**CSIS 2200:**

**Slide 01: Information System an Overview**

**Computer literacy:**

Skill in using productivity software, as well as having a basic knowledge of hardware and software, the internet, and collaboration tools and technologies.

**Information Literacy:**

Understanding the role of information in generating and using business intelligence (BI). Provides historical, current and predictive views of business operations and environments and gives organizations a competitive advantage in the marketplace.

**Transaction-Processing Systems (TPS):**

* Focus on data collection and processing.
* Used for cost reduction.
* Applied to structured task.
* Require minimal human involvement when automated.

**Management Information System (MIS):**

* Organized integration of hardware and software technologies, data, processes, and human elements.
* Designed to produced timely, integrated, relevant, accurate, and useful information for decision-making.
* Designing tasks
  + Define the system’s objectives
  + Collect and analyze data
  + Provide information in a useful format for decision-making purposes.
* MIS applications
  + Used in both private and public sectors.

**Components of an Information System:**

*Data | Database | Process | Information*

1. Data:
   1. Considered the input of a system.
   2. Sources of data:
      1. Internal: records.
      2. External: customers, suppliers, government agencies, labor…
   3. Has a time orientation:
      1. Past data: performance reports.
      2. Current data: operational reports.
   4. Can be collected in different forms:
      1. Disaggregated data: helps analyze sales by product, territory, or salesperson.
      2. Aggregated data: useful for reporting overall performance during a sales quarter.
2. Database:
   1. Collection of relevant data organized in a series of integrated files.
      1. Essential for the success of any information system.
   2. Database management System (DBMS)
      1. Used to create, organize, and manage databases
      2. Reduces personnel time needed to gather, process, and interpret data manually.
3. Process:
   1. Generates the most useful type of information for making decisions
      1. Transaction-processing reports.
      2. Models for decision analysis that can be built into the system or accessed from external sources.
4. Information:
   1. Consists of facts analyzed by the process component and is an output of an information system.
      1. Usefulness qualities
         1. Timelines
         2. Integration with other data and information
         3. Consistency and accuracy
         4. Relevance
   2. Needs to provide either a base for users to explore different options or insight into tasks.
   3. Usefulness is affected by the information system’s user interface.
      1. Graphical user interfaces (GUIs) are used because they are flexible and easy.
   4. Systems should produce information in different formats, including graphics, tables and exception reports.
      1. Increases likelihood of users understanding and being able to use the information.
   5. Users need to be able to make use of informal information when solving problems.

**Information technologies:**

* The internet.
* Computer networks.
* Database systems.
* Point-of-sale (POS) systems.
* Radio-frequency-identification (RFID)

**Importance of Information Systems:**

* Timely, relevant, and accurate information is a critical tool.
  + Enhance competitive position in the marketplace.
* Manage the four Ms of resources:
  + Manpower
  + Machinery
  + Materials
  + Money
* Personal Information System (PIS) or Human Resource Information System (HRIS)
  + Designed to provide information that helps decision makers in personnel carry out tasks effectively.
* Logistic Information System (LIS)
  + Designed to reduce the cost of transporting materials while maintaining safe and reliable delivery.
* Manufacturing information system (MIS)
  + Used to manage manufacturing resources.
  + Reduce manufacturing costs.
  + Increase product quality.
  + Improve inventory decisions.
* Financial information system (FIS)
  + Used to provide information to financial executives in a timely manner.
* Marketing information system (MKIS)
  + Used to improve marketing decisions.
  + Provides timely, accurate, and integrated information about the marketing mix.
  + Price, promotion, place and product.
* Marketing technology tools
  + Business, web, and mobile analytics
  + E-mail marketing
  + Search engine marketing
  + Mobile technologies
  + Marketing automation

**Using Information Technologies for a Competitive Advantage:**

* Michael Porter: Three strategies for successfully competing in the marketplace.
  + Overall cost leadership.
  + Differentiation.
  + Focus.
* Information systems
  + Help organizations reduce the cost of products and services.
  + Help bottom-line and top-line strategies.
  + Use enterprise systems to create an efficient and effective link between suppliers and consumers.
* Differentiation strategies
  + Making products and services different from competitors.
* Focus strategies
  + Focusing on specific market segments to achieve a cost or differentiation advantage.

**Porter’s Five Forces Model: Understanding the Business Environment**

* Analyzes a firm’s position in the marketplace and how information systems can make it more competitive.
  + Five forces
    - Buyer power.
      * **High** when customers have **many** choices.
      * **Low** when customers have **few** choices.
    - Supplier power.
      * **High** when customers have **fewer** options
      * **Low** when customers have **more** options
    - Threat of substitute products or services.
      * **High** when **many alternatives** to an organization’s products and services are available.
    - Threat of new entrants.
      * **Low** when duplicating a company’s product or service is difficult.
      * **Focus** strategies are used to ensure that the threat remains low.
    - Rivalry among existing competitors.
      * **High** when competitors occupy the same marketplace position.
      * **Low** when there are few competitors.

