Subject: MCA DMBI 1003

Unit 5

Business Intelligence

Made By: Priyanka Karale

Subject: MCA DMBI 1003

Introduction to Business Intelligence (BI)

Business intelligence (BI) is a technology-driven process for analyzing data and delivering actionable information that helps executives, managers and workers make informed business decisions.

Business intelligence combines business analytics, data mining, data visualization, data tools and infrastructure, and best practices to help organizations make more data-driven decisions.

Made By: Priyanka Karale

- BI represents the technical infrastructure that collects, stores, and analyzes company data.
- BI parses data and produces reports and information that help managers to make better decisions.
- Software companies produce BI solutions for companies that wish to make better use of their data.
- BI tools and software come in a wide variety of forms such as spreadsheets, reporting/query software, data visualization software, data mining tools, and online analytical processing (OLAP).
- Self-service BI is an approach to analytics that allows individuals without a technical background to access and explore data.

Types of BI Tools and Software

BI tools and software come in a wide variety of forms. Let's take a quick look at some common types of BI solutions.

- **Spreadsheets:** Spreadsheets like Microsoft Excel and Google Docs are some of the most widely used BI tools.
- **Reporting software:** Reporting software is used to report, organize, filter, and display data.
- **Data visualization software:** Data visualization software translates datasets into easy-to-read, visually appealing graphical representations to quickly gain insights.
- **Data mining tools:** Data mining tools "mine" large amounts of data for patterns using things like artificial intelligence, machine learning, and statistics.
- **Online analytical processing (OLAP):** OLAP tools allow users to analyze datasets from a wide variety of angles based on different business perspectives.

Definition of problem

- A problem is generally considered to be a task, a situation, or a person which is difficult to deal with or control due to complexity and in transparency.

- In everyday language, a problem is a question proposed for solution, a matter stated for examination or proof.

BI Problems

1. A lack of communication during the implementation process
2. Lack of support from the BI provider
3. Staff don’t want to adopt the solution
4. The BI software provider does not understand your industry
5. The solution is not in the cloud
6. Limited access to information due to dependence on technical staff
7. Poor data quality
8. Manual data entry
9. Integrating Data from Various Source Systems

Physical Database Design

- Physical database design in the context of business intelligence (BI) refers to the process of translating a logical data model into a physical structure that can be implemented and optimized for efficient data storage, retrieval, and processing.
- It involves making decisions about how to store and organize the data in a database management system (DBMS) to ensure optimal performance for BI applications.

Structure of Physical Database Design

1. Data Storage and Organization:

- **Table Structures**: Define tables based on the logical data model, considering factors such as normalization, denormalization, and indexing strategies.
- **Partitioning**: Divide large tables into smaller, manageable partitions to improve manageability and query performance.
- **Indexes**: Identify and create indexes on columns frequently used in queries to speed up data retrieval.
- **Clustering**: Arrange the physical order of data in a table to match the order of the clustered index, optimizing range query performance.
- **Materialized Views**: Precompute and store aggregated or complex query results to accelerate query response times.

2. File and Storage Management:

- **File Organization:** Determine how data files and indexes are organized on storage devices for optimal read/write performance.
- **Compression:** Implement data compression techniques to reduce storage space and improve I/O performance.
- **Archiving**: Move historical or infrequently accessed data to archival storage to optimize active database performance.
- **Buffer Management**: Manage a buffer pool in memory to cache frequently accessed data, reducing the need for disk I/O operations.

3. Concurrency Control:

- **Locking Mechanisms**: Implement appropriate locking mechanisms to manage concurrent access to data, ensuring data integrity and consistency.

4. Backup and Recovery:

- **Backup Strategies**: Design backup and recovery procedures to ensure data integrity and availability in case of system failures or data corruption.
- **Point-in-Time Recovery:** Implement mechanisms to recover the database to a specific point in time, allowing for accurate historical analysis.

5. Hardware Considerations:

- **Storage Devices:** Choose appropriate storage devices (SSD, HDD) based on performance requirements and budget constraints.
- **Parallel Processing:** Utilize multi-core processors and parallel processing capabilities to enhance query performance.
- **Memory:** Ensure sufficient RAM to accommodate the database cache and optimize query performance.

6. Data Security:

- **Encryption**: Implement data encryption techniques to secure sensitive data at rest and during transmission.
- **Access Controls**: Define and enforce user access controls and privileges to restrict unauthorized access to data.

