Normalizacja relacyjnych baz danych

Sebastian Ernst

- Zależność funkcyjna pomiędzy zbiorami atrybutów X oraz Y oznacza, że każdemu zestawowi wartości atrybutów X odpowiada dokładnie jeden zestaw wartości atrybutów Y.
- Zależności: trywialne, proste, tranzytywne.

$$X \to Y$$

- Zależność funkcyjna pomiędzy zbiorami atrybutów X oraz Y oznacza, że każdemu zestawowi wartości atrybutów X odpowiada dokładnie jeden zestaw wartości atrybutów Y.
- Zależności: trywialne, proste, tranzytywne.

$$X \to Y$$
 $X \to A_i$

- Zależność funkcyjna pomiędzy zbiorami atrybutów X oraz Y oznacza, że każdemu zestawowi wartości atrybutów X odpowiada dokładnie jeden zestaw wartości atrybutów Y.
- Zależności: trywialne, proste, tranzytywne.

$$X \to Y$$
 $X \to A_i$ $X \to Y$ $Y \to Z$

$$X \to Y$$

- "Jeżeli znam X, to znam Y."
- Przykłady:
 - PESEL → Nazwisko
 - Kod paskowy → Nazwa towaru, cena

Po co normalizacja?

- Redukcja nadmiarowości.
- Uniknięcie anomalii.

Postacie normalne – po co są?

- I. Brak powtarzających się elementów bądź grup elementów.
- 2. Brak częściowych zależności od kluczy złożonych.
- 3. Brak zależności od atrybutów niekluczowych.

Baza faktur i zamówień

International Widgets

DESCRIPTION

56" Blue Freen

3" Red Freen

750

Spline End (Xtra Large)

692

3" Red Freen

742 Evergreen Terrace Springfield, MO INVOICE NO: 125 DATE: September 13, 2002

To:

QUANTITY

4

32

Foo, Inc. 23 Main St. Thorpleburg, TX

ITEM ID

563

851

692

Customer No. __56

AMOUNT

\$14.00

\$8.00

\$60.00

UNIT

PRICE

3.50

12.00

.25

MONOR

VOICE NO: 126 ember 14, 2002

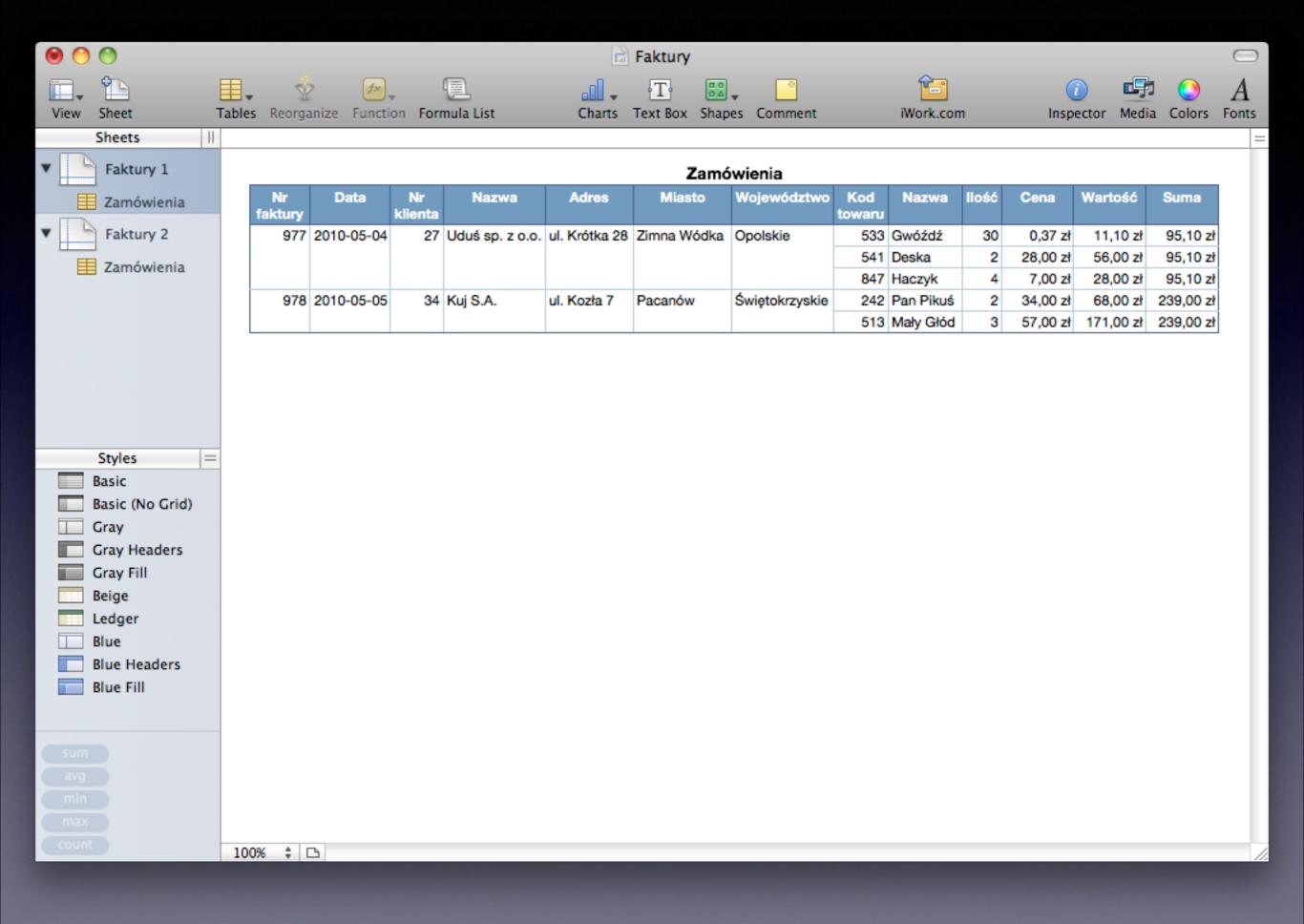
omer No. 2

TOTAL DUE \$82.00

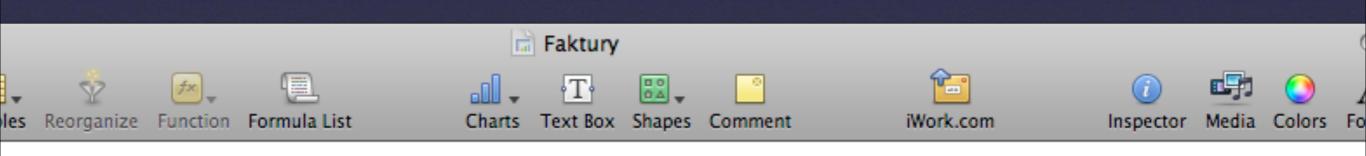
50 \$1,750.00

\$9,000.00

TOTAL DUE \$10,750.00



- Ile Panów Pikusiów zamówił Kuj S.A. w 2010 roku?
- Ile sprzedano Małych Głodów w świętokrzyskiem?
- Jakie towary sprzedano 6 grudnia 2009?

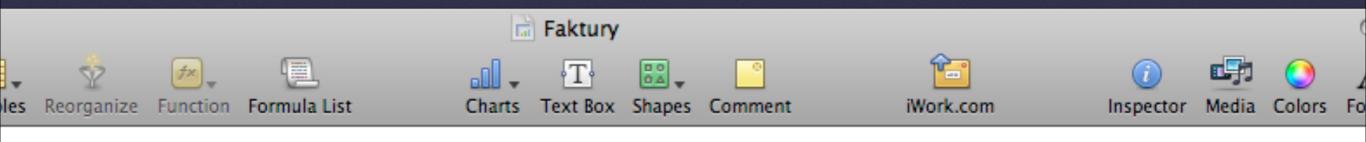


Zamówienia

Nr	Data	Nr	Nazwa	Adres	Miasto	Województwo	Kod	Nazwa	llość	Cena	Wartość	Suma
faktury		klienta					towaru					
977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka Opolskie		533	Gwóźdź	30	0,37 zł	11,10 zł	95,10 zł
							541	Deska	2	28,00 zł	56,00 zł	95,10 zł
							847	Haczyk	4	7,00 zł	28,00 zł	95,10 zł
978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	242	Pan Pikuś	2	34,00 zł	68,00 zł	239,00 zł
							513	Mały Głód	3	57,00 zł	171,00 zł	239,00 zł

INF:

Brak powtarzających się elementów bądź grup elementów.



Zamówienia

Nr	Data	Nr	Nazwa	Adres	Miasto	Województwo	Kod	Nazwa	llość	Cena	Wartość	Suma
faktury		klienta					towaru					
977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	533	Gwóźdź	30	0,37 zł	11,10 zł	95,10 zł
							541	Deska	2	28,00 zł	56,00 zł	95,10 zł
							847	Haczyk	4	7,00 zł	28,00 zł	95,10 zł
978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	242	Pan Pikuś	2	34,00 zł	68,00 zł	239,00 zł
							513	Mały Głód	3	57,00 zł	171,00 zł	239,00 zł

INF:

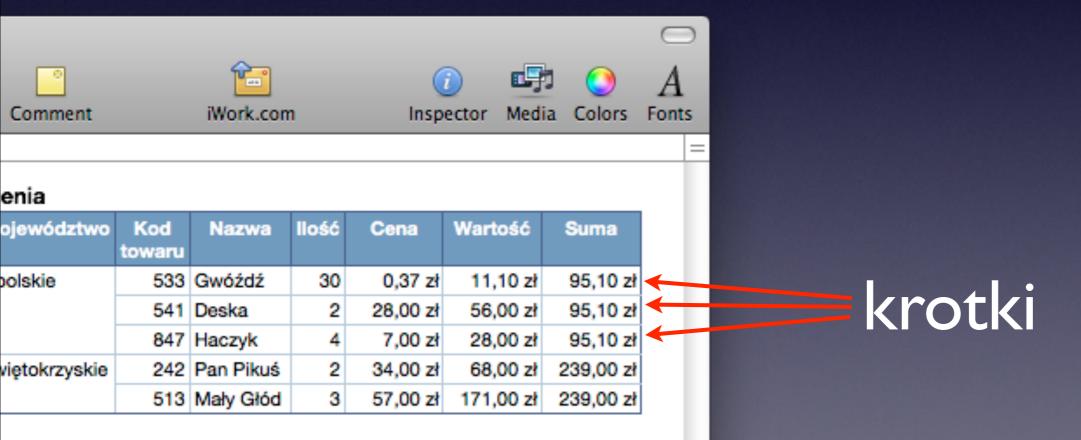
Brak powtarzających się elementów bądź grup elementów.



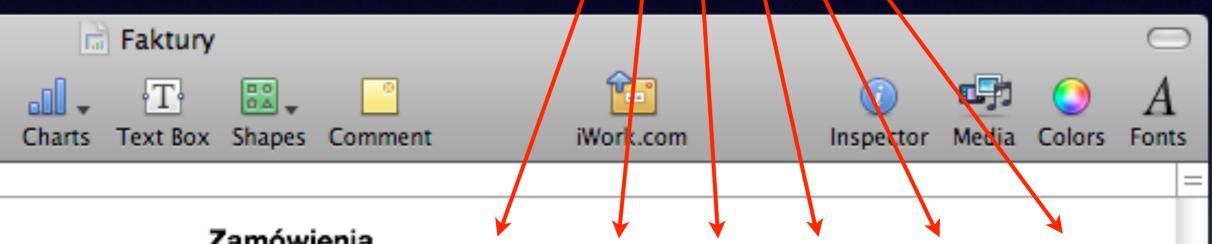
dane jednej faktury

INF:

Brak powtarzających się elementów bądź grup elementów.

