MIRPUR UNIVERSITY OF SCIENCE AND TECHNOLOGY DEPARTMENT OF SOFTWARE ENGINEERING

Business Intelligence

(Lecture # 1)
Introduction to Business Intelligence

Saman Fatima
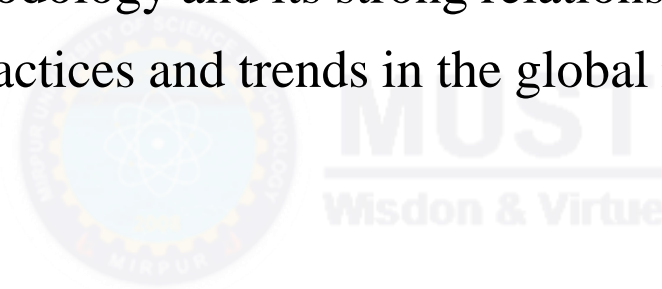*(Lecturer)*

Date:

Course Objectives

- To impart knowledge of data warehousing to students

- To give practical knowledge to students regarding data cleaning/ETL (Extract ,transform,Load)and EDA(Electronic design automation) techniques

- To teach the standard BI methodology to the students, i.e., how to solve business problems using BI techniques and tools

- To impart the skill of data-driven decision making through interactive dashboards with hands-on activities on a BI tool of the instructor's choice

- To convey knowledge about the BI practices and trends being followed in the global industry
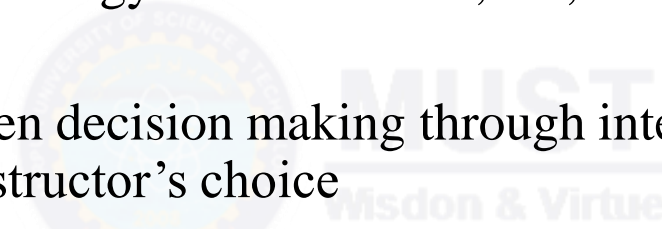
Course Learning Outcomes

1. Skills to execute Data-driven decision making through BI, through the use of specific charts and designing interactive dashboards and stories

2. Skills to execute Data cleaning/ETL and EDA activities for a BI project

3. Knowledge of core BI methodology and its strong relationship with data warehousing

4. Knowledge of current BI practices and trends in the global industry

LECTURE CONTENTS

1. What is Business Intelligence

2. What is warehouse

What is Business Intelligence

- **Definition:** Business Intelligence (BI) refers to the **skills, processes, technologies, applications, and practices** used to collect, store, analyze, and provide access to data for decision-making.

- Example:A retail company like Walmart collects sales data from all its stores. BI tools analyze this data to identify trends, such as which products sell best during specific seasons, helping management make informed inventory decisions.

- **Purpose:** Helping organizations make data-driven decisions by analyzing past and present data.

- **Key Features:**
  - Data Collection
  - Data Storage
  - Data Analysis
  - Data Visualization
  - Decision Making

- **Example:**
  - A retail store analyzes its sales data to manage inventory efficiently.
  - A bank detects fraudulent transactions using BI tools.

Evolution of Business Intelligence

Early Stage: Manual data collection and reporting.

Decision Support Systems (DSS) - 1970s-1980s: Automated data analysis and support for managers.

Modern BI: AI-powered analytics, real-time reports, and advanced BI tools.

Example:

Traditional businesses used spreadsheets, but today they use **Power BI** and **Tableau** for advanced reporting.

Key Components of Business Intelligence

- **Data Collection**
  - Gathering data from multiple sources (databases, spreadsheets, websites, customer transactions, etc.)
  - **Example:** Amazon collects user browsing history, purchase records, and customer reviews.

- **Data Storage**
  - Storing collected data in databases, data warehouses, or cloud storage.
  - **Example:** Google Cloud or Microsoft Azure stores large datasets for companies.

- **Data Analysis**
  - Analyzing raw data to find useful patterns.
  - **Example:** A bank analyzes transaction data to detect fraudulent activities.

- **Data Visualization**
  - Presenting analyzed data using graphs, charts, and dashboards for better understanding.
  - **Example:** Power BI and Tableau create sales performance dashboards for executives.