This image depicts the five forces model. There is rectangular box positioned at the center of the image, and four boxes have been positioned around this box. The box on the left is labeled suppliers, the box on the right is labeled buyers, the box at the top is labeled potential entrants, and the box at the bottom is labeled substitutes.  
The box at the center contains two points that read industry competitors and rivalry among existing competitors. A U-shaped arrow has been positioned between these points. The arrowhead points at the first point. 
An arrow arises from the right side of the box labeled suppliers that is positioned on the left and points at the box at the center of the image. This arrow is labeled bargaining power of suppliers. 
An arrow arises from the left side of the box labeled buyers that is positioned on the right and points at the box at the center of the image. This arrow is labeled bargaining power of buyers. 
An arrow arises from the bottom of the box labeled potential errants that is positioned at the top and points at the box at the center of the image. This arrow is labeled threats of new entrants. 
An arrow arises from the top of the box labeled substitutes that is positioned at the bottom and points at the box at the center of the image. This arrow is labeled threat of substitute products or services.  

**The IT Job Market**

* Operations and help desk.
* Programming.
* Systems design.
* Web design and web hosting.
* Network design and maintenance.
* Database design and maintenance.
* Robotics and artificial intelligence.
* Chief of Technology Officer (CTO) / Chief of Information Officer (CIO)
  + Oversees long-range planning and monitors new developments that can affects a company’s success.
* Chief Privacy Officer (CPO)
  + Responsible for managing risks and business impacts of privacy laws and policies.
* Manager of information systems services
  + Responsible for managing hardware, software, and personnel in the information systems department.
* Systems analyst
  + Responsible for the design and implementation of information systems.
  + Should have a sound understanding of business systems and functional areas within a business organization.
* Network administrator
  + Oversees a company’s internal and external network systems.
  + Provides network and cybersecurity.
* Database administrator (DBA)
  + Responsible for database design and implementation.
  + Required to have knowledge and understanding of data warehouses and data-mining tools.
* Computer programmer
  + Writes programs or software segments that allow the information system to perform a specific task.
* Webmaster
  + Designs and maintain the organization’s web site.
  + Have been in high demand owing to the popularity of e-commerce applications.

**Outlook for the Future**

* Hardware and software costs will decline
* Artificial intelligence and related technologies will improve and expand.
* Computer literacy and networking technology will improve.
* Personal computer will improve in power and quality.
* Internet growth will continue.
* Computer criminals will become more sophisticated.
* Protecting personal information will become more difficult.
* Ubiquitous computing and the Internet of Things (IoT).
* 3D printing, pervasive analytics, context aware computing, smart machines and devices, and cloud computing.
* Software defined applications and infrastructures.
* Security.
* Increased applications of augmented and virtual reality.

Check computer: Core, ram, cache, # of usb port

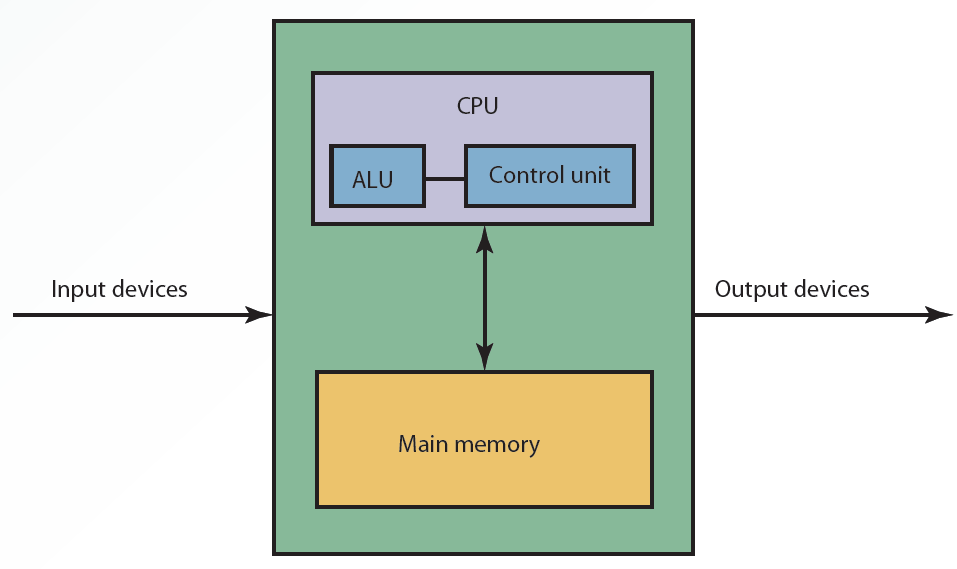
**Slide 02: Computers: The Machines Behind Computing**