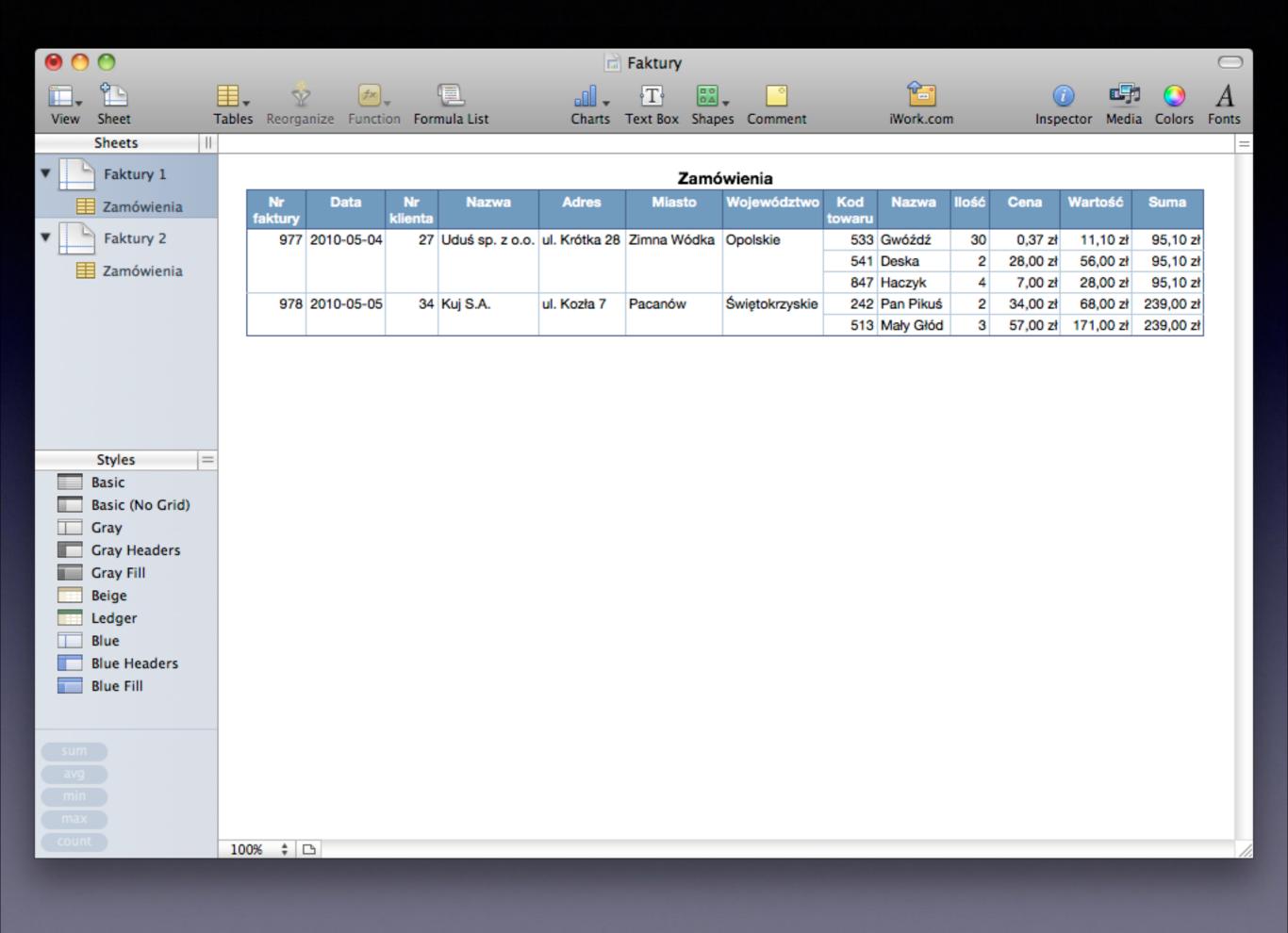


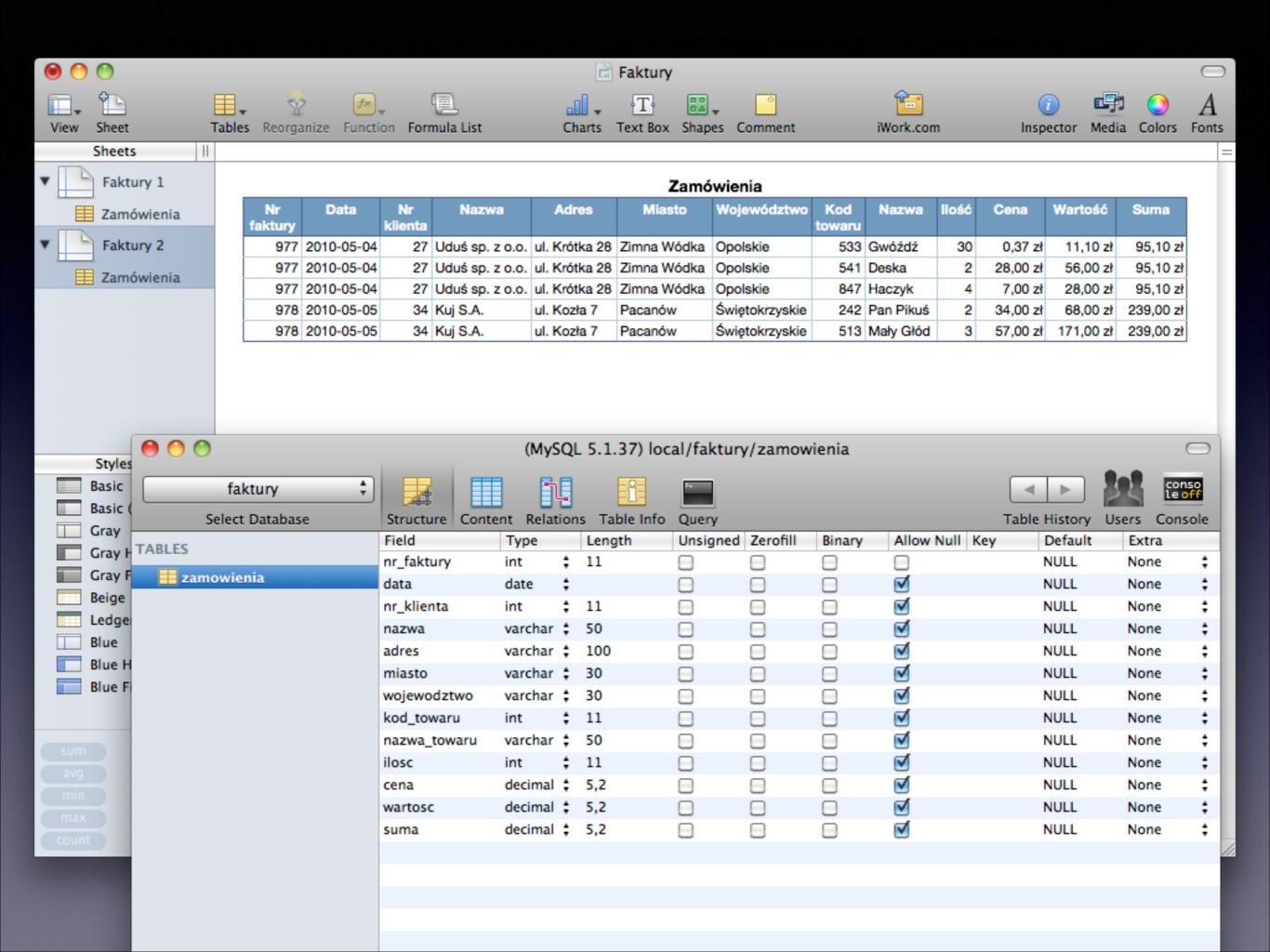


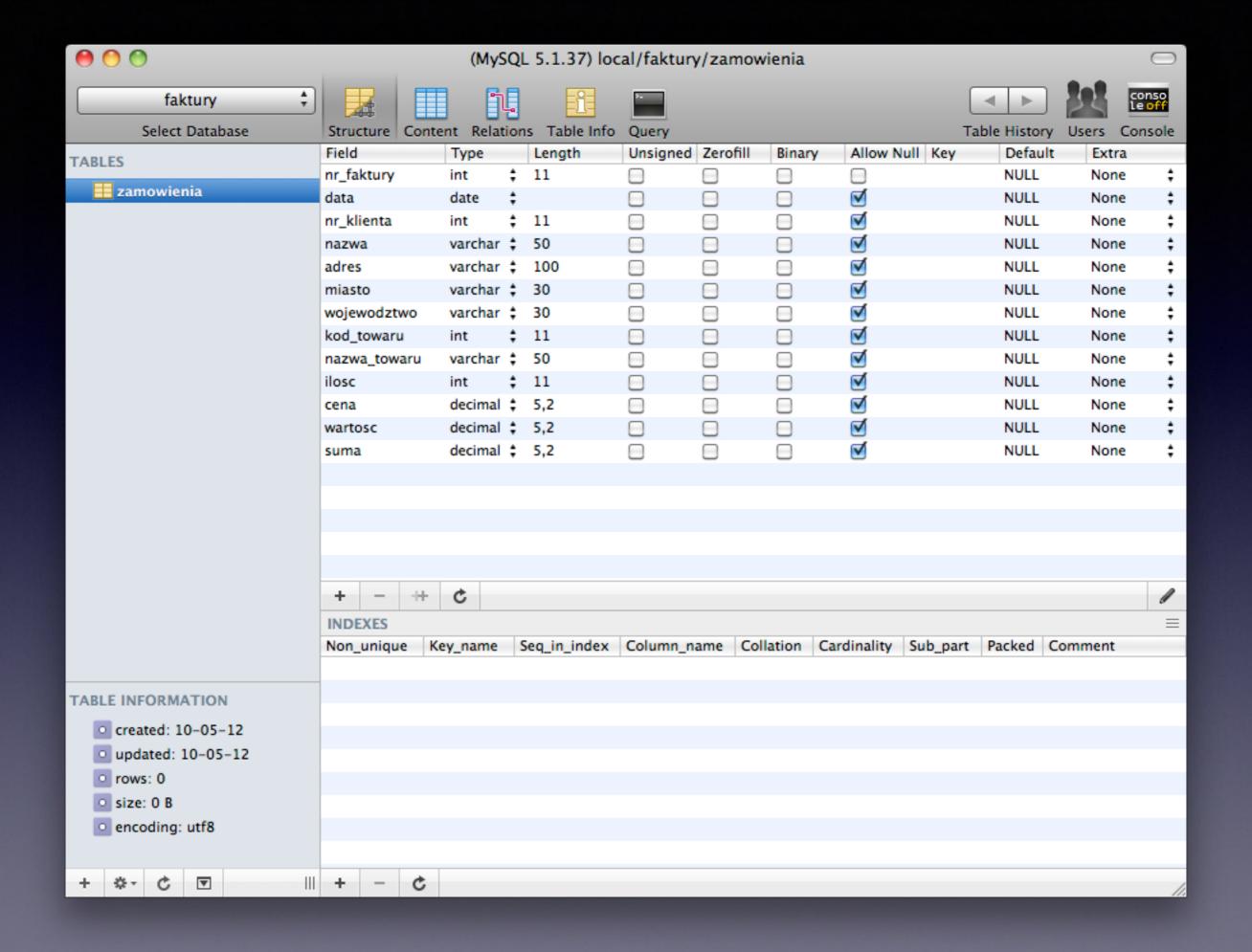


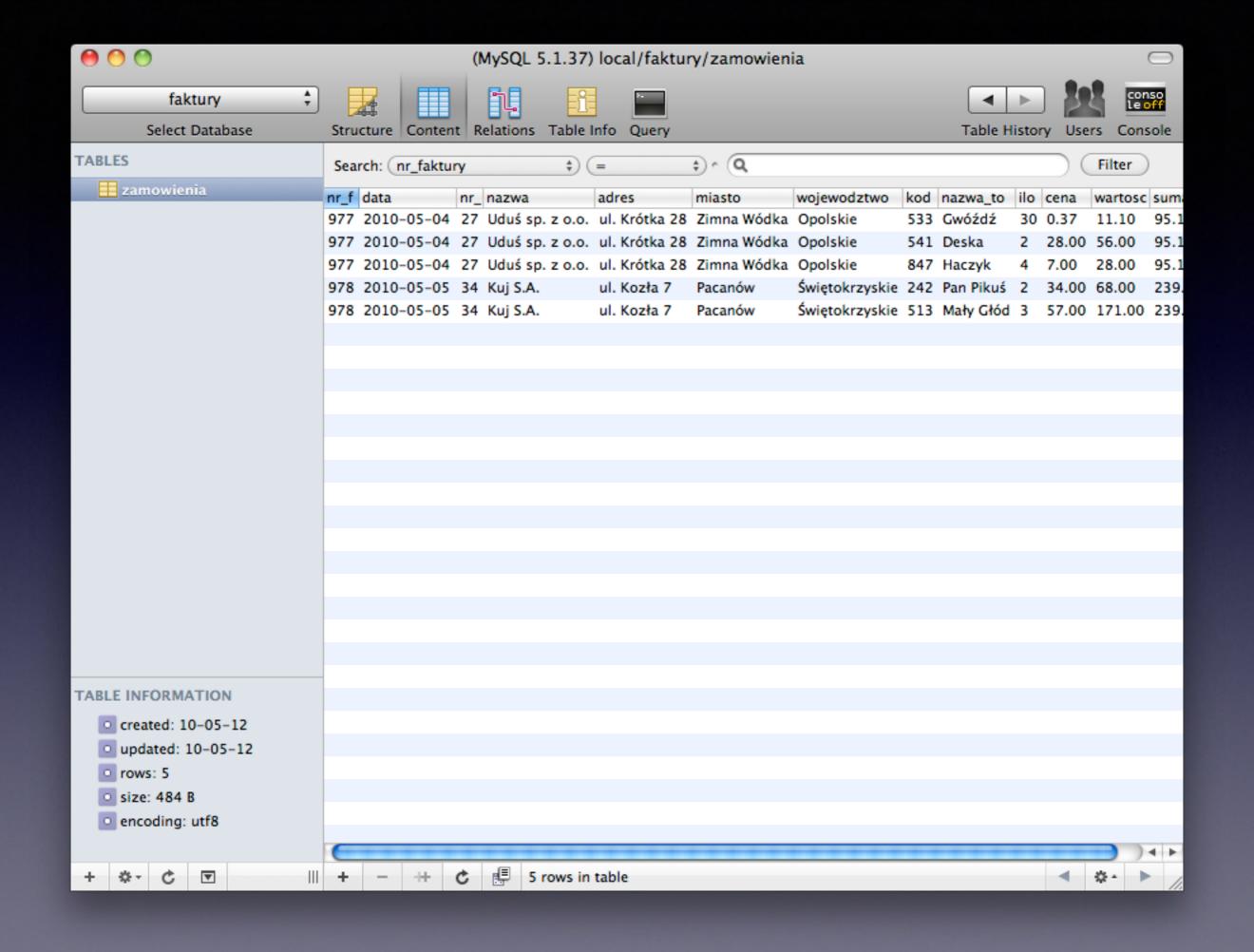
Zamówienia

Adres	Miasto	Województwo	Kod towaru		llość	Cena	Wartość	Suma
Krótka 28	Zimna Wódka	Opolskie	533	Gwóźdź	30	0,37 zł	11,10 zł	95,10 zł
			541	Deska	2	28,00 zł	56,00 zł	95,10 zł
			847	Haczyk	4	7,00 zł	28,00 zł	95,10 zł
Kozła 7	Pacanów	Świętokrzyskie	242	Pan Pikuś	2	34,00 zł	68,00 zł	239,00 zł
			513	Mały Głód	3	57,00 zł	171,00 zł	239,00 zł



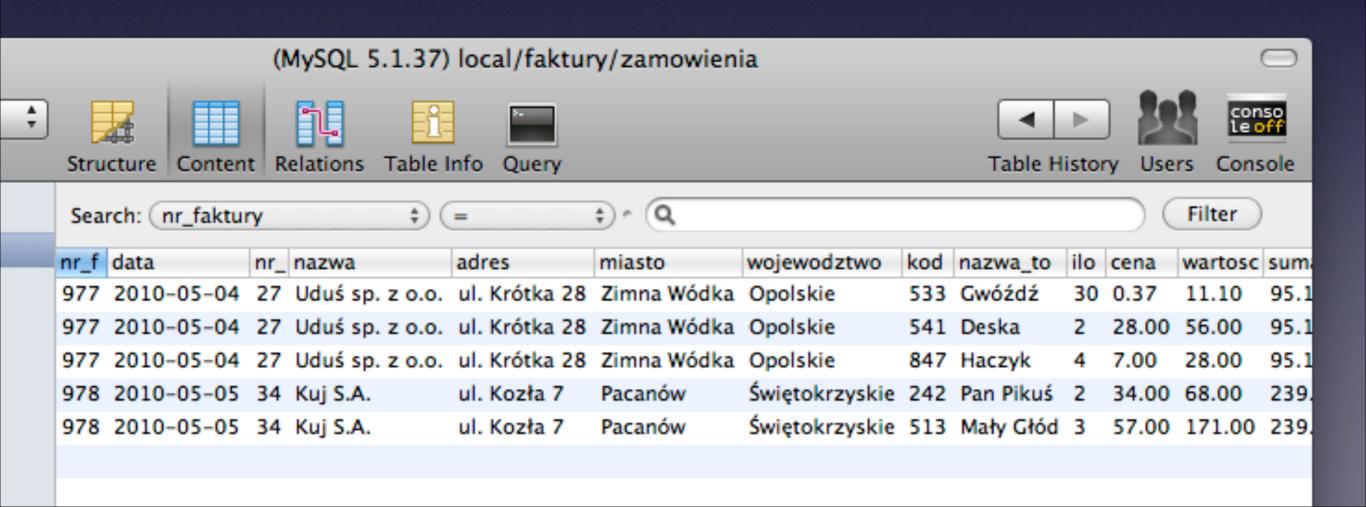






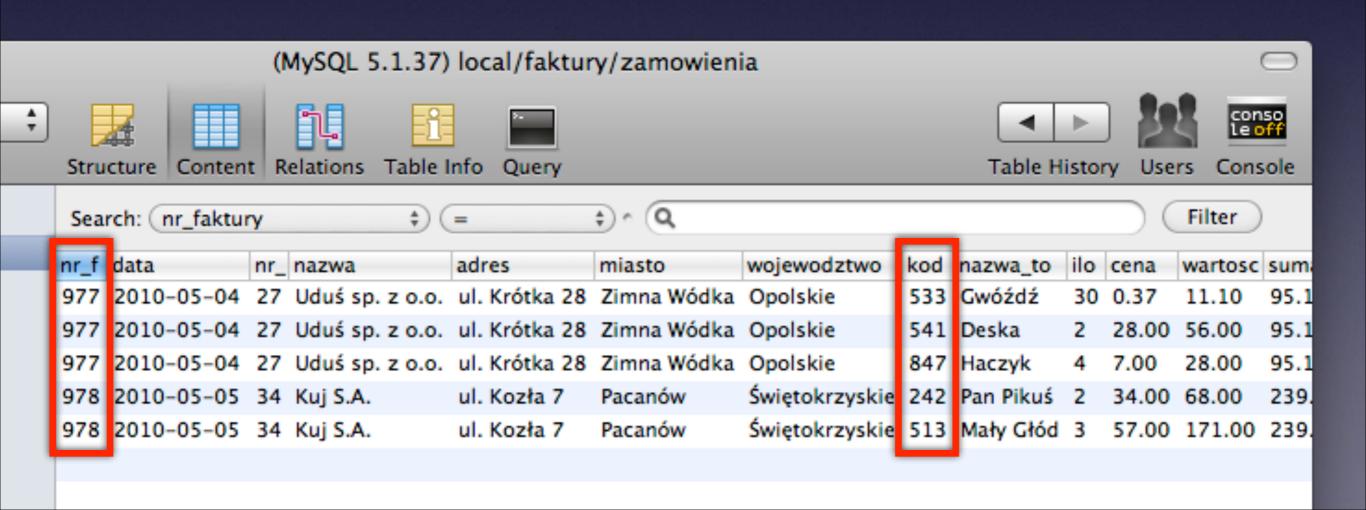
Wymagania INF:

