Data Warehousing Concepts

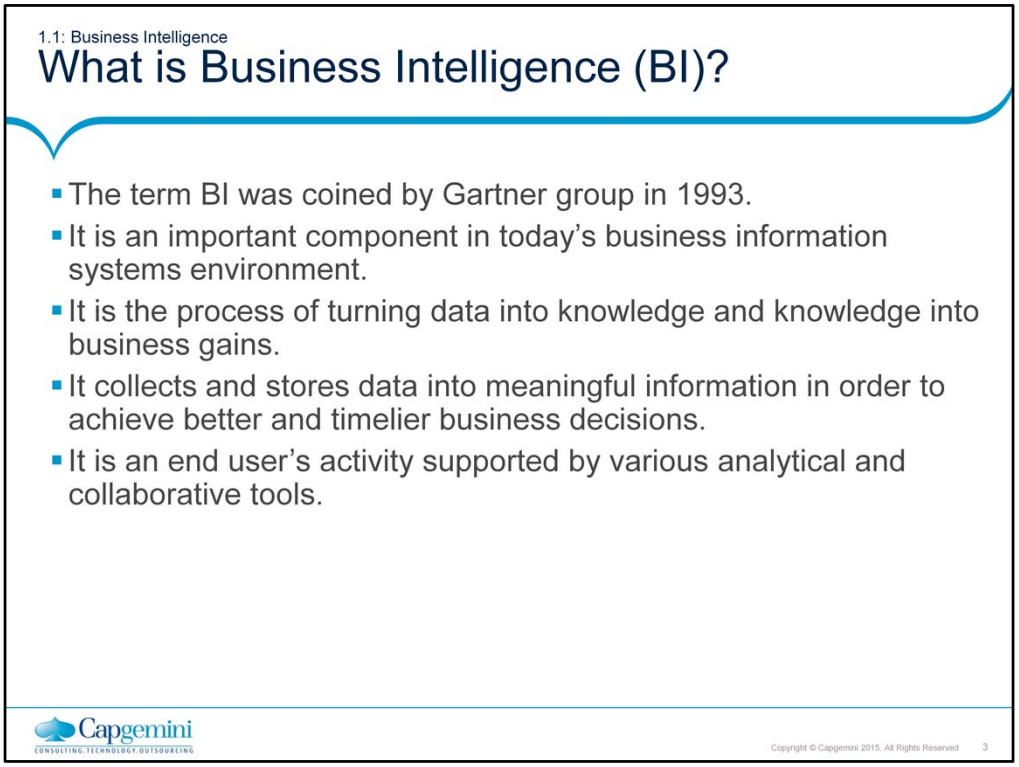
Lesson 1: Business Intelligence

Lesson Objectives

- In this lesson, you will learn:
  - What is Business Intelligence?
  - Need of Business Intelligence
  - Terms used in Business Intelligence
  - Components of Business Intelligence

