

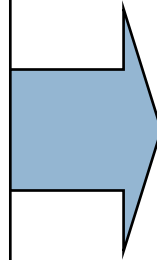
WYKŁAD 12: OLAP

Plan

- Inteligencja biznesowa
(*Bussiness Intelligence*)
- Hurtownia danych
- OLAP

Motywacja:

- **Zaawansowane metody ekstrakcji danych i techniki przechowywania danych**
- **Rozwój wielu dziedzin zastosowań**



- **Więcej danych:**
 - Bank, telecom, inne transakcje biznesowe ...
 - Dane naukowe: astronomia, biologia, ...
 - Web, tekst, oraz e-commerce

Ogromne bazy danych

- **Duża liczba rekordów:**
 - 10^6 - 10^{12} w przypadku baz danych o obiektach niebieskich (astronomia)
- **Duża liczba atrybutów** (cechy, pomiary, kolumny):
 - Setki zmiennych opisujących medyczne pomiary pacjentów

Motywacja

- „Jesteśmy zatopieni w morzu danych, podczas gdy pragniemy wiedzę”
- **PROBLEM:**
Jak zdobyć użyteczną wiedzę (informację) z dużych baz danych?
- **ROZWIĄZANIE:**
 - **Hurtownie danych:** Zbieranie danych (w czasie rzeczywistym)
 - **OLAP:** Przetwarzanie analityczne
 - **Data mining:** Odkrywanie interesującej wiedzy (reguł, regularności, wzorców, modeli) z dużych zbiorów danych



Ewolucja w technologii baz danych

- W latach 60-tych:
 - ▣ Kolekcja danych, tworzenia baz danych, IMS oraz sieciowe DBMS
- W latach 70-tych:
 - ▣ Relacyjny model danych, implementacja relacyjnych DBMS
- W latach 80-tych:
 - ▣ RDBMS, zaawansowane modele danych (extended-relational, OO, deductive, ...) oraz aplikacyjno-zorientowane DBMS
- Od 90-tych — obecnie:
 - ▣ Data mining, hurtownia danych, multimedialne bazy danych oraz „Web databases”

Business Intelligence

□ Co to jest Business Intelligence?

szeroka dziedzina obejmująca aplikacje i technologie służące gromadzeniu i analizie danych w celu wspomagania procesu podejmowania decyzji biznesowych.

□ Skuteczność działania w obszarze Business Intelligence wymaga dogłębnej znajomości wszystkich czynników wpływających na biznes.

□ Efektem zastosowania narzędzi BI jest dostępność do szybkiej informacji na temat najważniejszych wskaźników firmy, takich jak dane o klientach, konkurencji, partnerach biznesowych, sytuacji ekonomicznej i operacjach wewnętrznych.

Obszary związane z BI

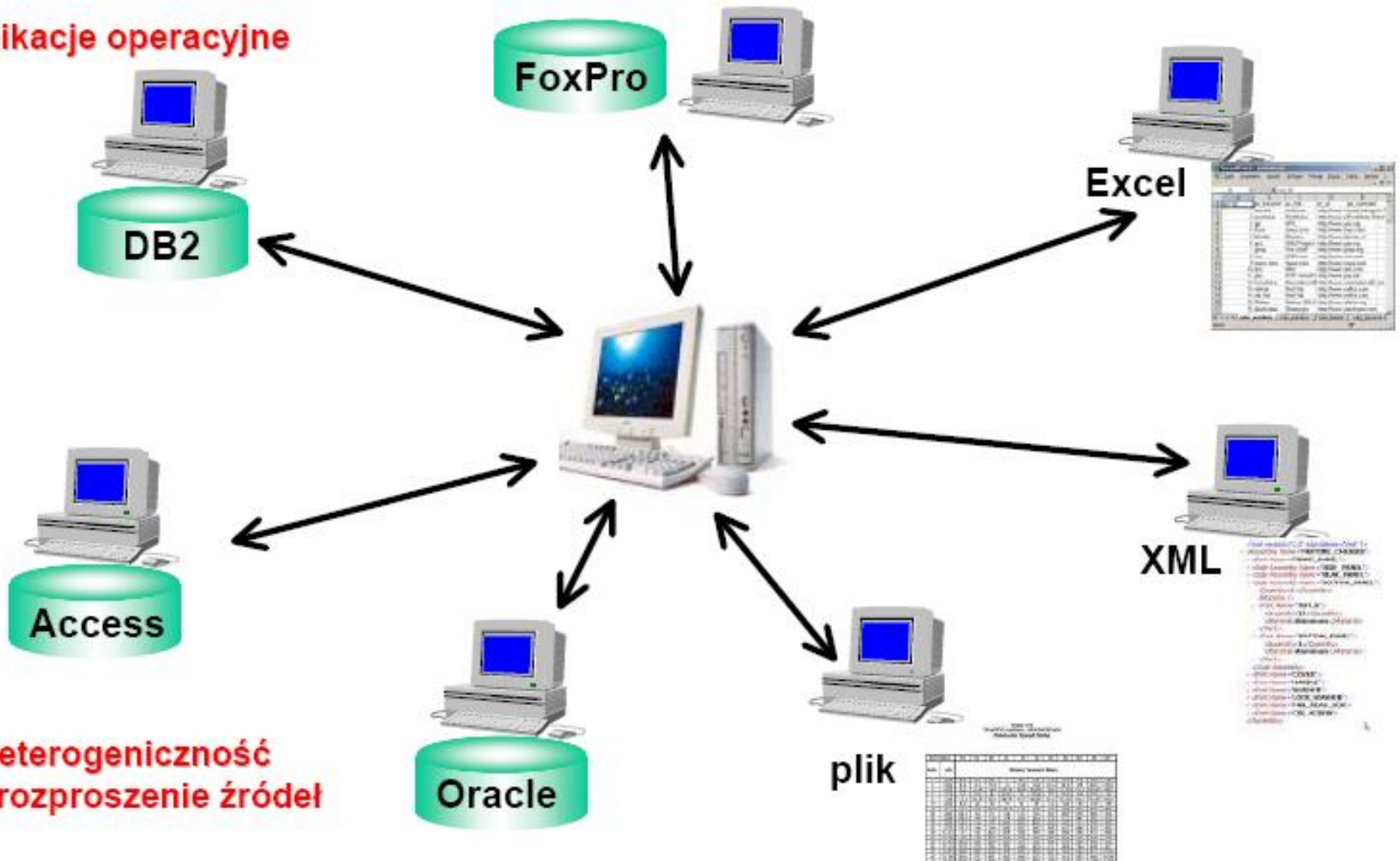
- **DW** - Hurtownie danych – ładowanie, przetwarzanie
- **Data mining** - Eksploracja danych, drążenie danych
- **OLAP** - Online Analytical Processing
- Czyszczenie danych i zarządzanie jakością danych
- **MIS**(Management Information Systems) - Systemy Informowania Kierownictwa
- **Raportowanie** - Wizualizacja informacji i panele dla kierownictwa
- Prognozowanie, finanse i budżetowanie
- Statystyki i techniczna analiza danych
- **CRM**(Customer Relationship Management) – Zarządzanie Relacjami z Klientami
- **DSS**(Decision Support Systems) – systemy wspomaganie decyzji

Problem Integracja danych

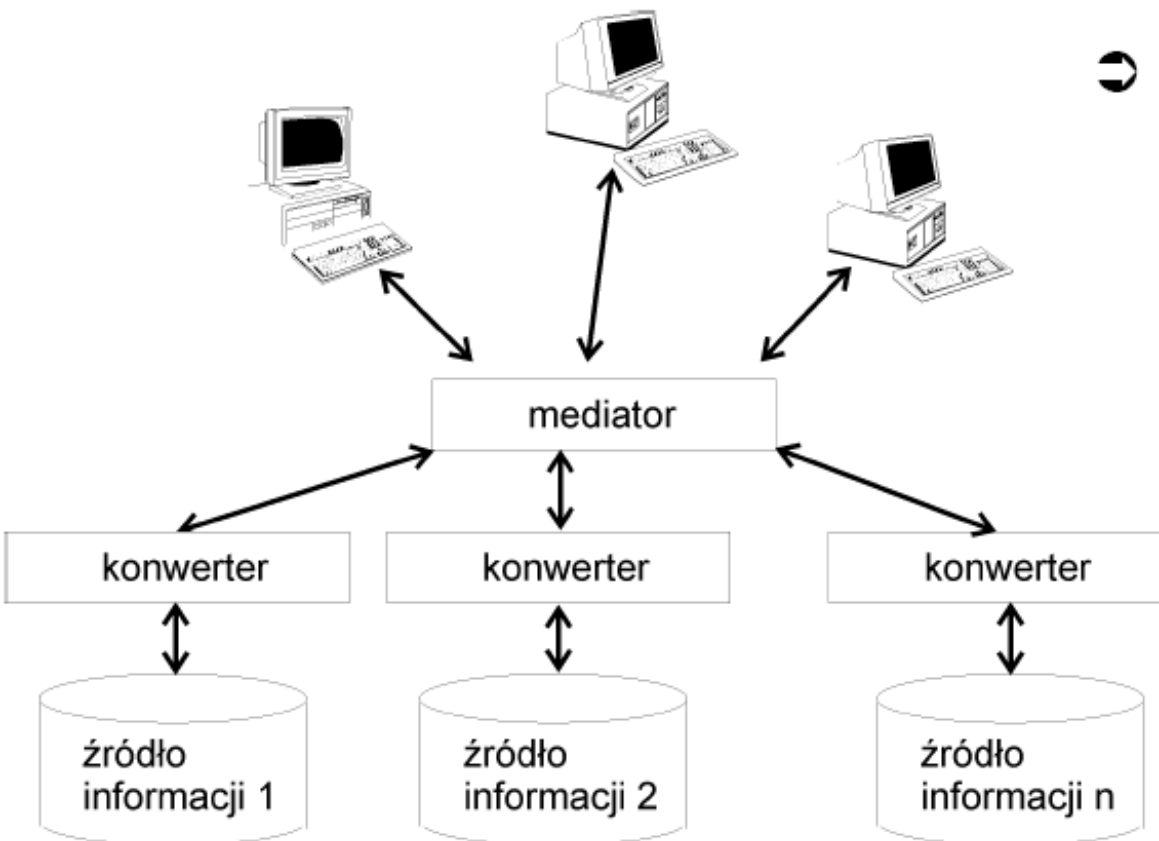
- **Różni producenci/technologie**
- **Różna funkcjonalność**
 - bazy danych / nie bazy danych
 - dialekty SQL
 - sposoby dostępu i przetwarzania danych
- **Różne modele danych**
 - hierarchiczne, sieciowe
 - relacyjne
 - obiektowe
 - obiektowo-relacyjne
 - wielowymiarowe
 - semistrukturalne
- **Architektury integracyjne**
 - system mediacyjny
 - hurtownia (magazyn) danych

