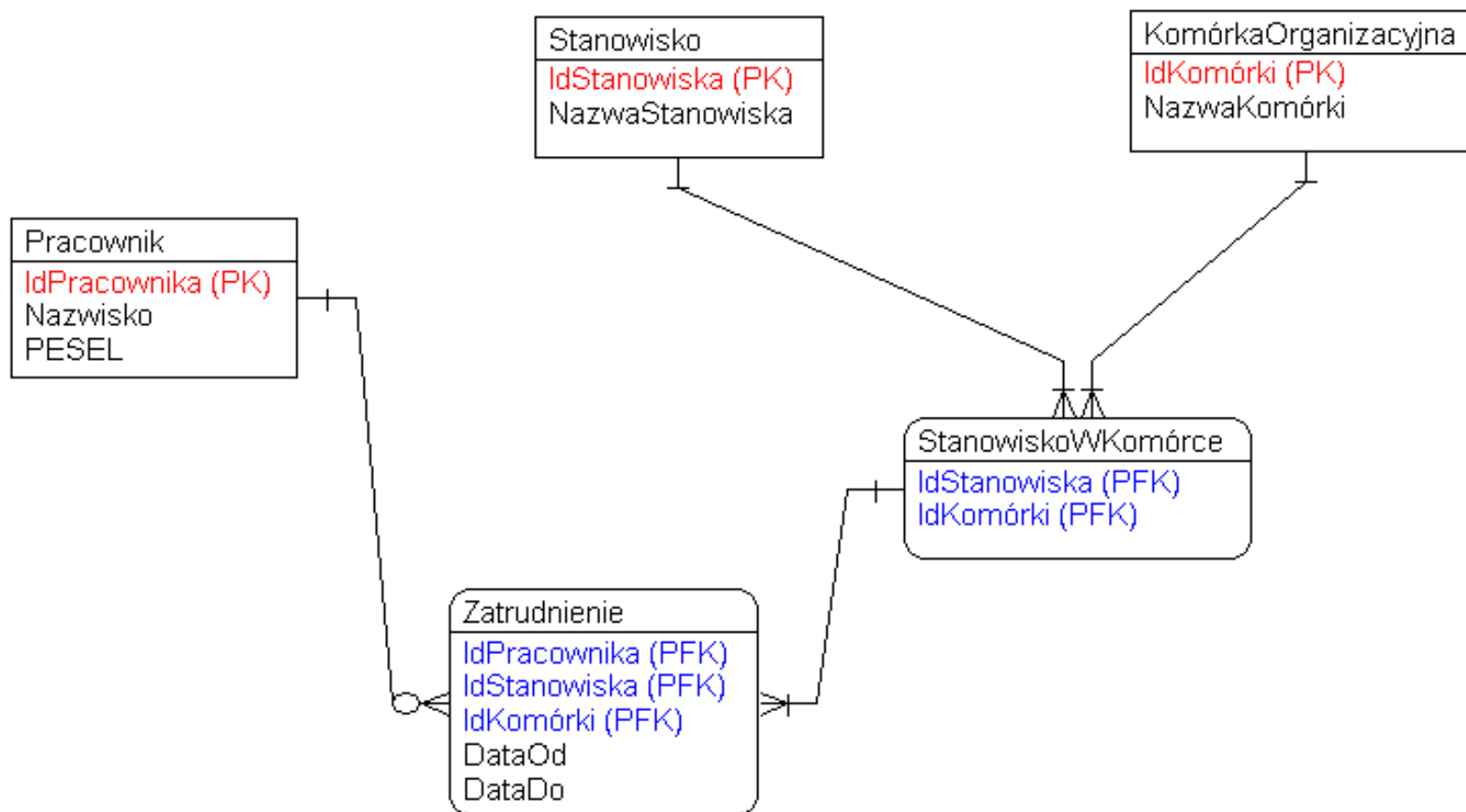


## Klucz sztuczny

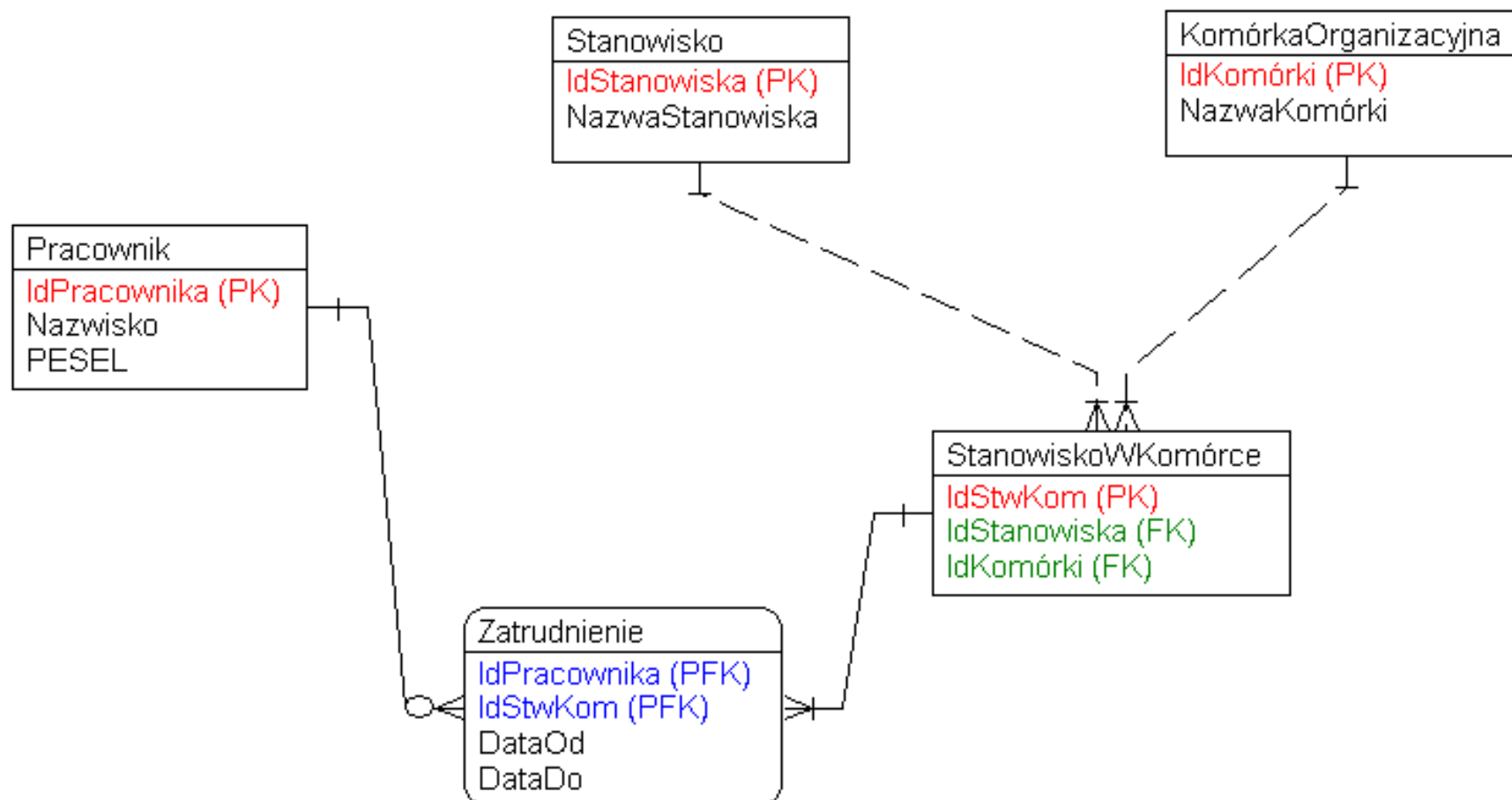
---

- Klucz stworzony wyłącznie dla potrzeb więzi w celu zastąpienia złożonego klucza głównego

