



 KATEDRA
INŻYNIERII
OPROGRAMOWANIA

Hurtownie danych

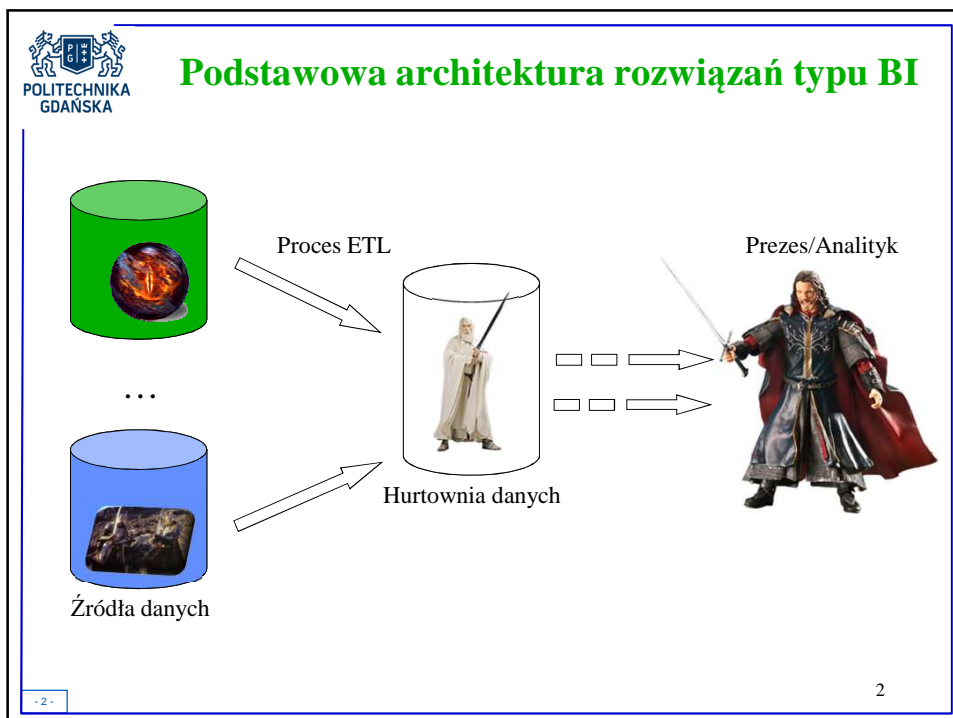
Rola hurtowni danych w systemach typu *Business Intelligence*

Krzysztof Goczyła
Teresa Zawadzka

*Katedra Inżynierii Oprogramowania
 Wydział Elektroniki, Telekomunikacji i Informatyki
 Politechnika Gdańska
 {kris, tegra}@eti.pg.gda.pl*



-1-





Hurtownia danych (ang. *data warehouse*)

Tematyczny (ang. *subject-oriented*), **zintegrowany** (ang. *integrated*), **zależny od czasu** (ang. *time-variant*), **trwały** (ang. *non-volatile*) **zbiór danych**, ukierunkowany na wspomaganie procesów podejmowania decyzji.

Bill Inmon

Kopia danych transakcyjnych zapisanych w **strukturach ukierunkowanych na zapytania analityczne**.

Ralph Kimball

- 3 -



Cechy hurtowni danych (1)

Zbiór danych

Hurtownia danych to bardzo duży zbiór danych / baza danych – (setki GB, pojedyncze TB).

Taka baza danych optymalizowana jest pod kątem przetwarzania **analitycznego**, a nie **transakcyjnego**.

- 4 -



Cechy hurtowni danych (2)

Tematyczny zbiór danych

Hurtownia danych może być używana do analizy konkretnego obszaru tematycznego np.: sprzedaży.

Zbiór danych zależny od czasu

Hurtownia danych przechowuje dane historyczne.

Trwały zbiór danych

Dane zapisane w hurtowni danych pozostają niezmiennicze.

