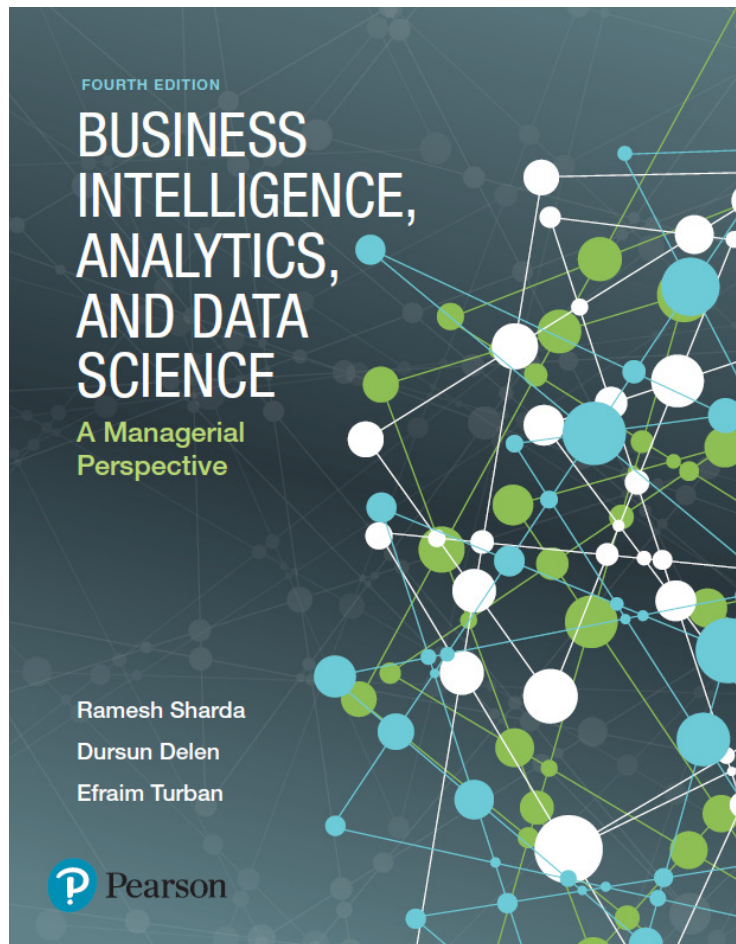


# Business Intelligence, Analytics, and Data Science: A Managerial Perspective

Fourth Edition



## Chapter 3 – Part A

Descriptive Analytics II:  
Business Intelligence and  
Data Warehousing

# Learning Objectives (1 of 2)

- 3.1** Understand the basic definitions and concepts of data warehousing
- 3.2** Understand data warehousing architectures
- 3.3** Describe the processes used in developing and managing data warehouses
- 3.4** Explain data warehousing operations
- 3.5** Explain the role of data warehouses in decision support