# Klucz złożony...

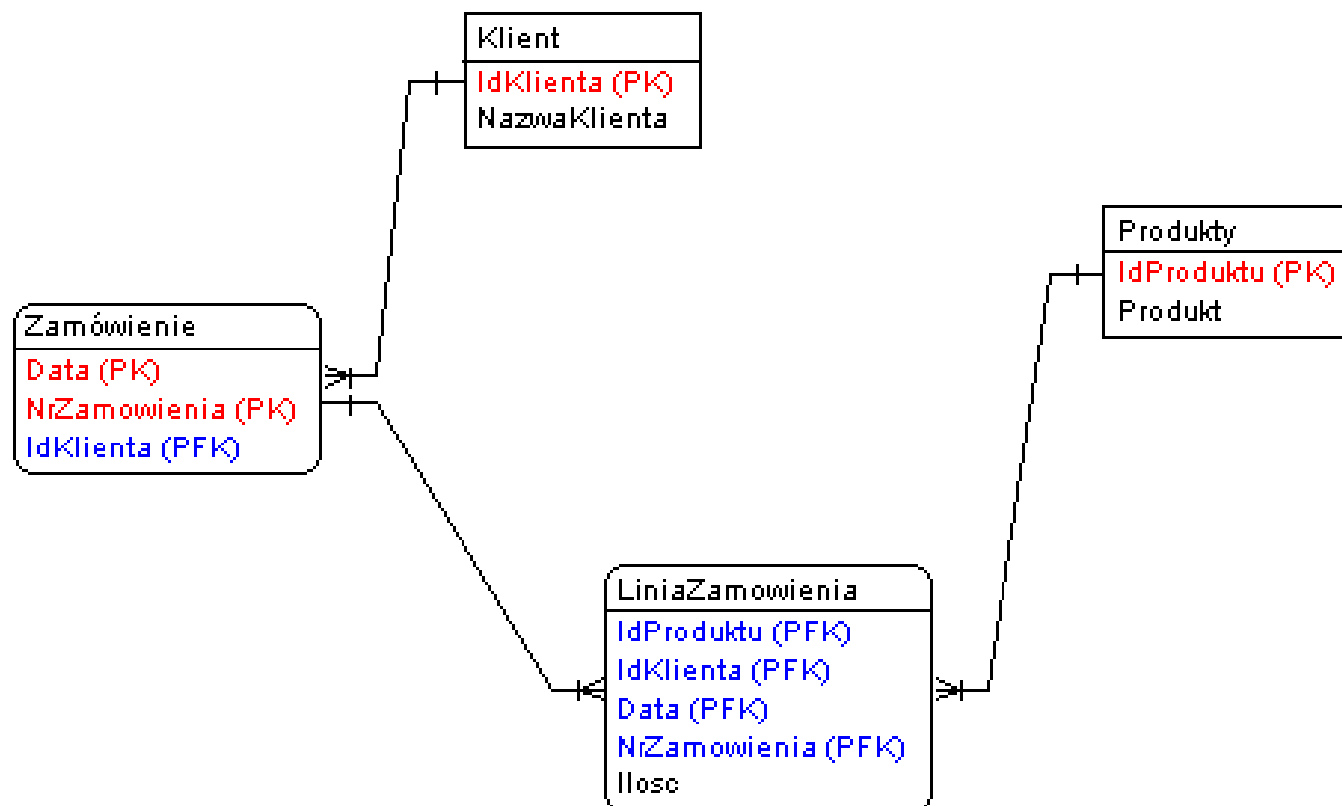


## ...zastąpiony kluczem sztucznym



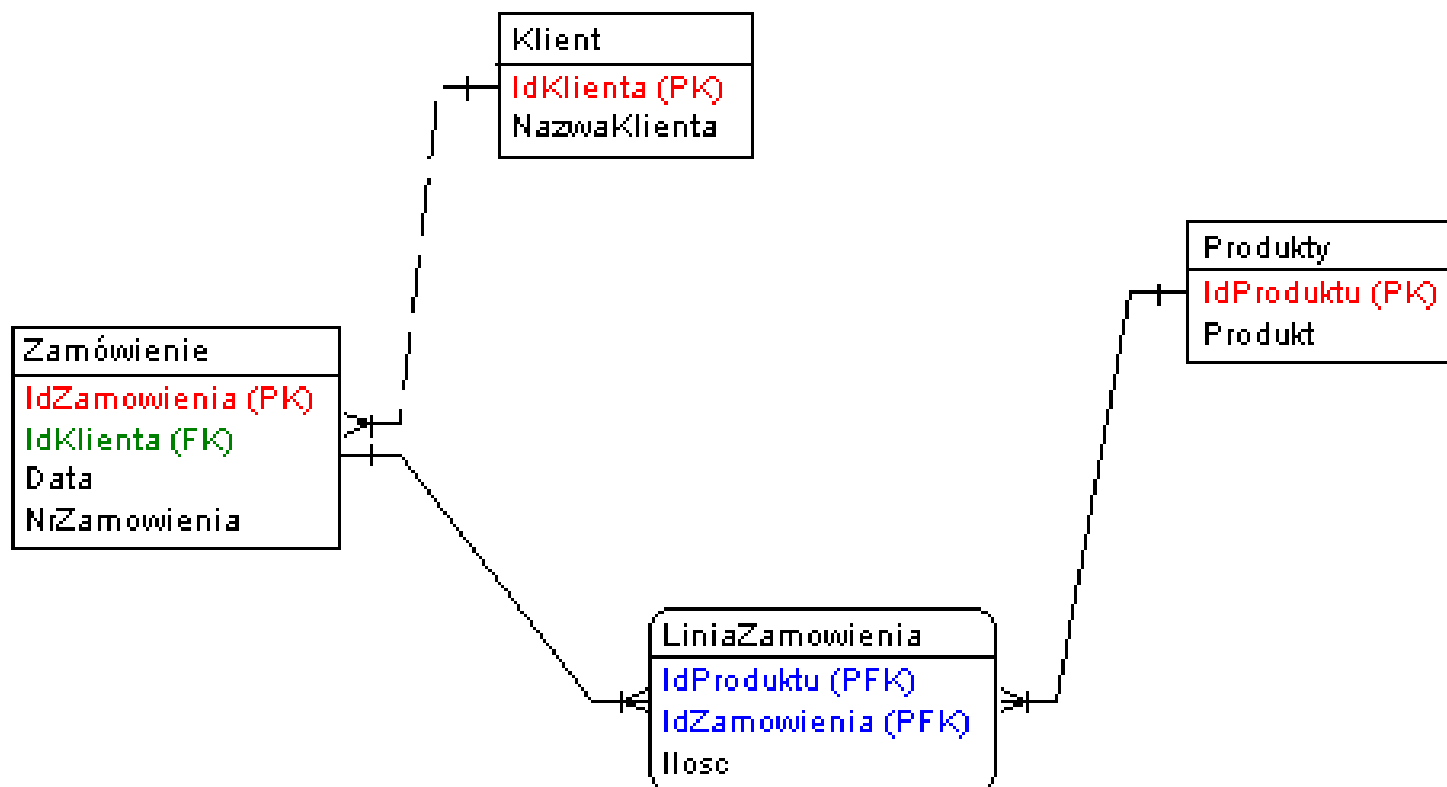
# Klucz złożony...