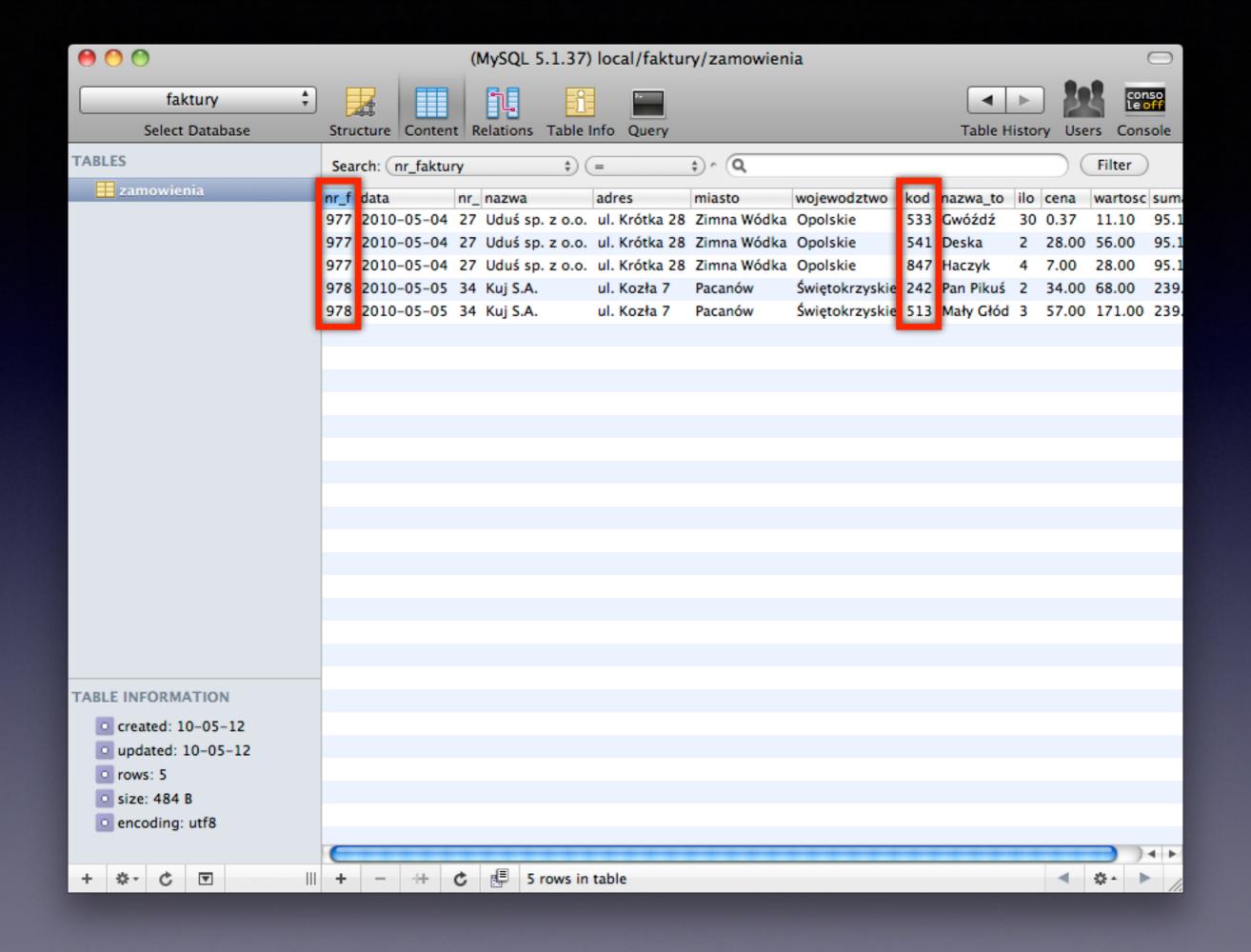
- Atomiczność: wiersz nie może zawierać powtarzających się grup o podobnej strukturze.
- Klucz główny: każdy wiersz musi posiadać unikalny identyfikator.

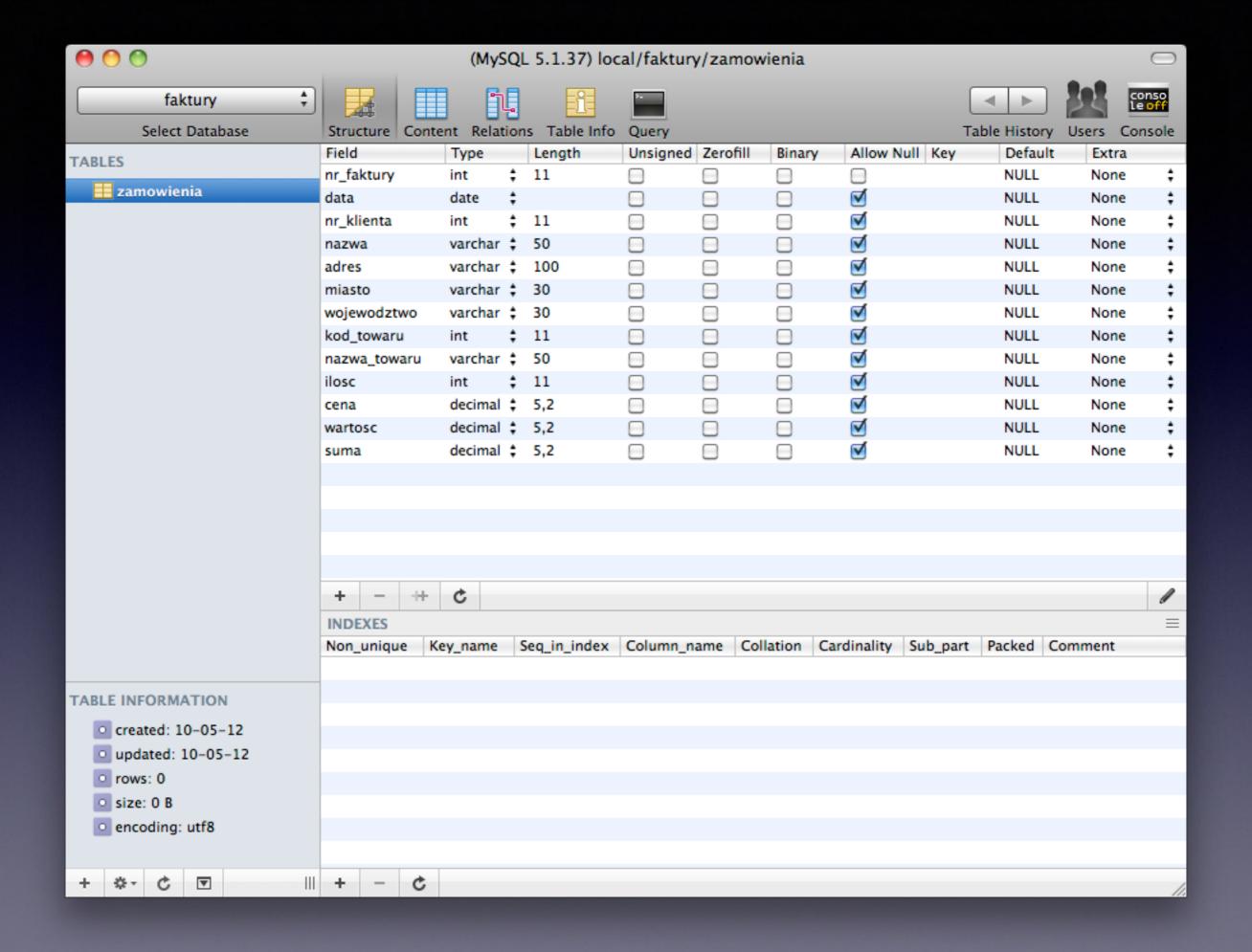


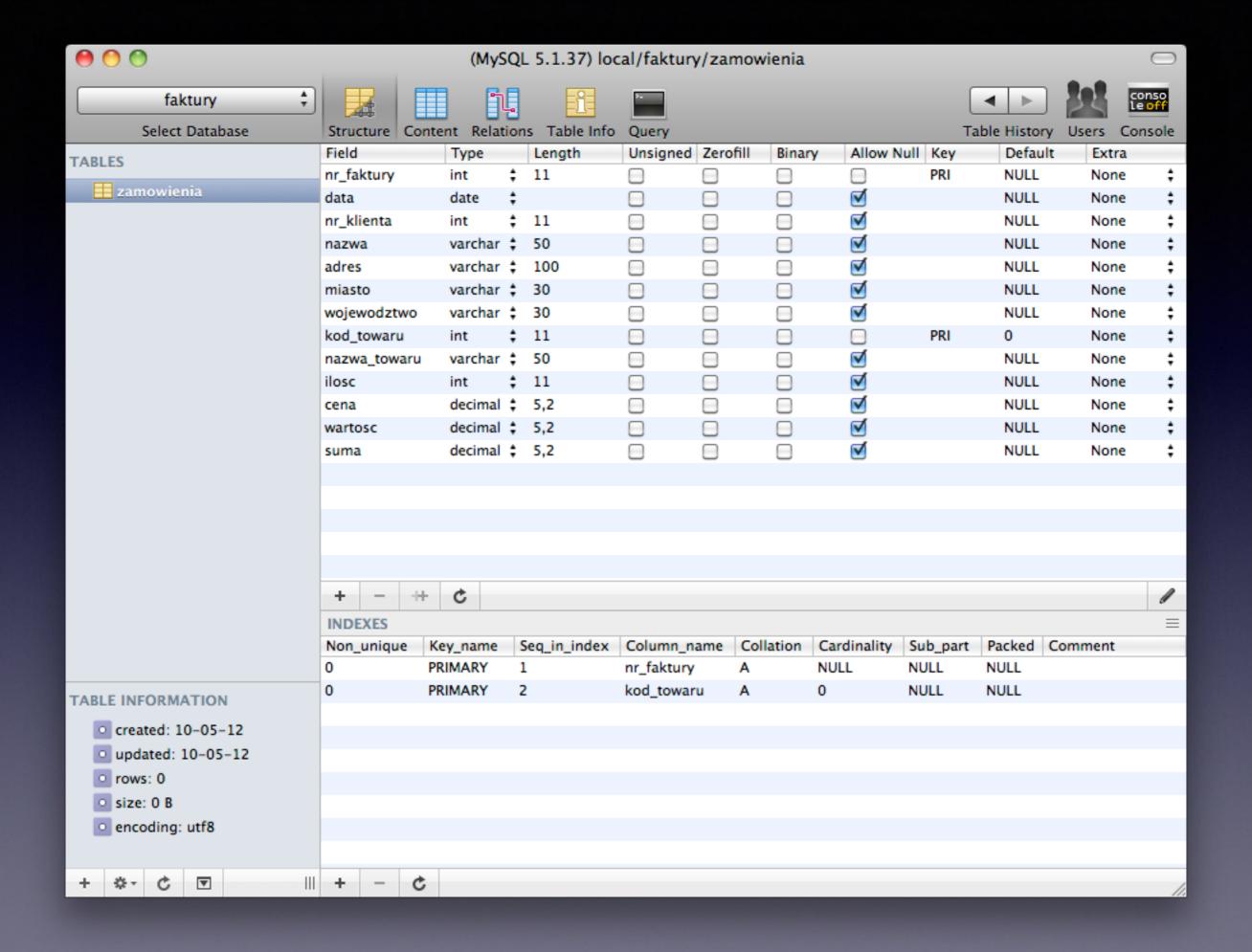
Wymagania INF:

- Atomiczność: wiersz nie może zawierać powtarzających się grup o podobnej strukturze.
- Klucz główny: każdy wiersz musi posiadać unikalny identyfikator.