## Learning Objectives (2 of 2)

**3.6** Explain data integration and the extraction, transformation, and load (ETL) processes

**3.7** Understand the essence of business performance management (BPM)

**3.8** Learn balanced scorecard and Six Sigma as performance measurement systems

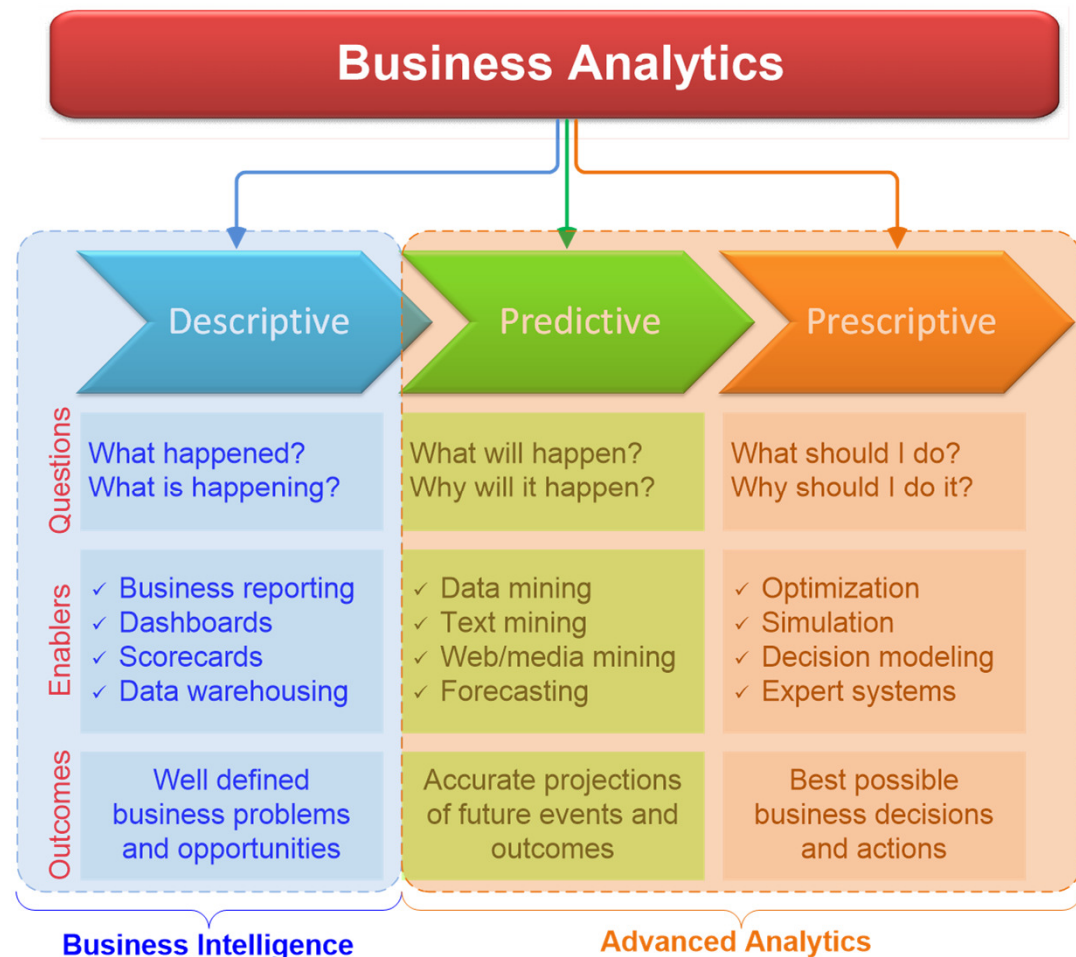
# OPENING VIGNETTE

## Targeting Tax Fraud with Business Intelligence and Data Warehousing

1. Why is it important for IRS and for U.S. state governments to use data warehousing and business intelligence (BI) tools in managing state revenues?
2. What were the challenges the state of Maryland was facing with regard to tax fraud?
3. What was the solution they adopted? Do you agree with their approach? Why?
4. What were the results that they obtained? Did the investment in BI and data warehousing pay off?
5. What other problems and challenges do you think federal and state governments are having that can benefit from BI and data warehousing?