Problem integracja danych

aplikacje operacyjne



Rozwiązanie 1: System mediacyjny



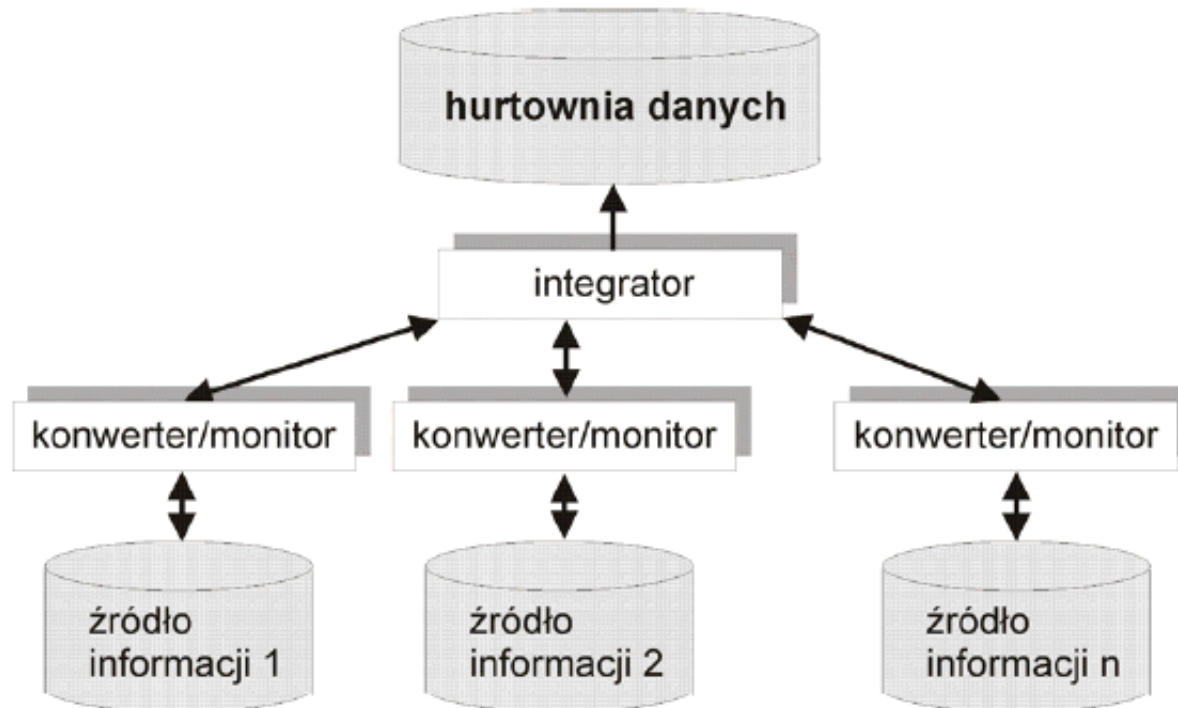
➤ Wady

- czas dostępu do danych
- niedostępność źródeł
- konwersja zapytań i danych

➤ Zalety

- brak redundancji danych
- dostęp do danych aktualnych

Rozwiązanie 2: Hurtownia danych



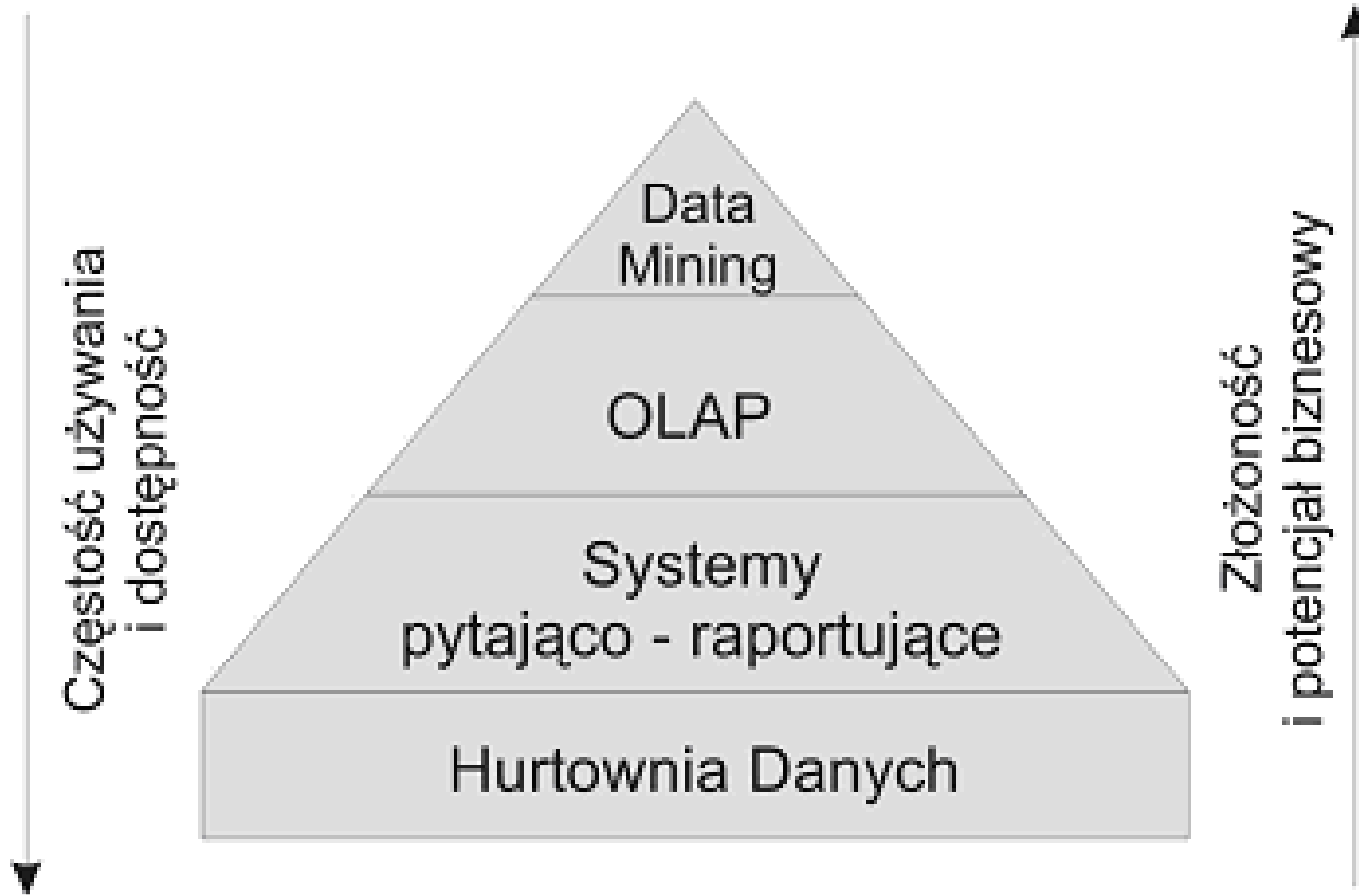
➤ Zalety

- dane zintegrowane (spójna struktura i wartości)
- szybkość dostępu do danych
- niezależność od awarii źródeł

➤ Wady

- redundancja danych
- odświeżanie danych

Pyramida BI



HURTOWNIA DANYCH



Co to jest hurtownia danych

- Hurtownia danych powinna być miejscem przechowywania

- historycznych,
- "nieulotnych",
- zorientowanych tematycznie,
- zintegrowanych

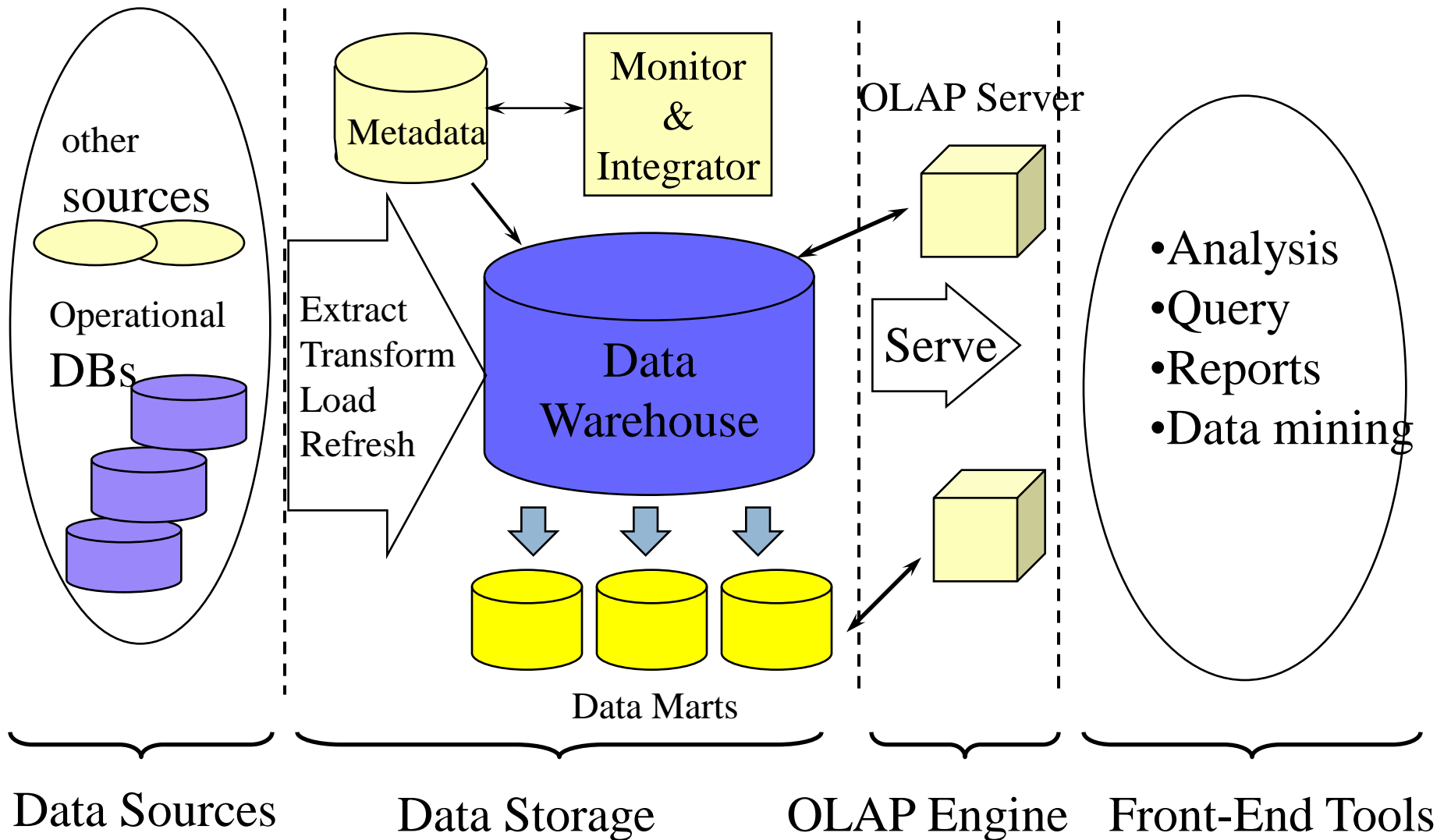
danych pochodzących z różnych rozproszonych baz źródłowych.