Capgemini
CONSULTING.TECHNOLOGY.OUTSOURCING

 2



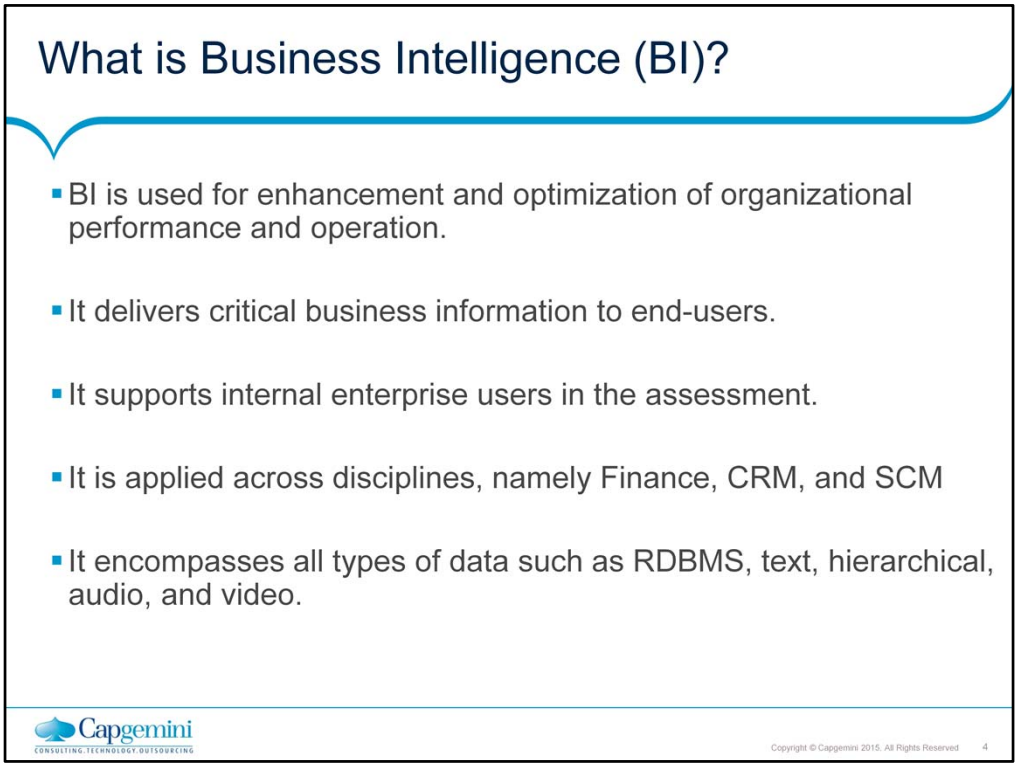
Business Intelligence:

- **Business Intelligence (BI)** is the process of getting useful information from data.

  BI is an important component in today's business information systems environment.

- As the business environment has become increasingly competitive, the need to use corporate data as a strategic resource has intensified. However, most of the organizations in technology based businesses are **data rich** and are **information poor**. Much of the essential information that is needed to anticipate changing market conditions and customer preferences is locked in various transactional systems, spread sheets, and log files. So without the ability to deliver the right information to the right people at the right time, companies cannot stay competitive in this fast changing economy. So the **BI value proposition** is a term for the ability to navigate complex sales channels by maximizing knowledge about the customer base and developing strategies that leverage that knowledge from decision to action.

- **BI applications** are decision support tools that enable real-time, interactive access, analysis, and manipulation of mission-critical corporate information.



Business Intelligence:

- Business Intelligence gives answers to the questions such as given below:
  - Who are my top ten customers?
  - How effective was my last sales campaign?
  - Who is my best sales person by volume, and by dollar revenue, per region, during the last week of each month? How does that compare with last year?
  - How much more intelligent can you make your business processes?
  - How much more insight can you gain into your business?
  - How much more integrated can your business processes be?
  - How much more interactive can your business be with customers, partners, employees and managers?
  - BI solutions answer all these questions.