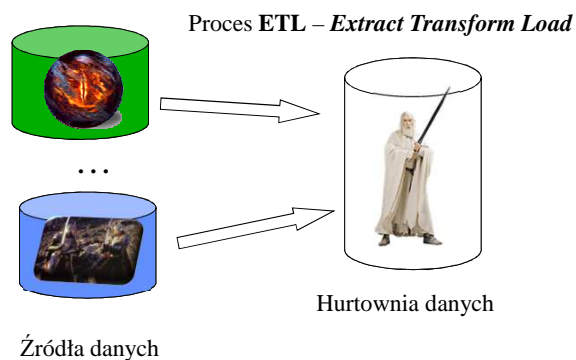
- 5 -



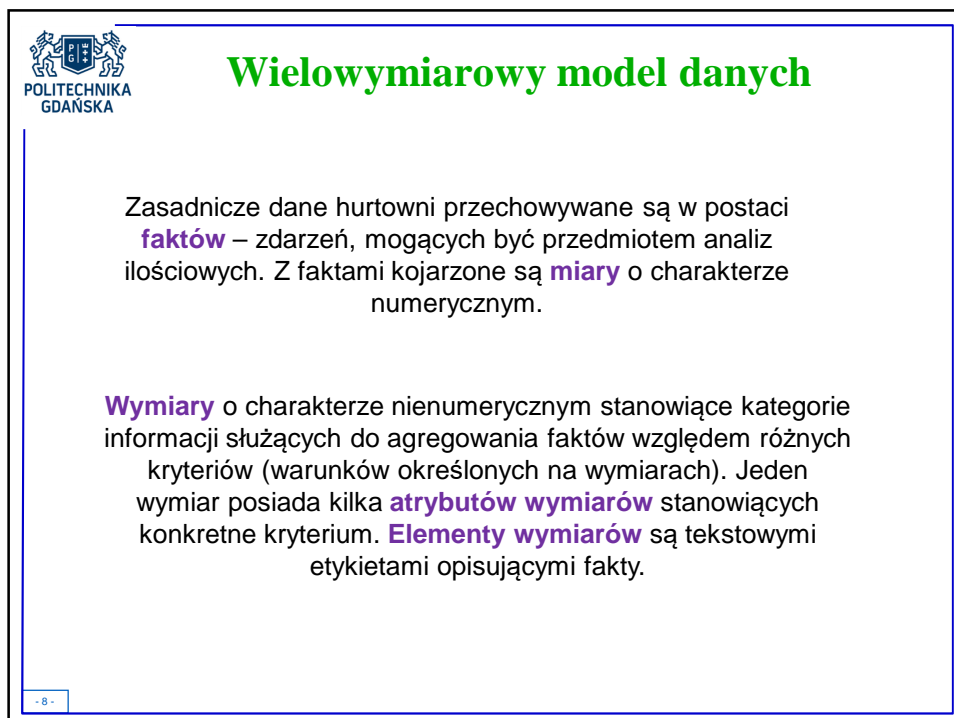
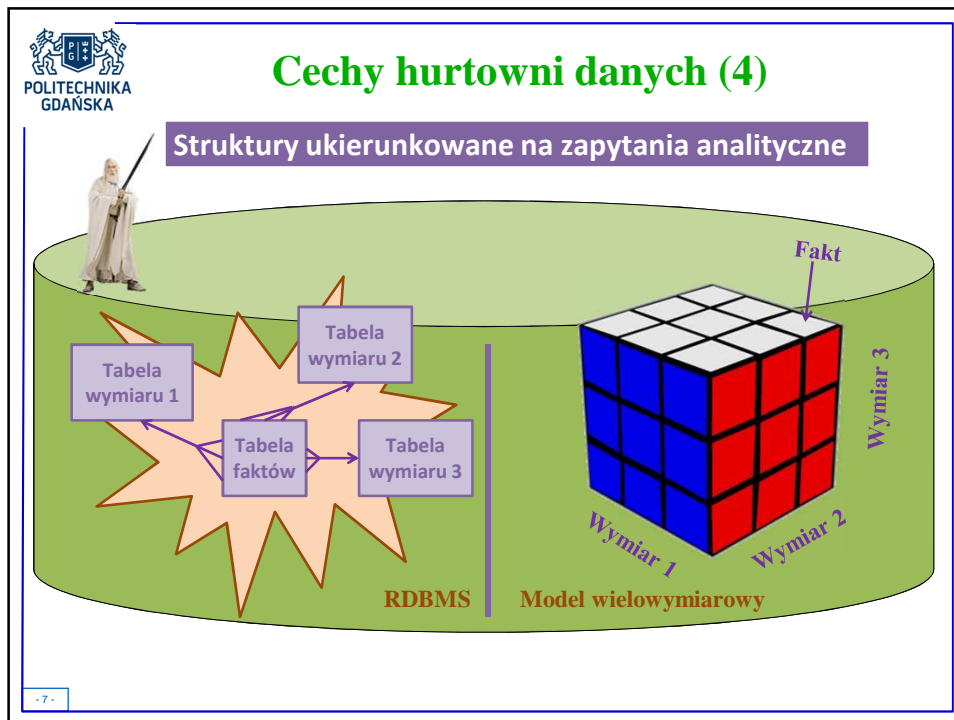
Cechy hurtowni danych (3)