**Defining a Computer:**

* A computer is a machine that:
  + Accept data as input.
  + Processes data without human intervention using stored instructions.
  + Outputs information.
* Instructions:
  + Step-by-step directions for performing a specific task.
  + Written in a language the computer can understand.
* Garbage in, garbage out (GIGO):
  + If the input data is erroneous, the information provided by the computer is also erroneous.
* Writing a computer program:
  + Identify what needs to be done.
  + Write the algorithm to achieve the goal.
  + Select a programming language for it. (depends on the problem and the computer)
* A program is also called a source code (a sequence of 0’s and 1’s).

**Components of a Computer System:**

* Hardware: physical devices such as keyboard, monitors, processing units.
* Software: programs written in computer languages.
* Central Processing Unit (CPU): the heart of the computer
  + Arithmetic Logic Unit (ALU): perform the arithmetic operations.
  + Control Unit: tells the computer what to do such as instructing the computer which device to read or send output to.
* Computers can have a single processor or multiple ones (two or more CPUs).



* Bus: link between devices connected to the computer.
  + Parallel or serial.
  + Internal (local) or external.
* Disk drive: peripheral device for recording, storing, and retrieving information.
* CPU case: enclosure containing the computers main components (computer chassis or tower).
* Motherboard: main circuit board containing connectors for attaching additional boards.

**Hardware Generations:**

**Table

Description automatically generated**

**The Power of Computers:**

* Speed.
  + Measured by the number of instructions executed per fractions of a second:
    - Millisecond: 1/1,000
    - Microsecond: 1/1,000,000
    - Nanosecond: 1/1,000,000,000
    - Picosecond: 1/1,000,000,000,000
* Accuracy.
* Storage and retrieval capabilities.
  + Save data in computer memory.
  + Access data from memory.
  + Data stored in bits.
  + American Standard Code for Information Interchange (ASCII).
    - Defines up to 128 characters.

Table

Description automatically generated

**Computer Operations:**

* Computer functions:
  + Three basics tasks:
    - Arithmetic operations.
    - Logical operations.
    - Storage and retrieval operations.
  + Input, Output and Memory
  + Main Memory:
    - Volatile and nonvolatile.
    - Random Access Memory (RAM).
    - Cache RAM: stores recently accessed memory.
      * Resides on the processor.
      * Made of silicon.
    - ROM (Read-only Memory).
      * Nonvolatile memory.
      * Data cannot be written to it.
      * Includes BIOS information and the computer system’s clock.
      * Programable read-only memory (PROM).
      * Erasable Programable read-only memory (EPROM).
  + Secondary Memory:
    - Magnetic disks.
      * Made of Mylar or metal.
      * Used for random-access processing.
    - Magnetic tape:
      * Made of a plastic material.
      * Stores data sequentially.
    - Optical disks:
      * Use lasers beams to access and store data.
      * CD-ROMs, WORM discs, and DVDs.
      * Include hard disks, USB flash drives, and memory cards.
    - Redundant Array of Independent Disks (RAID):
      * Collection of disk drives used for fault tolerance and improved performance.
      * Found in large network systems.
    - Cloud storage:
      * Involves multiple virtual servers that are hosted by third parties.

Table

Description automatically generated

* Storage Area Network:
  + Dedicated high-speed network consisting of both hardware and software.
  + Connect and manage shared storage devices
  + Disk arrays, tape libraries, and optical storage devices.
* Network-Attached Storage (NAS)
  + Network-connected computer dedicated to provide file based data storage services to other network devices.

**Classes of Computer:**

* Based on cost, amount of memory, speed and sophistication.
* Subnotebooks and notebooks
* Personal and minicomputers
* Mainframes
* Supercomputers

**Server Platforms: An Overview**

* Server: computer and all the software for managing network resources and offering services to a network.
* Available server platforms: applications servers, database servers, disk servers, fax servers, file servers, mail servers, print servers, remote access servers (RAS), and web servers.

**What is a Software?**

* Programs that run a computer system.
* Classification:
  + System software: work in the background and takes care of tasks, such as deleting waste files.
  + Application software: performs specialized tasks.