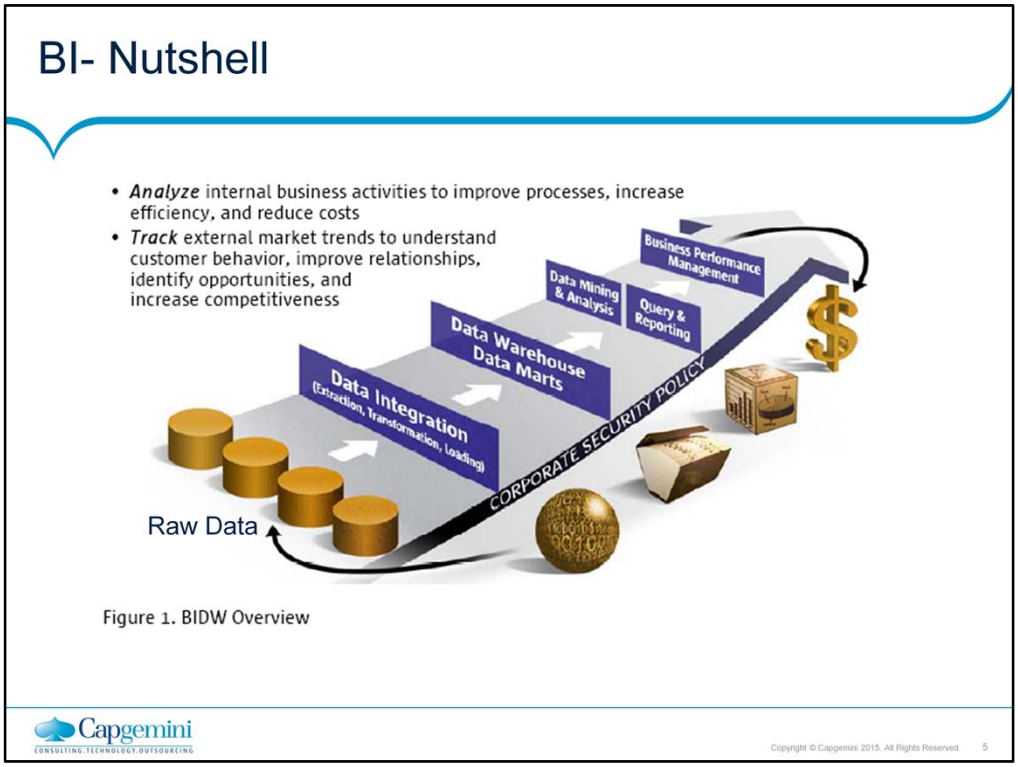


Figure 1 illustrates the major components of a BI system and the process of generating business results from raw data (the operational data that is used to run the business).

The BIDW process can be broken down into the following steps:
• Raw data is stored: Raw data is typically stored, retrieved, and updated by an organization's on-line transaction
processing (OLTP) system. Additional data that feeds into the data warehouse may include
external and legacy data that is useful to analyze the business.
• Information is cleansed and optimized: The information is then cleansed (for example, all duplicate items are removed) and optimized for decision support applications (i.e. structured for queries and analysis vs. structured for transactions). It is usually "read only" (meaning no updates allowed) and stored on separate systems to lessen the impact on the operational systems.

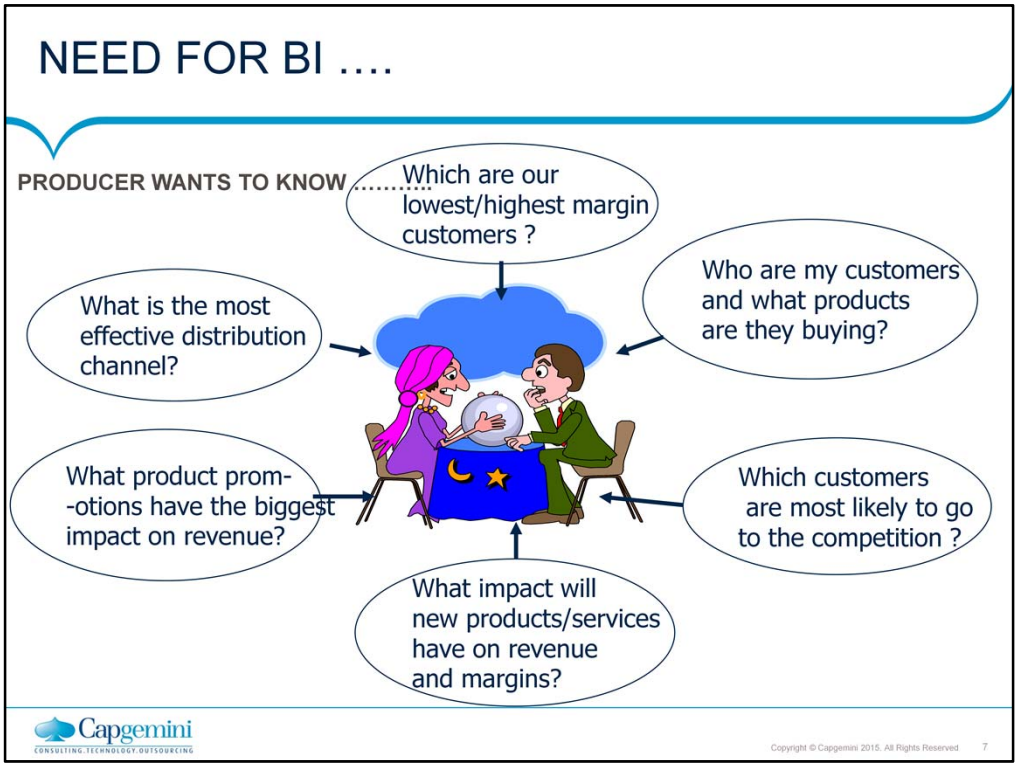
• Data mining, query and analytical tools generate intelligence: Various data mining, query and analytical tools generate the intelligence that enables companies to spot trends, enhance business relationships, and create new opportunities.