- Jej struktura projektowana jest niezależnie od struktury danych źródłowych, a dane przetwarza się na podstawie różnych aplikacji analitycznych.
- Przetwarzanie danych w hurtowni danych na ogół przebiega wielowarstwowo. Oznacza to, że dane najpierw podlegają czyszczeniu, standaryzacji, a następnie sklejanii i agregowaniu.
- W praktyce najczęściej wdrażane są funkcjonalne elementy hurtowni (*data marts*) dla poszczególnych segmentów działalności organizacji, a dopiero później podlegają one łączeniu w całość.

Hurtownia danych – tematycznie zorientowana

- Organizowana **tematycznie**, np. klient, produkt, sprzedaż.
- Zorientowana na modelowaniu i analizie danych dla kadr kierowniczych (podejmujących decyzję)
Nie jest przydatne dla codziennego użytku
- **Data mart**: Prosty i zwięzły widok na poszczególne tematy (aspekty) poprzez usunięcie danych zbędnych dla procesu podejmowania decyzji

Struktura tematyczna



Dane rozległe w czasie

- Horyzont czasowy w hurtowni danych jest zdecydowanie dłuższy niż w bazach danych operacyjnych:
 - **Bazy danych operacyjnych:** utrzymuje bieżące dane
 - **Hurtownia danych:**
 - udziela informacje z pewnej historycznej perspektywy, np. z ostatnich 5-10 lat.
Every key structure in the data warehouse
 - Hurtownia może zawierać różne aspekty czasowe: explicite (jawne) lub implicite (niejawne) o danych mimo, że operacyjne dane mogą nie zawierać informacji czasowych

Cele hurtowni danych

- 1. Zapewnienie jednolitego dostępu do wszystkich danych gromadzonych w ramach przedsiębiorstwa**
- 2. Dostarczenie technologii (platformy) przetwarzania analitycznego – technologii OLAP**
 - ▣ **wykonywanie zaawansowanych analiz, wspomagających zarządzanie przedsiębiorstwem, np.**
 - analiza trendów sprzedaży
 - analiza nakładów reklamowych i zysków
 - analiza ruchu telefonicznego
 - ▣ **Eksploracja danych (Data mining)**
 - ▣ **analiza rozwiązań alternatywnych (what-if analysis)**
 - symulowanie i przewidywanie przyszłości w MD

OLAP

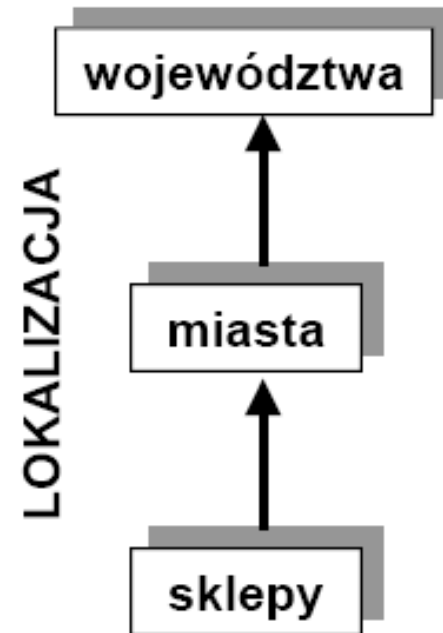
Online Analytical Processing

Co to jest OLAP

- **Zadanie:** dostarczanie informacji strategicznej i prezentowanie jej zgodnie ze schematem poznawczym człowieka.
- **Narzędzia:** bazy danych + zaawansowane modele matematyczne.
- Podstawą modelu OLAP są
 - **Fakty**
 - **informacje podlegające analizie**
 - sprzedaż, rozmowy telefoniczne
 - charakteryzowane ilościowo za pomocą miar
 - Liczba sprzedanych jednostek towaru, czas trwania rozmowy
 - **Wymiary**
 - **ustalają kontekst analizy**
 - sprzedaż czekolady (produkt) w Auchan(sklep)
 - w poszczególnych miesiącach roku (czas)
 - **składają się z poziomów, które tworzą hierarchię**

Hierarchia pojęć

- Konceptyjnie model OLAP można przedstawić jako hiperkostkę, która w swoim wnętrzu zawiera miary, natomiast wymiary stanowią jej brzegi.
- Najczęściej wszelkie analizy danych dotyczą różnych poziomów szczegółowości, dlatego wymiary posiadają wewnętrzną strukturę, ułatwiającą przechodzenie od ogółu do szczegółu.



Hurtownia danych a ODBMS

- **OLTP (on-line transaction processing)**
 - ▣ Utrzymuje informacje o codziennych operacjach: sprzedaż, remanent, bankowość, produkcja, lista płac, rejestracja, księgowość, itp..
 - ▣ Jest to główne zastosowania tradycyjnych systemów zarządzania bazami danych (relational DBMS)
- **OLAP (on-line analytical processing)**
 - ▣ Analiza danych i wspomaganie podejmowania decyzji
 - ▣ Jest głównym zadaniem dla hurtowni danych.

OLTP a OLAP

	OLTP	OLAP
użytkownik	"zwykły"	analityk
funkcja	bieżące operacje, kluczowe dla działania firmy	wspomaganie decyzji
dane	bieżące, elementarne	elementarne, zagregowane, historyczne
aplikacje	powtarzalność działań	ad hoc
dostęp	odczyt/zapis	odczyt
transakcja	krótka	długa (godziny)
l. przetwarzanych rek.	kilka, kilkadziesiąt	miliony lub więcej
l. użytkowników	kilkudzies., tysiące, setki tys.	kilku, kilkunastu
DB size	setki GB	dziesiątki TB
metric	przepustowość (l. transakcji w jednostce czasu)	czas odpowiedzi

Hurtownia danych a OLAP

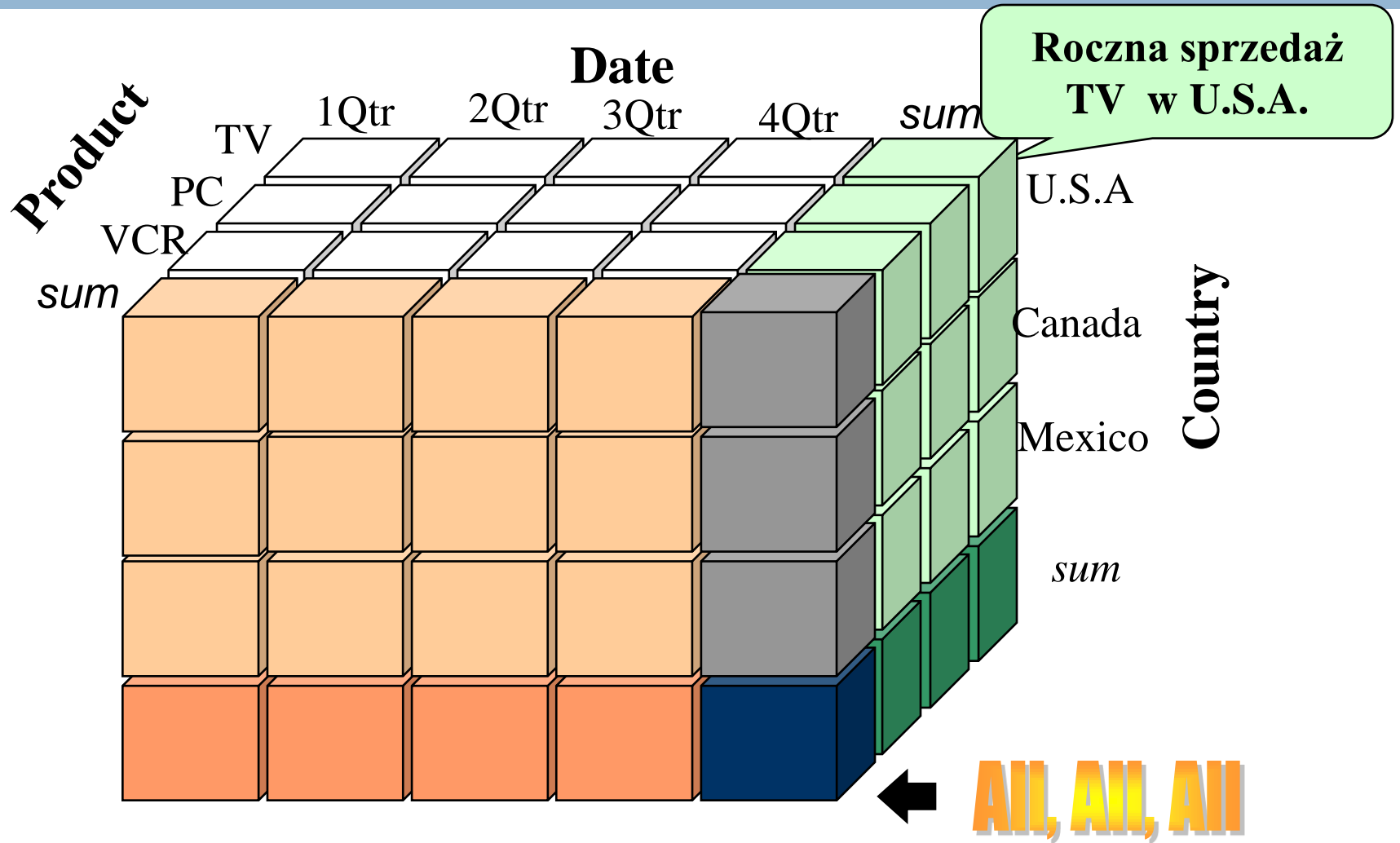
Dlaczego są to oddzielne systemy?