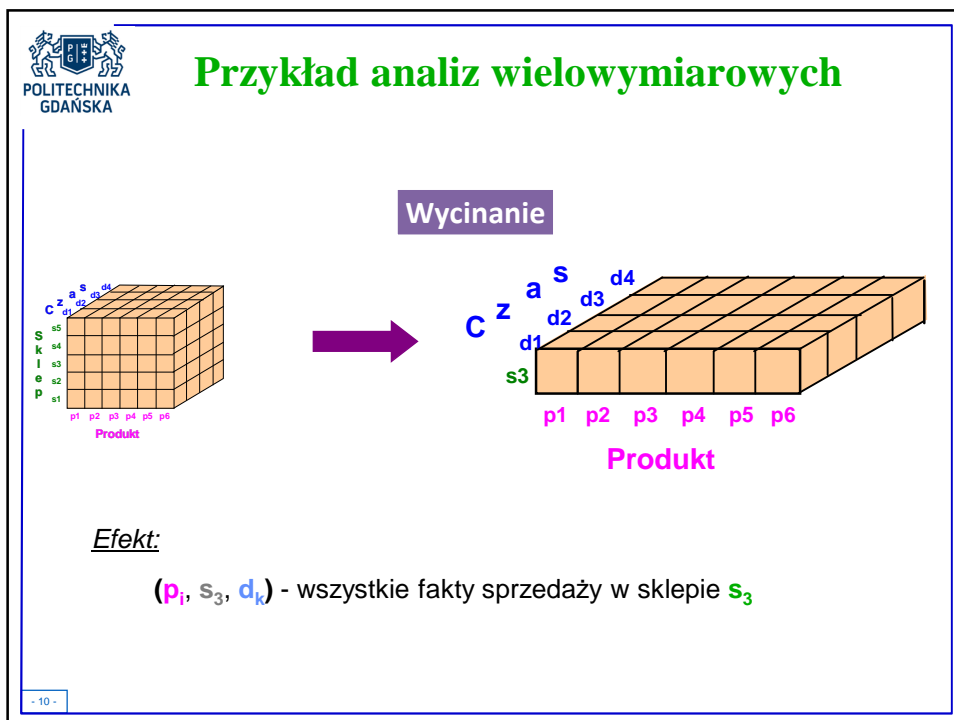
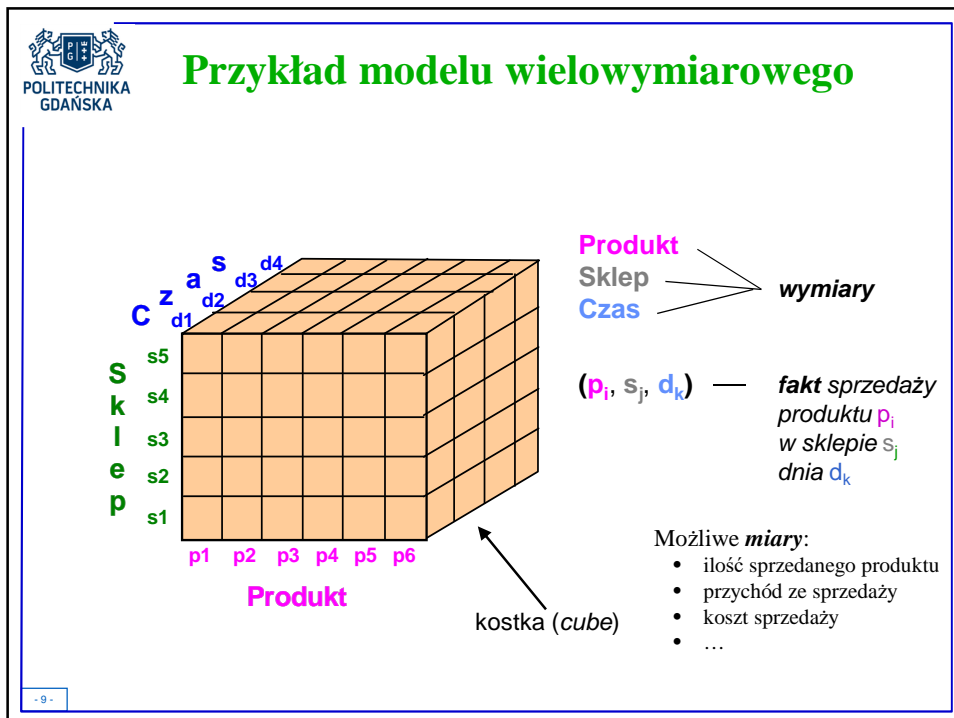
Zintegrowany zbiór danych


