Dimensional modeling

Kristo Raun

Data Engineering 2022

UNIVERSITY OF TARTU
1632

What is dimensional modeling?

- Dimensional modeling is widely accepted as the preferred technique for presenting analytic data because it addresses two simultaneous requirements:

- Source: Ross, Margy, and Ralph Kimball. The data warehouse toolkit: the definitive guide to dimensional modeling. John Wiley & Sons, 2013.

What is dimensional modeling?

- Dimensional modeling is widely accepted as the preferred technique for presenting analytic data because it addresses two simultaneous requirements:
  - Deliver data that's understandable to the business users.

- Source: Ross, Margy, and Ralph Kimball. The data warehouse toolkit: the definitive guide to dimensional modeling. John Wiley & Sons, 2013.

What is dimensional modeling?

- Dimensional modeling is widely accepted as the preferred technique for presenting analytic data because it addresses two simultaneous requirements:
  - Deliver data that's understandable to the business users.
  - Deliver fast query performance.

- Source: Ross, Margy, and Ralph Kimball. The data warehouse toolkit: the definitive guide to dimensional modeling. John Wiley & Sons, 2013.

What is dimensional modeling?

- Dimensional modeling is widely accepted as the preferred technique for presenting analytic data because it addresses two simultaneous requirements:
  - Deliver data that's understandable to the business users.
  - Deliver fast query performance.
- Dimensional modeling is a longstanding technique for making databases simple.

- Source: Ross, Margy, and Ralph Kimball. The data warehouse toolkit: the definitive guide to dimensional modeling. John Wiley & Sons, 2013.

Main flow of dimensional modeling

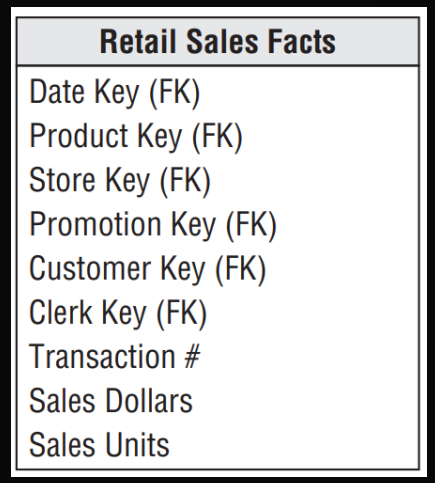
- Start from **business requirements**
  - What needs to be done? Why?

Main flow of dimensional modeling

- Start from **business requirements**
  - What needs to be done? Why?
- Design facts and dimensions