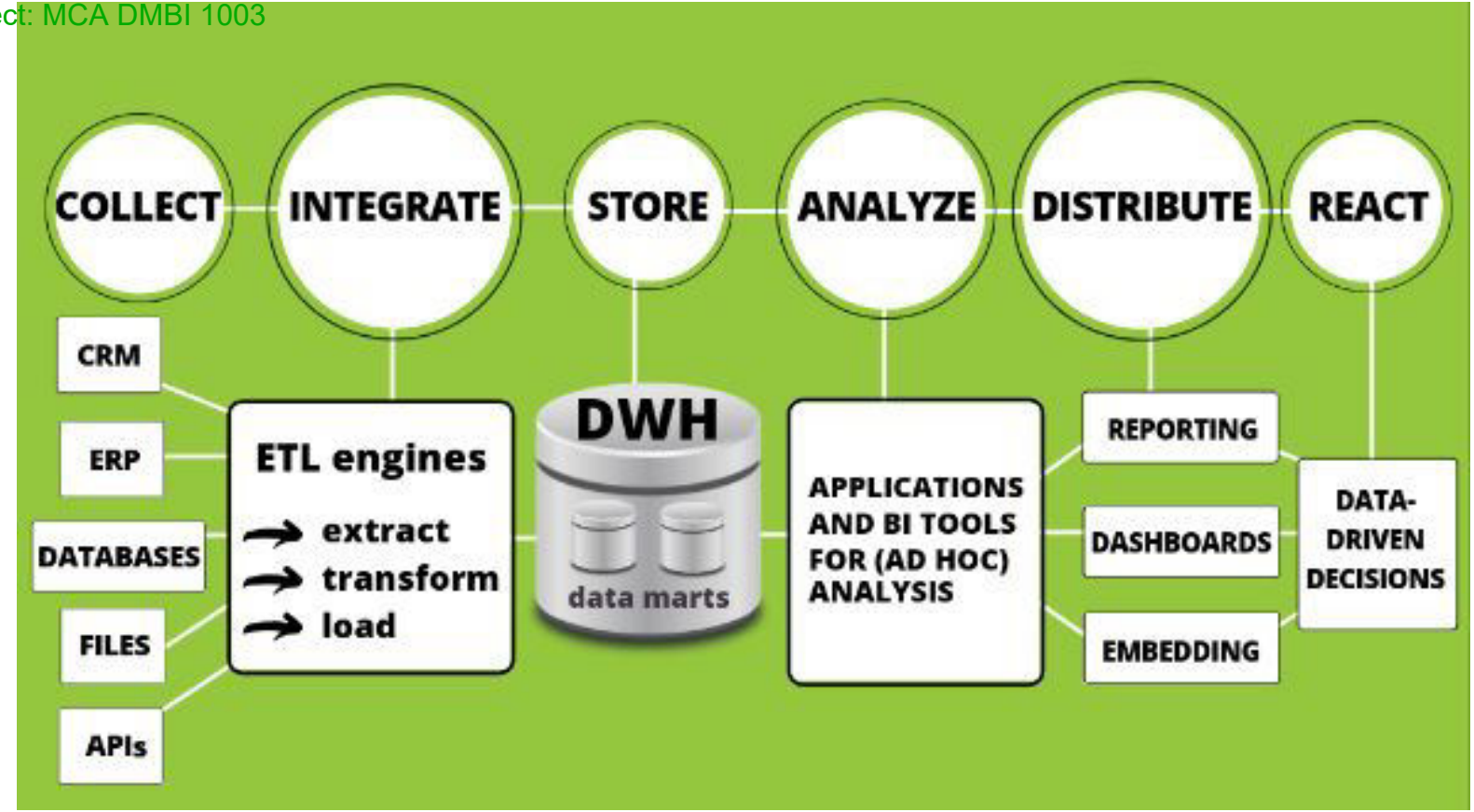
7. Performance Monitoring and Tuning:

- **Monitoring Tools**: Use monitoring tools to track database performance, identify bottlenecks, and optimize query execution.
- **Query Optimization**: Analyze query execution plans, rewrite complex queries, and create appropriate indexes to optimize query performance.

BI **architecture**

- Before you build a skyscraper, you need an architect to draw up the plans.
- Without a blueprint, construction workers would be left wandering around the site, unsure of what to do next. Situations can quickly go from bad to worse when everyone is working without a clear direction.
- The same can be said of business intelligence (BI) architecture.
- Business intelligence architecture is the framework that organizations deploy for collecting data, administering of information and supporting all the technologies of **business intelligence**.
- A successful business intelligence tool depends majorly upon a sustainable architecture.