# Business Intelligence and Data Warehousing

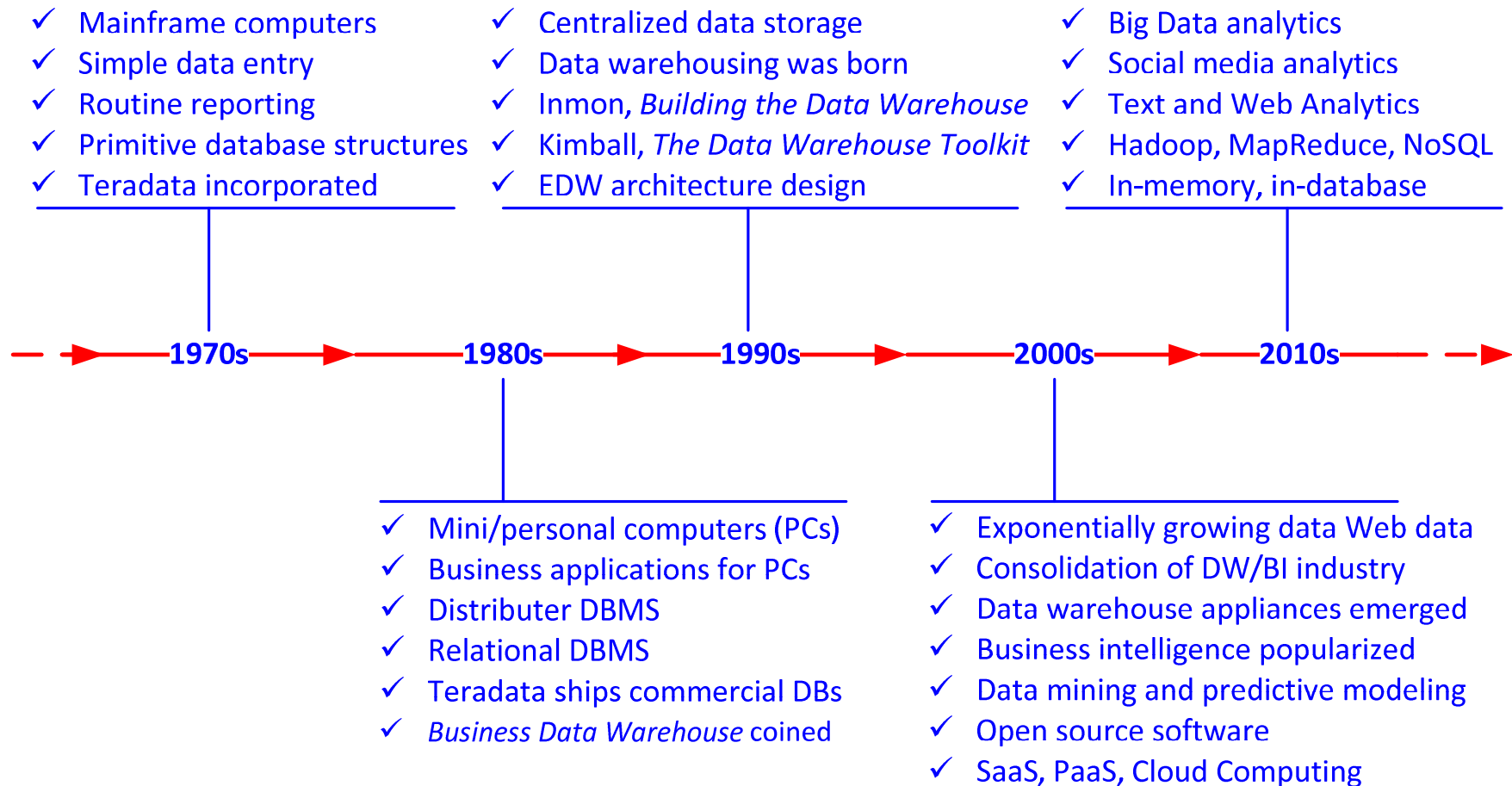
- BI used to be everything related to use of data for managerial decision support
- Now, it is a part of Business Analytics
  - BI = Descriptive Analytics



# What is a Data Warehouse?

- A physical repository where relational data are specially organized to provide enterprise-wide, cleansed data in a standardized format
- A relational database? (so what is the difference?)
- “The data warehouse is a collection of integrated, subject-oriented databases designed to support DSS functions, where each unit of data is non-volatile and relevant to some moment in time”

# A Historical Perspective to Data Warehousing



# Characteristics of DWs

- Subject oriented
- Integrated
- Time-variant (time series)
- Nonvolatile
- Summarized
- Not normalized
- Metadata
- Web based, relational/multi-dimensional
- Client/server, real-time/right-time/active...



# Data Mart

A departmental small-scale “DW” that stores only limited/relevant data

- **Dependent data mart**

A subset that is created directly from a data warehouse

- **Independent data mart**

A small data warehouse designed for a strategic business unit or a department

# Other DW Components

- **Operational data stores (ODS)**
  - A type of database often used as an interim area for a data warehouse
- **Oper marts**
  - An operational data mart
- **Enterprise data warehouse (EDW)**
  - A data warehouse for the enterprise
- **Metadata** – “data about data”
  - In DW metadata describe the contents of a data warehouse and its acquisition and use