Dane pochodzące z wielu różnych systemów baz danych i innych źródeł danych i są zbierane do jednego miejsca: zintegrowanej bazy danych - hurtowni danych.



- 6 -

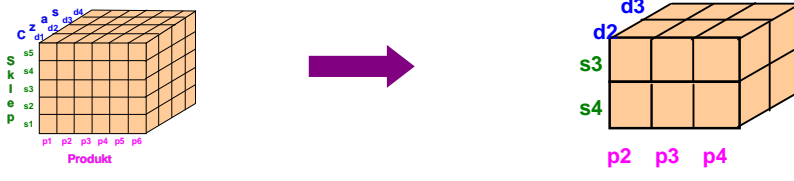




 POLITECHNIKA
GDAŃSKA

Przykład analiz wielowymiarowych


Wycinanie



Efekt:

(p_i, s_j, d_k) - fakty sprzedaży w sklepie s_3 i s_4 produktów p_2, p_3, p_4 w dniach d_2, d_3

- 11 -

 POLITECHNIKA
GDAŃSKA

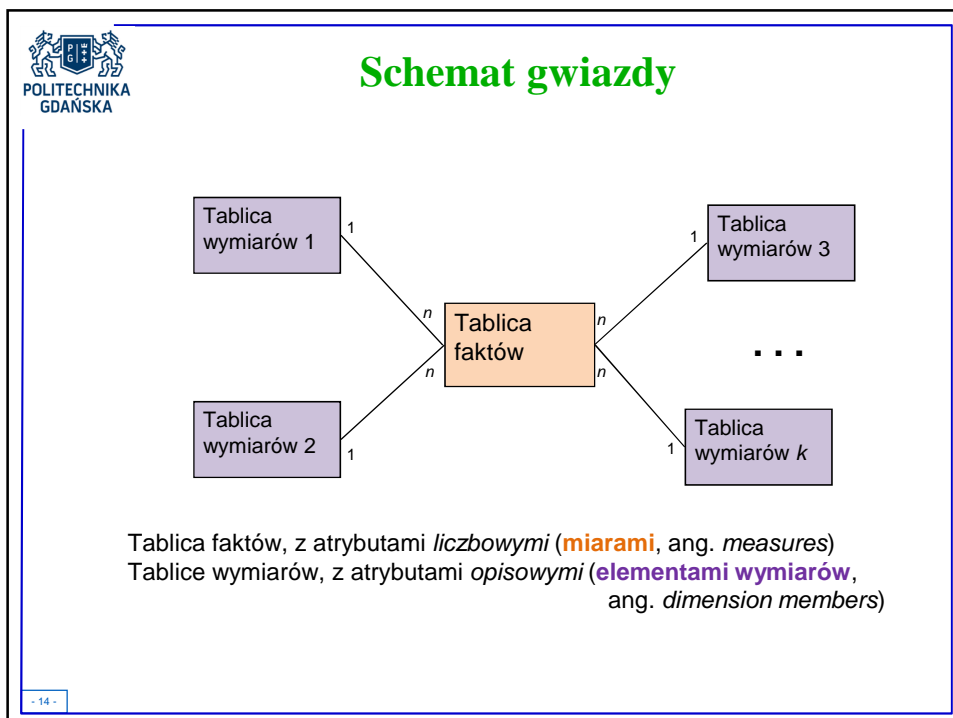
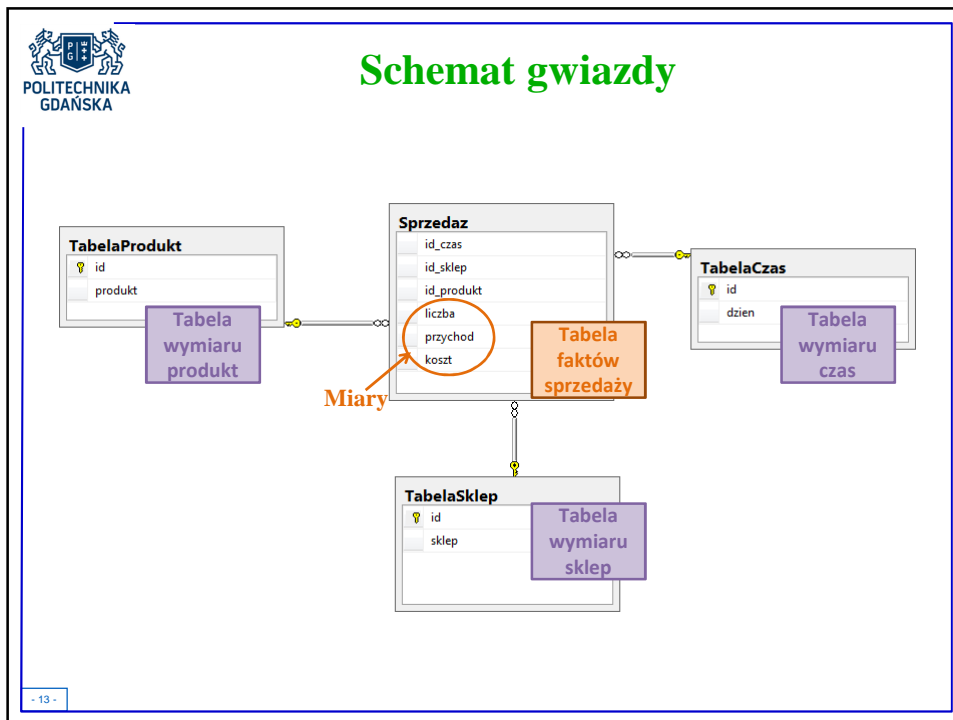
Agregacje