- **Wysoka wydajność dla obu dwóch systemów**
 - ▣ DBMS — ustawiony dla OLTP: metody dostępu, indeksowania, sterowania współbieżności, odzyskiwania
 - ▣ Hurtownia — ustawiony dla OLAP: złożone zapytania OLAP'owe, wielowymiarowe widoki, konsolidacja.
- **Różne funkcje i różne dane:**

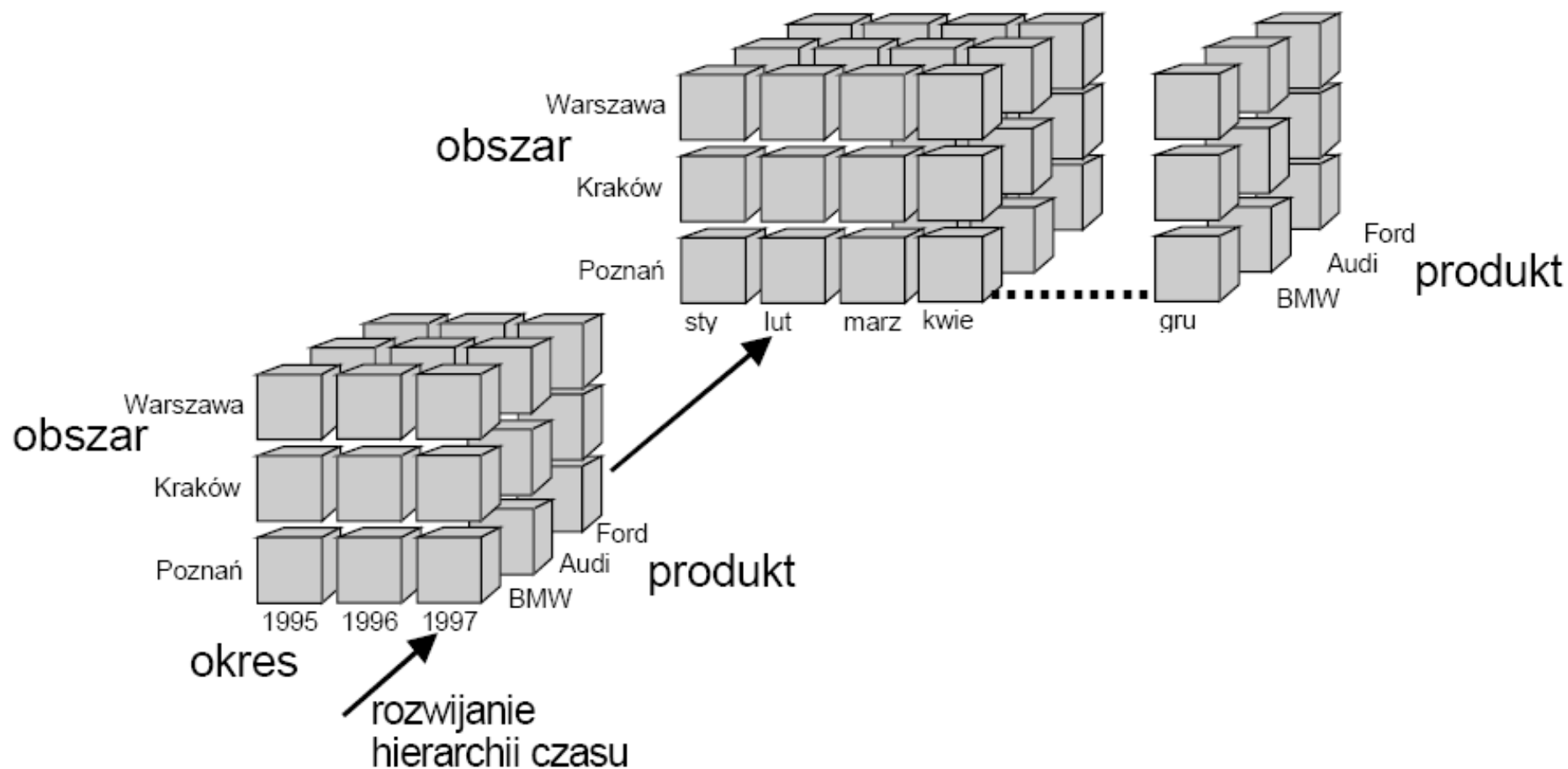
Typowe operacje OLAP

- **Roll up (drill-up):** podsumowanie
 - ▣ *Przejdźcie do wyższego poziomu w hierarchii lub redukcja wymiarów*
- **Drill down (roll down):** rozwinięcie (odwrotnie do roll-up)
 - ▣ *Przejdźcie do niższego poziomu w hierarchii lub wprowadzanie nowych wymiarów*
- **Slice and dice:**
 - ▣ *Rzut i selekcja*
- **Pivot (rotate):**
 - ▣ *Zmiana orientację kostki, wizualizacja,*
- **INNE:** *drill across, drill through*

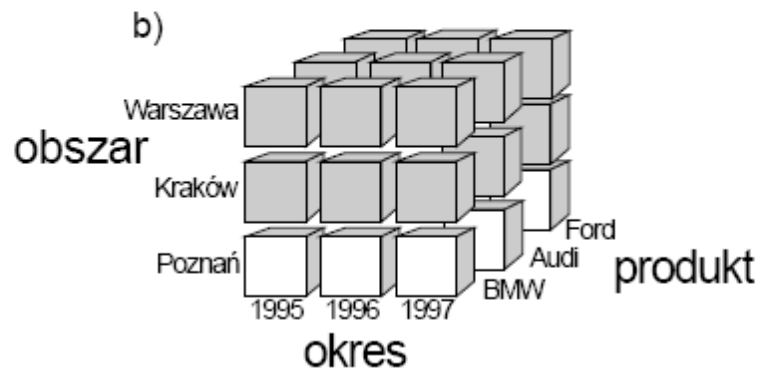
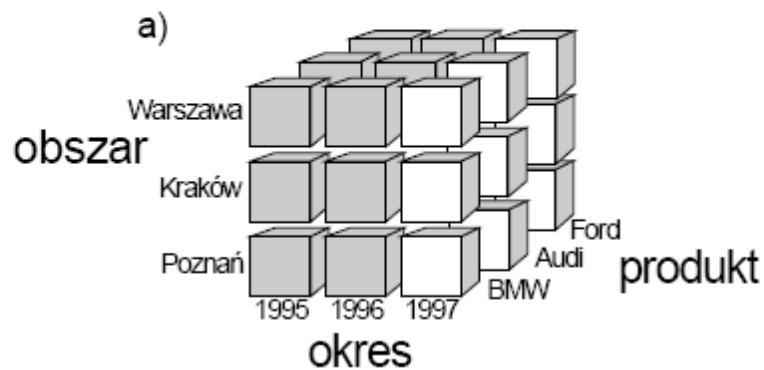
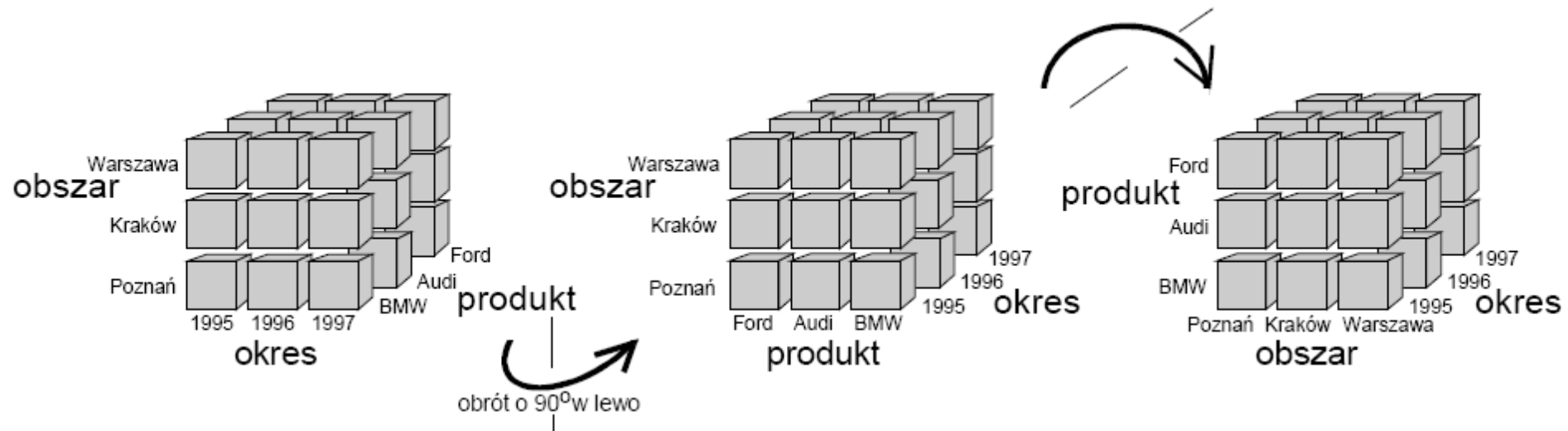
Przykład hyperkostki



Drill-down/ Roll-up



Rotating, slicing and dicing



Różne modele OLAP

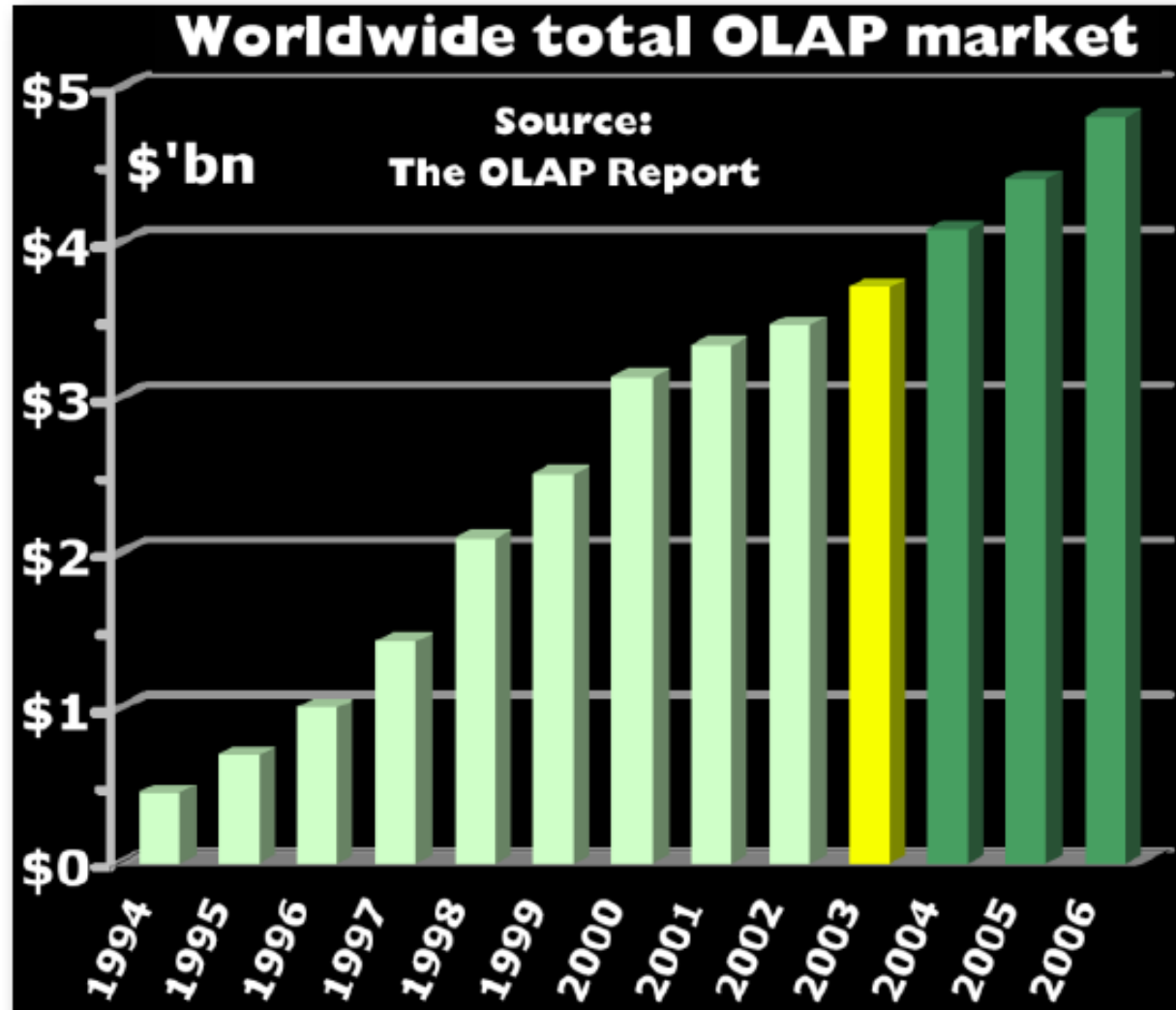
- **Model relacyjny (ROLAP)**
 - **schemat gwiazdy (ang. starschema)**
 - **schemat płatka śniegu (ang. snowflake schema)**
 - **schemat konstelacji faktów (ang. fact constellation schema)**
 - **schemat gwiazda-płatek śniegu (ang. starflake schema)**
- **Model wielowymiarowy (MOLAP, MDOLAP)**
- **Model hybrydowy (HOLAP)**