- **Decision-Making Support**
  - Using insights from BI reports to make strategic business decisions.
  - **Example:** Netflix uses BI analytics to recommend personalized movies and TV shows to users.

Data Sources for Business Intelligence

Internal Data Sources:

ERP Systems, CRM, HR Databases, Financial Records.

External Data Sources:

Market Trends, Social Media, Competitor Analysis.

**Data Formats:**

Structured Data: Databases (SQL, NoSQL)

Semi-Structured Data: JSON, XML files

Unstructured Data: Emails, PDFs, Images

- .

- **Visualization Tools:** Power BI, Tableau, Google Data Studio.

- **ETL (Extract, Transform, Load) Tools:** Talend, Apache Nifi, Informatica.

- **Data Warehousing:** Amazon Redshift, Google BigQuery, Snowflake.

- **Example:** A marketing team uses Power BI to analyze customer engagement trends.
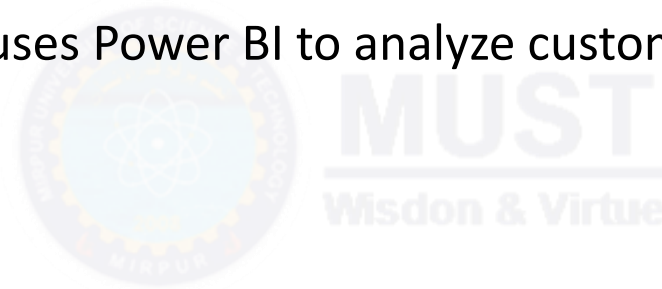
Business Intelligence vs. Data Analytics

Feature	Business Intelligence	Data Analytics
Focus	Past and present data	Future prediction
Tools	Dashboards, Reports	AI, Machine Learning
Usage	Decision Making	Trend Analysis

Introduction to Data Warehousing

- **Definition:** A central database where business data is stored for analysis.

- **Role:** Helps in complex queries, reporting, and analytics.

- **ETL Process:**
  - **Extract:** Collecting data from different sources.
  - **Transform:** Cleaning and standardizing data.
  - **Load:** Storing data in a data warehouse.

- **Example:** Amazon stores customers' purchase history to offer personalized product recommendations.

Example

- **Data Extraction (Extract)**

- Amazon collects data from various sources, including:

- **Purchase history** (What products did the user buy?)

- **Browsing history** (What products did the user view?)

- **Reviews and ratings** (How did the user rate a product?)

- **Payment methods** (Did the user pay with a credit card, debit card, or cash on delivery?)

- **Location and device information** (Where is the user logging in from, and what device is being used?)

Data Cleaning and Transformation (Transform)

- The collected data is cleaned and standardized for analysis.

- **Fixing incomplete or incorrect data** (e.g., correcting spelling errors in names or addresses)

- **Converting data into a common format** (e.g., standardizing different currencies)

- **Categorizing data** (e.g., grouping products into categories like clothing, electronics, footwear, etc.)

- **Storing Data in the Warehouse (Load)**

- The cleaned and transformed data is then loaded into a central **data warehouse**, where it is stored for analysis and reporting.

- **Data Analysis and Personalized Recommendations**

- Amazon's algorithms analyze a user's past data to identify **similar products**.

- If a user frequently buys electronics, Amazon will recommend **new and related electronic products**.

- If a user recently purchased a **mobile phone**, Amazon will suggest **mobile cases, screen protectors, and other accessories**.

- Amazon also analyzes what other customers who bought the same product have purchased and recommends those products.

Online Analytical Processing (OLAP)

- **Concept:** Enables fast queries and multidimensional analysis of data.