# Application Case 3.1

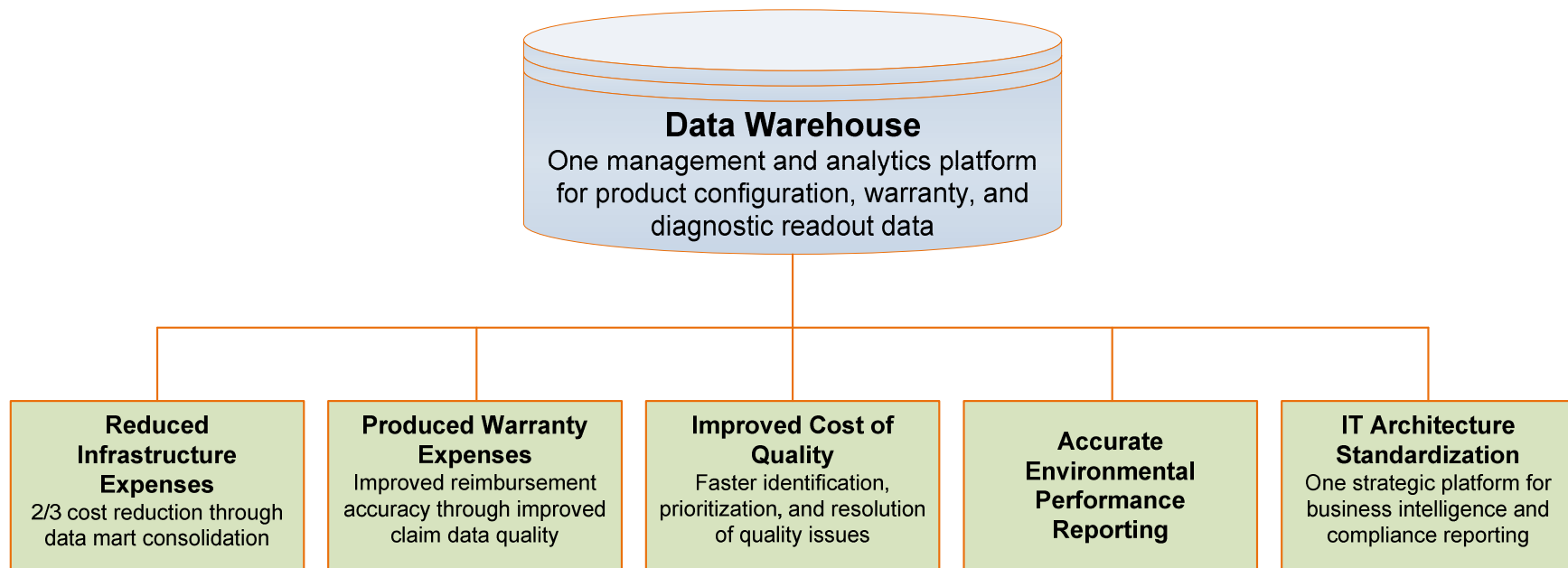
## A Better Data Plan: Well-Established TELCOs Leverage Data Warehousing and Analytics to Stay on Top in a Competitive Industry

### Questions for Discussion

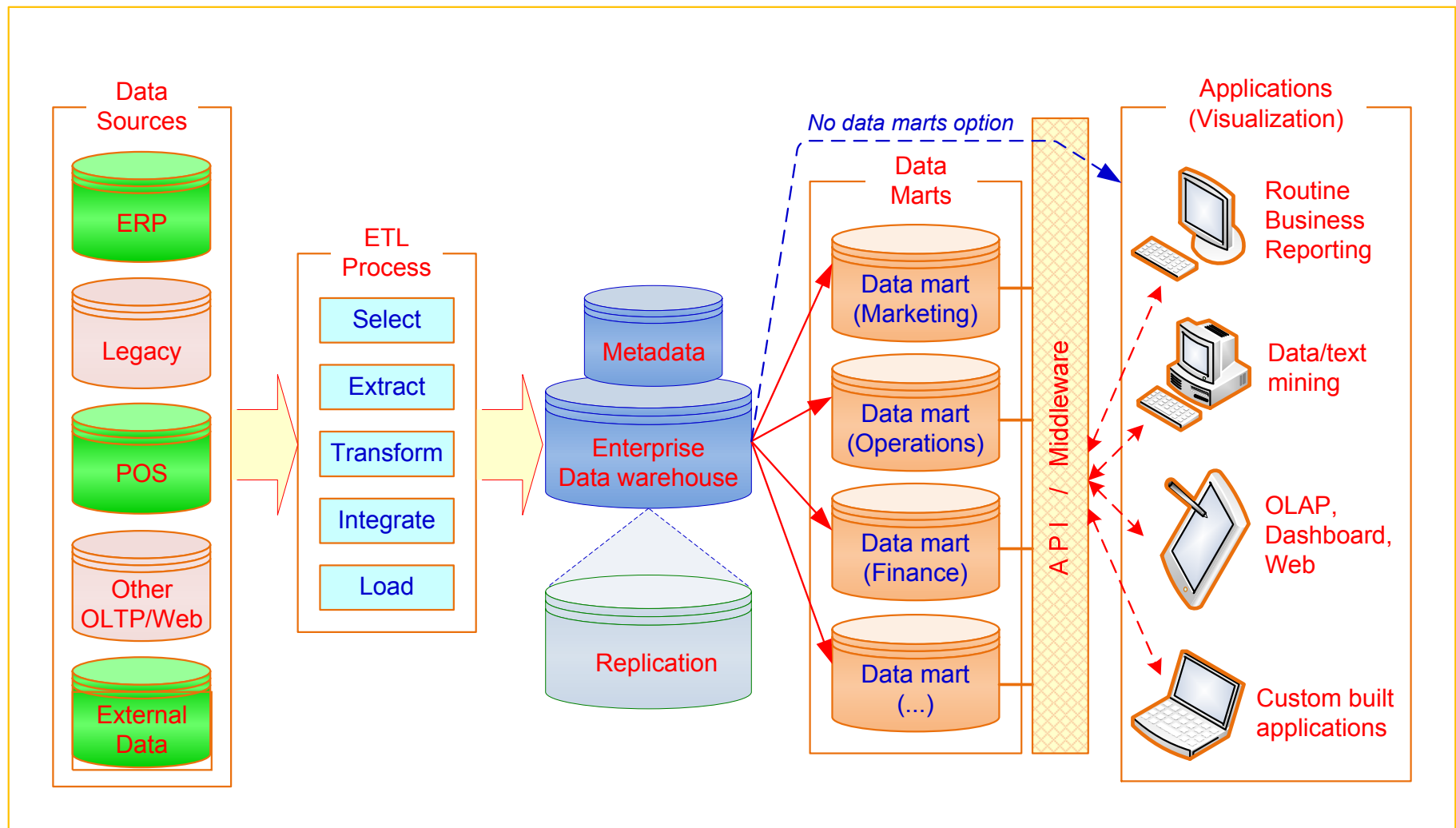
1. What are the main challenges for TELCOs?
2. How can data warehousing and data analytics help TELCOs in overcoming their challenges?
3. Why do you think TELCOs are well suited to take full advantage of data analytics?