Narzędzia

- ❑ **DB2 DataWarehouse Center,**
- ❑ **Sybase WarehouseStudio,**
- ❑ **Microsoft DataWarehousing Framework,**
- ❑ **SAP Datawarehouse management,**
- ❑ **NCRTeradata Warehouse Builder,**
- ❑ **Oracle Designer6i/9i**

Globalny rynek OLAP

- Przewidywane w roku 2003



Rozwój technologii OLAP

Błyskawicznie
rozwijający się rynek
badawczy i
technologiczny

Przewidywano:

9.9 * 10⁹\$ w 2008

(METAGROUP)

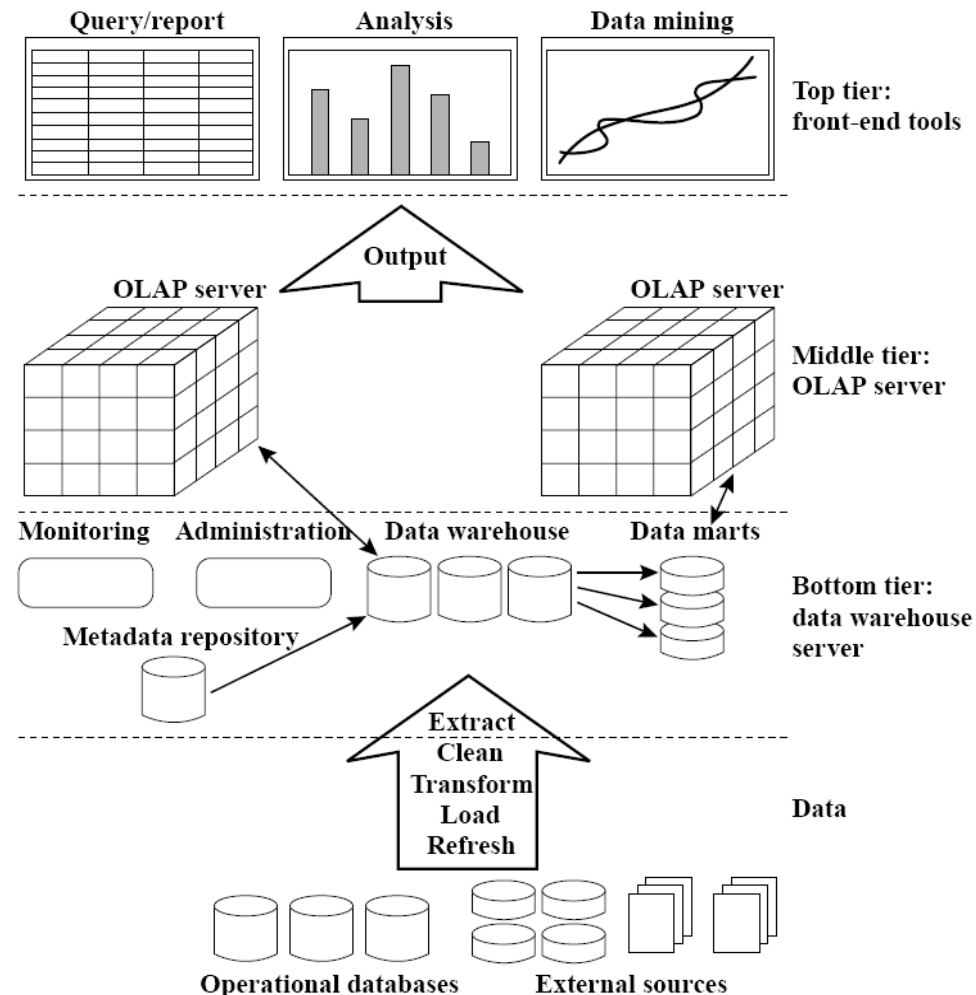


PROJEKTOWANIE HURTOWNI DANYCH NA POTRZEBĘ OLAP

- Modele wielowymiarowe
- Realizacje operacji OLAP

Motywacje

- **Cele:**
 - Wzrost sprzedaży
 - Zwiększenie udziału firmy w rynku, ...
- **Jak to robić:**
 - analiza zachowania klientów,
 - lokalizacja słabych punktów,
 - śledzenie trendów.
- **Problemy techniczne:**
 - Jakimi danymi dysponujemy?
 - Które dane powinniśmy gromadzić w hurtowni?
 - Wszystkie?
 - Tylko to, co niezbędne?
 - Jak schować dane w hurtowni?



Model wielowymiarowy

Baza zawiera fakty opisane przez wymiary i określające wartość miar.

- **Fakt** - pojedyncze zdarzenie będące podstawą analiz (np. sprzedaż).
Fakty opisane są przez wymiary i miary.
- **Wymiar** - cecha opisująca dany fakt, pozwalająca powiązać go z innymi pojęciami modelu przedsiębiorstwa (np. klient, data, miejsce, produkt).
Wymiary są opisane atrybutami.
- **Atrybut** - cecha wymiaru, przechowująca dodatkowe informacje na temat faktu (np. wymiar data może mieć atrybuty: miesiąc, kwartał, rok; wymiar klient może mieć atrybuty: nazwisko, region).
- **Miara** - wartość liczbowa przyporządkowana do danego faktu (np. wartość sprzedaży, liczba sztuk).

ROLAP

- Relacyjna implementacja dla OLAP;
- Problemy:
 - ▣ Zidentyfikowanie faktów
 - ▣ Zidentyfikowanie kluczowych wymiarów
 - ▣ Zaprojektowanie tabel faktów
 - ▣ Zaprojektowanie tabel wymiarów

ROLAP – zidentyfikowanie faktów

Należy wskazać kluczowe typy transakcji w systemie produkcyjnym, realizujące kluczowe akcje/operacje w obszarze działania przedsiębiorstwa, np.

- **handel:** transakcje sprzedaży
- **bankowość:** kursy walut, operacje na rachunkach
- **giełda:** wahania kursów akcji, operacje giełdowe
- **ubezpieczenia:** wykupienie polisy, zmiana warunków polisy, zgłoszenie szkody, wypłacenie odszkodowania
- **telekomunikacja:** zrealizowanie rozmowy przez abonenta, podłączenie telefonu, zawarcie umowy, zmiana abonamentu, płatności za abonament
- **opieka zdrowotna:** przyjęcie pacjenta do szpitala, forma leczenia, wynik leczenia

ROLAP – zidentyfikowanie wymiarów

Należy znaleźć kluczowe wymiary dla faktów
(określenie kontekstu analizy faktów)

- **handel:**
 - analiza sprzedaży w poszczególnych miastach i okresach czasowych
- **bankowość:**
 - Wahania kursów walut w poszczególnych dniach
 - analiza przyrostu ilości nowych rachunków w poszczególnych miesiącachz podziałem na rodzaje rachunków
- **giełda:**
 - Wahania kursów akcji poszczególnych firm w poszczególnych dniach
 - Ilość zawartych transakcji kupna lub sprzedaży w jednostce czasu i łączne kwoty tych operacji
- **ubezpieczenia:**
 - analiza przyrostu/spadku ilości polis poszczególnych rodzajów w miastach w poszczególnych miesiącach
- **telekomunikacja:**
 - analiza rozkładu czasu rozmów poszczególnych klientów w czasie doby

OLAP - Projektowanie tabeli faktów

- Poziom szczegółowości informacji → rozmiar tabeli faktów
 - ▣ rejestrowanie kwoty zakupu pojedynczego produktu
 - ▣ rejestrowanie sumarycznej kwoty zakupu całego koszyka
 - ▣ rejestrowanie sumarycznej kwoty zakupu w tygodniu
- Horyzont czasowy danych
 - ▣ jak długo przechowywać informacje na najwyższym poziomie szczegółowości?
 - ▣ opracowanie strategii agregowania danych starszych
 - **raporty roczne:** najczęściej wystarczają agregaty sumujące fakty z dokładnością do tygodnia
 - **raporty agregujące dane sprzed kilku lat:** najczęściej wystarczają agregaty sumujące fakty z dokładnością do miesiąca lub roku

OLAP - Projektowanie tabeli faktów

- Właściwy zbiór atrybutów tabeli faktów
- Usunięcie zbędnych atrybutów → rozmiar tabeli faktów
 - ▣ czy atrybut wnosi nową/niezbędną wiedzę o fakcie?
 - ▣ czy wartość atrybutu można wyliczyć?
- Minimalizacja rozmiarów atrybutów
 - ▣ **przykład: telekomunikacja**
 - tabela wymiaru Abonenci zawiera $8 \cdot 10^6$ abonentów
 - każdy abonent dzwoni średnio 2 razy dziennie
 - roczny horyzont czasowy tabeli faktów
 - zmniejszenie długości rekordów tabeli faktów o 10B zyskujemy 54GB

ROLAP – klucze podstawowe i sztuczne

- **Klucze naturalne**
 - ▣ nr rejestracyjny pojazdu, VIN, nr rachunku, NIP, PESEL
- **Klucze sztuczne –generowane automatycznie przez system**
 - ▣ nr klienta, id produktu, nr transakcji
- **Połączenie tabeli wymiaru i faktów za pomocą klucza podstawowego-obcego**
 - ▣ pokaż ilość szkód pojazdu o numerze rejestracyjnym xxx w ostatnim roku
 - zapytanie wyłącznie do tabeli faktów
- **Jeśli wartość klucza podstawowego może się zmienić wysoki koszt uaktualnienia faktów**
 - ▣ sytuacja mało prawdopodobna
 - ▣ uwaga: nr rachunku!