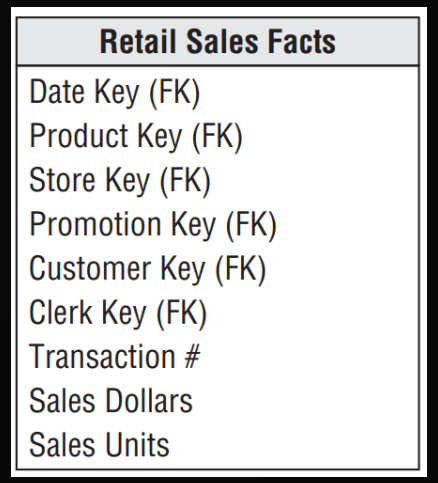
Facts and dimensions

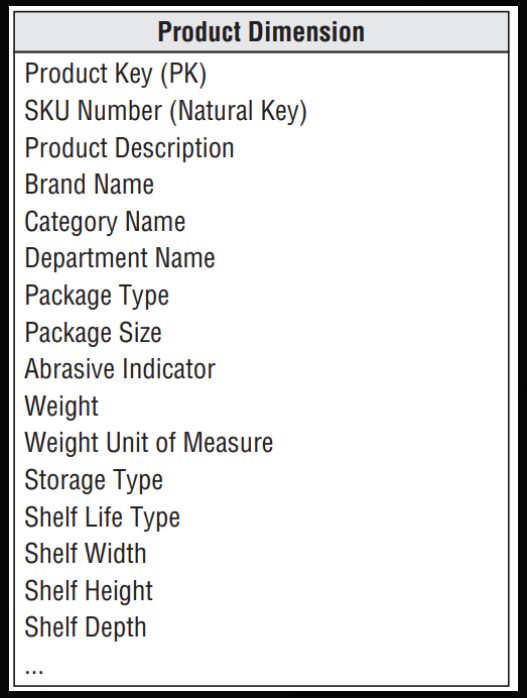
- Fact tables are for measurements



Facts and dimensions

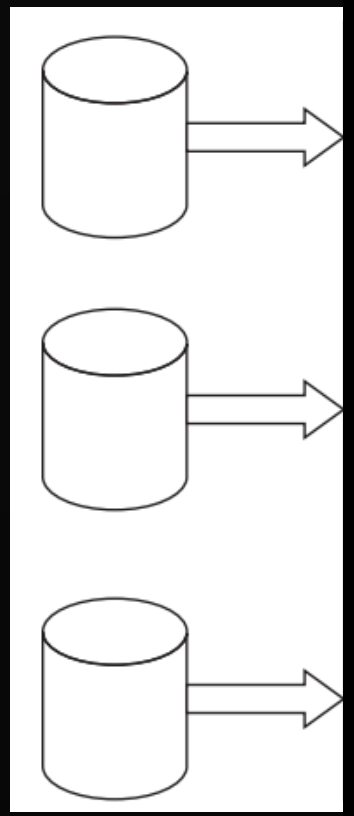
- Fact tables are for measurements
- Dimension tables are for descriptive context





General architecture

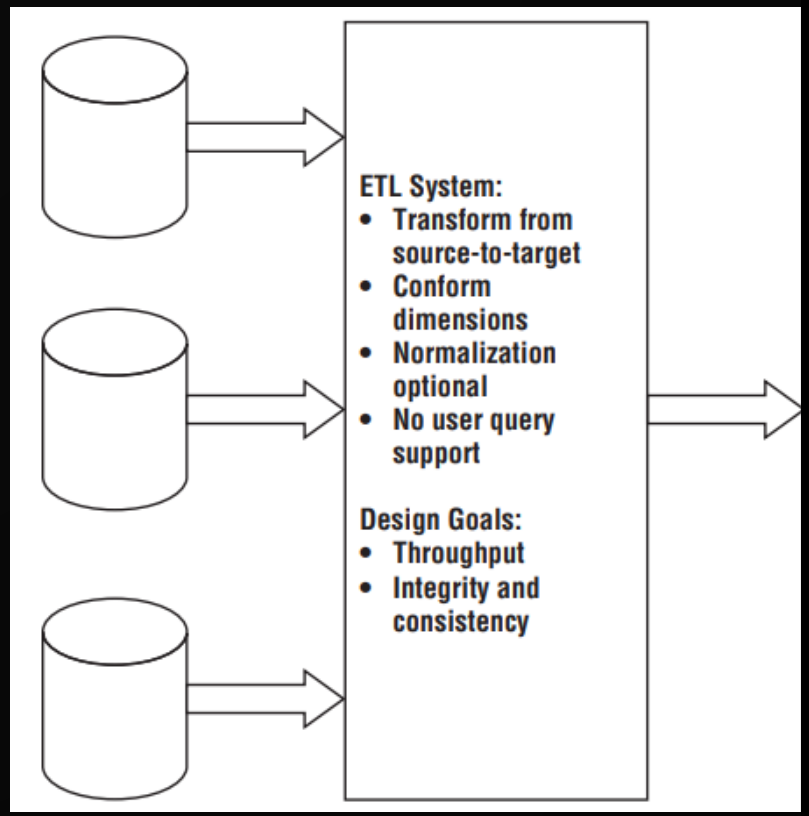
Sources



- Source: Ross, Margy, and Ralph Kimball. The data warehouse toolkit: the definitive guide to dimensional modeling. John Wiley & Sons, 2013.