[1.1]



## ...zastąpiony kluczem sztucznym

[1.1]



## Klucz sztuczny

---

- Klucz sztuczny może być wykorzystany do kodowania atrybutów tekstowych (w niektórych przypadkach także liczbowych) o powtarzających się wartościach, dla których można utworzyć listę
- Użycie klucza sztucznego wymaga stworzenia dodatkowej tabeli (słownika) pozwalającego na „rozkodowanie” klucza

Nr faktury	Za okres		Identyfikator nabywcy
	od	do	
21113332437	1.11.2007	30.11.2007	1
21218909871	1.12.2007	31.12.2007	1
21113332442	1.11.2007	30.11.2007	2
21218909900	1.12.2007	31.12.2007	2

Nr faktury	Liczba jednostek (czas połączenia)	Id usługi
21113332437	1	1
21113332442	1	1
21113332437	25,3	2
21113332442	15	2
21113332437	30	3
21113332442	28	3
21113332437	15	4
21113332442	12	4
21113332437	20	5
21113332442	15	5

Id rodzaju usługi	Usługa	Stawka VAT
1	Abonament	22
2	Połączenia krajowe	22
3	SMS	22

Id usługi	Id rodzaju usługi	Strefa czasowa	Kierunek	Cena
1	1			70,00
2	2	Szczyt	Era	2,00
3	2	Szczyt	Plus GSM	1,10
4	2	Poza szczytem	Plus GSM	1,00
5	3			0,50

Strefa czasowa
Szczyt
Poza szczytem

Kierunek
Era
Plus GSM

Nr faktury	Za okres		Identyfikator nabywcy
	od	do	
21113332437	1.11.2007	30.11.2007	1
21218909871	1.12.2007	31.12.2007	1
21113332442	1.11.2007	30.11.2007	2
21218909900	1.12.2007	31.12.2007	2

wartość netto \* stawka

wartość netto + kwota VAT

liczba jednostek \* cena

Nr faktury	Usługa	Strefa czasowa	Kierunek	Liczba jednostek (czas połączenia)	Wartość netto	Stawka VAT	Kwota VAT	Wartość brutto
21113332437	Abonament			1	70,00	22,00	15,40	85,40
21113332437	Połączenia krajowe	Szczyt	Era	25,3	50,60	22,00	11,13	61,73
21113332437	Połączenia krajowe	Szczyt	Plus GSM	30	33,00	22,00	7,26	40,26
21113332437	Połączenia krajowe	Poza szczytem	Plus GSM	15	15,00	22,00	3,30	18,30
21113332437	SMS			20	10,00	22,00	2,20	12,20
21113332437	Razem				186,10			186,10
21113332442	Abonament			1	70,00	22,00	15,40	85,40
21113332442	Połączenia krajowe	Szczyt	Era	15	30,00	22,00	6,60	36,60
21113332442	Połączenia krajowe	Szczyt	Plus GSM	28	30,80	22,00	6,78	37,58
21113332442	Połączenia krajowe	Poza szczytem	Plus GSM	12	12,00	22,00	2,64	14,64
21113332442	SMS			15	7,50	22,00	1,65	9,15
21113332442	Razem				156,30		33,07	183,37

tego nie trzeba pamiętać



Nr faktury	Usługa	Strefa czasowa	Kierunek	Liczba jednostek	Cena	Stawka VAT
21113332437	Abonament			1	70,00	22
21113332437	Połączenia krajowe	Szczyt	Era	25,3	2,00	22
21113332437	Połączenia krajowe	Szczyt	Plus GSM	30	1,10	22
21113332437	Połączenia krajowe	Poza szczytem	Plus GSM	15	1,00	22
21113332437	SMS			20	0,50	22
21113332442	Abonament			1	70,00	22
21113332442	Połączenia krajowe	Szczyt	Era	15	2,00	22
21113332442	Połączenia krajowe	Szczyt	Plus GSM	28	1,10	22
21113332442	Połączenia krajowe	Poza szczytem	Plus GSM	12	1,00	22
21113332442	SMS			15	0,50	22

to też jest powtarzająca się grupa danych, bo tabela może wyglądać tak:

Nr faktury	Usługa	Strefa czasowa	Kierunek	Liczba jednostek (czas połączenia)	Cena	Stawka VAT
21113332437	Abonament			1	70,00	22
21113332442				1		
21113332437	Połączenia krajowe	Szczyt	Era	25,3	2,00	22
21113332442				15		
21113332437	Połączenia krajowe	Szczyt	Plus GSM	30	1,10	22
21113332442				28		
21113332437	Połączenia krajowe	Poza szczytem	Plus GSM	15	1,00	22
21113332442				12		
21113332437	SMS			20	0,50	22
21113332442				15		