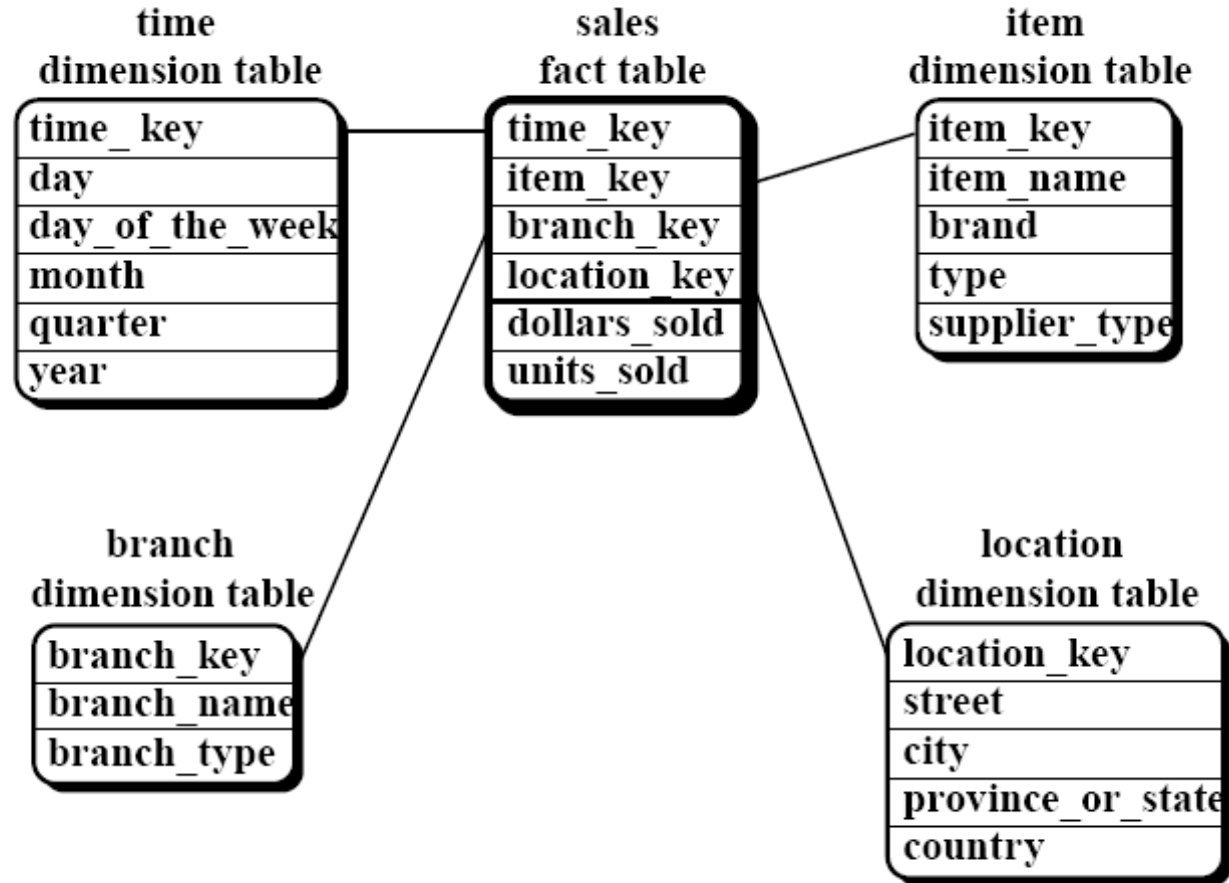
Schematy wielowymiarowych danych w ROLAP

- **schemat gwiazdy**
(ang. starschema)
- **schemat płatka śniegu**
(ang. snowflake schema)
- **schemat konstelacji faktów**
(ang. fact constellation schema)
- **schemat gwiazda-płatek śniegu**
(ang. starflake schema)

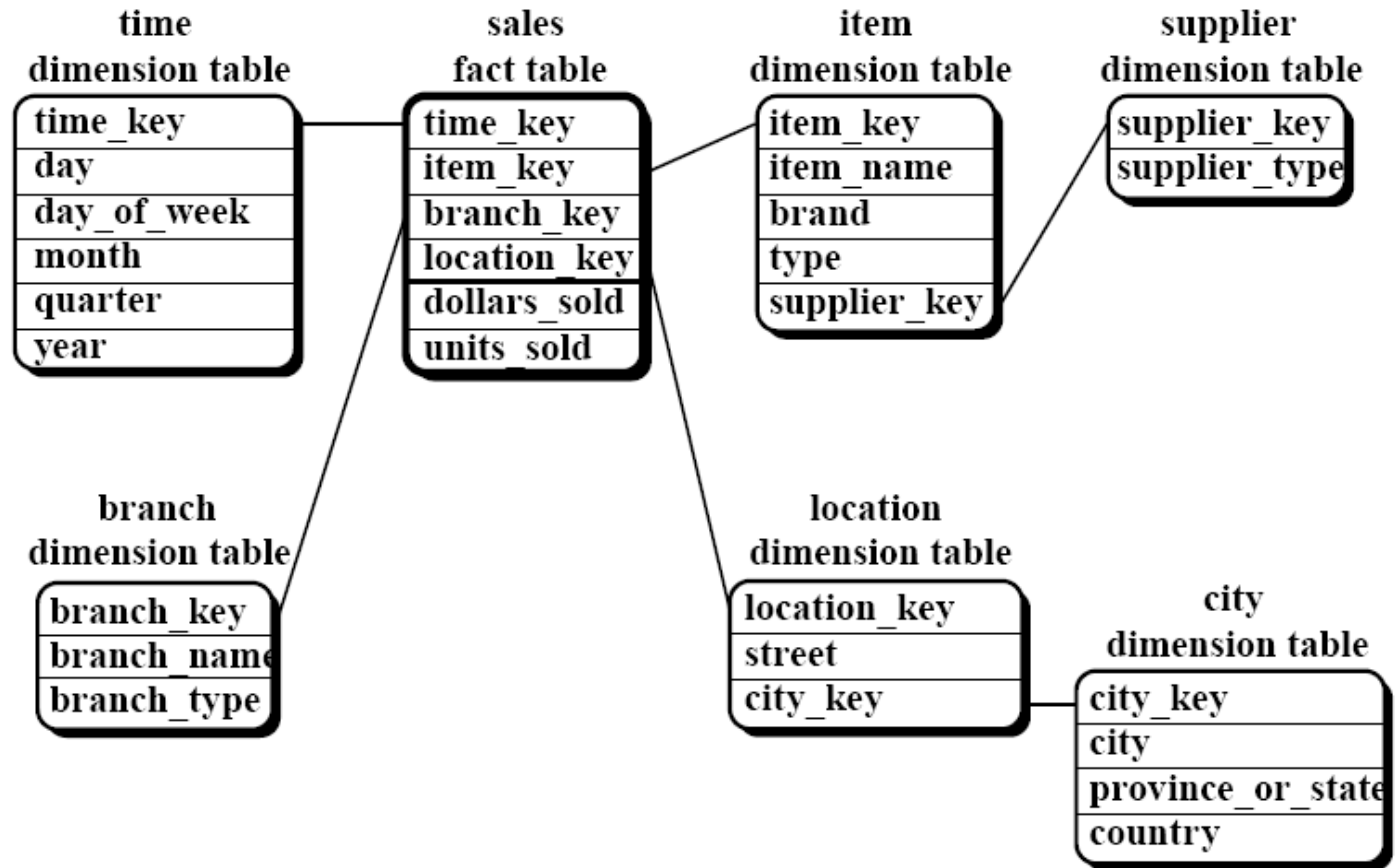
Schemat gwiazdy (stars schema)

- **Gwiazda: tabela wymiarów jest zdenormalizowana**
- **Zaleta:**
 - ▣ **operacja roll-up wykonywana szybko (bez konieczności łączenia z tabelą poziomą nadrzędnego)**
- **Wada:**
 - ▣ **na skutek redundancji danych rozmiar tabeli może być duży**
 - np. wymiar Czas z ziarnem 1 sek i horyzontem czasowym wymiaru 10 lat
→ 300 000 000 rekordów
 - astronomia, fizyka jądrowa, telekomunikacja

Schemat gwiazdy



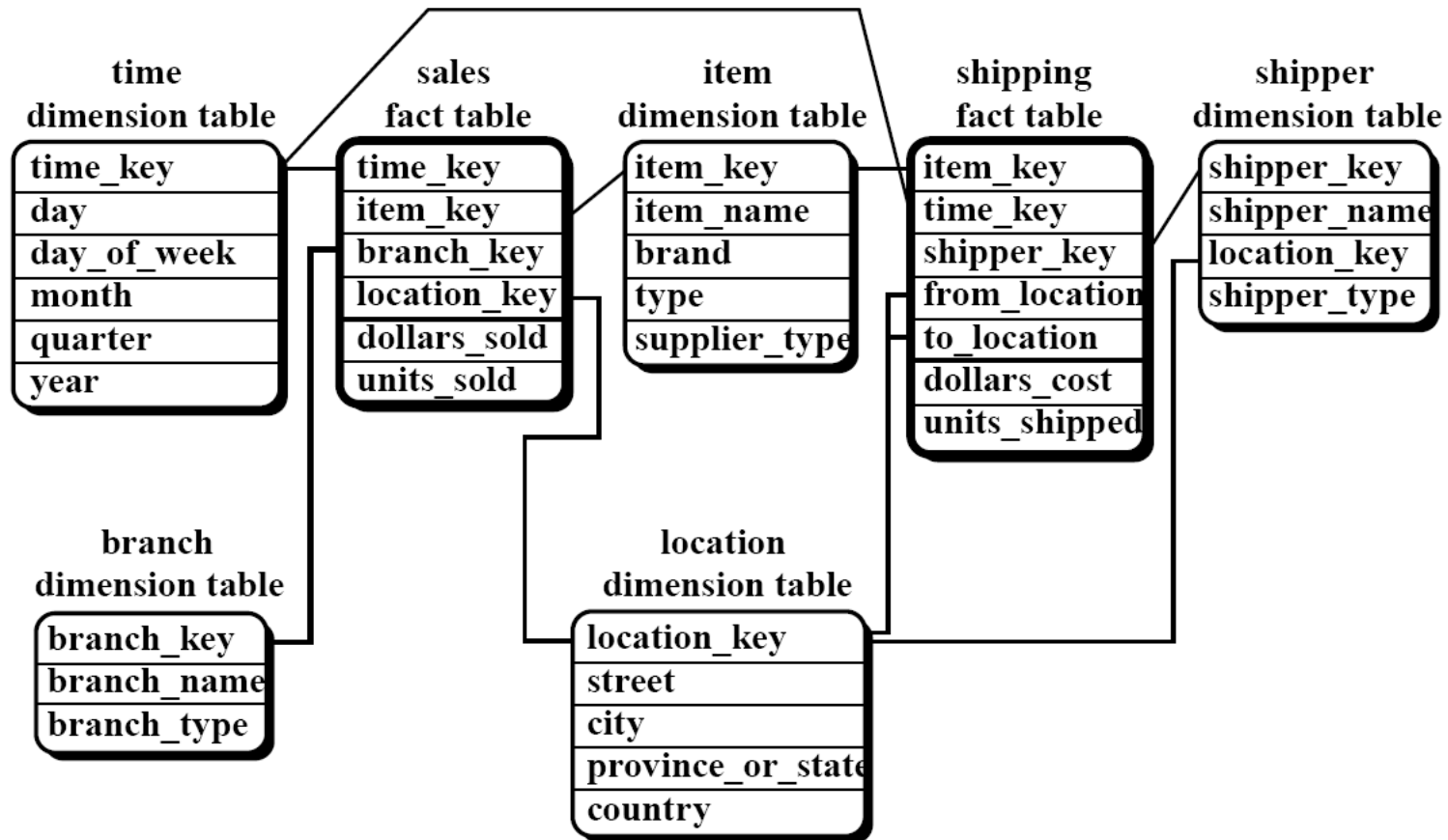
Schemat gwiazdy



Schemat platek sniegu (snowflake schema)

