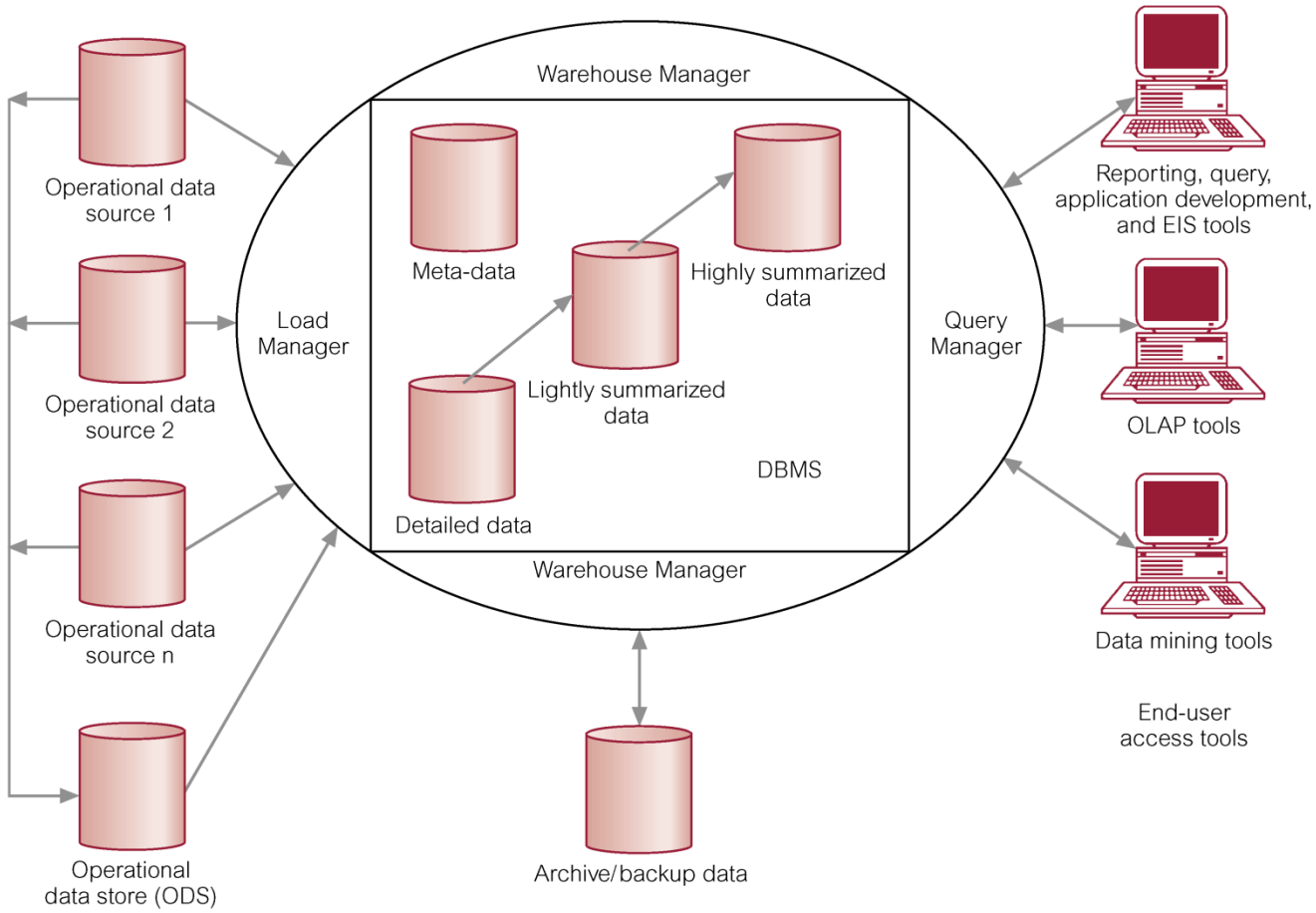
Data Warehouse Concepts

Example Data Warehouse Architecture



Data Warehouse Development Methodologies

Methodology	Main Advantage	Main Disadvantage
Inmon's Corporate Information Factory	Potential to provide a consistent and comprehensive view of the enterprise data.	Large complex project that may fail to deliver value within an allotted time period or budget.
Kimball's Business Dimensional Lifecycle	Scaled down project means that ability to demonstrate value is more achievable within an allotted time period or budget.	As data marts can potentially be developed in sequence by different development teams using different systems; the ultimate goal of providing a consistent and comprehensive view of corporate data may never be easily achieved.