# DW for Data-Driven Decision Making

- An example of a DW supporting data-driven decision making in automotive industry



# A Generic DW Framework

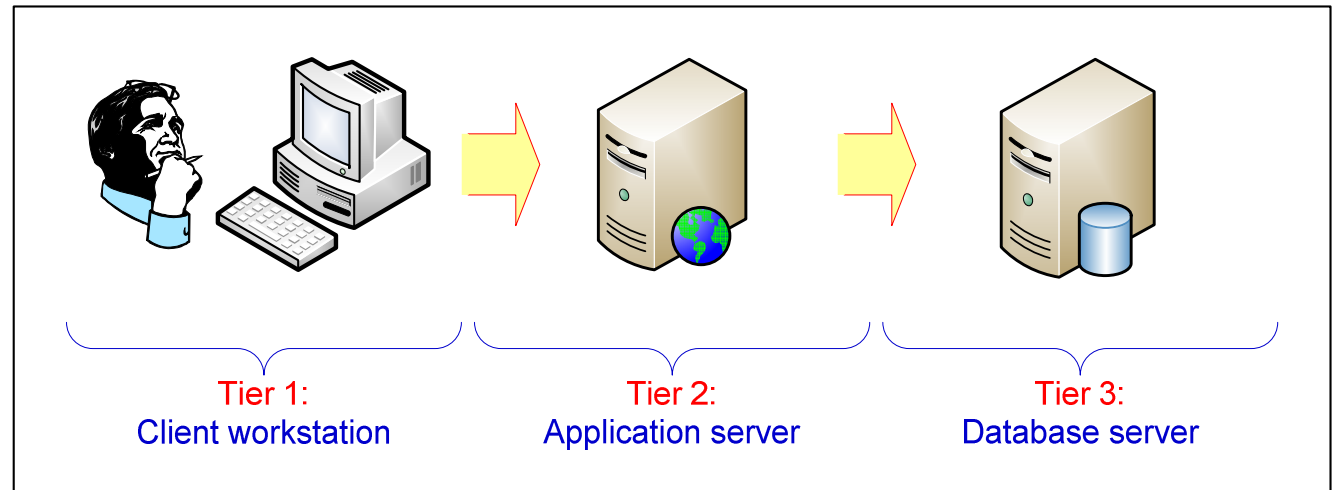


# DW Architecture

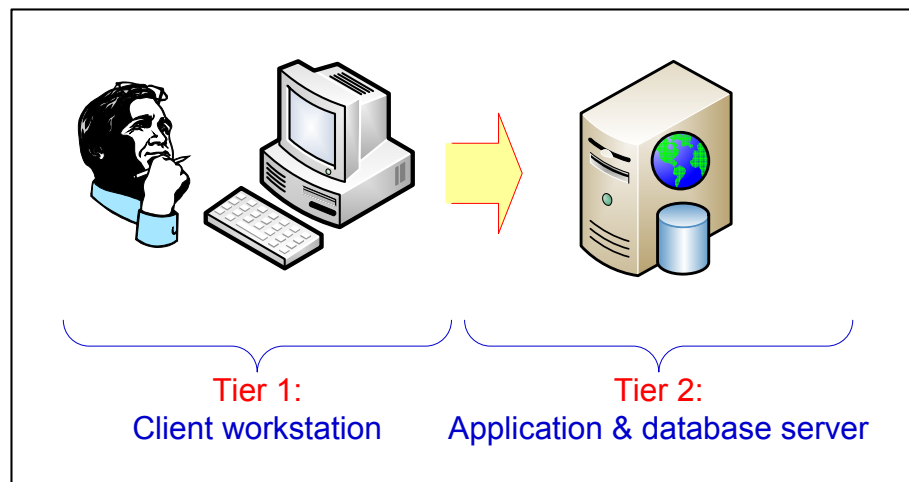
- **Three-tier architecture**
    1. Data acquisition software (back-end)
    2. The data warehouse that contains the data & software
