INTRODUCTION TO BUSINESS ANALYTICS

Module 1 Study Material

What is Business Intelligence?

Business Intelligence (BI) refers to the processes, technologies, and tools that organizations use to collect, integrate, analyze, and present business data in a way that provides actionable insights. The goal of business intelligence is to support better decision-making by transforming raw data into meaningful information that can guide strategic, tactical, and operational choices within a business.

Simple Example of BI

Let's consider a simple example of business intelligence (BI) using a retail store.

Imagine you're the manager of a small retail store that sells electronic gadgets. You've been collecting data on sales, inventory, and customer behaviour over the past few months. Now, you want to use business intelligence to gain insights and make informed decisions.

Below are the steps in BI for the retail store.

Step 1: Data Collection and Integration

You gather data from various sources, such as sales transactions, inventory records, and customer demographics. This data might be stored in different formats and locations. Using a BI tool, you integrate this data into a central repository or data warehouse.

Step 2: Data Analysis and Visualization

Once the data is collected and integrated, you can start analyzing it. For instance, you might want to know which products are the best-sellers, which days of the week have the highest sales, and how your sales are trending over time. You create various visualizations like bar charts, line graphs, and pie charts to represent this information.

Step 3: Extracting Insights

From the visualizations, you start to gain insights. For example, you might notice that sales of a particular gadget tend to spike during weekends, suggesting that it's popular among casual shoppers. Additionally, you might identify that sales have been steadily declining for another gadget, indicating a potential need for adjustments in pricing or marketing strategies.

Step 4: Decision Making

Armed with these insights, you can make informed decisions to improve your business. In response to the declining sales, you might decide to run a promotion on the struggling gadget to stimulate interest. For the popular gadget, you might allocate more shelf space and plan marketing campaigns around weekends to maximize its potential.

Step 5: Monitoring and Iteration

After implementing your decisions, you continue to monitor the impact. Did the promotion increase sales of the struggling gadget? Did the change in marketing strategy lead to even higher sales for the popular gadget? You can track these outcomes and refine your strategies based on the ongoing data you collect.

In this example, business intelligence helps you uncover patterns, trends, and insights from your data, enabling you to make more strategic decisions that can positively improve your business's performance. BI tools provide a way to transform raw data into actionable information, guiding your choices for growth and success.

Key components of business intelligence

Data Collection and Integration:

BI involves gathering data from various sources, such as internal systems (like sales, finance, and HR) and external sources (like market data and social media). This data might be in different

formats and locations, so a central repository or data warehouse is often used to consolidate and integrate the data.

Data Analysis

Once the data is collected and integrated, it's analyzed to uncover patterns, trends, and relationships. This analysis can involve complex queries, calculations, and statistical techniques to gain insights into business performance, customer behavior, market trends, and more.

Data Visualization

Visualizations like charts, graphs, maps, and dashboards are used to present the analyzed data in a more understandable and accessible format. Visual representations make it easier for business users to quickly grasp the insights and draw conclusions.

Reporting

BI tools generate regular reports that summarize important metrics and KPIs (Key Performance Indicators). These reports can provide managers and executives with snapshots of how the business is performing and help them identify areas that need attention.

Ad Hoc Querying

BI tools often allow users to perform ad hoc queries, which means they can explore the data and ask specific questions on the fly without needing to go through a predefined report.

Data Mining

This involves using advanced techniques to discover hidden patterns and insights in large datasets. Data mining can reveal correlations and trends that might not be apparent through traditional analysis.

Predictive Analytics

BI can also include predictive modeling, where historical data is used to build models that predict future outcomes. This can help

businesses anticipate customer preferences, demand patterns, and potential opportunities or risks.

Performance Management

BI can be used to track and measure key performance indicators (KPIs) against business goals and targets, allowing organizations to gauge their progress and make adjustments as needed.

Decision Support

The ultimate goal of BI is to provide decision-makers with the information they need to make informed choices. Whether it's optimizing supply chain management, refining marketing strategies, or allocating resources more effectively, BI assists in making data-driven decisions.

Overall, business intelligence empowers organizations to transform data into actionable insights, fostering a culture of data-driven decision-making that can lead to improved efficiency, competitiveness, and strategic growth.