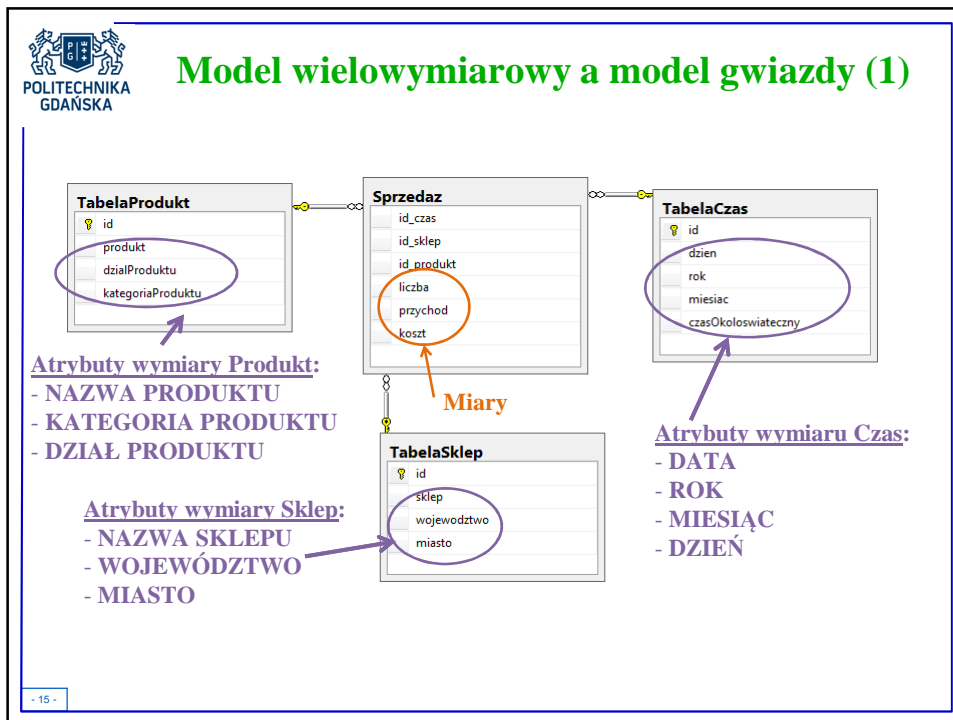
Dla wyciętego zbioru faktów wyliczane są **agregacje**.