    3. Client (front-end) software that allows users to access and analyze data from the warehouse
  - **Two-tier architecture**
    - First two tiers in three-tier architecture are combined into one
- ... sometimes there is only one tier?

# DW Architectures

## 3-tier architecture



## 2-tier architecture



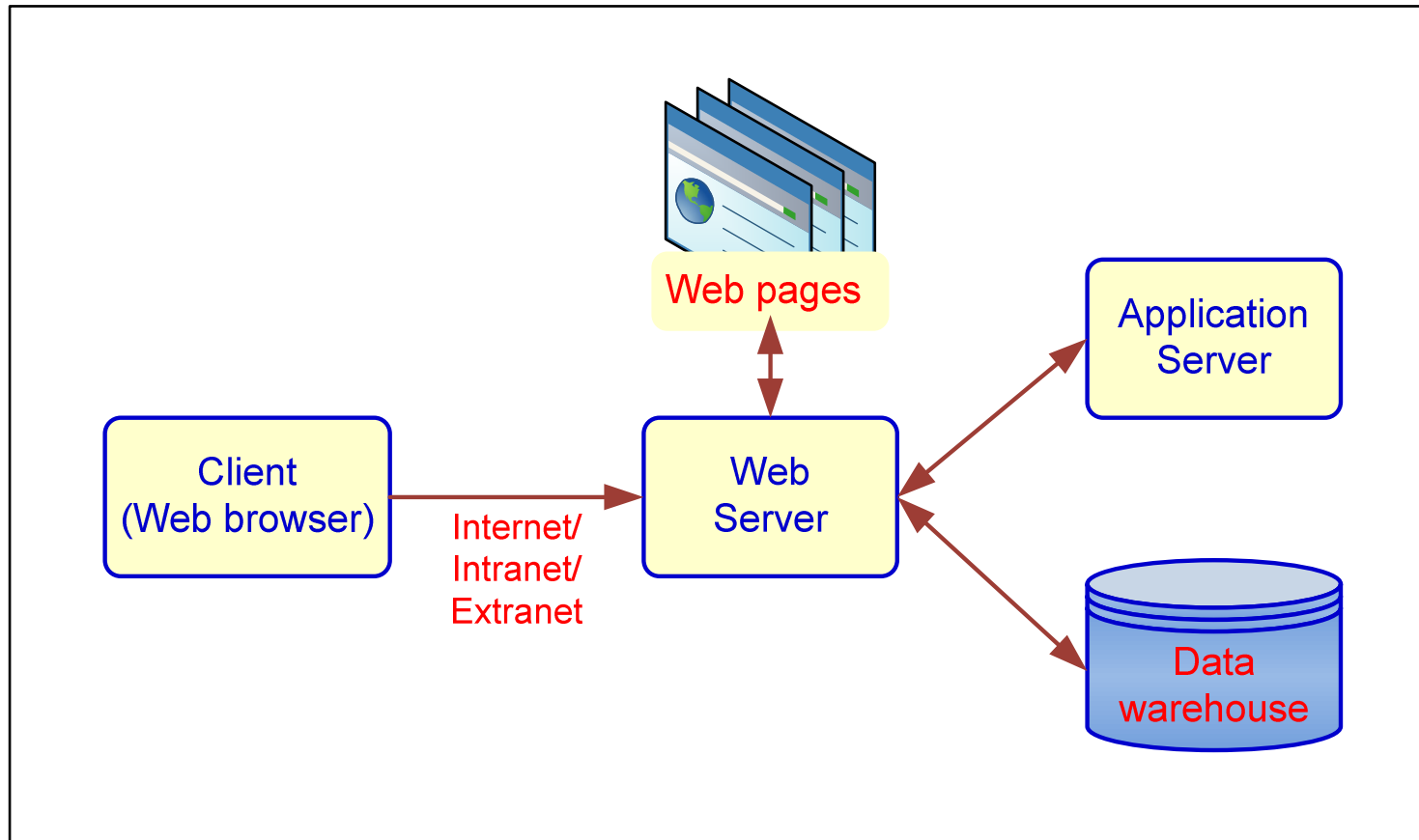
1-tier  
Architecture  
?