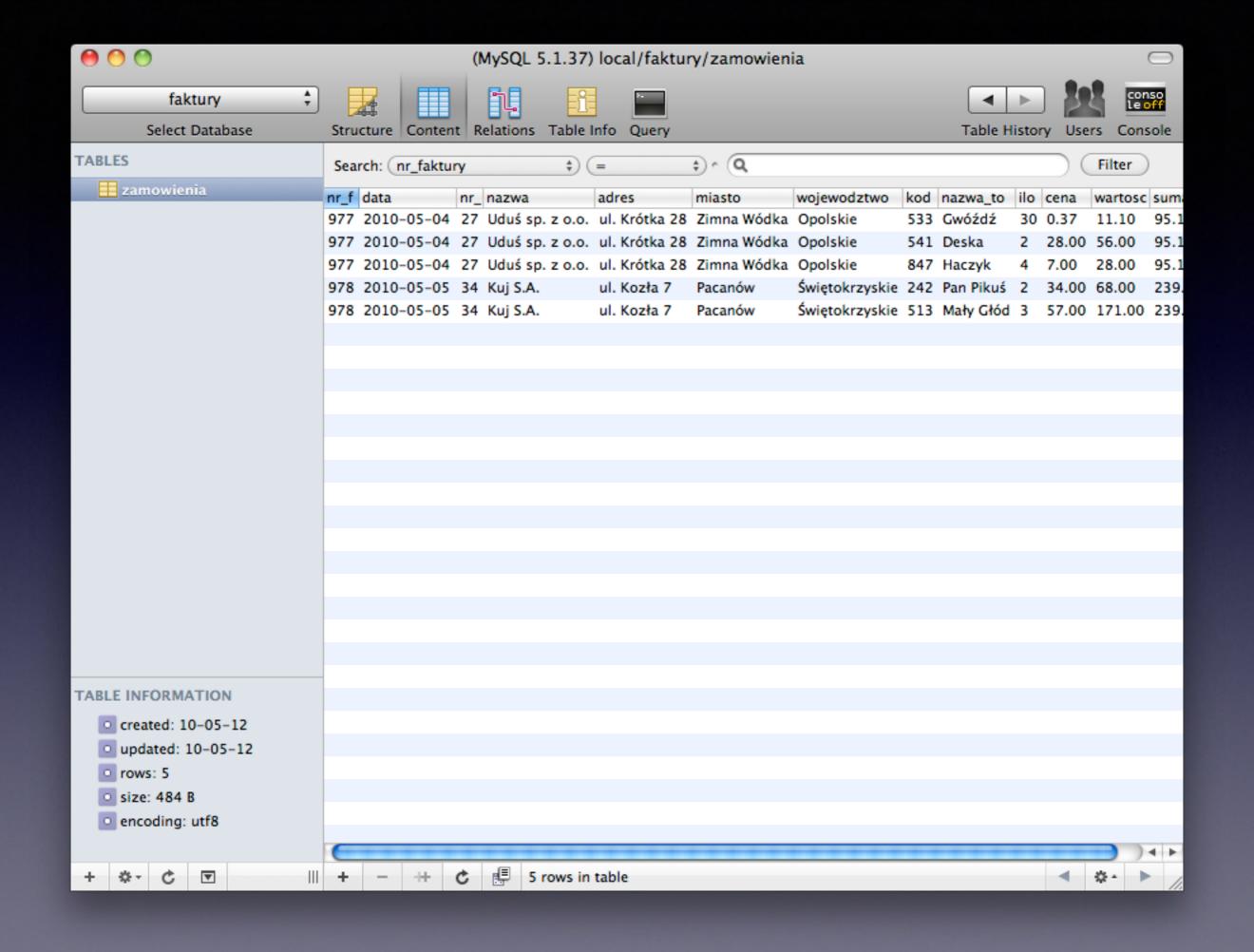


2NF

3NF

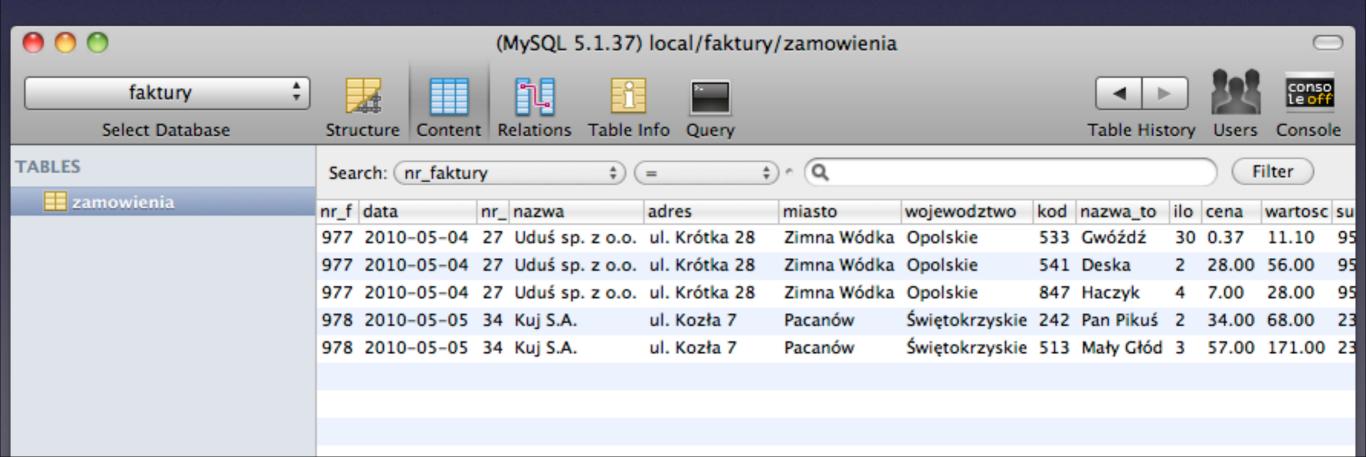
2NF

3NF



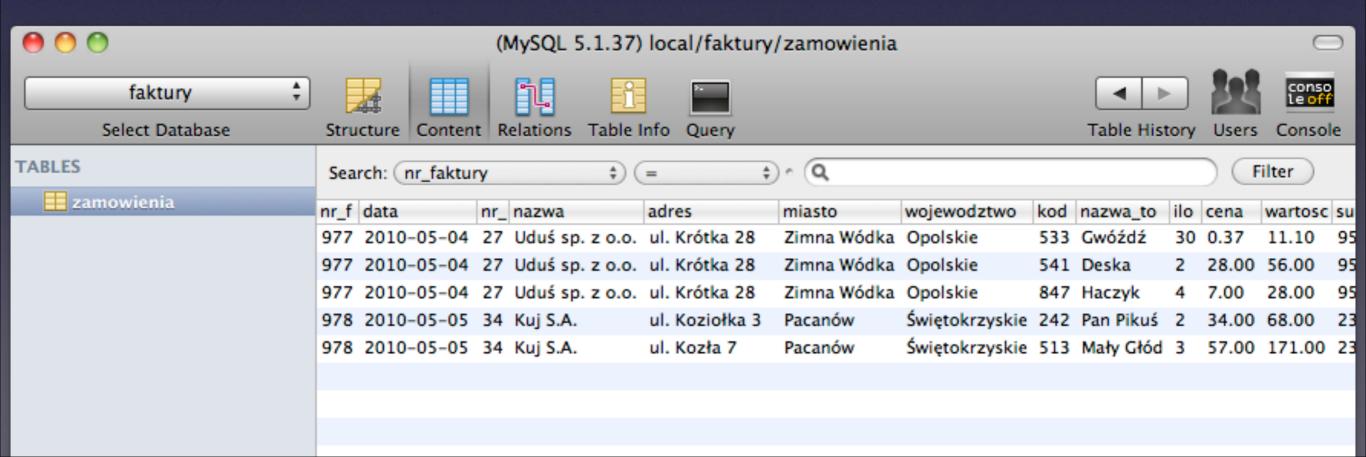
Anomalie modyfikacji

 Klient Kuj S.A. przeniósł się z ul. Kozła 7 na ul. Koziołka 3.



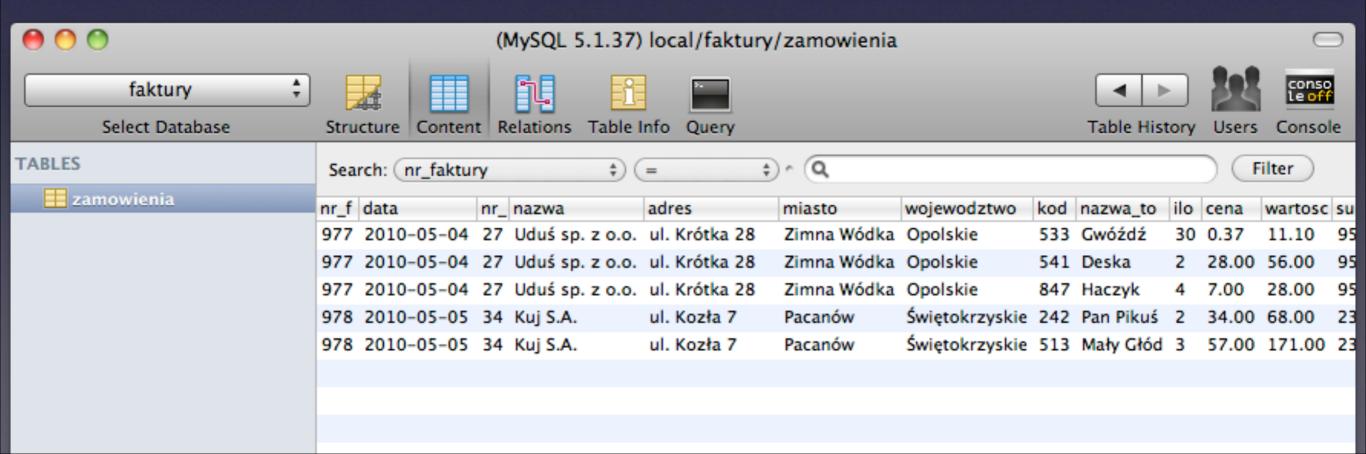
Anomalie modyfikacji

 Klient Kuj S.A. przeniósł się z ul. Kozła 7 na ul. Koziołka 3.



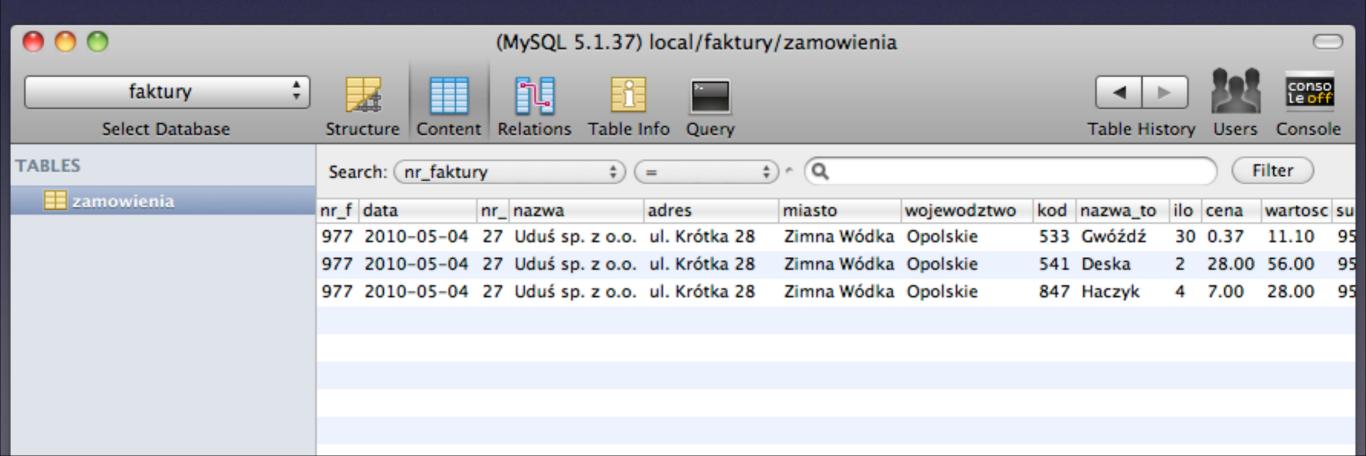
Anomalie usuwania

 Klient Kuj S.A. anulował swoje ostatnie zamówienie.



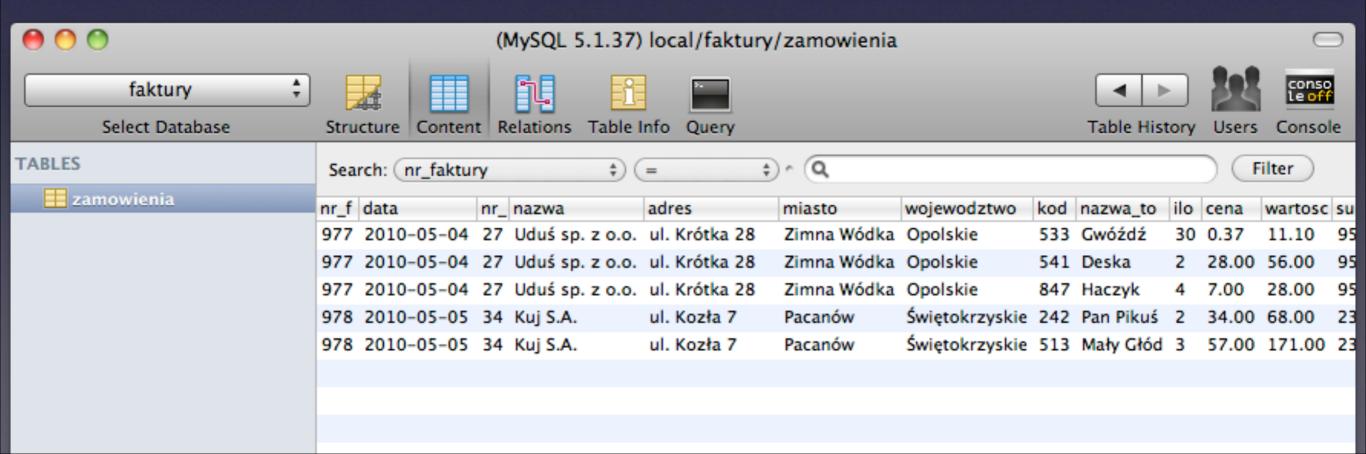
Anomalie usuwania

 Klient Kuj S.A. anulował swoje ostatnie zamówienie.



Anomalie dodawania

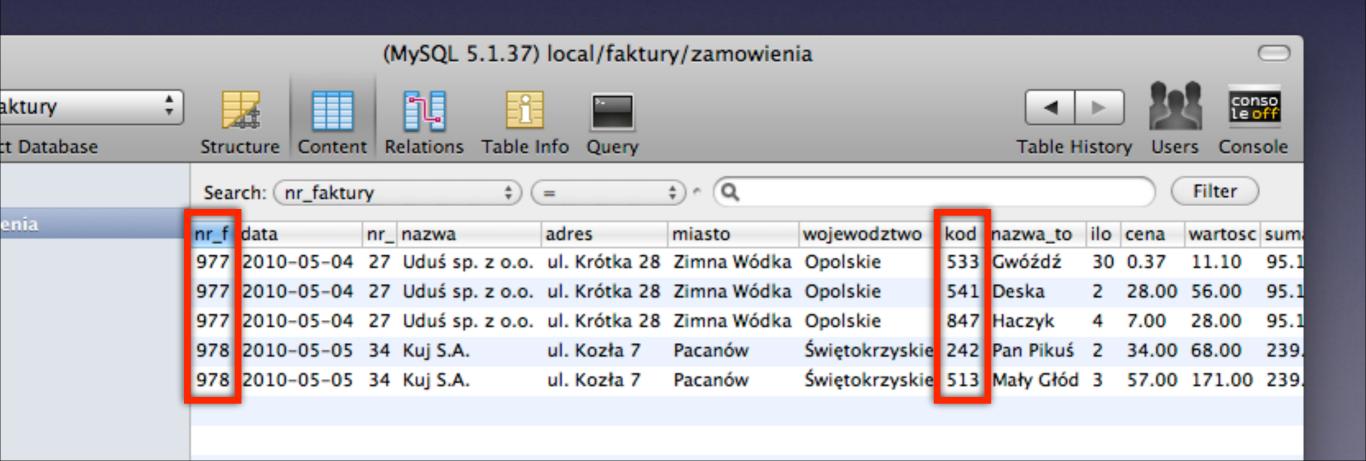
 Chcemy dopisać nowego klienta, który jeszcze nic u nas nie zamawiał.



2NF:

Brak częściowych zależności od kluczy złożonych.

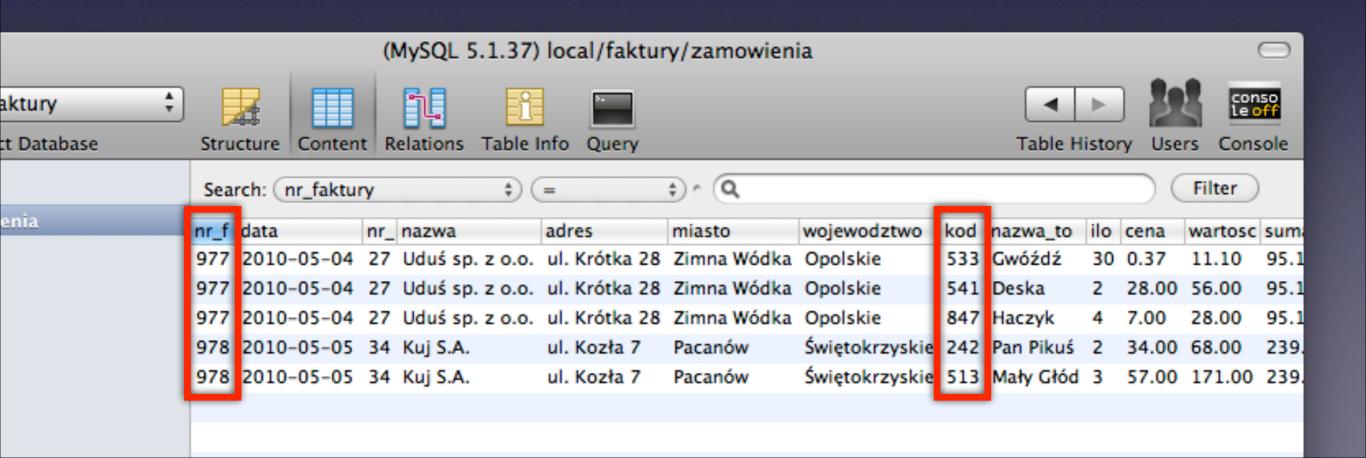
- Relacja w INF.
- Każdy atrybut spoza klucza głównego jest od niego w pełni funkcyjnie zależny.
- Czyli: nie jest zależny od jego fragmentu.



2NF:

Brak częściowych zależności od kluczy złożonych.

- Wartość każdej kolumny musi zależeć wyłącznie od całego klucza złożonego.
- "Czy ta kolumna może istnieć <u>bez</u> któregoś ze składników klucza?"