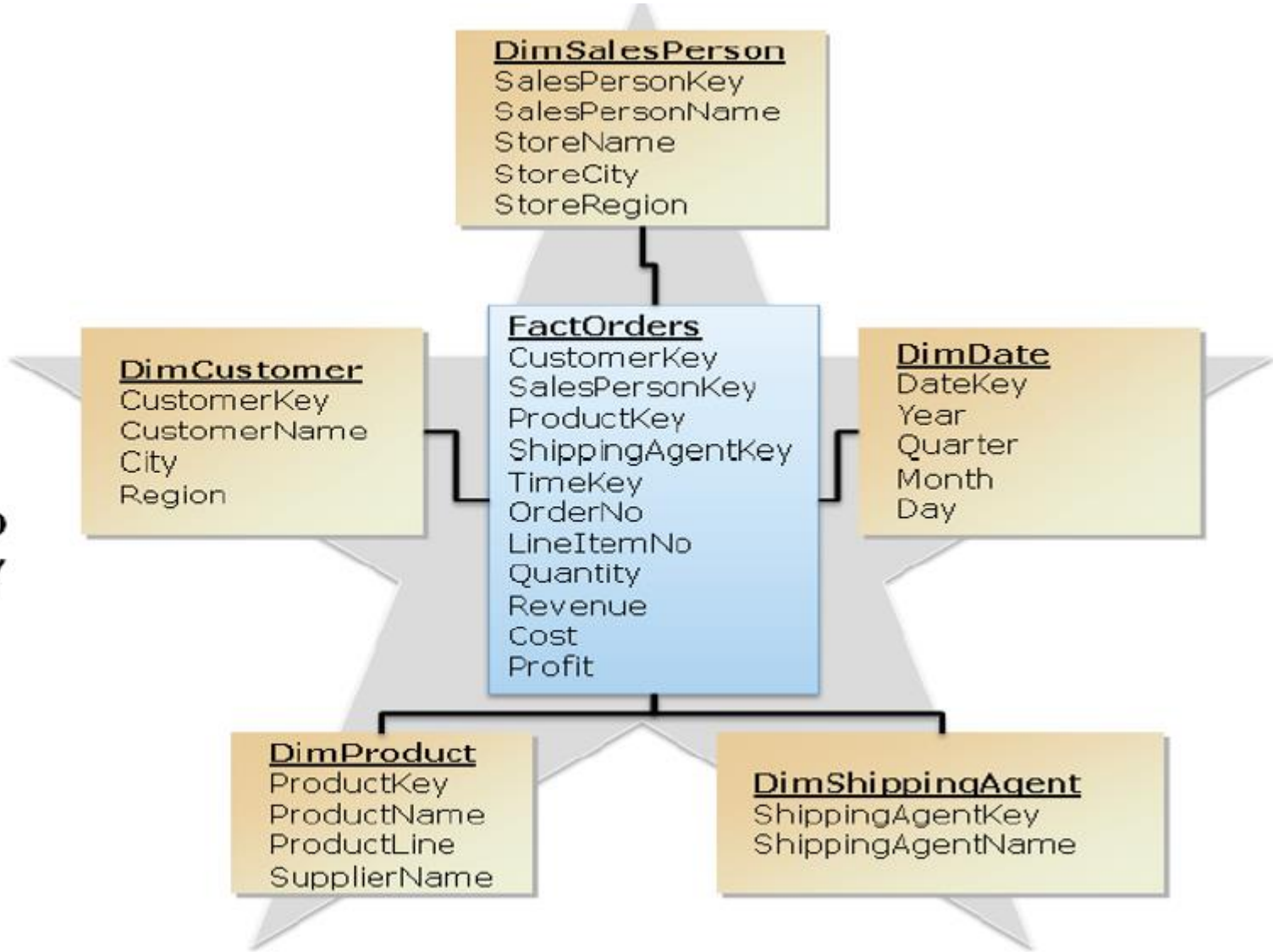
Modeling

- Inmon's Methodology
  - Traditional database techniques
  - ER Modeling
  - 3rd Normal Form
- Kimball's Methodology
  - Dimensionality Modeling
  - Dimensional Model
    - Star
    - Snowflake

Star Schema

Dimensional Data Model with Denormalized Dimension Tables

- Group related dimensions into dimension tables
- Group related measures into fact tables
- Relate fact tables to dimension tables by using foreign keys



Snowflake Schema

Dimensional Data Model with Normalized Dimension Tables

- Normalized dimension tables
- Consider when:
  - A subdimension can be shared between multiple dimensions
  - A well-understood hierarchy exists, and the dimension table contains a large amount of duplicated data
  - A sparse dimension has several different subtypes
  - Multiple fact tables of varying grain reference different levels in the dimension hierarchy