Nr faktury	Liczba jednostek (czas połączenia)	Id usługi	Usługa	Strefa czasowa	Kierunek	Cena	Stawka VAT
21113332437	1	1	Abonament			70,00	22
21113332442	1						
21113332437	25,3	2	Połączenia krajowe	Szczyt	Era	2,00	22
21113332442	15						
21113332437	30	3	Połączenia krajowe	Szczyt	Plus GSM	1,10	22
21113332442	28						
21113332437	15	4	Połączenia krajowe	Poza szczytem	Plus GSM	1,00	22
21113332442	12						
21113332437	20	5	SMS			0,50	22
21113332442	15						

brak dobrego kandydata na klucz grupy i dlatego wprowadzamy klucz sztuczny i wykonujemy dekompozycję

Nr faktury	Liczba jednostek (czas połączenia)	Id usługi
21113332437	1	1
21113332442	1	1
21113332437	25,3	2
21113332442	15	2
21113332437	30	3
21113332442	28	3
21113332437	15	4
21113332442	12	4
21113332437	20	5
21113332442	15	5

Id usługi	Usługa	Strefa czasowa	Kierunek	Cena	Stawka VAT
1	Abonament			70,00	22
2	Połączenia krajowe	Szczyt	Era	2,00	22
3	Połączenia krajowe	Szczyt	Plus GSM	1,10	22
4	Połączenia krajowe	Poza szczytem	Plus GSM	1,00	22
5	SMS			0,50	22

Id usługi	Usługa	Strefa czasowa	Kierunek	Cena	Stawka VAT
1	Abonament			70,00	22
2	Połączenia krajowe	Szczyt	Era	2,00	22
3	Połączenia krajowe	Szczyt	Plus GSM	1,10	22
4	Połączenia krajowe	Poza szczytem	Plus GSM	1,00	22
5	SMS			0,50	22

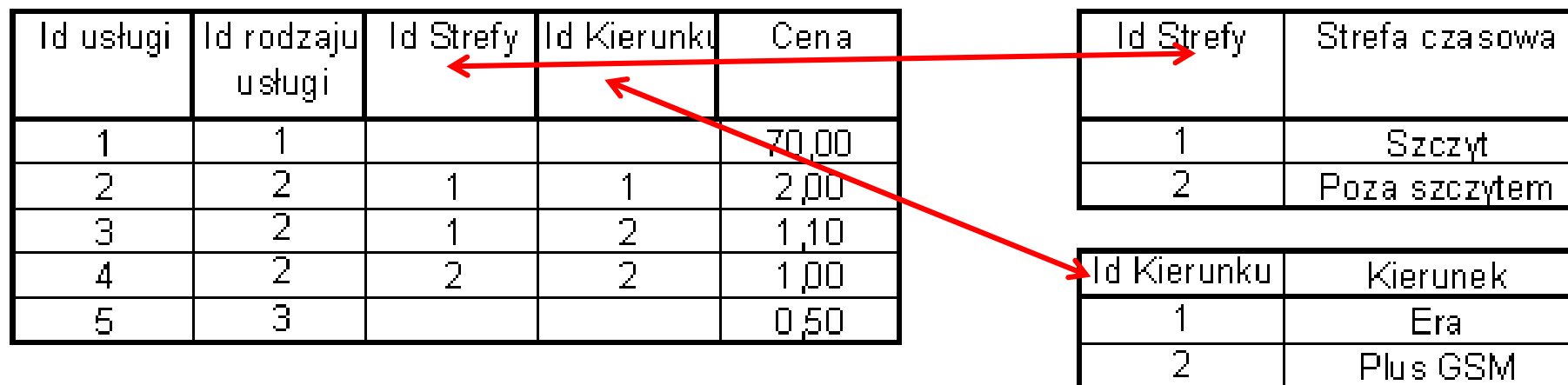
to też jest powtarzająca się grupa danych, w której kluczem jest rodzaj usługi i dlatego trzeba tablicę zdekomponować:

Id usługi	Id rodzaju usługi	Strefa czasowa	Kierunek	Cena
1	1			70,00
2	2	Szczyt	Era	2,00
3	2	Szczyt	Plus GSM	1,10
4	2	Poza szczytem	Plus GSM	1,00
5	3			0,50

Id rodzaju usługi	Usługa	Stawka VAT
1	Abonament	22
2	Połączenia krajowe	22
3	SMS	22

Id usługi	Id rodzaju usługi	Strefa czasowa	Kierunek	Cena
1	1			70,00
2	2	Szczyt	Era	2,00
3	2	Szczyt	Plus GSM	1,10
4	2	Poza szczytem	Plus GSM	1,00
5	3			0,50

to nie są powtarzające się grupy danych, ale powtarzające się dane, które warto przechowywać w słownikach:



Id usługi	Id rodzaju usługi	Strefa czasowa	Kierunek	Cena
1	1			70,00
2	2	Szczyt	Era	2,00
3	2	Szczyt	Plus GSM	1,10
4	2	Poza szczytem	Plus GSM	1,00
5	3			0,50