- **Płatek śniegu: tabele wymiaru są znormalizowane**
 - mniejszy rozmiar poszczególnych tabel poziomów
 - operacja roll-up wykonywana wolniej (konieczność łączenia tabel poziomów)
- **W praktyce:**
 - atrybuty różnych poziomów często wykorzystywane w operacji roll-up → umieszczenie w tej samej zdenormalizowanej tabeli poziomu najniższego
 - atrybuty rzadko wykorzystywane w roll-up → umieszczenie w znormalizowanych tabelach poziomów wyższych

Schemat kostelacji faktów



Schemat gwiazda-platek sniegu

- **Część wymiarów znormalizowanych**

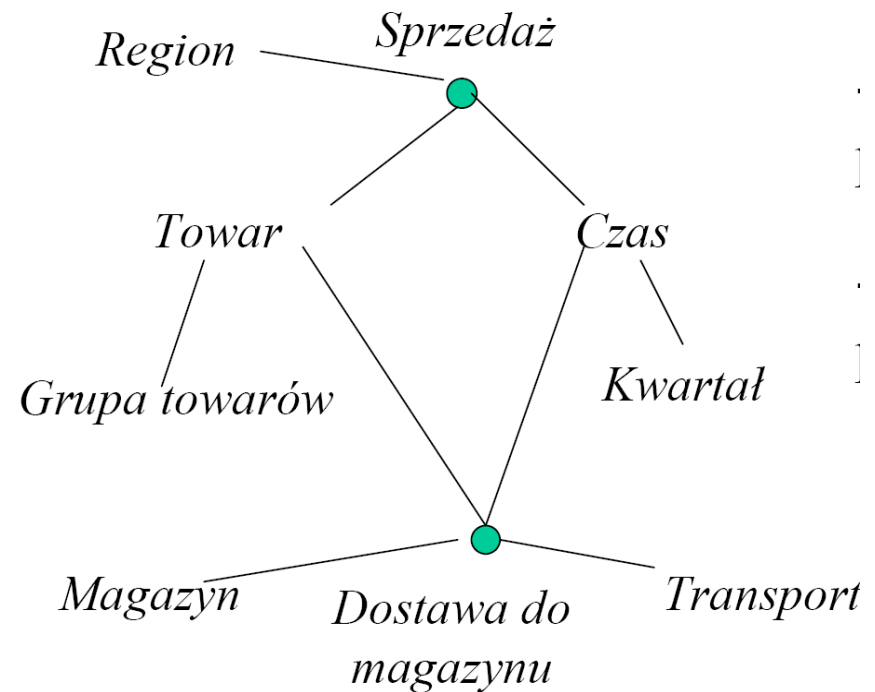
- poziomy wyższe wykorzystywane rzadko
- oszczędność miejsca

- **Część wymiarów zdenormalizowanych**

- wszystkie poziomy wymiarów wykorzystywane często
- efektywność zapytań

Model punktowy

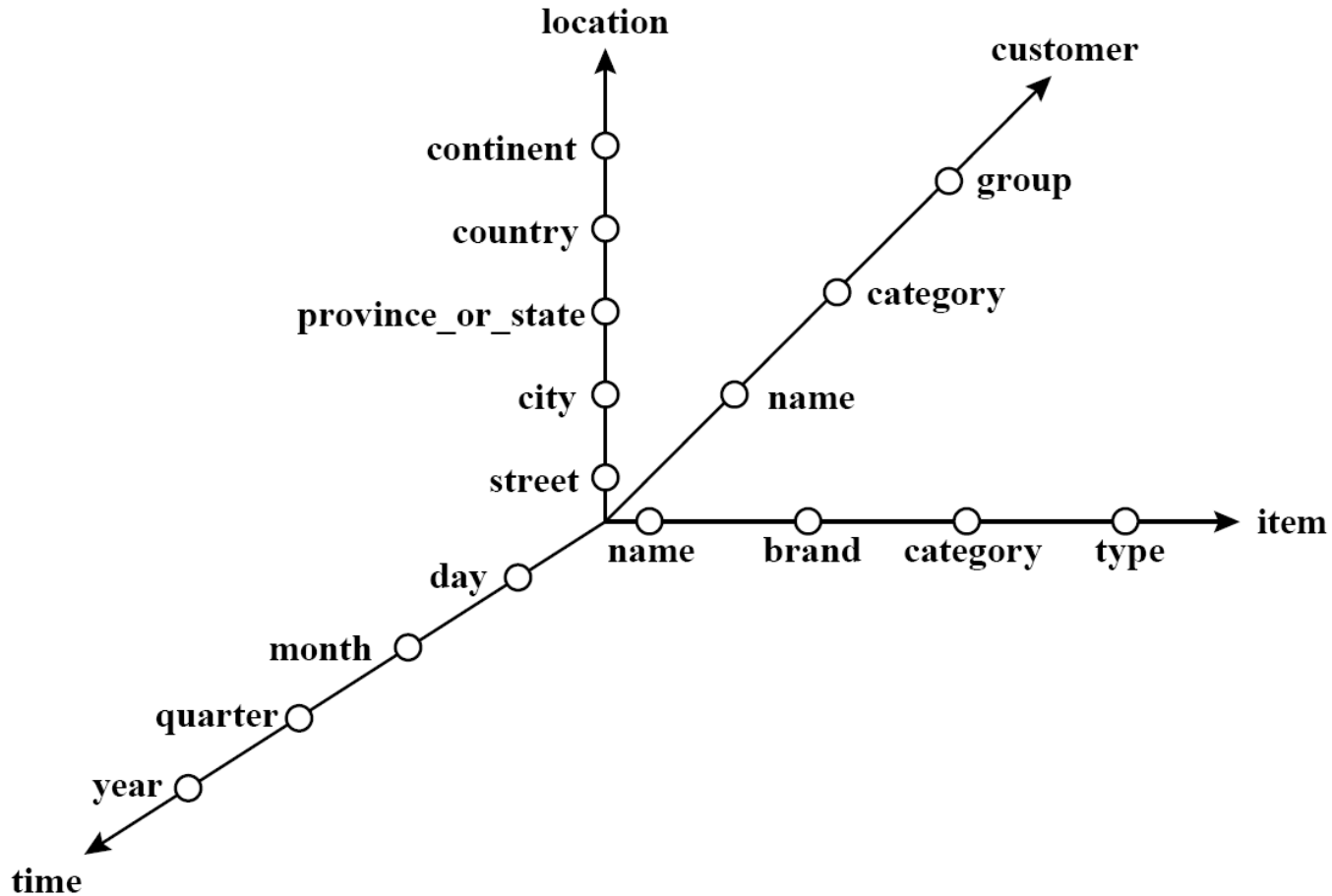
- Przed stworzeniem logicznego modelu danych (gwiazda, płatek sniegu) należy uzgodnić model pojęciowy.
- Przykładowa technika modelowania: model punktowy.
 - ▣ Fakty reprezentowane są punktami
 - ▣ Wymiary reprezentowane są przez nazwy
 - ▣ Podobnie reprezentujemy kolejne poziomy hierarchii
 - ▣ Model może obejmować wiele (konstelacje) faktów, korzystających częściowo ze wspólnej hierarchii wymiarów



Model punktowy

- W modelu punktowym zapisujemy ponadto informacje na temat:
 - Nazw atrybutów
 - Typów danych
 - Wierzów integralności
 - Retrospekcji (zmienności wartości atrybutów w czasie)
 - Częstotliwości odświeżania
 - Pochodzenia danych (źródło, transformacje)
 - Metadanych biznesowych
- **Retrospekcja** może być:
 - **prawdziwa** – zapisujemy wszelkie zmiany wartości wraz z dokładnym czasem
 - **fałszywa** – nowe wartości zastępują stare
 - **trwała** – nie przewidujemy zmiany wartości.

Model punktowy – przykład



Agregacja danych

- Agregacje to operacje zamieniające zbiór wartości (przeważnie liczbowych) miar opisujących fakty, na pojedynczą wartość.
- **Podstawowa operacja tworząca kostki danych.**
 - Przykłady:
 - Suma
 - Liczba rekordów
 - Średnia
 - Minimum, maksimum, mediana
 - Dodatkowy warunek (np. HAVING)
 - Specjalne (stored procedures)
 - **Przykładowe zapytanie:**

```
SELECT miesiac, SUM(kwota) FROM tablica_faktów
GROUP BY miesiac HAVING SUM(kwota)>500
```

Poziomy agregacji danych

- Nie zawsze potrzebujemy danych opisanych z pełną dostępną dokładnością.
- **Główny cel: wydajność**
- Agregacja wartości może dotyczyć pomijania pewnych wymiarów lub atrybutów w hierarchii.
- Hurtownia danych (lub hurtownie tematyczne – ang. *data mart*) mogą przechowywać zmaterializowane podkostki danych i korzystać z nich podczas analiz.

Przykład

- **Dane:** fakty sprzedaży (50 mln. rekordów, miara: wartość),
 - ▣ klienci identyfikowani kodem pocztowym (3000 różnych kodów),
 - ▣ towary (60 grup, 800 nazw indywidualnych),
 - ▣ czas: 3 lata (1000 dni),
 - ▣ sklepy (18 sztuk).

Model gwiazdy:

- tablice z ponad 50 mln. rekordów. (łącznie)
- Kostka danych:
 $3000 * 800 * 1000 * 18 = 43\,200\,000\,000$ komórek.

Wersja zagregowana:

- ignorujemy wymiar klientów, towary rozpatrujemy tylko w grupach, czas rozpatrujemy w skali miesięcy.
- Kostka danych: $60 * 36 * 18 = 38\,880$ komórek (sumy wartości).