• Organizations use intelligence to make strategic business decisions: With this intelligence, organizations can make effective decisions, and create strategies and programs for competitive advantage.

• The system is regulated by an overall corporate security policy o Information in a data warehouse is typically confidential and critical to a company's business operations. Consequently, access to all functions and contents of a data
warehouse environment must be secure from both external as well as internal threats and should be regulated by an overall, corporate security policy.

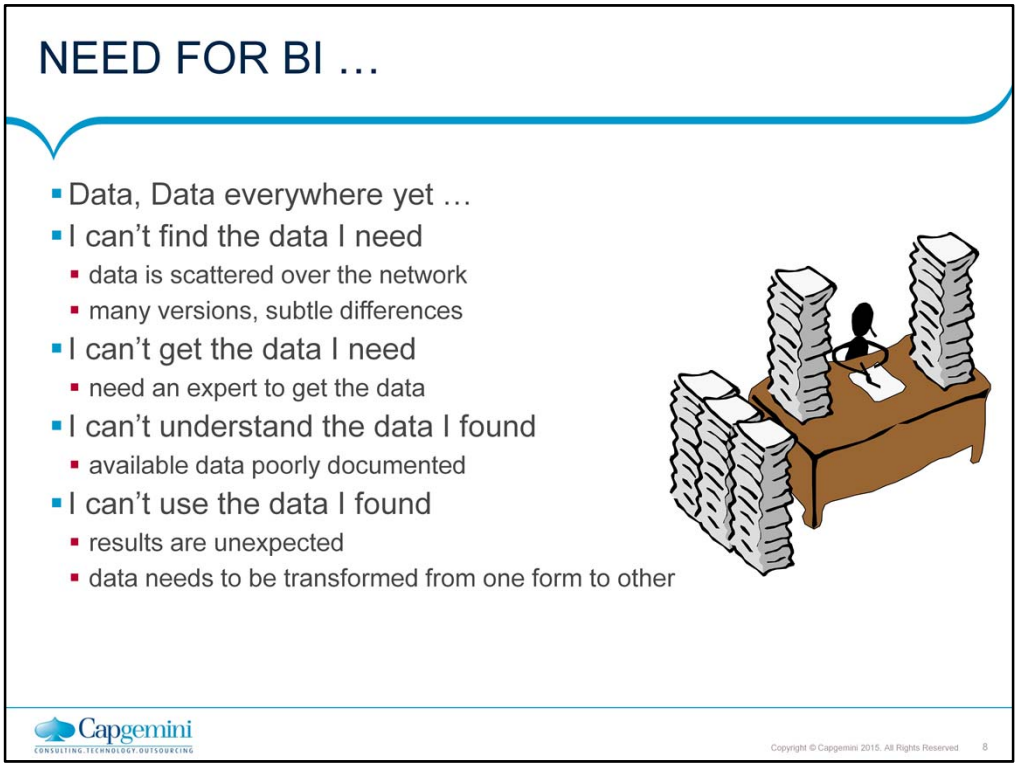
• Business performance management applications track results: A well-run BIDW operation also includes Business Performance Management (BPM) applications, which help track the results of the decisions made and the performance of the programs created.

integration.intelligence.insight



If u see some of the top management questions such as
what is the most effective distribution chennel? Or
who are our customers and what products are they buying.

These questions will help top management to take long range decision. Hence, regular operational system does not answer above questions because of regular database consists of only current information. So if at all we need to answer above question we need to have huge amount of historical and detailed data in the database which is integrated from several sources.



In most of the organization either it is big or small some of the problems are common such as;

May not be able to find required data since the data which is maintained are geographically scattered and whatever they are available it is in different version or format.
eg: The data might be maintained in excell sheet. Or The data might be maintained using any database such as oracle, db-2 or sql etc or It might be in just word documents.