General architecture

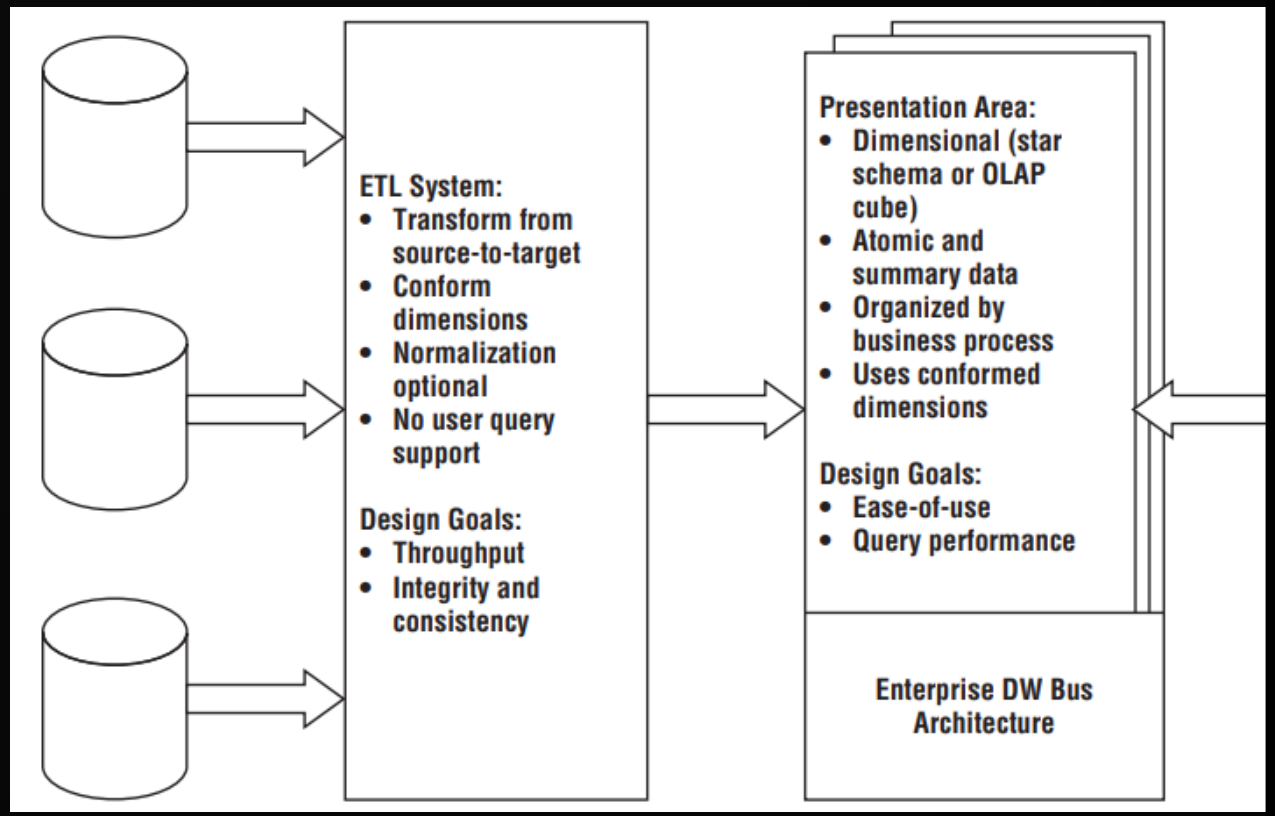
Sources Staging area



- Source: Ross, Margy, and Ralph Kimball. The data warehouse toolkit: the definitive guide to dimensional modeling. John Wiley & Sons, 2013.