Evolution of Business Intelligence (BI)

The evolution of Business Intelligence (BI) spans several decades and has been shaped by technological advancements, changing business needs, and the growing availability of data. Here's a brief overview of the key stages in the evolution of BI:

1960s - 1970s: Early Data Processing: The origins of BI can be traced back to the era of mainframe computers. Organizations began using basic data processing tools to manage and analyze business data. These early systems focused on basic data storage and retrieval rather than sophisticated analysis.

1980s - 1990s: Decision Support Systems (DSS): The concept of Decision Support Systems emerged during this period. DSS allowed users to interact with data through more user-friendly interfaces and conduct rudimentary analysis. These systems

facilitated simple querying and reporting, offering a step beyond basic data processing.

1990s - Early 2000s: Data Warehousing and OLAP: Data warehousing gained prominence in the 1990s, allowing organizations to centralize and store large volumes of data for analysis. Online Analytical Processing (OLAP) technology allowed for multidimensional analysis, enabling users to pivot, drill down, and slice data for deeper insights.

Late 1990s - Early 2000s: Emergence of Data Mining and Data Integration: The late 1990s saw a focus on data mining and advanced analytics. Organizations started using predictive and statistical modelling to uncover hidden patterns in data. Additionally, data integration tools became essential for merging data from disparate sources into a unified view for analysis.

Mid 2000s - 2010s: Self-Service BI and Visualization: The advent of user-friendly self-service BI tools allowed business users to create reports, visualizations, and dashboards without heavy reliance on IT departments. This democratization of data access empowered more employees to make data-driven decisions.

2010s - Present: Big Data and Advanced Analytics: The rise of big data brought about new challenges and opportunities. Organizations began harnessing technologies like Hadoop and NoSQL databases to handle and analyze vast amounts of data. Additionally, the integration of AI and machine learning into BI tools enabled predictive and prescriptive analytics on a larger scale.

Present and Beyond: AI-driven BI and Embedded Analytics: Today, AI-driven BI is becoming more prevalent, automating insights extraction and making recommendations based on data analysis. Embedded analytics is another trend, where BI capabilities are integrated directly into other applications, allowing users to access insights without switching tools.

Throughout this evolution, the focus of BI has shifted from basic data processing to advanced analytics and real-time insights. The landscape has also expanded to include cloud-based BI solutions,

mobile access, and the integration of BI into the fabric of business operations.

Overall, the evolution of BI reflects the ongoing quest to turn raw data into actionable insights, enabling businesses to make informed decisions and stay competitive in a data-driven world.

Different Levels of Analytics (Business Analytics)

Analytics encompasses a range of techniques and approaches for interpreting data to gain insights, make informed decisions, and uncover patterns and trends. Analytics can be classified into different levels based on the complexity of the techniques used and the depth of insights derived. Here are the commonly recognized levels of analytics:

Descriptive Analytics:
Descriptive analytics focuses on summarizing historical data to provide a clear understanding of what has happened in the past. It involves basic aggregation and reporting techniques to answer questions like "What happened?" and "How did it happen?" Key tools in descriptive analytics include charts, graphs, dashboards, and basic statistical measures like averages and totals.

Diagnostic Analytics:
Diagnostic analytics goes a step beyond descriptive analytics by aiming to understand the reasons behind certain outcomes. It involves analyzing historical data to identify patterns and correlations. Diagnostic analytics addresses questions like "Why did it happen?" and "What were the contributing factors?" Root cause analysis and trend analysis are common techniques used in diagnostic analytics.

Predictive Analytics:
Predictive analytics involves forecasting future outcomes based on historical data and statistical modeling. It uses techniques such

as regression analysis, time series analysis, and machine learning algorithms to predict trends, behaviors, and events. Predictive analytics seeks to answer questions like "What is likely to happen?" and "What might happen in the future?"

Prescriptive Analytics:

Prescriptive analytics takes a more advanced approach by not only predicting future outcomes but also suggesting optimal actions to achieve desired outcomes. It involves optimization algorithms and decision models that consider various possible scenarios and recommend the best course of action. Prescriptive analytics addresses questions like "What should we do?" and "How can we achieve the best results?"