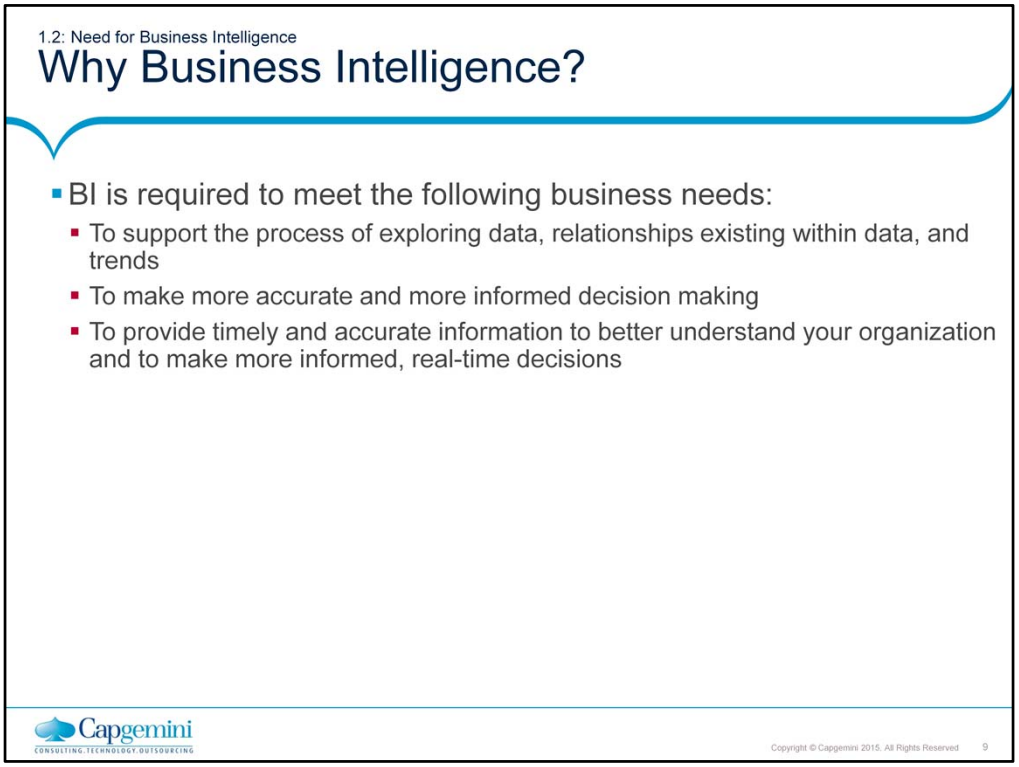
2. May not be able to analyze the data due to lack of expertise in the organization

3. May not be able to fetch properly because of data is maintained poorly (Might be maintained in unstructured way eg. using note pad or word document)

4. The result of all problems leads unexpected results hence these unstructured data, maintained in different formats need to transform into single format so that it will help top management to take strategic decision.

In order to resolve above problems the only solution is DataWarehouse. The Data warehouse is called single version of truth in which different sources of data is captured and stored in a single place.

For this, The Data Warehouse Motivation is
Huge amounts of data need to be summarized in various forms to enable data creators and data users to get quick overviews and dig into details as needed with high performance and flexibility



Need for Business Intelligence:

- **BI exhibits the following utility features:**
  - BI is a general term for applications, platforms, tools and technologies that support the process of exploring data, relationships existing within data, and trends.
  - BI is important in helping organizations to stay ahead of the competition by providing the means for quicker, more accurate, and more informed decision making.
  - BI provides timely and accurate information to better understand the organization and to make more informed, real-time decisions.