Subject: MCA DMBI 1003

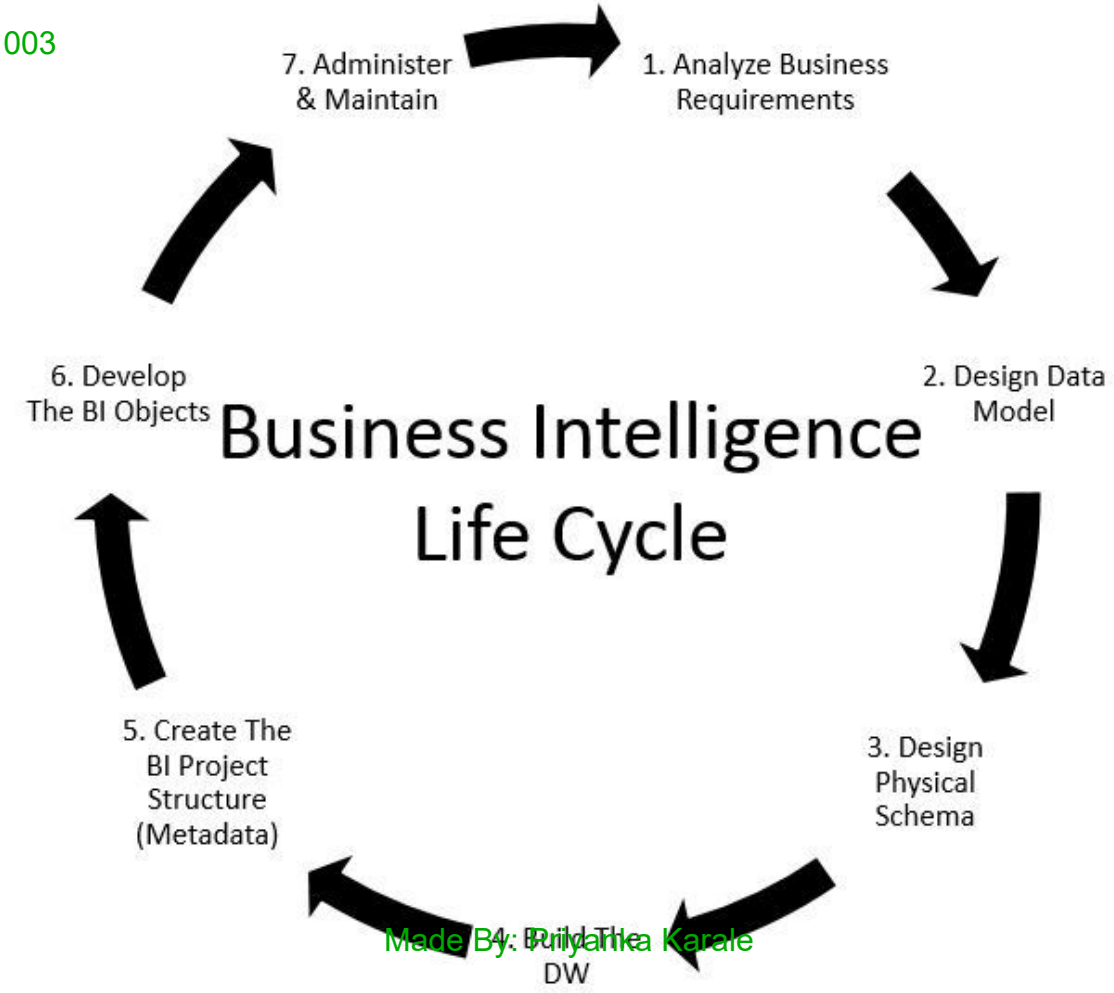


Made By: Priyanka Karale

- **Collection of data**: The first step is related to the collection of relevant data from various external and internal sources which can be databases, ERP- or CRM systems, flat files, or APIs, just to name a few.
- **Data integration**: At this stage, the data collected is integrated into a centralized system, often with the help of ETL processes. Here the data is also cleaned and prepared for analysis.
- **Storage of data:** This is where a DWH comes into the picture. A warehouse is a place in which structured data is stored. It makes it available for querying and analysis.
- **Data analysis**: After the information is processed, stored, and cleaned it is ready to be analyzed. With the help of the right tool, the data is visualized and used for strategic decision-making.

- **Distribution of data**: The data, now in the form of graphs and charts, is distributed in different formats. This can be online reporting, dashboarding, or embedding solutions.
- **Reaction based on insights**: The final stage of the architecture is to extract actionable insights from the data and use them to make improved decisions to ensure company growth.

Subject: MCA DMBI 1003



Made By: Priyanka Karale

MAJOR COMPONENTS OF BI

- DATA SOURCES
- DATA WAREHOUSES AND DATA MARTS
- BI METHODOLOGIES

Data sources:

In a first stage, it is necessary to gather and integrate the data stored in the various primary and secondary sources, which are heterogeneous in origin and type. The sources consist for the most part of data belonging to operational systems, but may also include unstructured documents, such as emails and data received from external providers.

Data warehouses and data marts:

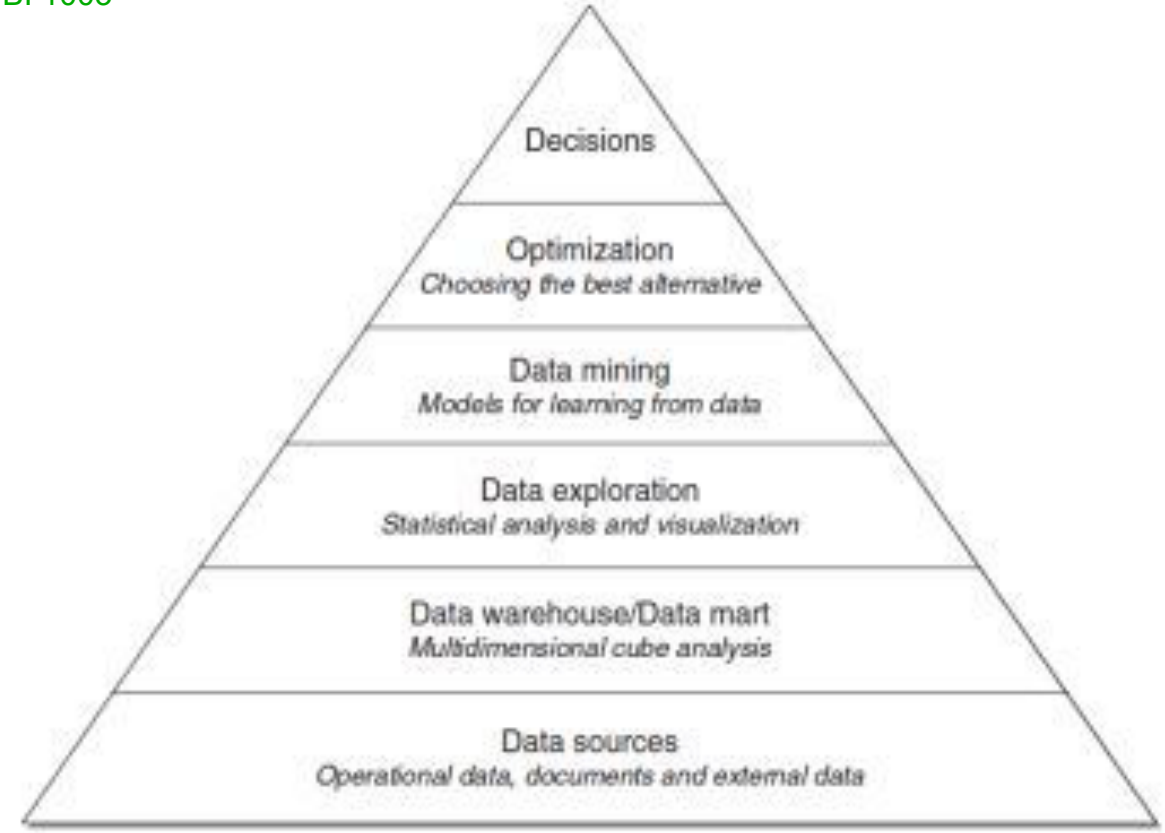
Using extraction and transformation tools known as extract, transform, load (ETL), the data originating from the different sources are stored in databases intended to support business intelligence analyses. These databases are usually referred to as data warehouses and data marts.

Business intelligence methodologies:

Data are finally extracted and used to feed mathematical models and analysis methodologies intended to support decision makers. In a business intelligence system, several decision support applications may be implemented, most of which will be described in the following sections:

- multidimensional cube analysis
- exploratory data analysis
- time series analysis
- optimization models.

Subject: MCA DMBI 1003



Made By: Priyanka Karale

Components of BI:

1. **OLAP (Online Analytical Processing):** It helps executives in sorting and selecting aggregate of data in order to monitor it strategically.
2. **Corporate Performance Management (CPM)/Advanced Analytics:** This tool helps with the statistics of certain product or service. It also helps in predicting the performance of particular product in market.
3. **Real-time BI:** Real time BI helps businesses in keeping updated with the changing marketing trends. While clients are engaged on the website, marketing team can announce special offers and grab client's attention. This is possible with real-time BI
4. **Data warehousing:** It involves storage of large amounts of data to the benefit of different divisions of an organization.
5. **Data sources:** This involves taking raw data and creating systematic data sources with the help of various software applications. BI tools put these datasets to create pie charts, graphs or tables etc.

Benefits of BI:

- Organizations can use **BI tools** at any stage to analyze, manage and visualize business data through which they can stay relevant and maximize their revenue streams. Following are the key benefits of BI:

  - Quick and accurate reporting
  - Delivering business insights that add value to the businesses
  - Insightful data analysis
  - Improving data quality
  - Enhancing operational data efficiency
  - Achieving customer satisfaction
  - Understanding market trends and improving decision making process
  - Lower margins and increasing business revenues .

Applications of Business Intelligence

Retail:

- Sales and Inventory Analysis: Analyzing sales data to optimize inventory levels and prevent overstocking or stockouts.
- Customer Behavior Analysis: Studying customer purchasing patterns to personalize marketing strategies and improve customer satisfaction.

Healthcare:

- Patient Data Analysis: Analyzing patient records and historical data to improve patient care and treatment outcomes.
- Operational Efficiency: Optimizing resource allocation, staff scheduling, and supply chain management.

Finance:

- Risk Management: Analyzing market trends and customer behavior to assess financial risks and make informed investment decisions.
- Customer Profitability Analysis: Analyzing customer data to understand the profitability of different customer segments and tailor services accordingly.

Education:

- Student Performance Analysis: Analyzing student data to identify learning patterns, assess teacher effectiveness, and improve educational outcomes.
- Institutional Planning: Analyzing enrollment trends and student demographics to plan resources and courses effectively.