# Data Warehousing Architectures

- Issues to consider when deciding which architecture to use:
  - Which database management system (DBMS) should be used?
  - Will parallel processing and/or partitioning be used?
  - Will data migration tools be used to load the data warehouse?
  - What tools will be used to support data retrieval and analysis?

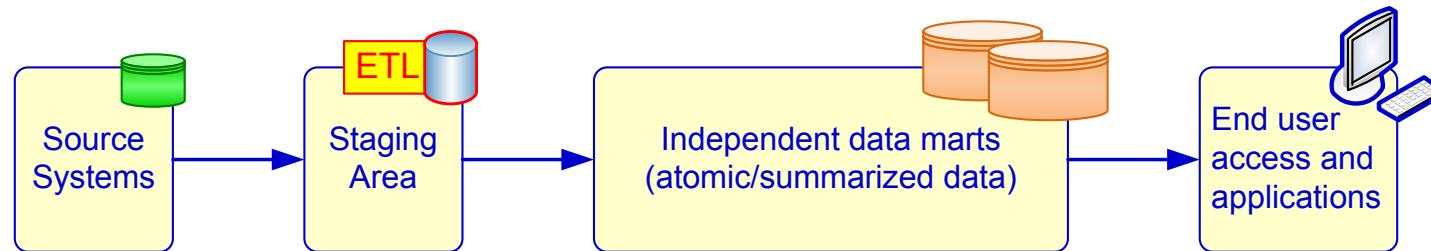


# A Web-based DW Architecture

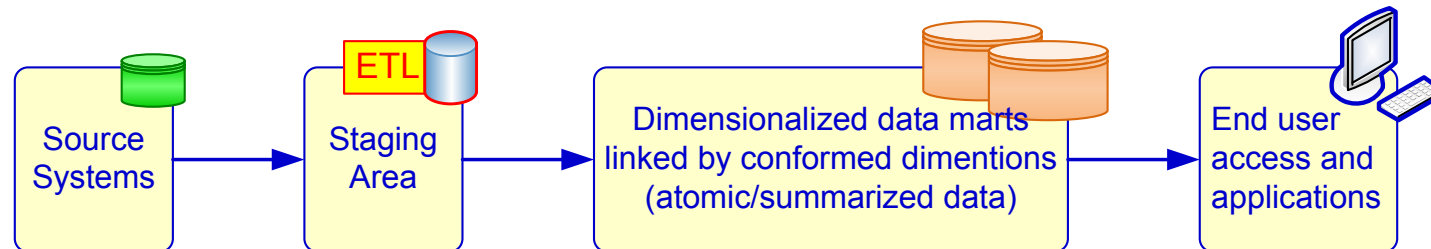


# Alternative DW Architectures

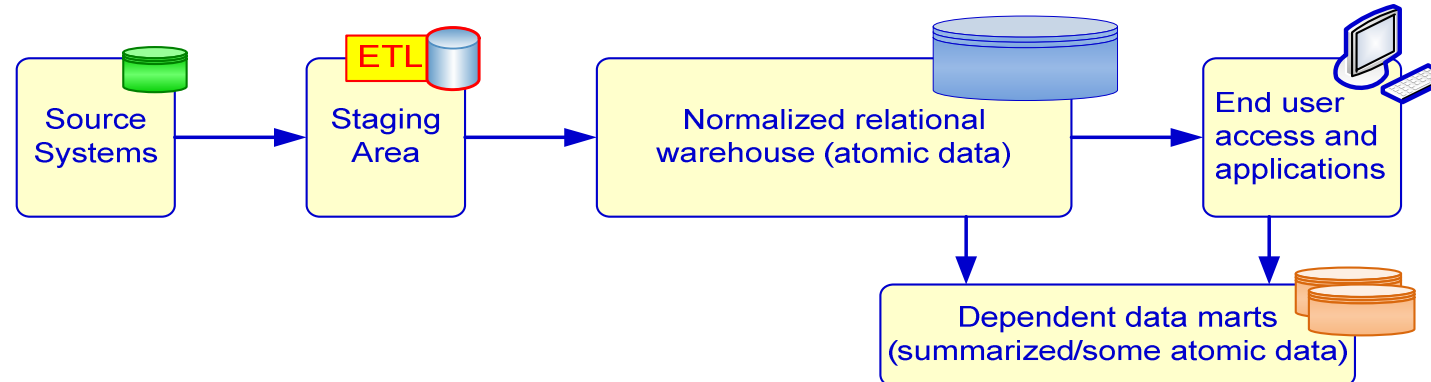
(a) Independent Data Marts Architecture



(b) Data Mart Bus Architecture with Linked Dimensional Datamarts

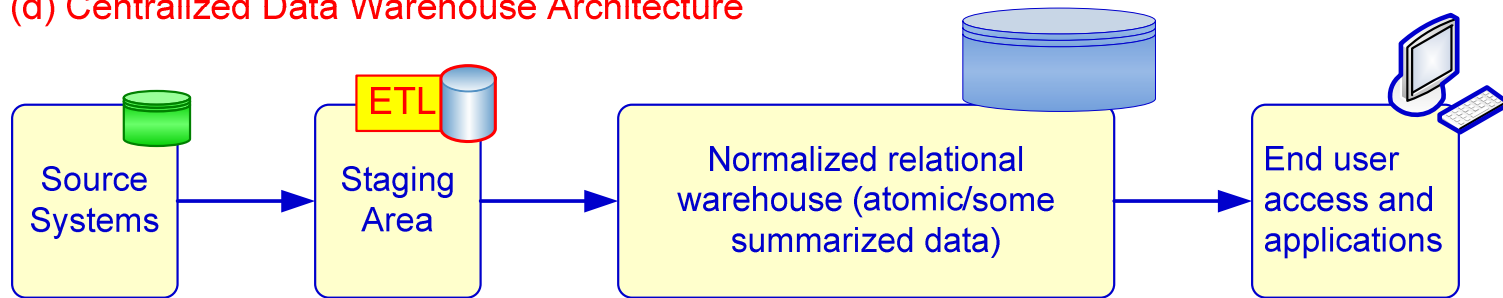


(c) Hub and Spoke Architecture (Corporate Information Factory)

