Business Intelligence

• Business intelligence (BI) is a set of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information…

  data → information → decisions → actions

• Term coined by the Gartner Group – early 1990s

How is this different from everyday business decision making?
Managing information

• Master Data Management (MDM)
  – Provides for a comprehensive and consistent definition of all data in an organization
  – Ensures uniform and consistent views of all data

• Supports proper Governance
  – For controlling and monitoring business health
  – Creates accountability

• What is needed in order to monitor a business’ health?

Monitoring business health

• Setting goals

• Defining Key Performance Indicators (KPI)
  – Quantifiable measurements that assess a company’s effectiveness in reaching its goals

Examples:
  (General):
  (Financial):
  (Human resources):

Business Intelligence components

• What tools are needed (in general) to monitor business health and support strategic decision making?

• Typical major components:
  – Data extraction, transformation, and loading tools
  – Data store (data warehouse or data mart)
  – Data query and analysis tools (OLAP)
  – Data presentation and visualization tools (dashboard)

• Major software vendors
  – Microsoft, Oracle, SAP, IBM
Different types of data

- **Operational data (Transactional databases)**
  - Stored in highly normalized tables in a relational database
  - Dynamically updated
  - Focus of traditional information systems

- **Decision support data (Data warehouses)**
  - Stored in formats that facilitate data extraction, data analysis, and decision making
  - Often aggregated
  - Often with redundancies

General steps of BI

1. Collecting and storing operational data
2. Aggregating the operational data into decision support data
3. Analyzing the decision support data to generate information
4. Presenting the information to the end user to support decision-making
5. Making business decisions (and generating more data)
6. Monitoring results to evaluate outcomes of the business decisions

Data warehouse

- A **data warehouse** is a database optimized for data analysis and read-only query processing
  - The prime objective of a data warehouse is to provide fast and accurate answers to data analysis queries

- Data are typically extracted from various sources and then transformed and integrated into the data warehouse

- Development of a data warehouse requires time, money, and considerable managerial effort
Data warehouse data characteristics

• Integrated
  – Consistent format and meaning

• Subject-oriented
  – Organized and summarized to answer questions

• Time-variant
  – Captures and represents the flow of data over time

• Nonvolatile
  – Once data enters the warehouse, it is never removed

Data mart

• A data mart is a small, single-subject data warehouse subset that provides decision support to a small group of people
  – Less organizational commitment required
    • Lower cost
    • Shorter implementation time
      – Data warehouse: ~ 1 to 3 years
      – Data mart: ~ 6 months to 1 year

Online Analytical Processing (OLAP)

• Graphical user interface
• Analytical processing logic
• Data-processing logic

• Advanced functionality for:
  – Data presentation
  – Data aggregation
  – Data modeling and analysis

• Capacity for multi-dimensional analysis
• Used with both transactional DBs and data warehouses

Differentiating characteristics
Multi-dimensional DBMS

- Data is stored in multi-dimensional arrays
  - Typically visualized as being stored as a data cube
- Data retrieval is much quicker than with standard relational databases
- Provides opportunity to “slice and dice” data
- Foundation for multi-dimensional OLAP
Multi-dimensional DBMS - structure

- Star schema
  - Fact tables
    - Associated with a particular type of (aggregated) data
      - Ex: sales table
  - Dimension tables
    - Attributes provide descriptive information about the facts within a given dimension
      - Ex: product, time, and location tables

Example – Multi-dimensional DB

- Fact tables
  - Sales: Sales.Quantity, Sales.Price, Sales.Total
  - Orders: Order.Quantity, Order.Price, Order.Amount

- Dimension tables
  - Location: Region.ID, Loc.State, Loc.City, Loc.Office