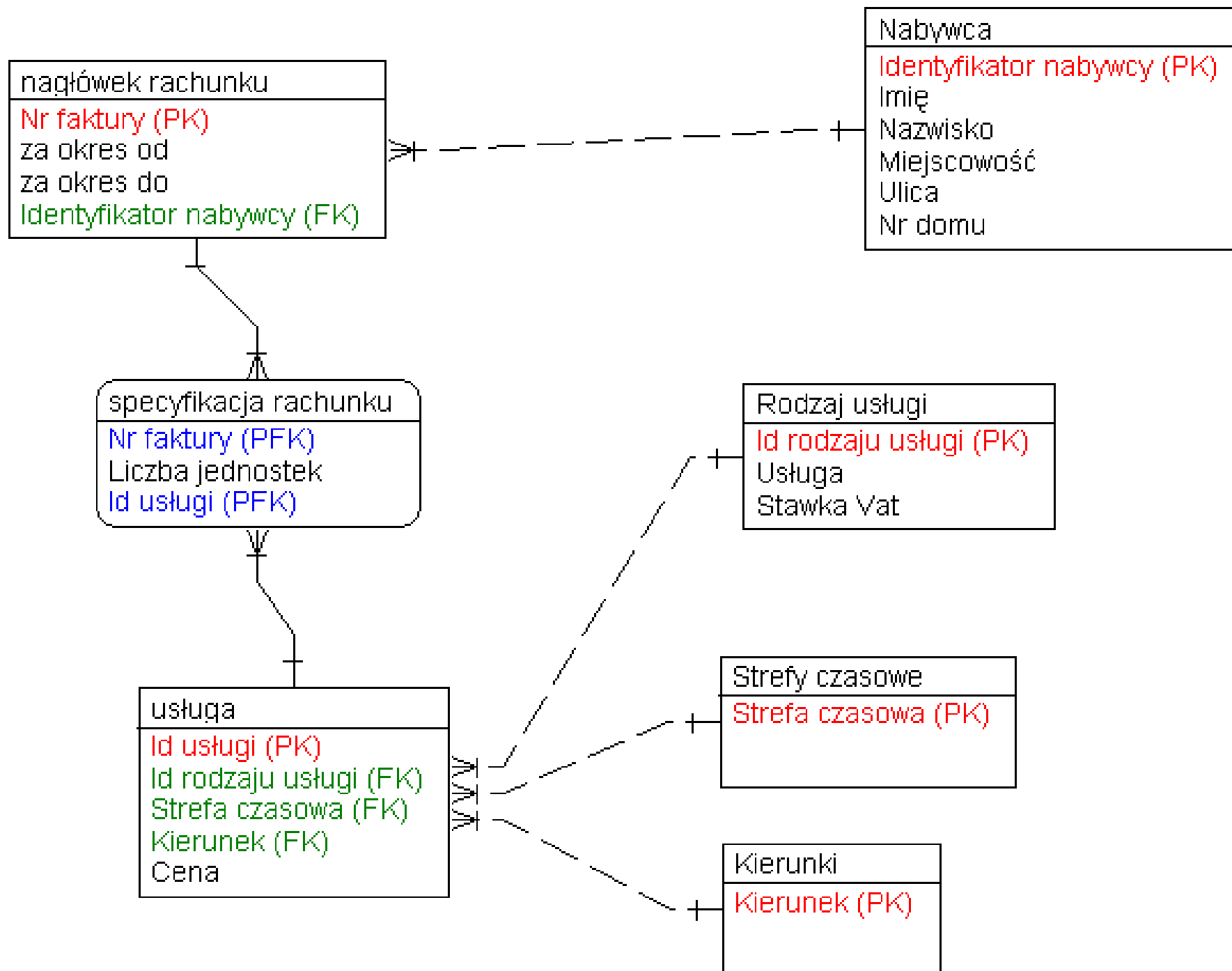
  

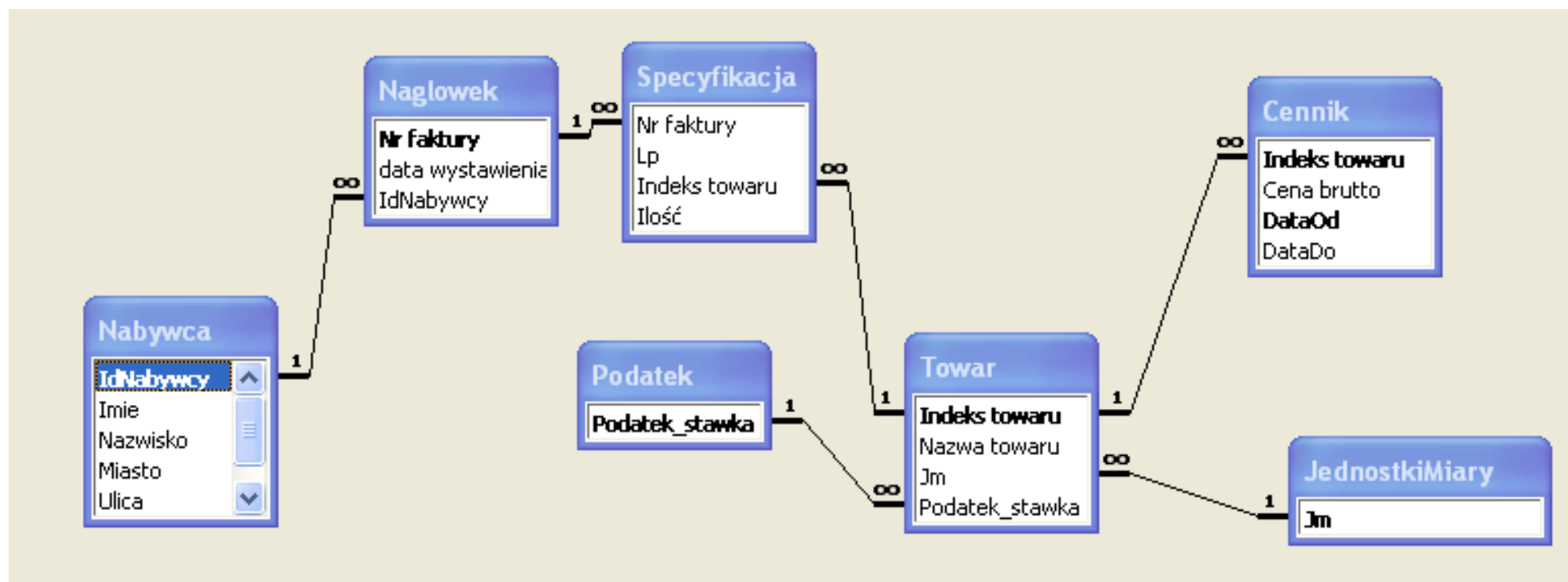
Strefa czasowa
Szczyt
Poza szczytem

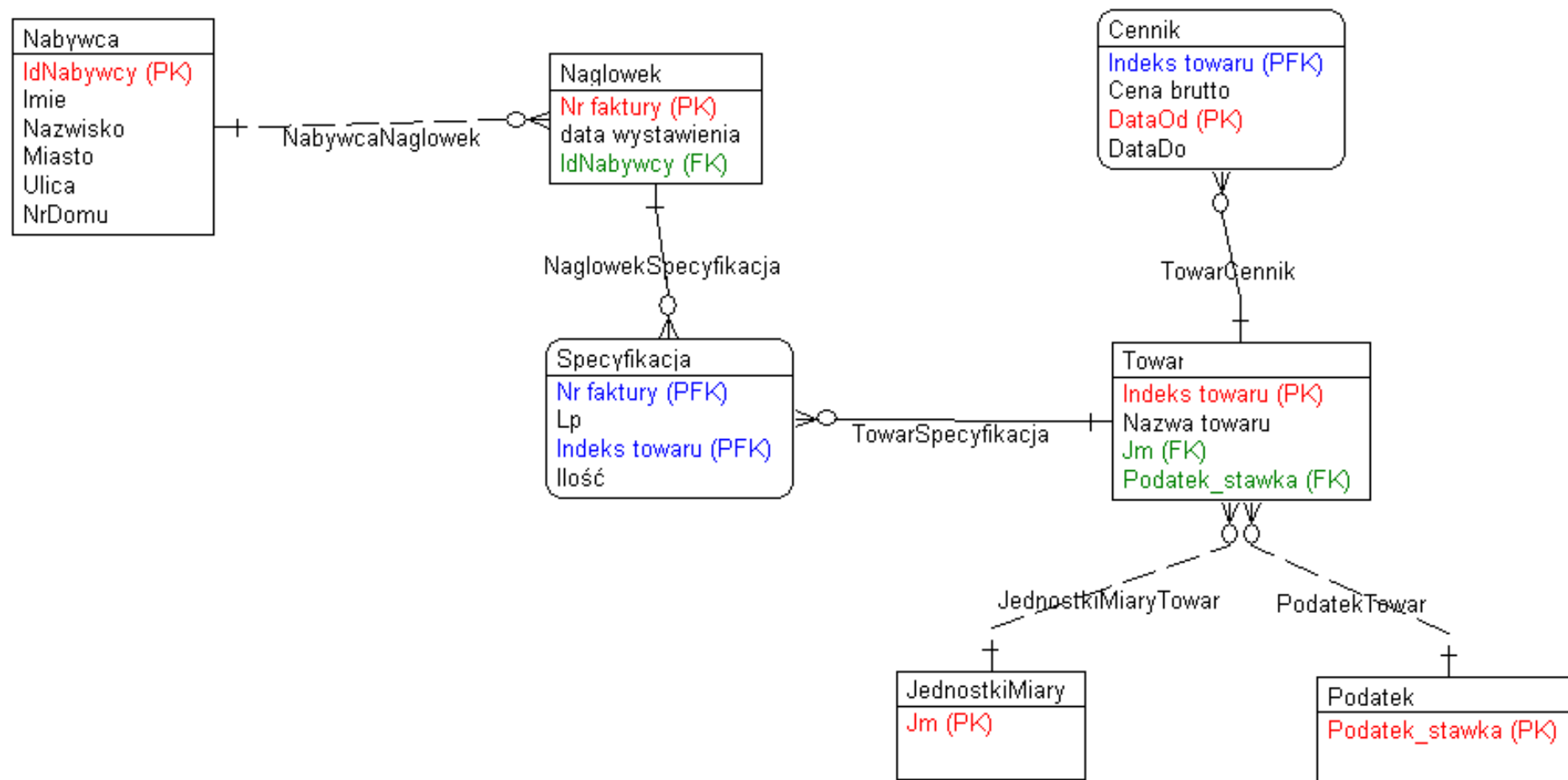
  

Kierunek
Era
Plus GSM

można nie używać sztucznych kluczy ale należy wówczas zadbać o