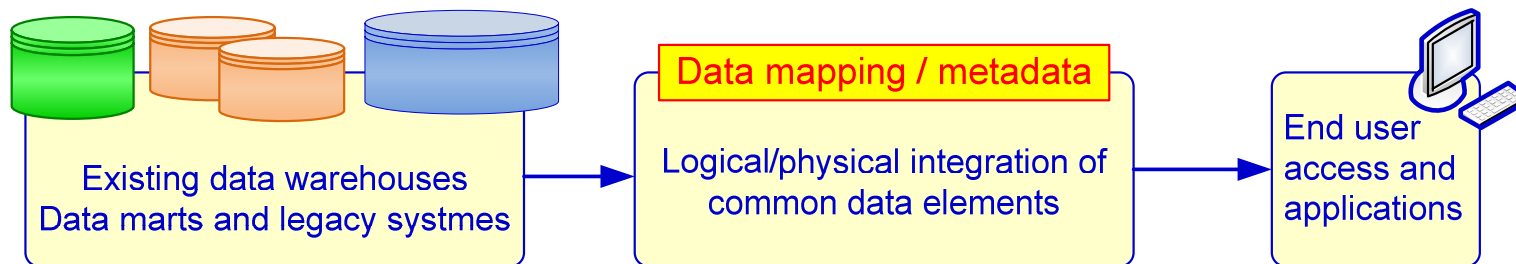


# Alternative DW Architectures

(d) Centralized Data Warehouse Architecture



(e) Federated Architecture



- Each architecture has advantages and disadvantages!
- Which architecture is the best?

# Ten Factors that Potentially Affect the Architecture Selection Decision

1. Information interdependence between organizational units
2. Upper management's information needs
3. Urgency of need for a data warehouse
4. Nature of end-user tasks
5. Constraints on resources
6. Strategic view of the data warehouse prior to implementation
7. Compatibility with existing systems
8. Perceived ability of the in-house IT staff
9. Technical issues
10. Social/political factors