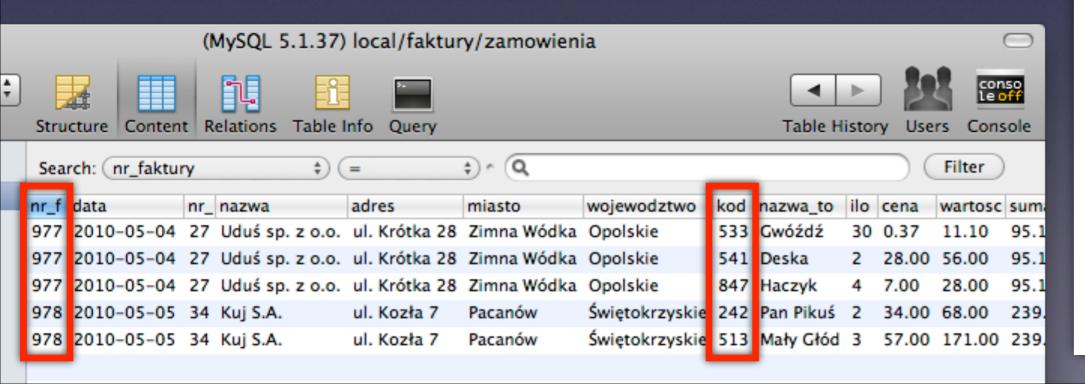
2NF:

Brak częściowych zależności od kluczy złożonych.

 Wartość każdej kolumny musi zależeć wyłącznie od całego klucza złożonego.

"Czy ta kolumna może istnieć <u>bez</u> któregoś ze

składników klucza?"



ER Diagram: faktury

zamowienia								
nr_faktury:///								
data: <i>DATE</i>								
nr_klienta:///								
nazwa: <i>VARCHAR</i>								
adres: VARCHAR								
miasto:VARCHAR								
wojewodztwo:VARCHAR								
kod_towaru:///								
nazwa_towaru:VARCHAR								
ilosc:INT								
cena:DECIMAL								
wartosc:DECIMAL								
suma:DECIMAL								

"Czy ta kolumna może istnieć bez któregoś ze składników klucza?"

(MySQL 5.1.37) local/faktury/zamowienia











Structure Content Relations Table Info Query **ER Diagram: faktury**

zamowienia
nr_faktury:///
data: <i>DATE</i>
nr_klienta: <i>INT</i>
nazwa: <i>varchar</i>
adres: <i>varchar</i>
miasto: VARCHAR
wojewodztwo:varchar
kod_towaru:///
nazwa_towaru:varchar
ilosc:///
cena:DECIMAL
wartosc:DECIMAL
suma:DECIMAL



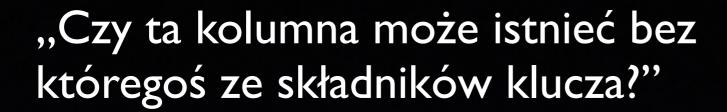


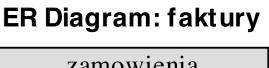


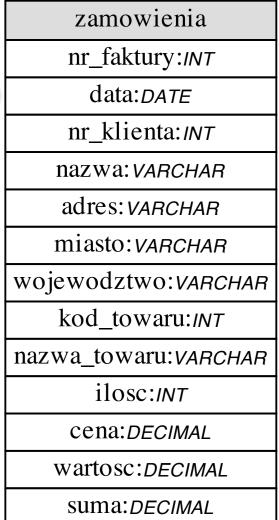
Table History

Users Console

	Search: (nr_faktury												
	nr_f	data	nr_	nazwa	adres	miasto	wojewodztwo	kod	nazwa_to	ilo	cena	wartosc	suma
1	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	533	Gwóźdź	30	0.37	11.10	95.1
1	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	541	Deska	2	28.00	56.00	95.1
1	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	847	Haczyk	4	7.00	28.00	95.1
1	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	242	Pan Pikuś	2	34.00	68.00	239.
1	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	513	Mały Głód	3	57.00	171.00	239.
	_							_					

















Relations Table Info Query







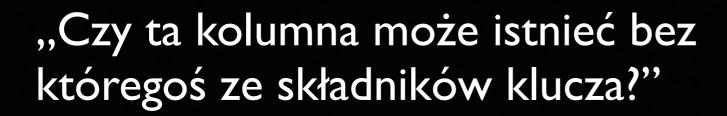
Table History

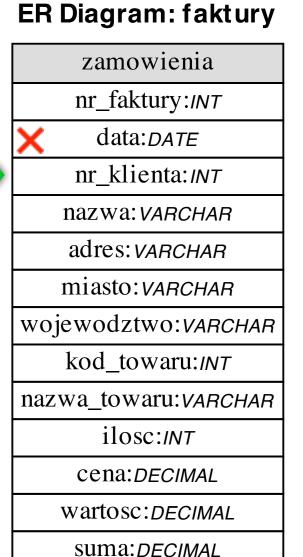
Users

Console

Filter

	Sear	cn:ktui	у		=	•) · (4						riitei	
	nr_f	data	nr_	nazwa	adres	miasto	wojewodztwo	kod	nazwa_to	ilo	cena	wartosc	sum
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	533	Gwóźdź	30	0.37	11.10	95.1
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	541	Deska	2	28.00	56.00	95.1
١	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	847	Haczyk	4	7.00	28.00	95.1
П	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	242	Pan Pikuś	2	34.00	68.00	239.
	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	513	Mały Głód	3	57.00	171.00	239.







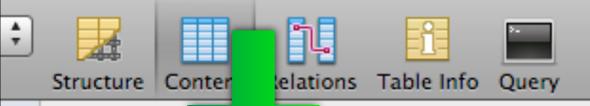




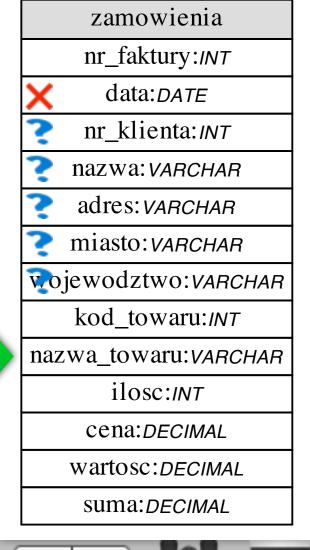




Table History Console

	Sear	rch: nr_fakt		\$	=	‡) ↑ (Q) (Filter	
ı	nr_f	data	nr_	nazwa	adres	miasto	wojewodztwo	kod	nazwa_to	ilo	cena	wartosc	suma
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	533	Gwóźdź	30	0.37	11.10	95.1
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	541	Deska	2	28.00	56.00	95.1
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	847	Haczyk	4	7.00	28.00	95.1
ı	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	242	Pan Pikuś	2	34.00	68.00	239.
	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	513	Mały Głód	3	57.00	171.00	239.

"Czy ta kolumna może istnieć bez któregoś ze składników klucza?"



le History

conso

Console

ER Diagram: faktury











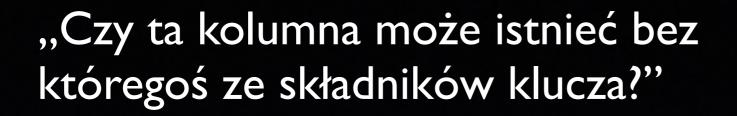


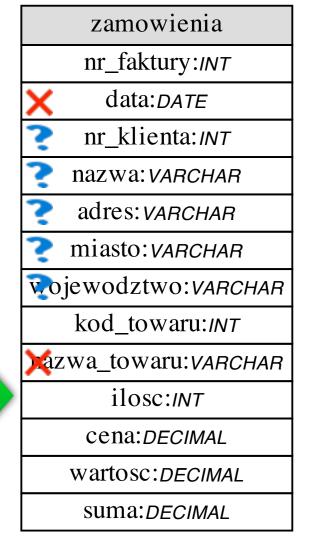
Table Info

Query

Search: (nr_	faktury	\$)(=	÷) ↑ (Q				Filter
r f data	nr nazwa	adres	miasto	wojewodztwo	kod nazwa to	ilo cena	wartosc

	_							_					
			_	nazwa	adres		wojewodztwo	kod	nazwa_to	ilo	cena	wartosc	suma
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	533	Gwóźdź	30	0.37	11.10	95.1
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	541	Deska	2	28.00	56.00	95.1
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	847	Haczyk	4	7.00	28.00	95.1
ı	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	242	Pan Pikuś	2	34.00	68.00	239.
ı	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	513	Mały Głód	3	57.00	171.00	239.
	_							_					







(MySQL 5.1.37) local/faktury/zamowienia





Search: (nr_faktury







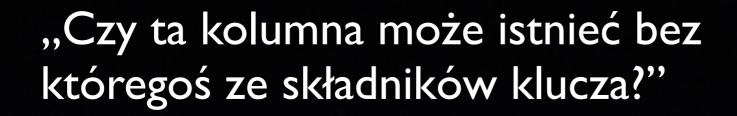
Structure Content Relations Table Info Query

Table H	y	Users
		7

conso

_			•										
	nr f	data	nr	nazwa	adres	miasto	wojewodztwo	kod	nazwa to	ilo	cena	wartosc	sum
П	_		_					_					_
1	9//	2010-05-04	27	Uduś sp. z o.o.	ul. Krotka 28	Zimna Wodka	Opolskie	533	Gwóźdź	30	0.37	11.10	95.1
1	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	541	Deska	2	28.00	56.00	95.1
1	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	847	Haczyk	4	7.00	28.00	95.1
1	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	242	Pan Pikuś	2	34.00	68.00	239.
1	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	513	Mały Głód	3	57.00	171.00	239.

\$) n (Q





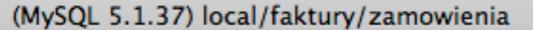




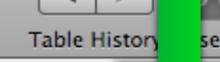




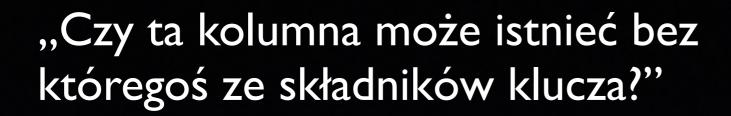




Table Info Query



	Sear	rch: (nr_faktur	у	\$	=	∌ ↑ (Q						Filter	
I	nr f	data	nr	nazwa	adres	miasto	wojewodztwo	kod	nazwa to	ilo	cena	wartosc	Sum
_			_	Uduś sp. z o.o.			-	_	Gwóźdź				95.1
_				Uduś sp. z o.o.					Deska				95.1
_				Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka			Haczyk				95.1
ı	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	242	Pan Pikuś	2	34.00	68.00	239.
	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	513	Mały Głód	3	57.00	171.00	239.



zamowienia nr_faktury:/// data: DATE nr_klienta:/// nazwa: VARCHAR adres: VARCHAR miasto: VARCHAR wojewodztwo: varchar kod_towaru:/// nazwa_towaru:*varchar* ilosc:INT cena: DECIMAL wartosc: DECIMAL suma: DECIMAL