General architecture

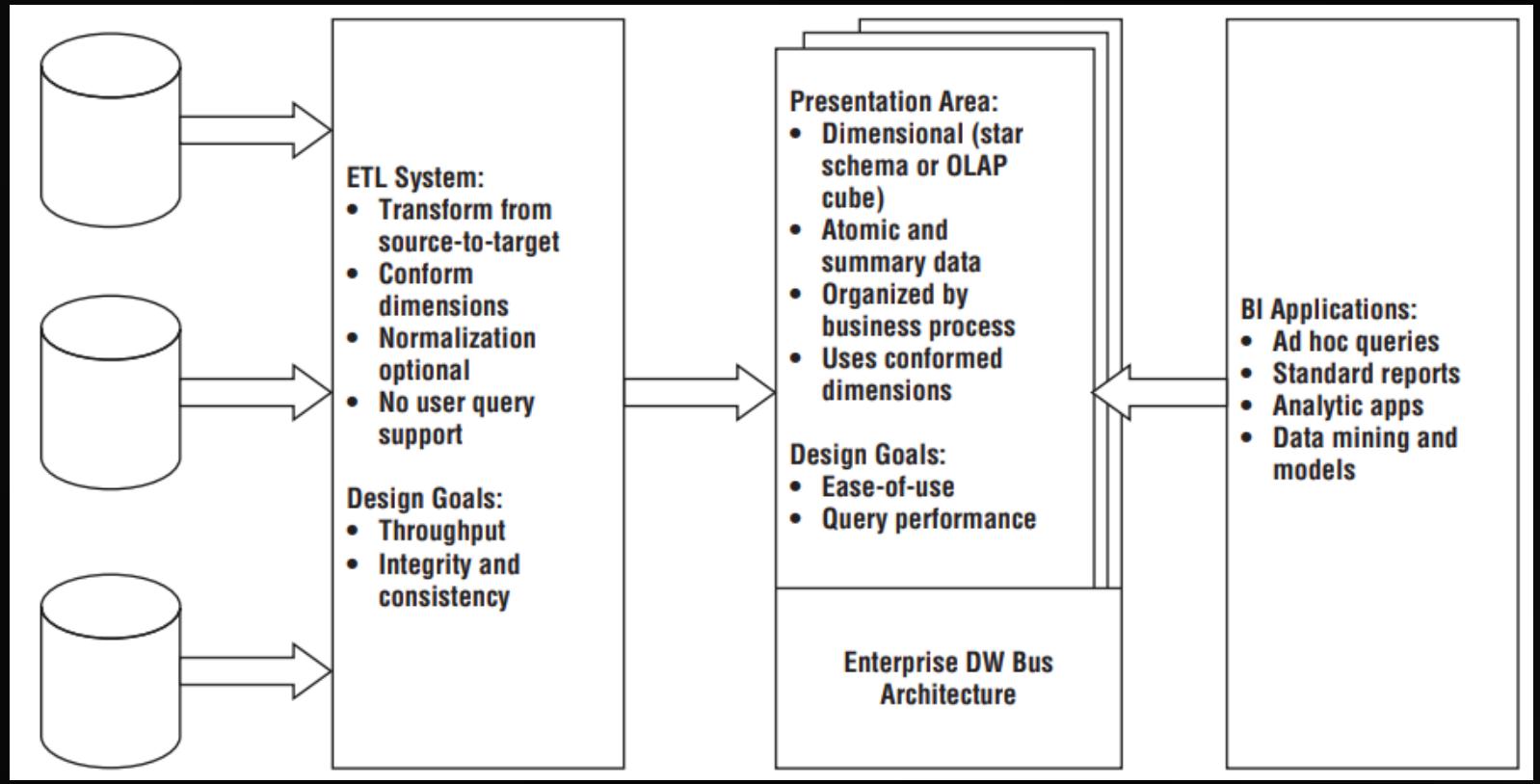
Sources Staging area Presentation layer



- Source: Ross, Margy, and Ralph Kimball. The data warehouse toolkit: the definitive guide to dimensional modeling. John Wiley & Sons, 2013.

General architecture

Sources Staging area Presentation layer BI/ML



- Source: Ross, Margy, and Ralph Kimball. The data warehouse toolkit: the definitive guide to dimensional modeling. John Wiley & Sons, 2013.

Final words…

Final words…

- Dimensional modeling is fast and efficient way of modeling your analytical data.

Final words…

- Dimensional modeling is fast and efficient way of modeling your analytical data.
- But there are many more methods. And technology keeps on advancing.

Final words…

- Dimensional modeling is fast and efficient way of modeling your analytical data.
- But there are many more methods. And technology keeps on advancing.
  - DWH design
    - CIF/EDW (Bill Inmon)
    - Data Vault

Final words…

- Dimensional modeling is fast and efficient way of modeling your analytical data.
- But there are many more methods. And technology keeps on advancing.
  - DWH design
    - CIF/EDW (Bill Inmon)
    - Data Vault
  - DWH vs
    - Data Lake
    - Data Lakehouse