Powyższa agregacja nie pozwala na wygenerowanie wszystkich tych raportów, które mogły by tworzone oryginalnie, ale za to dla pozostałych raportów może działać o 3 rzędy wielkości szybciej.

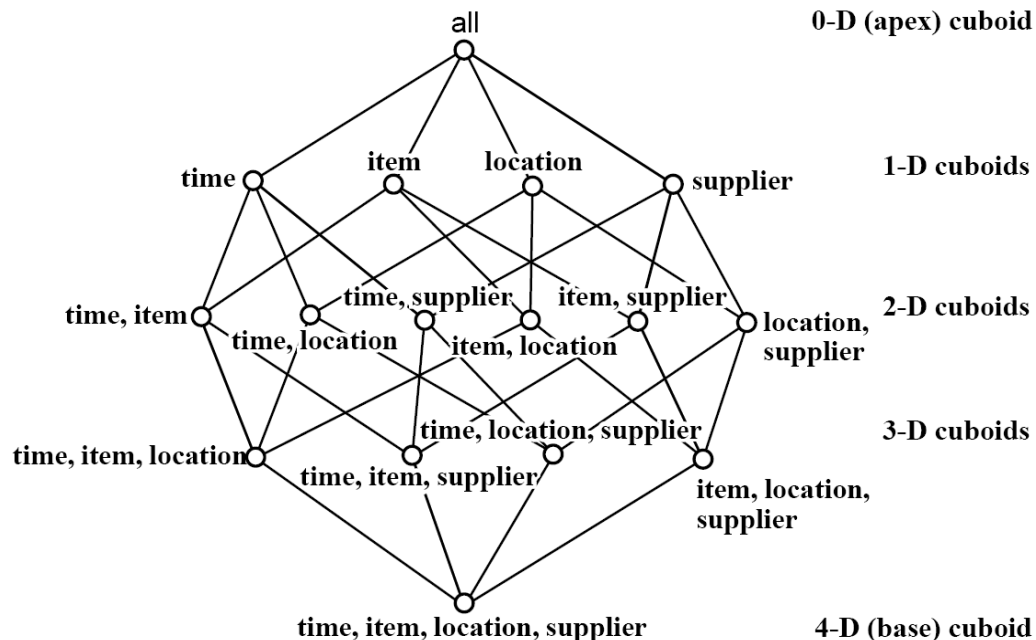
Optymalizacja agregacji

□ Problem:

- Które wymiary możemy pominąć
- Jak agregować dane (*sum, min, max, avg, count*)?
- Które atrybuty hierarchii agregujemy, a które pozostawiamy?
- Ile powinno powstać zmaterializowanych kostek pomocniczych?
- Jak optymalnie obliczać konkretne zapytanie?

□ Cel:

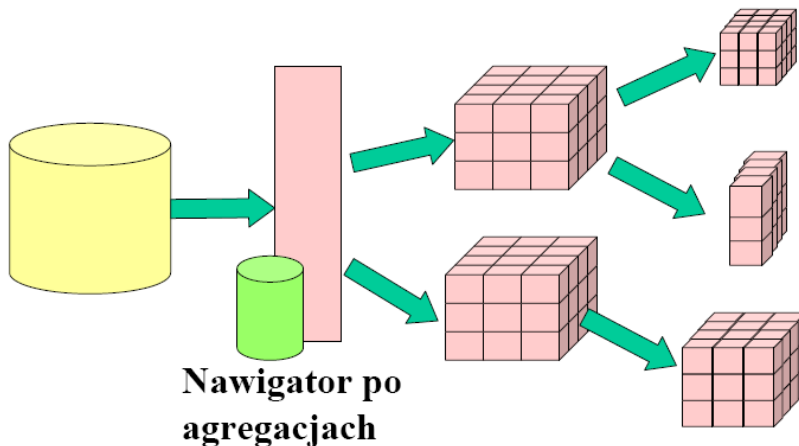
- np. minimalizacja średniego czasu przetwarzania zadawanych w praktyce zapytań OLAP.



Musimy wybrać takie kostki, z których później najłatwiej będzie generować raporty.

Metadane = dane o danych

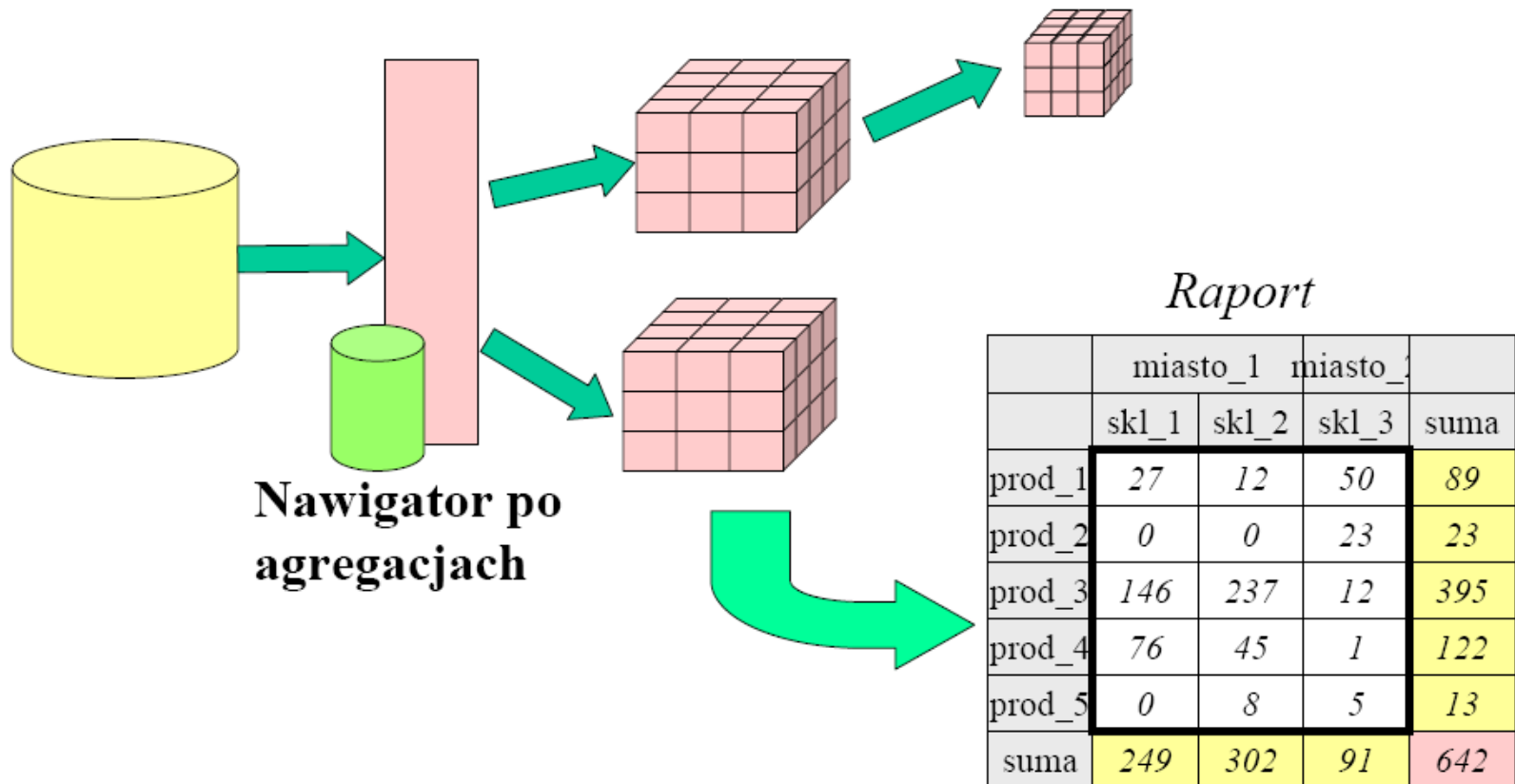
- Skąd wiadomo, które zapytania są często wykonywane?



- Informacje o zadawanych zapytaniach OLAP są gromadzone jako **metadane** w specjalnej bazie- repozytorium.
- Mogą być wykorzystane do optymalizacji zapytań (np. materializacji niektórych agregacji).
- Metadane zawierają też wiele innych informacji, np.
 - ▣ opis pojęciowy i logiczny danych,
 - ▣ informacje o źródłach i ich integracji,
 - ▣ dziennik aktualizacji itd.

Nawigator agregacji

- Nawigator po agregacjach wspomaga też przeglądanie danych i przygotowywanie raportów.
- Zapewnia wygodny wgląd w wielowymiarowy model danych.



Prezentowanie czasu w tabeli faktów

- **Sztuczny identyfikator (data_id)**
 - ▣ konieczność łączenia z tabelą wymiaru czasu
- **Naturalny identyfikator (data, timestamp) – składowanie fizycznej daty**
 - ▣ sposób bardziej efektywny
 - ▣ większość analiz wykonuje się w wymiarze czasu
 - ▣ zapytanie nie zawiera połączenia z tabelą wymiaru czasu
- **Składowanie przesunięcia czasowego**
- **Składowanie zakresów dat**

Skladowanie przesunięcia czasowego

partycjonowane tabele faktów

klient_id	kwota	nr_dnia
100	57.60	1
203	123.90	1
4005	99.00	2
205	79.40	3
111	205.90	3
5008	432.00	13
23	332.40	14
567	87.00	31

klient_id	kwota	nr_dnia
100	57.60	1
203	123.90	2
4005	99.00	3
205	79.40	4
111	205.90	5
5008	432.00	28
23	332.40	29
567	87.00	29

- **Wada:**
 - konieczność konwersji daty z zapytania użytkownika do postaci przesunięcia czasowego
 - perspektywa
- **Zalety:**
 - podział dużej tabeli na mniejsze, z których każda może być adresowana w zapytaniu niezależnie
 - mniejszy rozmiar atrybutu reprezentującego datę (1 B)

Skladowanie zakresu czasowego

- **Np.**
 - ▣ atrybuty: data_od, data_do
 - ▣ stan magazynu supermarketu
 - ▣ sprzedaż w okresie czasowym
- **Wada: bardziej złożone zapytanie testujące warunki początku i końca okresu**
- **Zaletą: rozszerzenie zakresu ważności rekordu poprzez zmodyfikowanie wartości data_do**
 - ▣ np. liczba produktów w magazynie nie ulega zmianie w danym dniu → zmodyfikuj wartość data_do dla tego produktu

Nawigacja po agregacjach

- **Drilling down / rolling up**
- **Slicing and dicing**
- **Rotating**
- **Pivoting**

Literatura

- *Ch. Todman. Projektowanie hurtowni danych. WNT, Warszawa 2003.*
- *M. Jarke, M. Lenzerini, Y. Vassiliou, P. Vassiliadis. Hurtownie danych. Podstawa organizacji i funkcjonowania, WSiP, Warszawa 2003.*
- *V. Poe, P. Klauer, S. Brobst. Tworzenie hurtowni danych. WNT, Warszawa 2000.*