What is business analytics?

A subset of BI, business analytics (BA) refers to the process of taking your company's raw data and turning it into useful information, including identifying trends, predicting outcomes, and more.

Business intelligence focuses on descriptive analytics

BI prioritizes descriptive analytics, which provides a summary of historical and present data to show what has happened or what is currently happening. BI answers the questions "what" and "how" so you can replicate what works and change what does not.

Business analytics focuses on predictive analytics

Business analytics, however, prioritizes predictive analytics, which uses data mining, modeling, and machine learning (ML) to determine the likelihood of future outcomes. BA answers the question “why” so it can make more educated predictions about what will happen. With BA, you can anticipate developments and make the changes necessary to succeed.

Subject: MCA DMBI 1003

Difference Between BI and Data Mining

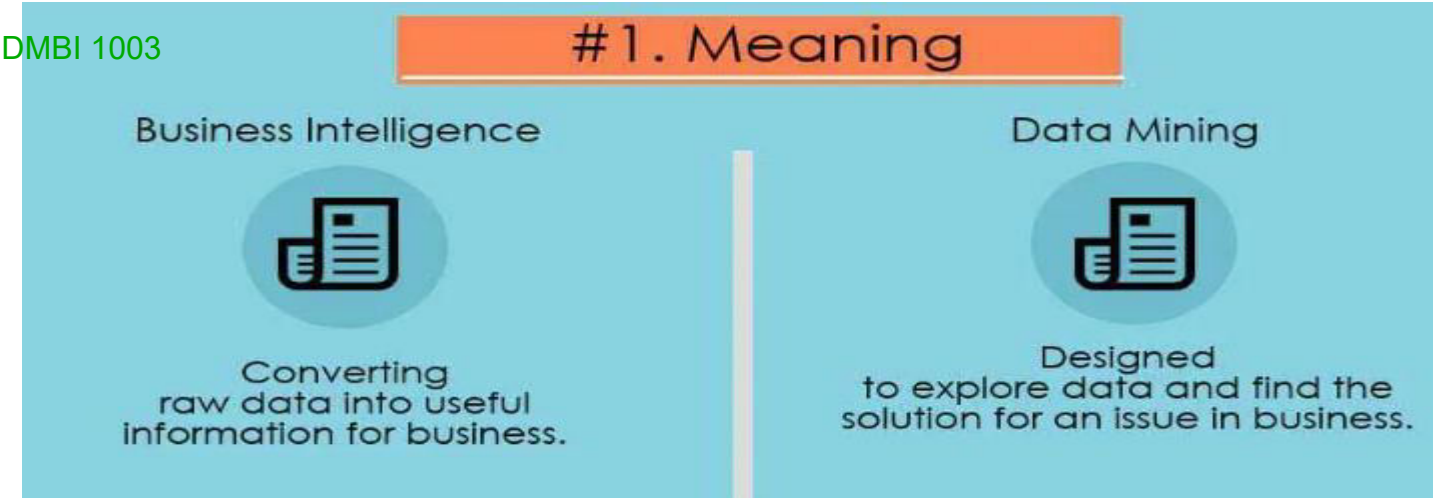
Business Intelligence and Data Mining are two different approaches to analyzing data that are used for different purposes.

BI is used for monitoring and improving business operations, while Data Mining is used for discovering patterns and relationships in data to predict future outcomes.

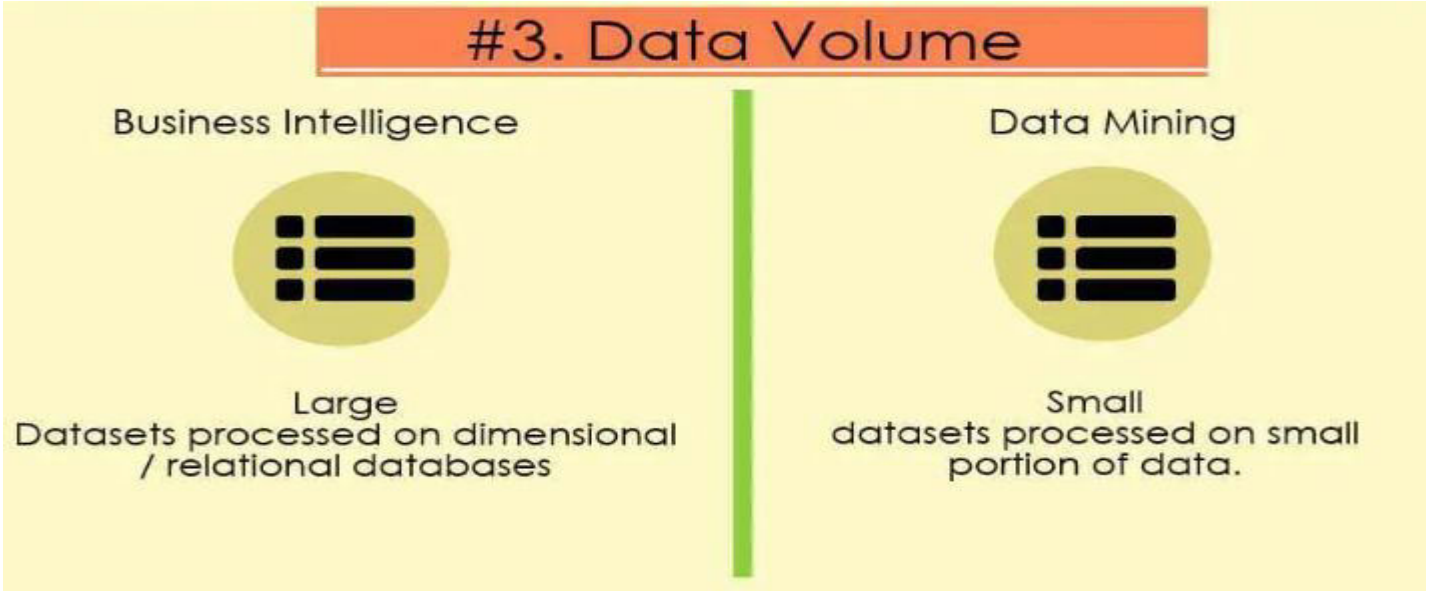
Organizations should choose the approach that best suits their needs and goals.