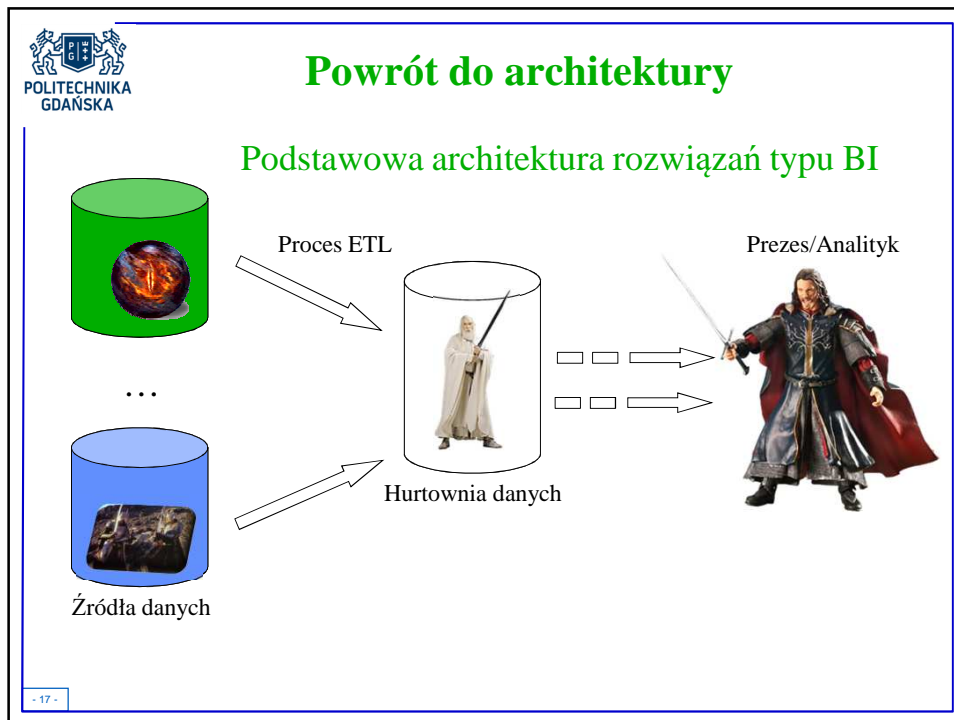
Miary mają przypisane **funkcje agregujące**. Funkcja agregująca określa jaką **operacja matematyczna** jest wykonywana na wartościach miar.

Prawidłową funkcją agregującą dla miar: ilość sprzedanego produktu, przychód ze sprzedaży i koszt sprzedaży jest funkcja SUM.

- 12 -







POLITECHNIKA GDAŃSKA

Podejście z dołu do góry

Kopia danych transakcyjnych zapisanych w strukturach ukierunkowanych na zapytania analityczne.

Ralph Kimball

Data Sources

Staging

Star Schema

Data Access

ERP

CRM

SCM

SFA

Others

Analyst

Reporting

Data Mining

DT – Dimension Table
FT – Fact Table

Zaczerpnięte z: <http://searchbusinessintelligence.techtarget.in/tip/Inmon-vs-Kimball-Which-approach-is-suitable-for-your-data-warehouse>

- 19 -

POLITECHNIKA GDAŃSKA

Minihurtownia (ang. *data mart*)

Tematyczna hurtownia danych

Gdy dane w hurtowni obejmują tylko pewien wycinek danych zintegrowanych:

↓

minihurtownia
(podhurtownia, zbiorcza baza danych, ang. *data mart*)

Bill Inmon

Ralph Kimball

- 20 -



Co każdy student wiedzieć powinien...

1. Znać definicje hurtowni danych i rozumieć między nimi różnicę?
2. Rozumieć pojęcia modelu wielowymiarowego.
3. Rozumieć pojęcie schematu gwiazdy.
4. Znać architekturę B. Inmona i Ralpa Kimballa.
5. Umieć wyjaśnić różnice pomiędzy podejściem z góry do dołu i z dołu do góry.