Each level of analytics builds upon the previous one, with increasing complexity and potential for actionable insights. Organizations often progress through these levels as they develop their analytical capabilities and strive to make data-driven decisions that have a positive impact on their operations, strategies, and customer interactions.

Applications of Business Analytics in Industries

Business analytics is widely used in various industries to extract insights from data and make informed decisions. Here are some applications of business analytics in different industries:

Retail and E-Commerce:

Customer Segmentation: Analyzing customer data to segment customers based on behavior, demographics, and preferences.

Demand Forecasting: Predicting demand for products to optimize inventory management.

Price Optimization: Adjusting pricing strategies based on market conditions and consumer behavior.

Market Basket Analysis: Identifying product affinities to improve product recommendations and cross-selling.

Finance and Banking:

Risk Management: Analyzing data to assess credit risk, fraud detection, and market risk.

Customer Churn Prediction: Identifying customers at risk of leaving the bank or switching products.

Investment Analysis: Analyzing financial markets and investment opportunities using predictive modeling.

Compliance and Fraud Detection: Monitoring transactions and identifying suspicious activities for regulatory compliance.

Healthcare:

Patient Outcome Analysis: Using data to predict patient outcomes, readmissions, and treatment effectiveness.

Disease Surveillance: Tracking the spread of diseases and identifying potential outbreaks.

Healthcare Resource Optimization: Optimizing hospital resources, staff scheduling, and bed management.

Drug Discovery: Analyzing genomic and clinical data to discover new drugs and treatments.

Manufacturing:

Supply Chain Optimization: Analyzing supply chain data to improve inventory management and reduce lead times.

Quality Control: Identifying defects and quality issues through data analysis.

Predictive Maintenance: Using IoT data to predict when equipment and machinery need maintenance.

Production Planning: Optimizing production schedules to meet demand while minimizing costs.

Telecommunications:

Customer Churn Analysis: Identifying factors leading to customer churn and implementing retention strategies.

Network Performance Analysis: Monitoring network quality and resolving issues in real-time.

Pricing and Packaging Optimization: Analyzing customer usage data to create tailored pricing plans.

Market Segmentation: Segmenting the customer base for targeted marketing campaigns.

Transportation and Logistics:

Route Optimization: Finding the most efficient routes for delivery and transport.

Fleet Management: Analyzing vehicle data for maintenance, fuel efficiency, and performance.

Demand Forecasting: Predicting shipping and logistics demands to optimize resources.

Supply Chain Visibility: Tracking the movement of goods throughout the supply chain.

Energy and Utilities:

Smart Grid Analytics: Analyzing data from smart meters and sensors to optimize energy distribution.

Asset Performance Management: Predicting when equipment and infrastructure need maintenance.

Energy Consumption Analysis: Identifying opportunities for energy efficiency and cost reduction.

Demand Response: Using analytics to balance energy supply and demand in real-time.

Hospitality and Tourism:

Revenue Management: Adjusting pricing and availability based on demand and booking patterns.

Customer Experience Analysis: Gathering and analyzing feedback and reviews to improve service.

Marketing Personalization: Creating personalized marketing campaigns based on customer preferences.

Market Trend Analysis: Monitoring trends and competition to make strategic decisions.

These are just a few examples, and business analytics has applications in virtually every industry. The common goal is to use data to gain insights, optimize operations, and make informed decisions that drive business success.

OLAP

OLAP, which stands for Online Analytical Processing, is a category of computer processing that enables interactive and complex data analysis. OLAP systems are designed for querying, reporting, and analyzing multidimensional data models, allowing users to gain insights from data quickly and efficiently. OLAP technology is commonly used in business intelligence and data warehousing applications.

There are two primary types of OLAP systems:

MOLAP (Multidimensional OLAP): In MOLAP systems, data is stored in multidimensional cubes, and the processing is performed directly on these cubes. Popular MOLAP solutions include Microsoft Analysis Services and IBM Cognos TM1.

ROLAP (Relational OLAP): ROLAP systems store data in relational databases and perform OLAP operations by generating SQL queries. These systems provide greater flexibility but may sacrifice some query performance. Examples of ROLAP systems include Oracle OLAP and SAP BW (Business Warehouse).