integralność poprzez zapewnienie kaskadowej aktualizacji:  
***on update cascade on delete cascade***

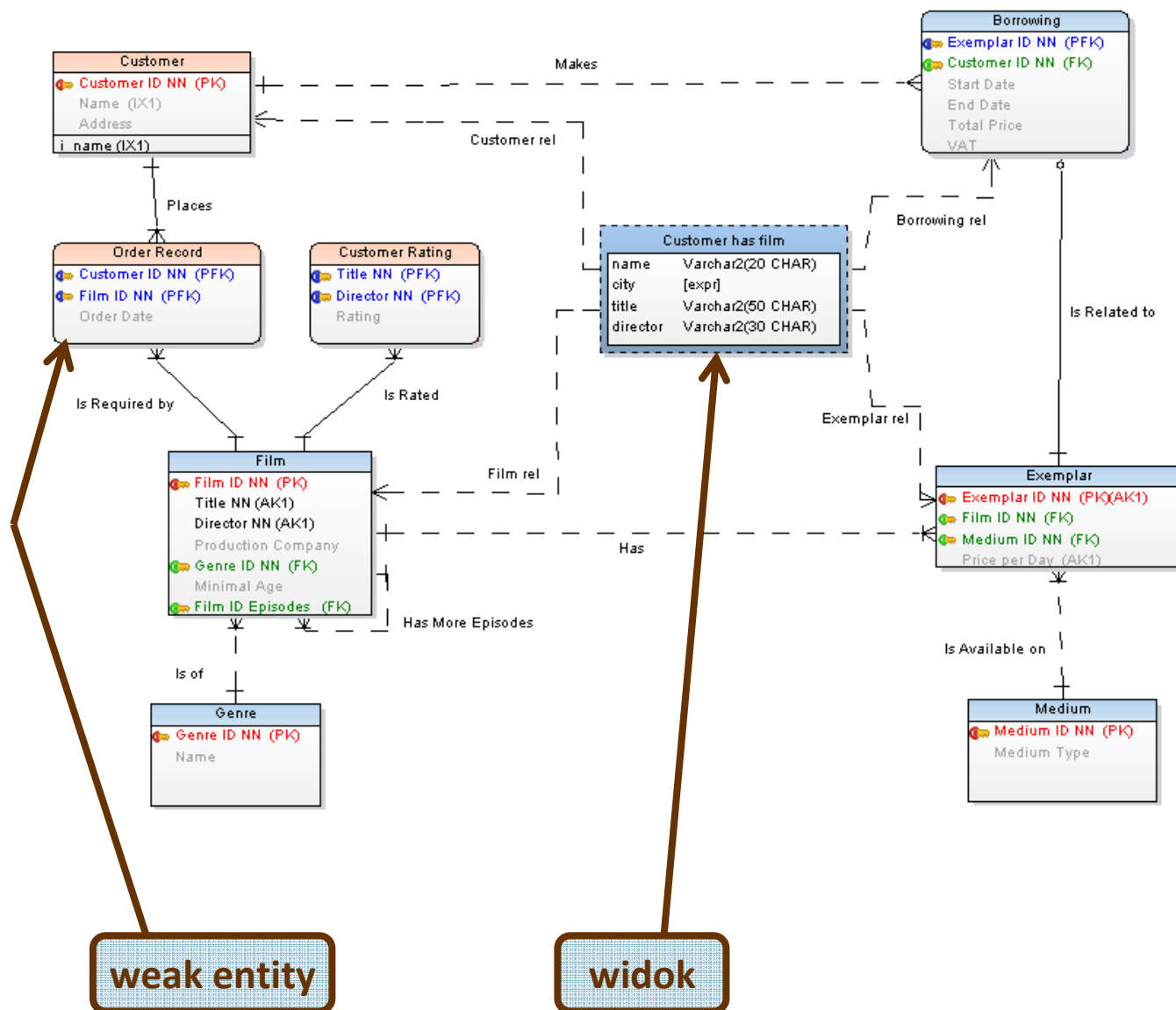








# Toad



Project	Videorental Project
Model	Videorental
Author	Radim Mario Tkaick
Company	Quest Software, Inc.
Version	
Date of Creation	2007-03-15 09:50
Last Change	2011-06-01 12:55

Display notes:

- IE notation
- Customer and Customer Rating tables in special Category
- Indexes displayed
- Logical names displayed
- View and View relationships

Category
Customer info