Table History

ER Diagram: faktury







Sparch: (nr faktury



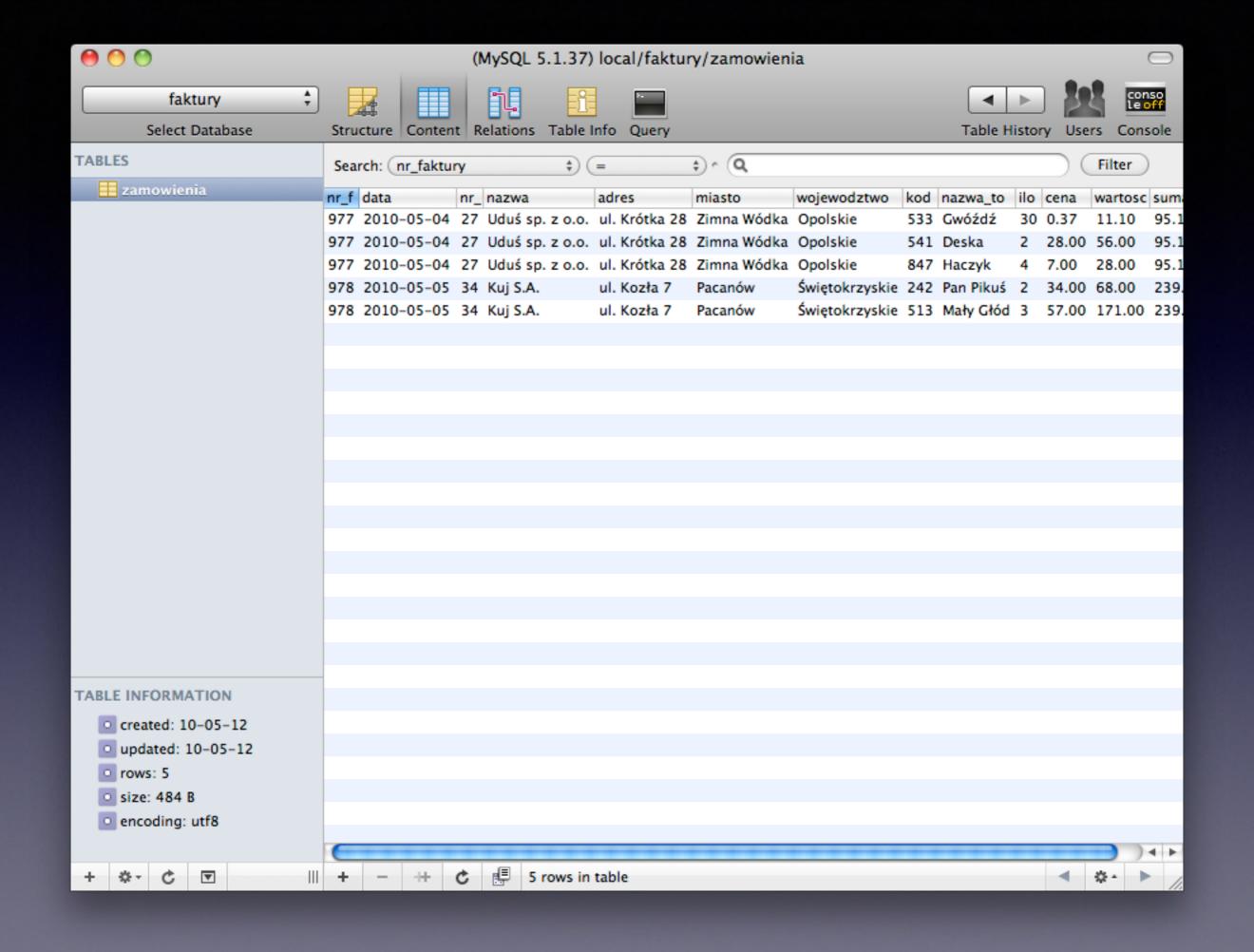


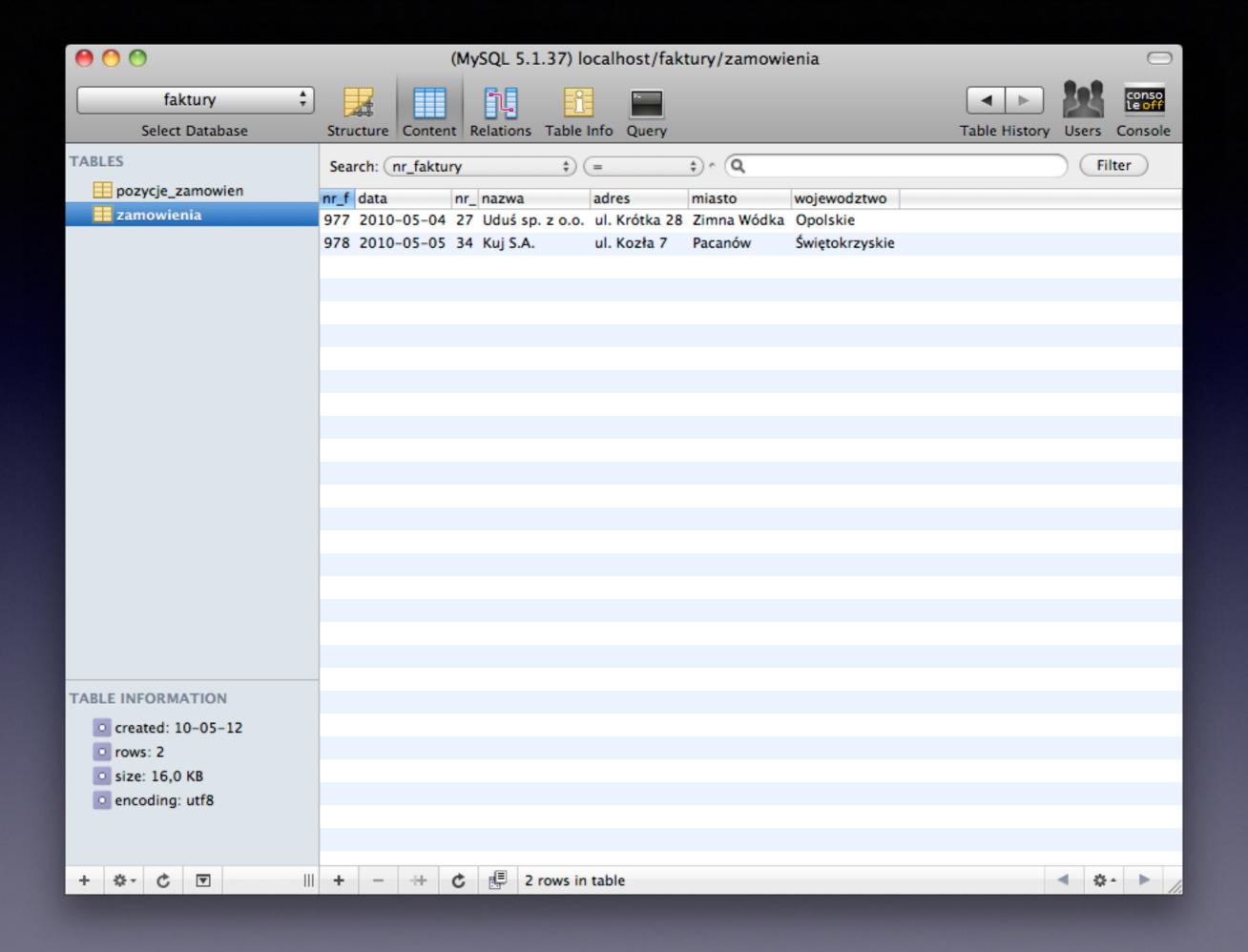


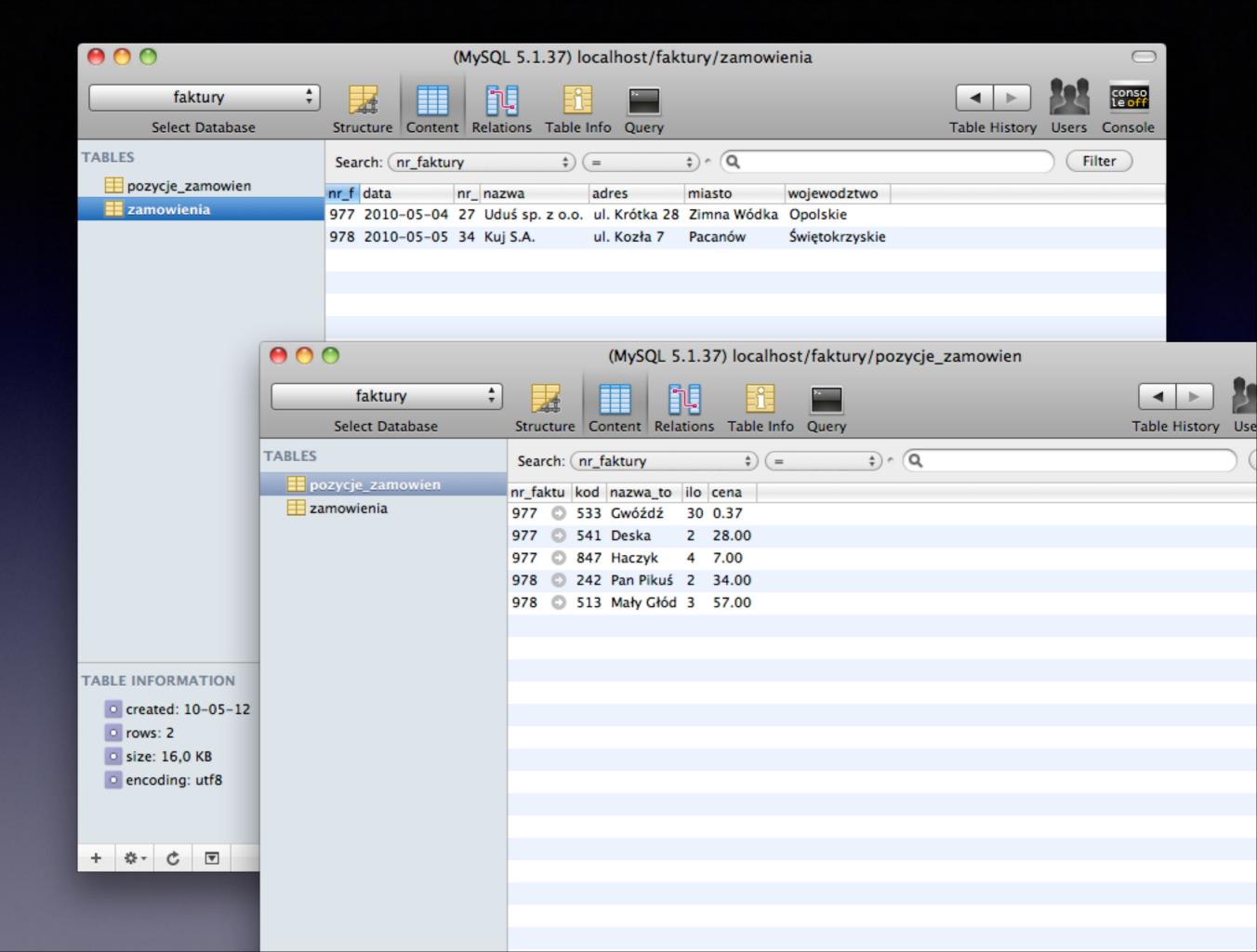
Table Info

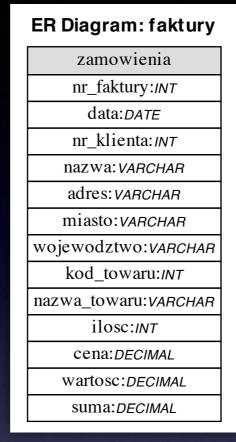
Query

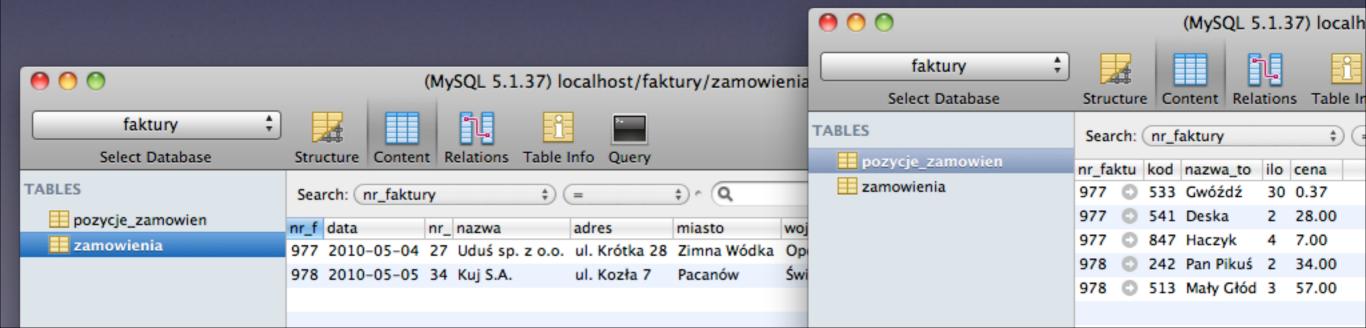
	Sear	cn: (nr_raktur	У	• •		•) · (4						riitei	
ı	nr_f	data	nr_	nazwa	adres	miasto	wojewodztwo	kod	nazwa_to	ilo	cena	wartosc	sum
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	533	Gwóźdź	30	0.37	11.10	95.1
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	541	Deska	2	28.00	56.00	95.1
ı	977	2010-05-04	27	Uduś sp. z o.o.	ul. Krótka 28	Zimna Wódka	Opolskie	847	Haczyk	4	7.00	28.00	95.1
ı	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	242	Pan Pikuś	2	34.00	68.00	239.
1	978	2010-05-05	34	Kuj S.A.	ul. Kozła 7	Pacanów	Świętokrzyskie	513	Mały Głód	3	57.00	171.00	239.

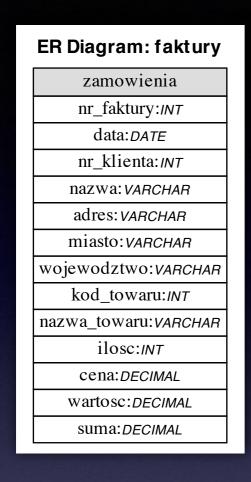


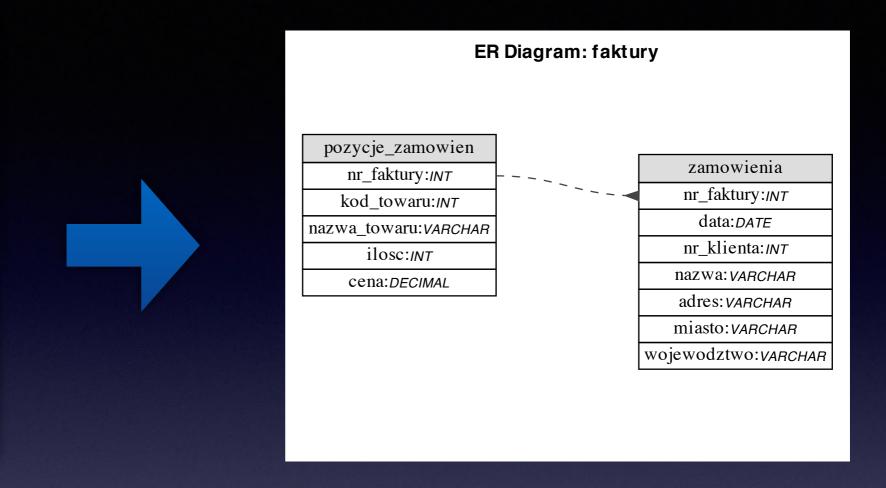


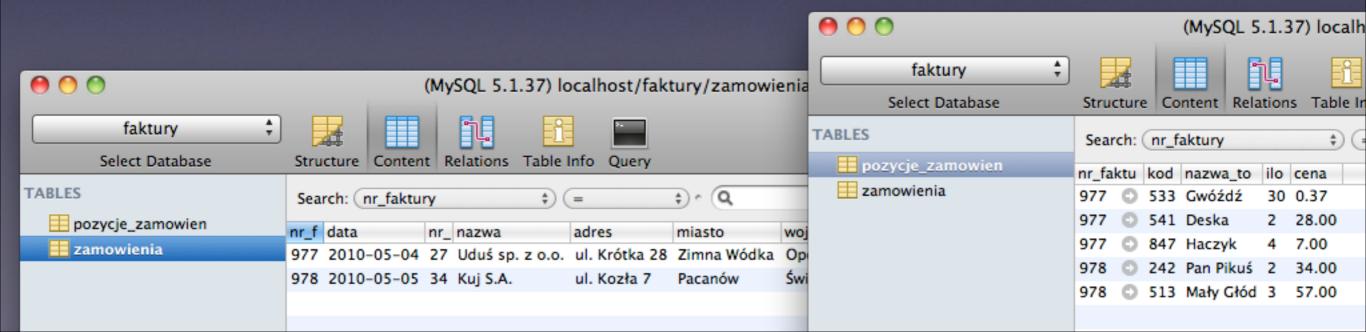




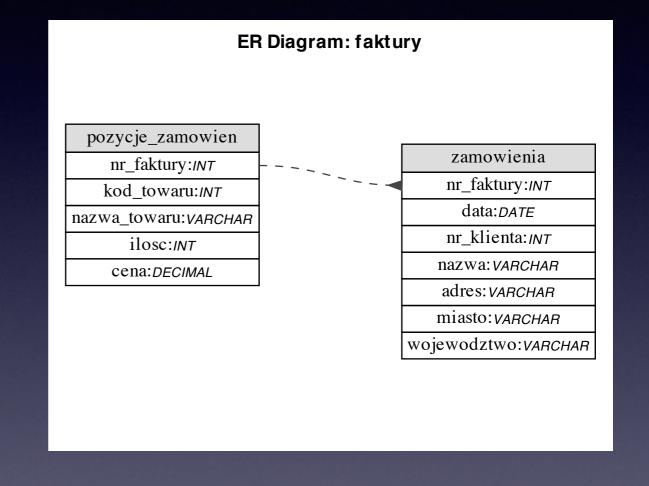








- Reguły 2NF dotyczą
 tylko tabel ze złożonym
 kluczem głównym, więc
 tabela zamowienia spełnia
 2NF automatycznie.
- Nowo utworzona tabela pozycje_zamowien musi być sprawdzona pod kątem zgodności z 2NF.