OLAP systems are widely used in business intelligence to support tasks such as data exploration, reporting, and data visualization. They provide decision-makers with the ability to quickly access and analyze data to make informed business decisions.

OLTP

OLTP, which stands for Online Transaction Processing, is a type of database system and processing that manages and facilitates the execution of day-to-day transactional operations in an organization. These transactions typically involve the insertion, updating, and retrieval of relatively small amounts of data, often involving individual records. OLTP systems are optimized for rapid and concurrent access by multiple users or applications.

Examples of OLTP applications include:

E-commerce: Managing product inventory, processing customer orders, and handling online payments.

Banking: Managing customer accounts, processing transactions, and providing real-time account balance information.

Healthcare: Managing patient records, scheduling appointments, and updating medical histories.

Retail: Handling point-of-sale transactions, tracking inventory, and managing customer data.

Airline Reservations: Booking and managing flight reservations, seat assignments, and ticketing.

Human Resources: Managing employee records, processing payroll, and tracking leave balances.

In contrast to OLTP systems, OLAP (Online Analytical Processing) systems are designed for complex data analysis and reporting, using aggregated data from OLTP systems. Both OLTP and OLAP systems play crucial roles in an organization's data infrastructure, serving different purposes to support transactional and analytical needs, respectively.

Sources of Big Data

Big data is generated from a wide range of sources, but here are six major sources of big data:

Social Media: Social media platforms like Facebook, Twitter, Instagram, and LinkedIn produce massive amounts of data daily. This includes text, images, videos, likes, shares, comments, and user interactions. Analyzing this data can provide valuable insights into user behavior and sentiment.

IoT Devices: The Internet of Things (IoT) encompasses a vast network of interconnected devices, sensors, and equipment that continuously generate data. This data includes information from smart homes, industrial sensors, wearables, and more. It often relates to temperature, humidity, location, machine performance, and other variables.

E-commerce: Online shopping platforms and retail websites generate significant data on customer behavior, product views, purchase history, and transaction details. This data helps businesses understand customer preferences and optimize their offerings.

Web and Clickstream Data: Websites and online services record data on user interactions, including page views, clicks, navigation paths, and the time spent on webpages. Clickstream data provides insights into user behavior and can help improve website design and user experience.

Mobile Apps: Mobile applications collect data on user interactions, location, app performance, and more. Mobile app data is valuable for understanding how users engage with apps and services on smartphones and tablets.

Healthcare: The healthcare industry generates large volumes of data from electronic health records (EHRs), medical imaging (e.g., X-rays and MRIs), patient histories, and clinical trials. This data is essential for patient care, medical research, and health system optimization.

These sources are just a subset of the many data generators in the big data landscape. Businesses and organizations across various industries leverage big data from these sources to gain insights, make data-driven decisions, and improve their products, services, and operations. Analyzing big data has become increasingly

critical for competitiveness and innovation in the modern data-driven world.

Characteristics of Big Data or 5Vs of Big Data

The "5 Vs of Big Data" is a framework that describes the key characteristics or dimensions of big data. These characteristics help us understand the challenges and opportunities associated with large and complex datasets. The 5 Vs of Big Data are:

Volume: Volume refers to the sheer size of the data. Big data involves vast amounts of information, often measured in terabytes, petabytes, or more. It's characterized by the ability to handle and process massive quantities of data.

Velocity: Velocity refers to the speed at which data is generated and how quickly it must be processed and analyzed. Big data often comes in real-time or near-real-time, and it requires systems that can handle high data ingestion rates and provide rapid insights.

Variety: Variety pertains to the diversity of data types and sources. Big data encompasses structured data (e.g., databases), unstructured data (e.g., text, images, videos), and semi-structured data (e.g., XML or JSON). It may also involve data from various sources and formats.

Veracity: Veracity relates to the quality and trustworthiness of the data. Big data sources can include errors, inconsistencies, and noise. Data veracity challenges the reliability and accuracy of the information and requires data cleaning and validation.

Value: Value is the ultimate goal of big data analysis. The value dimension emphasizes the importance of extracting meaningful insights, patterns, and actionable information from the data. The value obtained from analyzing big data can lead to better decision-making, improved processes, and innovation.