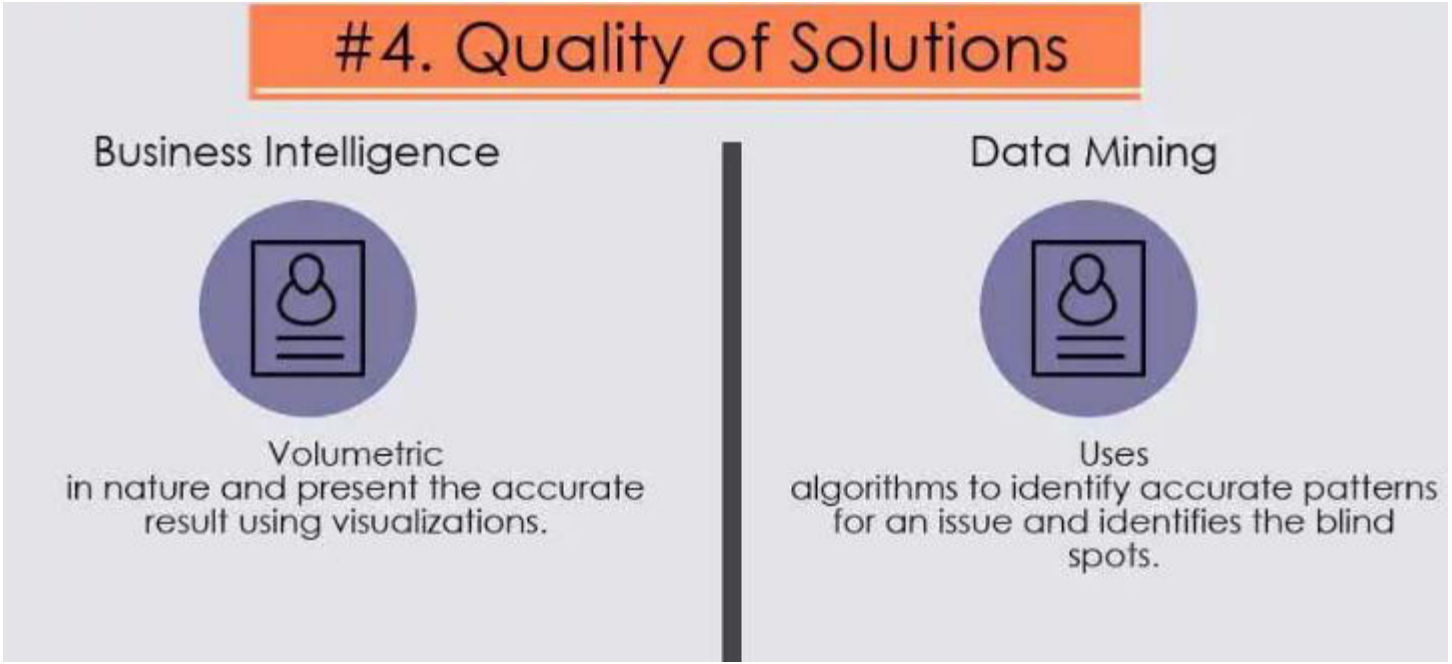
Made By: Priyanka Karale

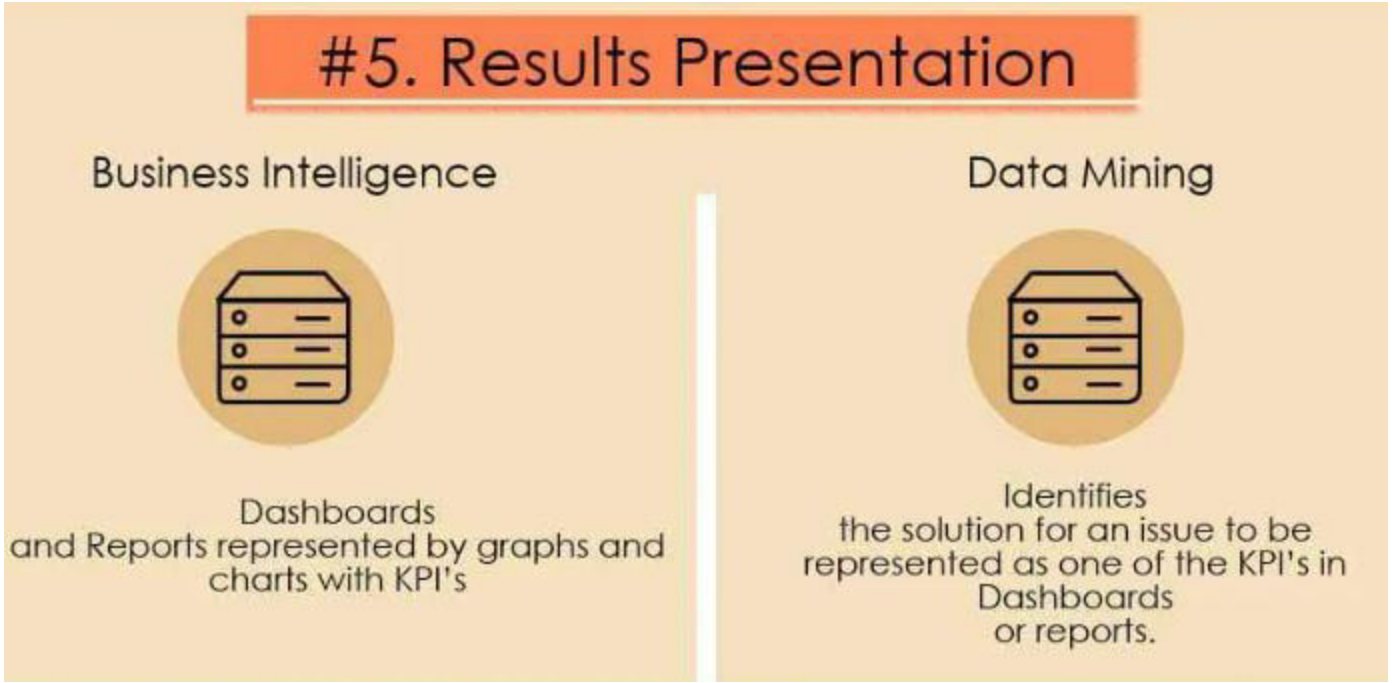
Subject: MCA DMBI 1003

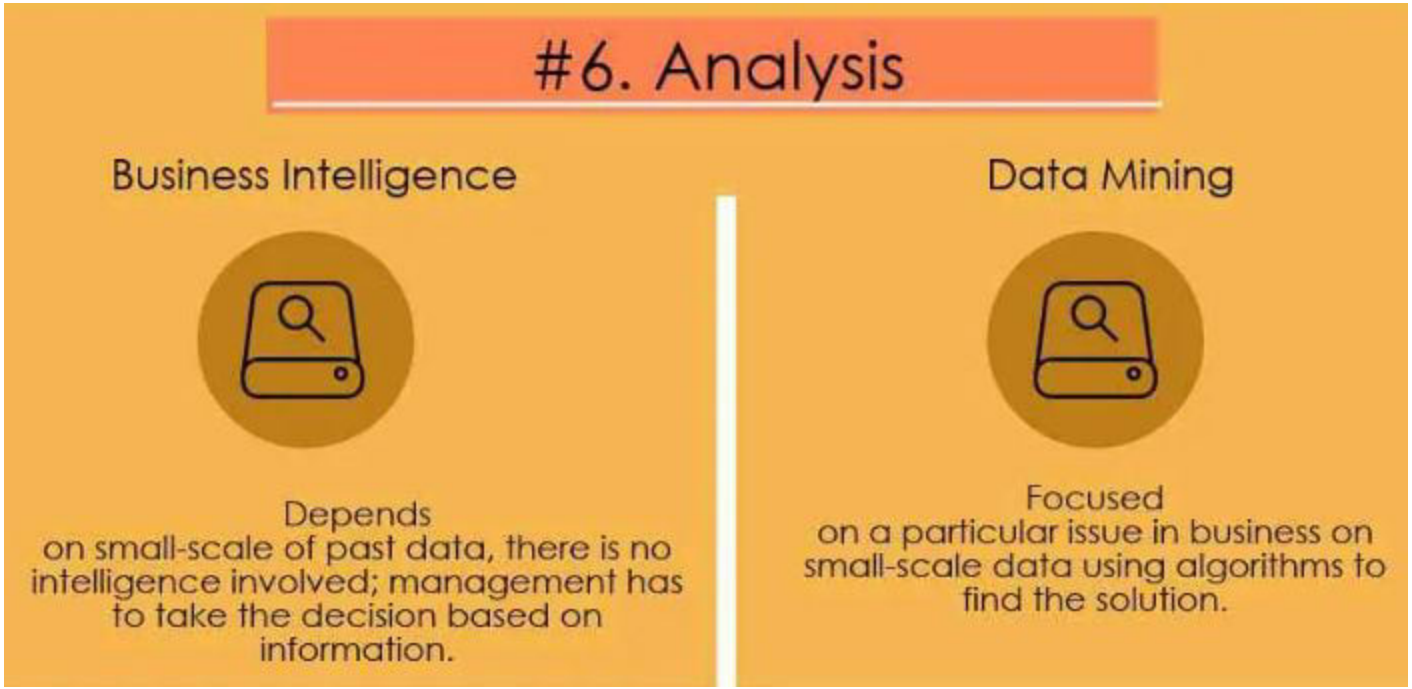


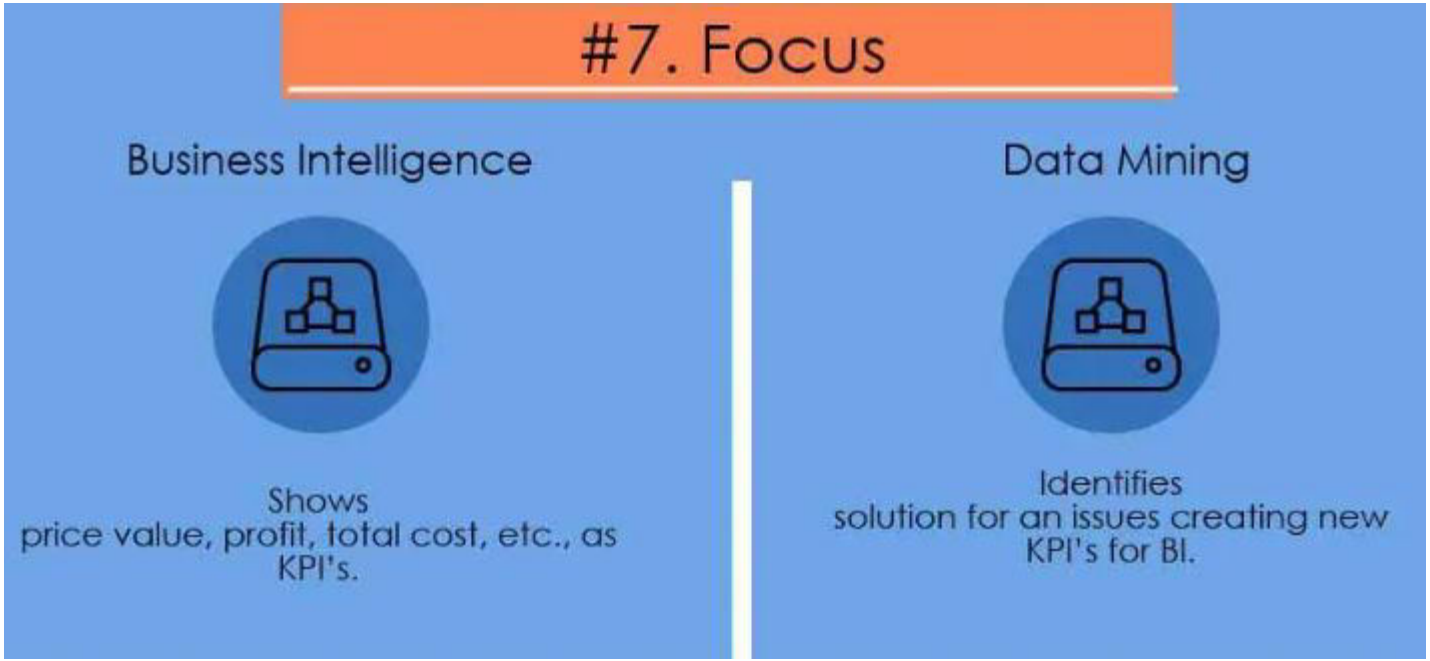
Made By: Priyanka Karale


#3. Data Volume
Business Intelligence
Large
Datasets processed on dimensional / relational databases
Data Mining
Small
datasets processed on small portion of data.


#4. Quality of Solutions
Business Intelligence
Volumetric
in nature and present the accurate
result using visualizations.
Data Mining
Uses
algorithms to identify accurate patterns
for an issue and identifies the blind
spots.


#5. Results Presentation
Business Intelligence
Dashboards
and Reports represented by graphs and
charts with KPI's
Data Mining
Identifies
the solution for an issue to be
represented as one of the KPI's in
Dashboards
or reports.


#6. Analysis
Business Intelligence
Depends on small-scale of past data, there is no intelligence involved; management has to take the decision based on information.
Data Mining
Focused on a particular issue in business on small-scale data using algorithms to find the solution.


#7. Focus
Business Intelligence
Shows
price value, profit, total cost, etc., as
KPI's.
Data Mining
Identifies
solution for an issues creating new
KPI's for BI.