pozycje_zamowien

nr_faktury:///

kod_towaru:///T

nazwa_towaru:*VARCHAR*

ilosc:INT

cena: DECIMAL

zamowienia

nr_faktury:///

data:DATE

nr_klienta:///

nazwa: VARCHAR

adres: VARCHAR

miasto: VARCHAR

wojewodztwo: VARCHAR

pozycje_zamowien

nr_faktury:///

kod_towaru:///T

azwa_towaru:*VARCHAR*

ilosc:INT

cena: DECIMAL

zamowienia

nr_faktury:///

data:DATE

nr_klienta:///

nazwa: VARCHAR

adres: VARCHAR

miasto: VARCHAR

wojewodztwo: VARCHAR

pozycje_zamowien

nr_faktury:///

kod_towaru:///T

azwa_towaru:*varchar*

ilosc:INT

cena: DECIMAL

zamowienia

nr_faktury:///

data:DATE

nr_klienta:///

nazwa: VARCHAR

adres: VARCHAR

miasto: VARCHAR

wojewodztwo: VARCHAR

pozycje_zamowien

nr_faktury:///
kod_towaru:///
ilosc:////
ilosc:////
cena:DECIMAL

zamowienia

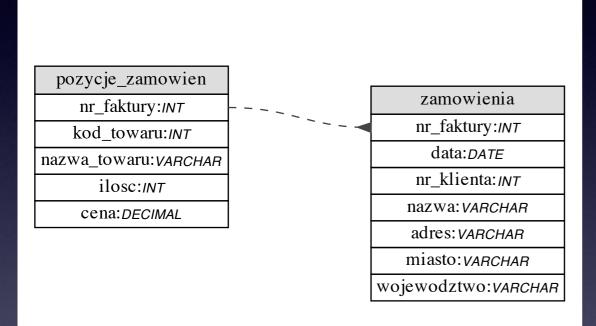
nr_faktury:///
data:DATE

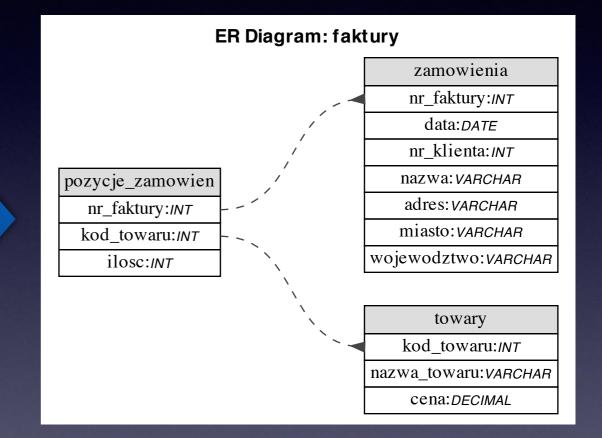
nr_klienta:///
nazwa:VARCHAR

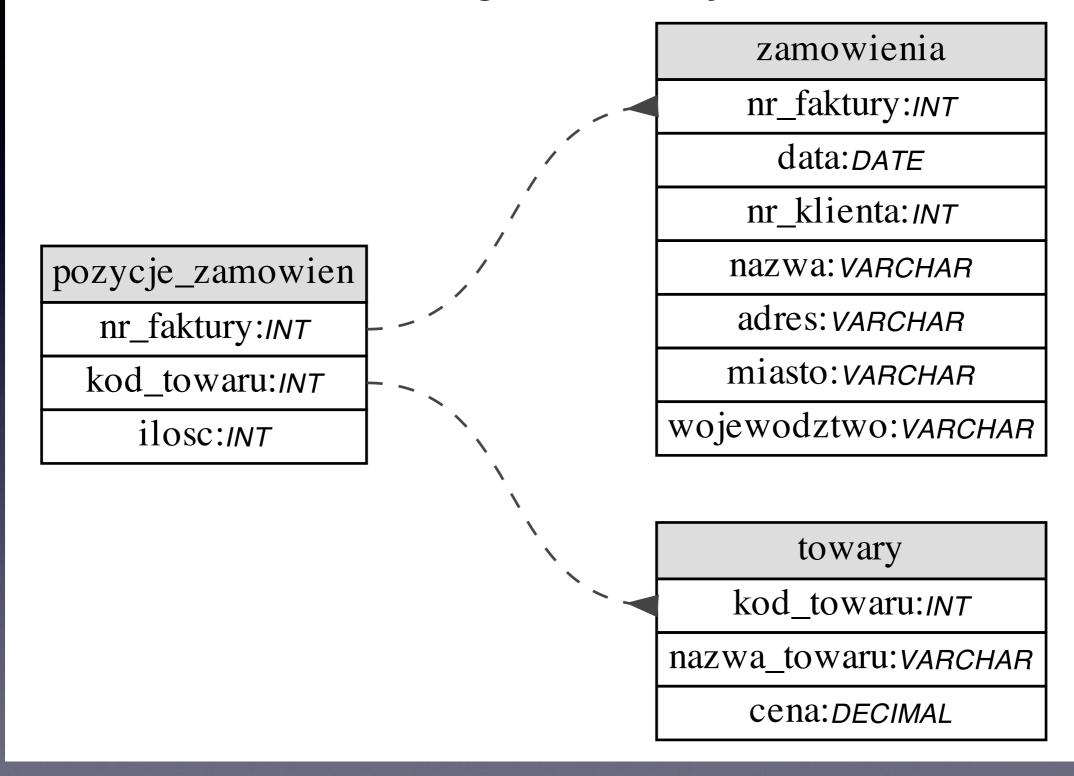
adres:VARCHAR

miasto:VARCHAR

wojewodztwo:VARCHAR







2NF

3NF

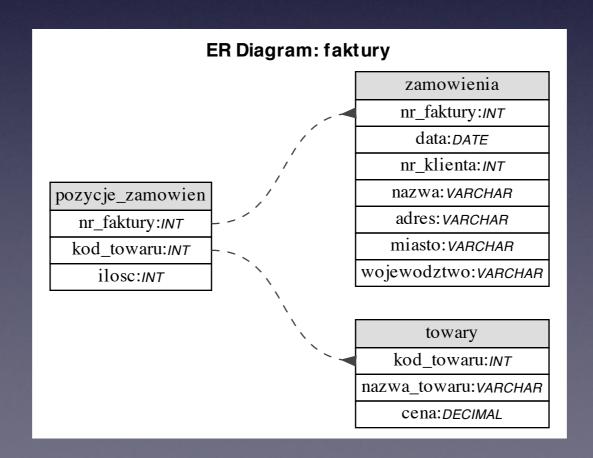
INF 2NF

3NF

3NF:

Brak zależności od atrybutów niekluczowych.

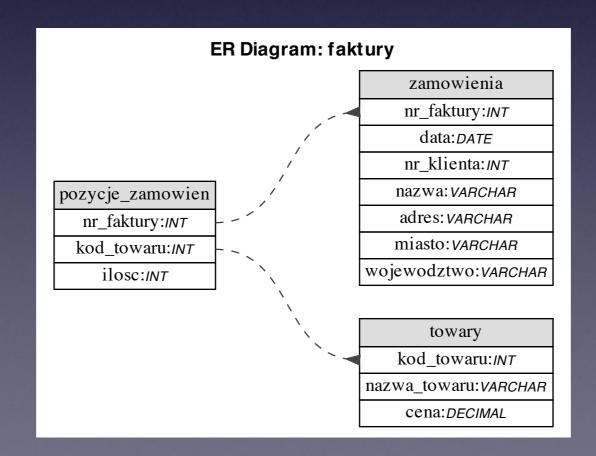
- Relacja w INF i 2NF.
- Żaden atrybut spoza klucza głównego nie jest od niego tranzytywnie zależny.
- Czyli: jest zależny bezpośrednio.

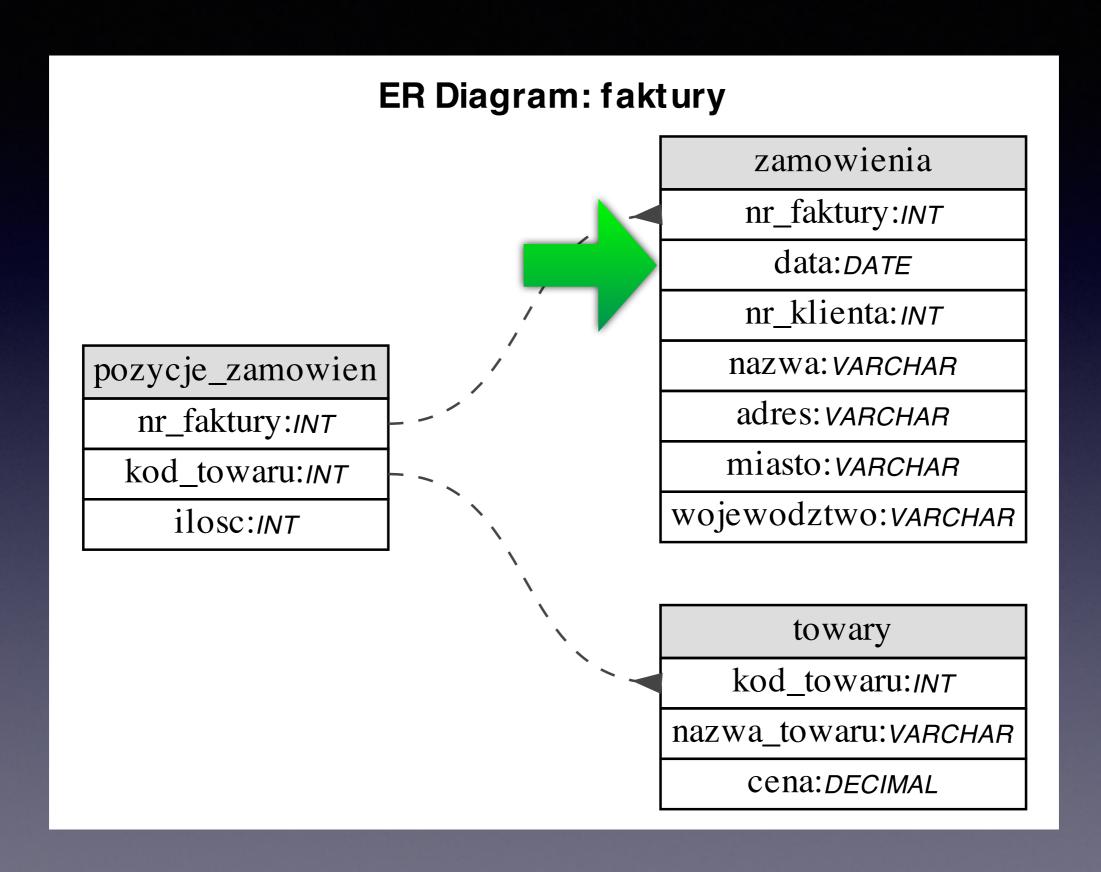


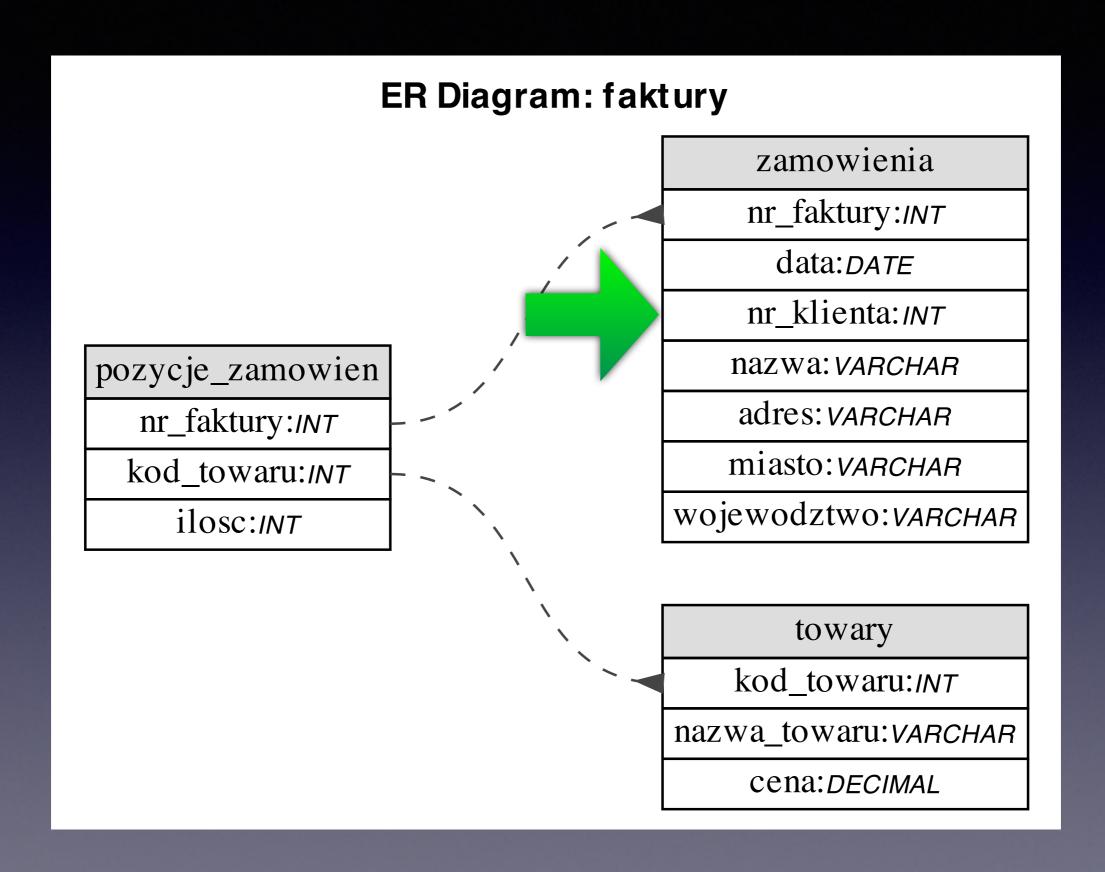
3NF:

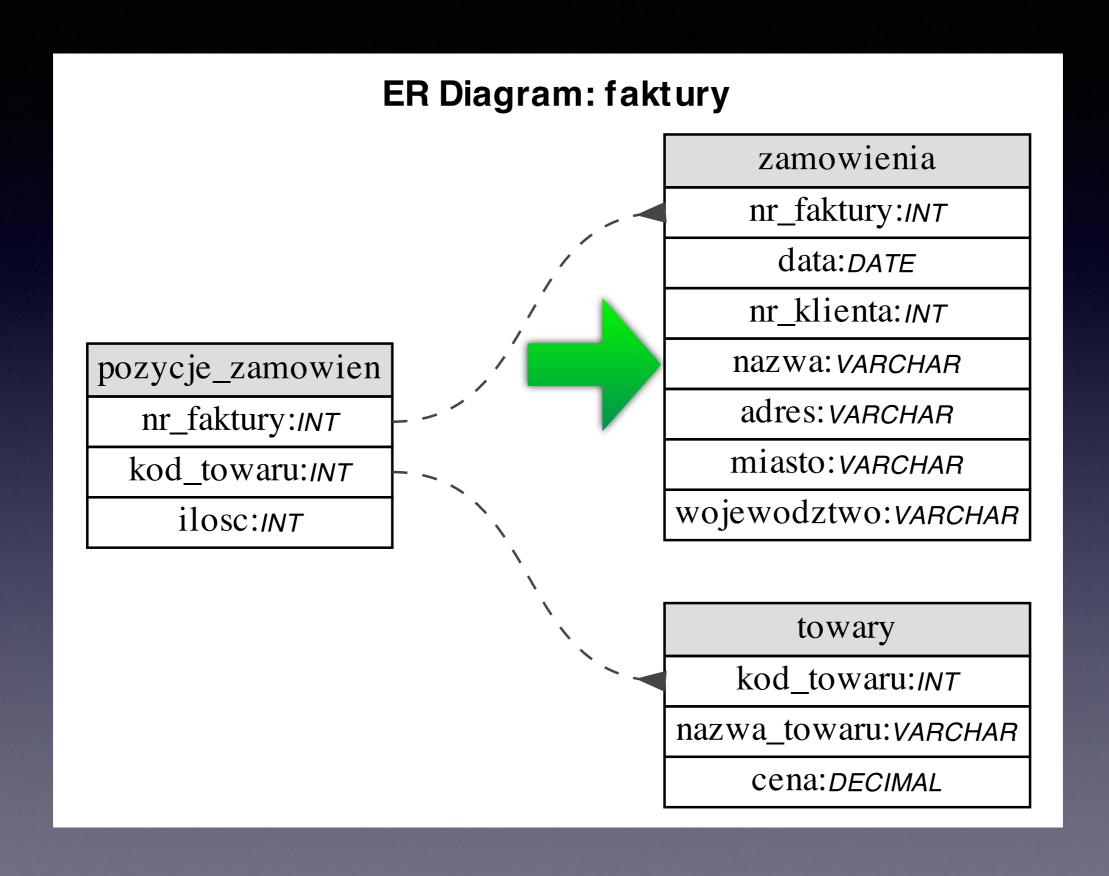
Brak zależności od atrybutów niekluczowych.

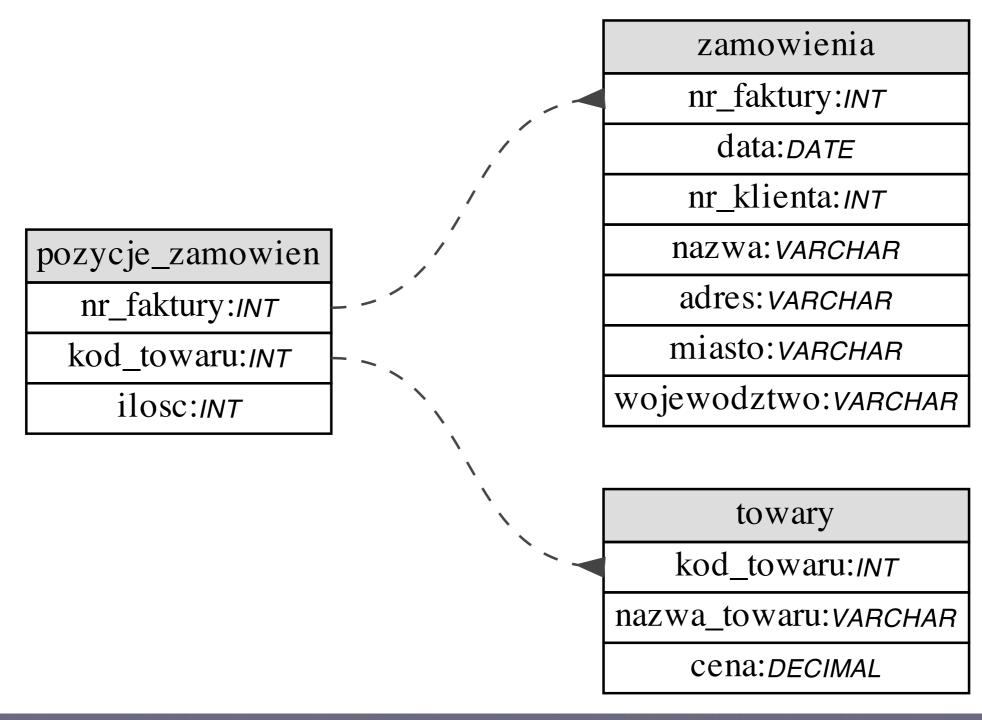
- "Czy ta kolumna jest zależna od innej kolumny niekluczowej?"
- 3NF ma na celu ekstrakcję danych o samodzielnych encjach do innej tabeli.