- **Types of OLAP:**
  - **Multidimensional OLAP (MOLAP)**
  - **Relational OLAP (ROLAP)**
  - **Hybrid OLAP (HOLAP)**

- **Uses:** Trend analysis, forecasting, and reporting.
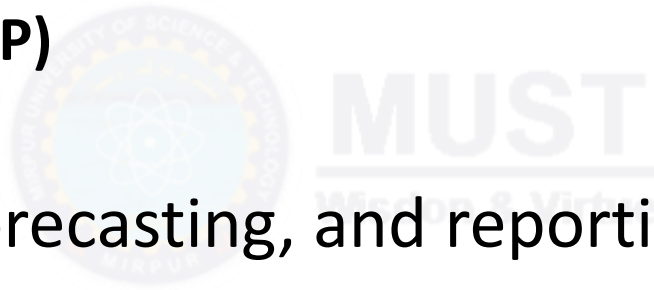
Online Analytical Processing (OLAP)

- **Types of OLAP:**

- **Multidimensional OLAP (MOLAP):**
  - Stores data in a **multidimensional cube format** for quick retrieval.
  - Provides **faster query performance** but requires more storage space.
  - Example: A retail store analyzing sales by **region, product category, and time** using pre-aggregated data cubes.

- **Relational OLAP (ROLAP):**
  - Stores data in **relational databases (SQL-based)** and processes queries dynamically.
  - **Slower query performance** but uses less storage compared to MOLAP.
  - Example: A banking system analyzing **customer transactions over different periods** using real-time calculations.

- **Hybrid OLAP (HOLAP):**
  - A combination of **MOLAP and ROLAP**, balancing **performance and storage efficiency**.
  - Stores frequently used data in cubes (**MOLAP**) and keeps detailed data in relational databases (**ROLAP**).
  - Example: An e-commerce company tracking **product sales trends** while allowing drill-down for specific transaction details.

Data Mining

- **Definition:** Extracting patterns from large datasets.

- **Techniques:**
  - **Classification (Grouping Customers)**
  - **Clustering (Market Basket Analysis)**
  - **Association Rules (Product Recommendations)**

- **Example:** Netflix analyzes viewing habits to recommend movies.

Data Mining

- **Techniques of Data Mining:**
- **Classification:**
  - Categorizes data into predefined groups.
  - Example: **Grouping customers** based on purchasing behavior (e.g., frequent buyers vs. occasional buyers).
- **Clustering:**
  - Identifies natural groupings within data.
  - Example: **Market Basket Analysis**, where customers with similar shopping habits are grouped together.
- **Association Rules:**
  - Discovers relationships between items in a dataset.
  - Example: **Product recommendations**—if a customer buys a laptop, they are likely to buy a laptop bag as well.

Benefits of Business Intelligence

- Improved decision-making.

- Better customer insights.

- Increased operational efficiency.

- Competitive advantage.
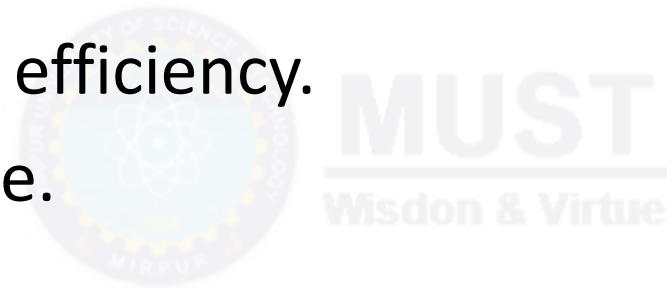
Challenges in Business Intelligence

- **Data Security and Privacy**
- Sensitive user data (e.g., personal information, financial records) can be misused if not protected.
- Risk of **data breaches and unauthorized access**.
- **Example:** Social media platforms must ensure that user data is not exposed to third parties without consent.
- **Integrating Data from Multiple Sources**
- Data comes from different formats and databases (structured and unstructured).
- **Challenges:** Handling inconsistencies, missing values, and duplicate records.
- **Example:** A multinational company collecting sales data from online platforms, physical stores,
- and mobile apps must **standardize** the data for analysis.
- **High Implementation Costs**
- Requires **advanced infrastructure**, skilled professionals, and software tools.
- Small businesses may struggle with **expensive hardware and licensing fees**.
- **Example:** A startup implementing AI-powered data mining may face high costs for cloud computing and data storage

Future of Business Intelligence

- AI and Machine Learning Integration.

- Automated BI Tools.

- Real-Time Data Analytics.

References

- Business Analytics -Taking Business Intelligence Beyond Reporting by Gert Laursen and Jesper Thorlund, Wiley 2010

- Business Analytics for Managers by Wolfgang Jank (Published by Springer)

- Business Analytics by James R. Evans (Published by Pearson)

THANKS