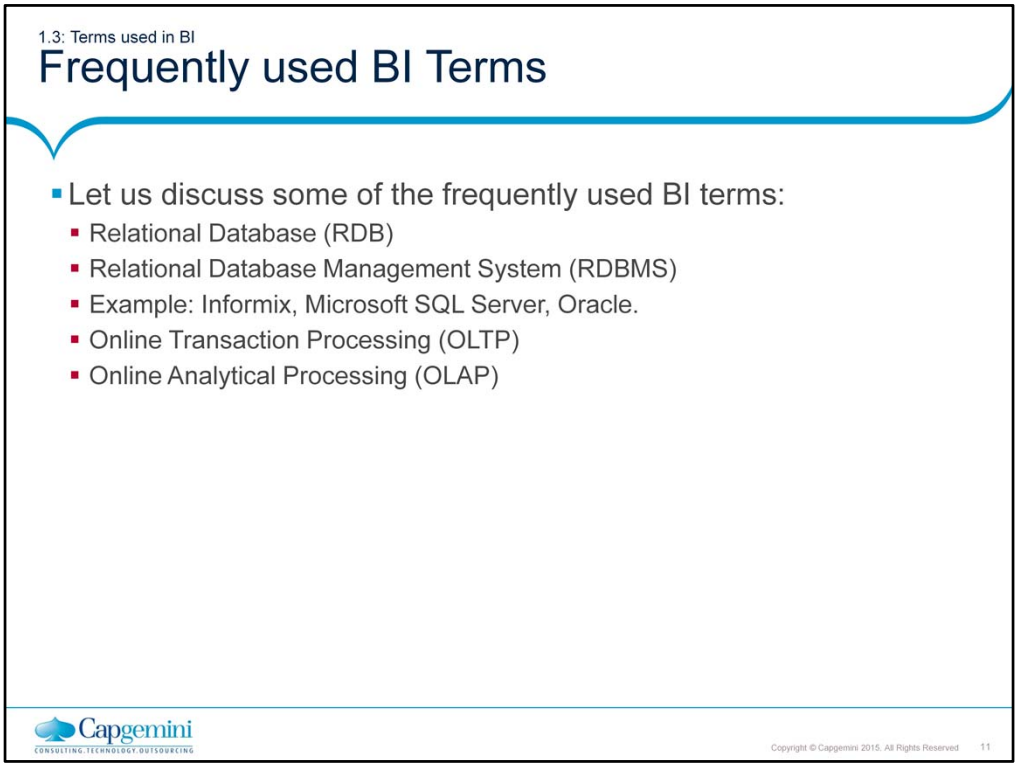
- **But why do you need Business Intelligence?**
  - For many years, database vendors have focused on getting data into a database. The emphasis has led to great achievements in **online transaction processing** and capacity. Many companies have accumulated data that can be measured in gigabytes, terabytes, and even petabytes.
  - **Transactional data,** which is the data that is used to run the business, is good for keeping track of what is happening in an organization. However, it is not well suited to finding out why things are happening or predicting future performance.
  - Hence there arises a strong need for BI applications.

Why Business Intelligence?

- Data Analysis is a huge and crucial part of Business Intelligence.
- Many organizations need to know the overall performance and the way its business is functioning.
- BI is used to gather past as well as present data.
- Modern BI systems are capable of managing large amount of unstructured data.

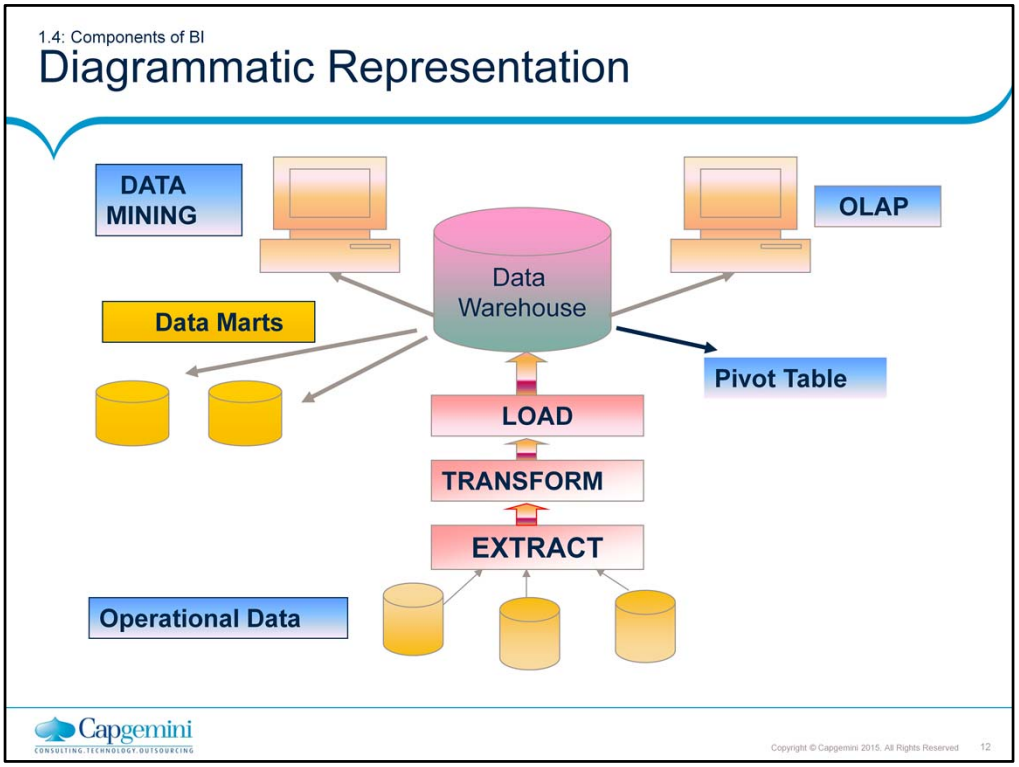
Capgemini
CONSULTING.TECHNOLOGY.OUTSOURCING