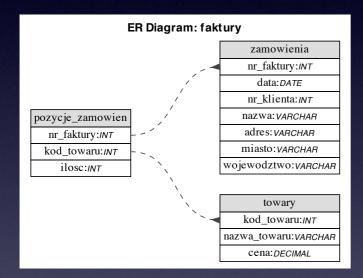




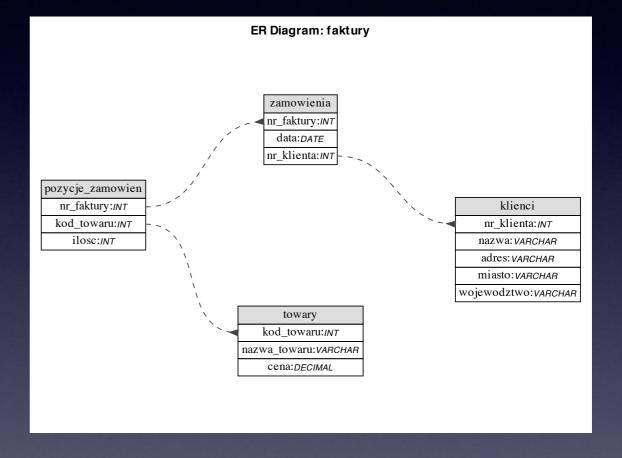


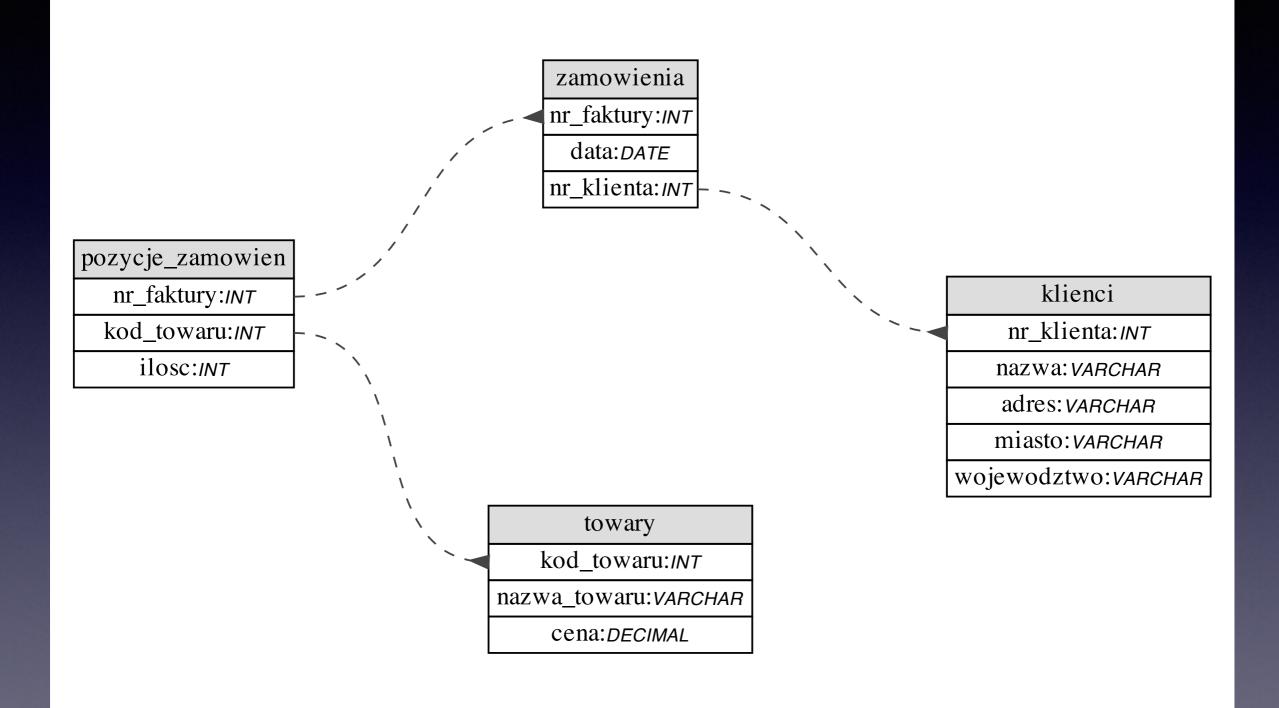


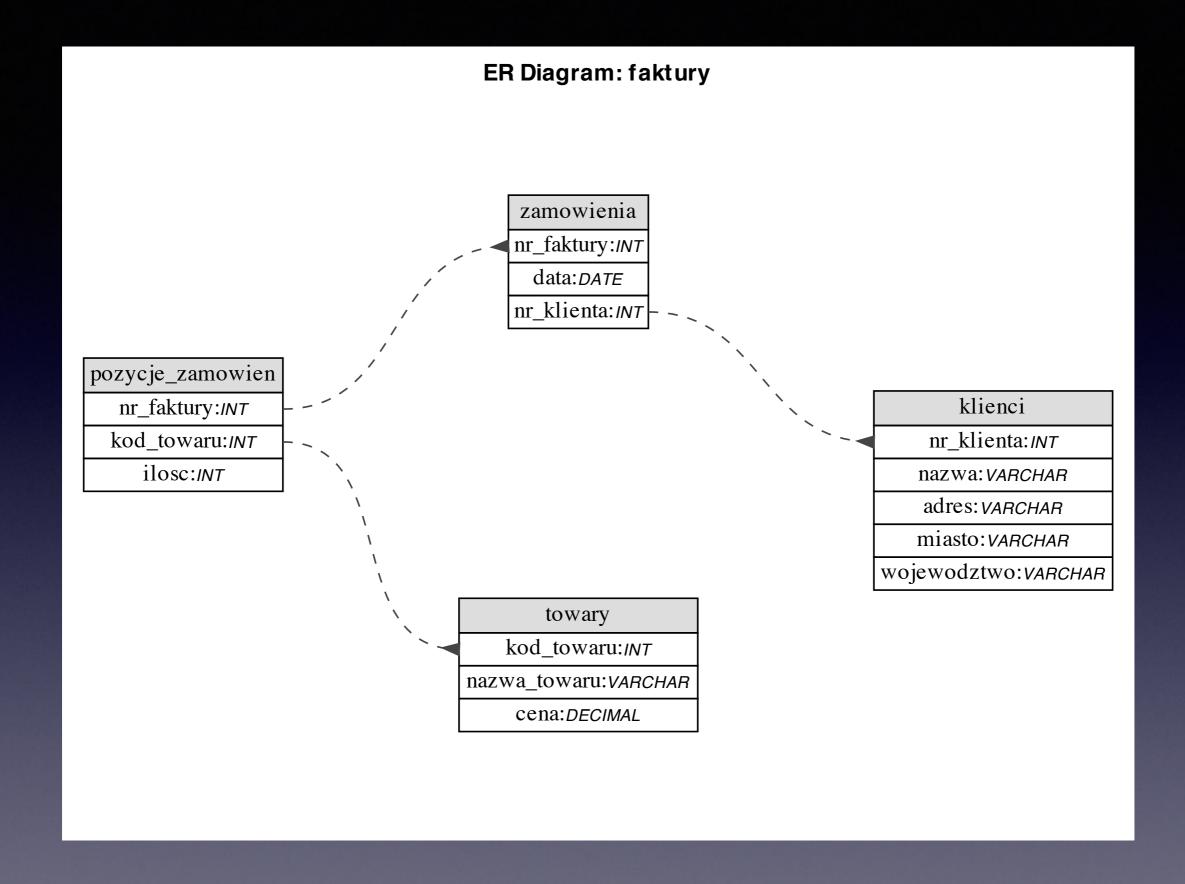












INF 2NF

3NF

2NF 3NF

- Boyce-Codd Normal Form.
- Każda nietrywialna zależność jest od klucza (niekoniecznie głównego).
- Nie istnieją w ogóle zależności tranzytywne.

- Zazwyczaj tabela w 3NF również spełnia BCNF.
- Tabela w 3NF nie posiadająca wielu nakładających się kluczy kandydujących zawsze będzie w BCNF.

Today's Court Bookings

Court	Start Time	End Time	Rate Type
1	09:30	10:30	SAVER
1	11:00	12:00	SAVER
1	14:00	15:30	STANDARD
2	10:00	11:30	PREMIUM-B
2	11:30	13:30	PREMIUM-B
2	15:00	16:30	PREMIUM-A

Today's Court Bookings

Court	Start Time	End Time	Rate Type
1	09:30	10:30	SAVER
1	11:00	12:00	SAVER
1	14:00	15:30	STANDARD
2	10:00	11:30	PREMIUM-B
2	11:30	13:30	PREMIUM-B
2	15:00	16:30	PREMIUM-A

- SAVER, for Court I bookings made by members
- STANDARD, for Court I bookings made by non-members
- PREMIUM-A, for Court 2 bookings made by members
- PREMIUM-B, for Court 2 bookings made by nonmembers

Today's Court Bookings

Court	Start Time	End Time	Rate Type							
1	09:30	10:30	SAVER							
1	11:00	12:00	SAVER							
1	14:00	15:30	STANDARD							
2	10:00	11:30	PREMIUM-B							
2	11:30	13:30	PREMIUM-B							
2	15:00	16:30	PREMIUM-A							

- SAVER, for Court I bookings made by members
- STANDARD, for Court 1 bookings made by nonmembers
- PREMIUM-A, for Court 2 bookings made by members
- PREMIUM-B, for Court 2 bookings made by nonmembers

{Court, Start Time} {Court, End Time} {Rate Type, Start Time} {Rate Type, End Time}

Today's Court Bookings

Court	Start Time	End Time	Rate Type
1	09:30	10:30	SAVER
1	11:00	12:00	SAVER
1	14:00	15:30	STANDARD
2	10:00	11:30	PREMIUM-B
2	11:30	13:30	PREMIUM-B
2	15:00	16:30	PREMIUM-A

BCNF

Today's Court Bookings

Court	Start Time	End Time	Rate Type
1	09:30	10:30	SAVER
1	11:00	12:00	SAVER
1	14:00	15:30	STANDARD
2	10:00	11:30	PREMIUM-B
2	11:30	13:30	PREMIUM-B
2	15:00	16:30	PREMIUM-A



Rate Types

Rate Type	Court	Member Flag
SAVER	1	Yes
STANDARD	1	No
PREMIUM-A	2	Yes
PREMIUM-B	2	No

Today's Bookings

Rate Type	Start Time	End Time
SAVER	09:30	10:30
SAVER	11:00	12:00
STANDARD	14:00	15:30
PREMIUM-B	10:00	11:30
PREMIUM-B	11:30	13:30
PREMIUM-A	15:00	16:30

BCNF

 Nie każdy schemat w 3NF da się rozłożyć do BCNF.

Przykład: 3NF, ale nie BCNF

Miasto Ulica Kod pocztowy

- Klucze: {M, U}, {U, K}
- Zależności:
 MU→K, K→M.
- Występują anomalie: usuwając ulicę możemy utracić informację o mieście.

- 3NF tak: brak atrybutów niekluczowych.
- BCNF nie:
 K nie jest kluczem.
- Relacja nierozkładalna do BCNF bez utraty zależności.

4NF

- Nowy typ zależności: zależności wielowartościowe.
- 4NF jest spełniona jeżeli: dla każdej zależności wielowartościowej X→→Y,
 X jest nadkluczem.
- Zbiór atrybutów Y jest zależny wielowartościowo od zbioru X gdy z każdą konfiguracją wartości atrybutów z X jest związany zbiór konfiguracji wartości z Y niezależnie od wartości pozostałych atrybutów.

Zależność wielowartościowa

Zajęcia	Wykładowca	Podręcznik
BD2	Parewicz	Ullman
BD2	Traczyk	Ullman
BD2	Parewicz	Barker
BD2	Traczyk	Barker
BDA	Parewicz	Date
BDA	Kowalski	Date

- Podręczniki do danego przedmiotu nie zależą od wykładowców.
- Nie ma zależności funkcyjnych.
- Zależności: Z→→W,
 Z→→P.
- Można zamienić części krotek
 ZW i P.

4NF

Zajęcia	Wykładowca	Podręcznik
BD2	Parewicz	Ullman
BD2	Traczyk	Ullman
BD2	Parewicz	Barker
BD2	Traczyk	Barker
BDA	Parewicz	Date
BDA	Kowalski	Date

Dla każdej zależności wielowartościowej X→→Y, X jest nadkluczem.

- Nie ma zależności funkcyjnych jest BCNF.
- Zależności: $Z \rightarrow \rightarrow W$, $Z \rightarrow \rightarrow P$; Z = 1 nadkluczem.
- Nadmiar informacji: powtórzone dane podreczników i wykładowców.
- Aby doprowadzić do 4NF rozkładamy na ZW i ZP.

4NF

- Nauka baz danych kończy się zazwyczaj po BCNF.
- Wynika to z przekonania, że w większości rzeczywistych aplikacji tabele niespełniające 4NF zdarzają się rzadko.
- W analizie wykonanej w 1992 roku, ponad 20% aplikacji zawierało jedną lub więcej tabel niespełniających 4NF, a spełniających BCNF.
 - źródło:Wu, M. S. (1992). The practical need for fourth normal form. ACM SIGCSE Bulletin, 24(1), 19-23. doi: 10.1145/135250.134515.

Algorytm normalizacji

- Identyfikacja zależności funkcyjnych.
- Generacja wszystkich kluczy, podział atrybutów na kluczowe i niekluczowe.
- Sprawdzenie 2NF: czy są atrybuty niekluczowe zależne od fragmentu klucza?
- Sprawdzenie 3NF: czy są atrybuty niekluczowe zależne tranzytywnie od klucza?
- Sprawdzenie BCNF: czy są zależności nie od klucza?

Rozpoznawanie 2NF i 3NF

- 2NF warunki wystarczające:
 - istnieją tylko klucze proste,
 - wszystkie atrybuty kluczowe.
- 3NF warunki wystarczające:
 - wszystkie atrybuty kluczowe,
 - brak zależności tranzytywnych.

Zasady projektowania

- Jeden obiekt jedna tablica.
- Atrybuty != obiekty.
- Atomiczność atrybutów.
- Atrybuty tylko raz, przy "swoim" obiekcie.
- Encje asocjacyjne osobne tablice.
- Dekompozycja M:N do I:N.
- Eliminacja redundancji.

Denormalizacja

- Optymalizacja wydajności operacji odczytu przez wprowadzenie danych nadmiarowych lub grup danych.
- Baza zdenormalizowana to nie to samo, co baza, której nie poddano normalizacji.

Denormalizacja

- Widoki zmaterializowane, m.in.:
 - przechowywanie liczników "N"
 obiektów jako atrybutów tabeli "I" w
 relacjach I:N,
 - dodawanie atrybutów z innej tabeli.
- Schematy gwiaździste.
- Prekompilacja kostek OLAP.