 10



Terms used in BI:

- **Relational Database (RDB):**
  - It is a database that conforms to the relational model.
- **Relational Database Management System (RDBMS):**
  - It refers to the software used to create a RDB.
  - Example: Informix, Microsoft SQL Server, Oracle
- **Online Transaction Processing (OLTP):**
  - OLTP is a process which is used for day to day transaction processing.
  - **Example:** Operational systems, High volume data collection
- **Online Analytical Processing (OLAP):**
  - This processing method provides fast access to shared multidimensional data.
  - It is used to generically refer to software and applications that provide users with the ability to store and access data multi-dimensionally.



Components of BI:

Following are the various components of BI:

1. **Operational Data:** Typically data is sourced from transaction processing systems. It is also called as Data Source. Typically data is sourced from transaction processing systems (Manufacturing, ERP, Sales). Example: Customer, Inventory, Credit, Sales, Operation and External are the data source.
2. **ETL Tools:**
   - **Extract:** It is the process of pulling the data from external and operational data sources in order to source data for the data warehouse.
   - Transform: It is the process that converts data to the format required by data warehouse. It cleanses data to ensure accuracy. It validates primary keys against defined owner. It converts to different numbering schema.
   - Load: It is the process that loads data to data warehouse. It follows guidelines as outlined by the data warehouse.
3. **DWH:** Data Warehouse integrates and aggregates data from various operational and external database maintained by different Business Units.
4. **Data Mart:** Data mart is a repository of data collection from operational data source and other sources that are designed to serve a particular community of knowledge workers.
5. **Reports:** A report presents the data in a format understandable by the end user.

Components of BI:

Following are the various components of BI (contd.):

6. **OLAP:** OLAP is a category of software technology that enables the users to gain insight into data through fast, consistent, interactive access to a wide variety of possible views of information.
7. **Pivot Table:** A pivot table is the simplest tool to aggregate data by creating a dimension for each field and grouping the same values in a field. A pivot table is a data summarization tool found in data visualization programs such as spreadsheets. It allows you to reorganize and summarize selected columns and rows of data in a spreadsheet or database.

Summary

- In this lesson, you have learnt:
  - BI helps to extract information from data.
  - BI helps organizations in making real time decisions.
  - Components of BI are given below:
    - Data Warehouse
    - OLTP
    - OLAP
    - ETL tools
    - Data marts
    - Reports
    - Pivot table

Capgemini
CONSULTING.TECHNOLOGY.OUTSOURCING

 14

Review Questions

- Question 1: This a huge and crucial part of Business Intelligence.
  - Option 1: Data collection
  - Option 2: Data analysis
  - Option 3: Data availability

- Question 2: OLAP Analysis is not the part of BI presentation.
  - True / False
- Question 3: ___ operation converts data to format required by data warehouse.

Capgemini
CONSULTING.TECHNOLOGY.OUTSOURCING

 15

Review Question: Match the Following

1. pulling the data from external and operational data sources	A. OLAP Analysis
2. part of BI presentation	B. RDB
3. Software for relational database	C. Extract
	D. OLTP
	E. RDBMS

Capgemini
CONSULTING.TECHNOLOGY.OUTSOURCING

 16