**Operating System Software:**

* Set of programs controlling and managing computer hardware and software.
  + Provides an interface between a computer and the user.
  + Increases computer efficiency by helping users share computer resources and performing repetitive tasks for users.
  + Increases computer efficiency by helping users share computer resources and performing repetitive tasks for users.
* Consists of control programs to manage hardware and resources by performing:
  + Job management.
  + Resource allocation.
  + Data management
  + Communication
* Supervisor program (i.e., the Kernel)
  + Responsible for controlling all other programs in th OS.
* Application Software:
  + Commercial software or software developed in house; used t perform variety of tasks on a personal computer.
    - Word processing, spreadsheet, database, presentation and graphics.
    - Desktop publishing.
    - Financial planning and accounting.
    - Computer-aided design (CAD).
* Computer Languages:
  + Machine Language:
    - First generation of computer languages.
    - Consists of a series of 0s and 1s representing data or instructions.
    - Dependent on the machine.
    - Time-consuming to write a program.
  + Assembly language:
    - Second generation of computer languages.
    - Machine dependent and a higher-level language than machine language.
    - Uses a series of short codes, or mnemonics, to represent data or instructions.
  + High-level languages:
    - Third generation of computer languages.
    - Machine independent and self-documenting.
    - Used for web development and internet applications.
  + Fourth-generation languages (4GLs):
    - Commands are powerful and easy to learn.
    - Use macro codes that can take the place of several lines of programming.
  + Fifth-generation languages:
    - Used artificial intelligence technologies.
    - Knowledge-based systems, natural language processing (NLP), visual programming, and a graphical approach to programming.
    - Designed to facilitate natural conversations between an individual and the computer.

**Slide 03: Database Systems, Database Workers and Data Marts**

**Databases:**

* Database:
  + Collection of related data that is stored in a central location or in multiple locations.
* Data Hierarchy:
  + Structure and organization of data, which involves fields, records, and files.
* Database Management System (DBMS)
  + Software for creating, storing, maintaining, and accessing database files.
  + Makes using databases more efficient.

**Interaction between the User, DBMS, and Databases:**

This illustration depicts the interaction between the User, DBMS, and Database. 
A rectangular box has been positioned at the center of the illustration. This box is labeled D B M S.
A circle has been placed on the left side of the illustration. This circle is labeled user. An arrow arises from the right side of this circle and points at the box labeled D B M S. A rectangular box has been aligned on the top of this arrow, and it contains text that reads user makes a request for information. An arrow arises from the left side of the box labeled D B M S and points at the circle. A rectangular box has been aligned at the bottom of this arrow, and it contains text that reads D B M S returns information to the user. 
A cylinder has been placed on the right side of the illustration. This cylinder is labeled database. An arrow arises from the right side of the box labeled D B M S and points at the cylinder. A rectangular box has been aligned on the top of this arrow, and it contains text that reads D B M S searches the database. An arrow arises from the left side of the cylinder and points at the box labeled D B M S. A rectangular box has been aligned at the bottom of this arrow, and it contains text that reads D B M S retrieves the information. 

**Types of Databases:**

* Internal data:
  + Collected from within an organization.
  + Stored in the organization’s internal databases and can be used by functional information systems.
* External data:
  + Comes from a variety of sources.
  + Stored in a data warehouse.

**Methods for Accessing Files**

* Sequential access file structure:
  + Records in files are organized and processed in numerical or sequential order.
  + Records are organized based on a primary key (e.g., Social Security numbers or Account numbers).
  + Used for backup and archive files because they rarely need updating.
* Random Access File Structure
  + Records can be accessed in any order, regardless of their physical locations in storage media.
  + Fast and very effective when a small number of records need to be processed daily or weekly.
  + Records are stored on magnetic disks to achieve speed.
* Indexed Sequential Access Method (ISAM)
  + Records accessed sequentially or randomly, depending on the number accessed.

**Random access** is used for a **small amount of data**. While the **sequential access** is used for a **larger set of data**. Indexed sequential one uses an index structure with two parts: index value and a pointer to the disk location of the record matching the indexed value.

**Logical Database Design**

* Information is viewed in a database in two ways
  + Physical view: how data is stored on and retrieved from storage media.
  + Logical view: how information appears to users and how it can be organized and retrieved.
    - Depending on the user, there can be more than one logical view of data.
* Data model determines how data is created, represented, organized, and maintained
  + Data structure.
  + Operations.
  + Integrity rules.
* Hierarchical model:
  + Relationships between records form treelike structure.
  + Records are called nodes, and relationships between records are called branches.

This flowchart depicts an example of a hierarchical model. It is divided into four levels. The content against the second, third, and fourth levels reads siblings.
A box has been positioned at the top center of the flowchart, and this box is labeled supplier A. Three lines arise from the bottom of this box and lead downward to three different boxes, which have been positioned horizontally. All three boxes have been labeled product line. 
Two lines arise from the bottom of the first box labeled product line. These lines lead downward to two different boxes. The first box is labeled P1, and the second box is labeled P2. Two lines arise from the bottom of the box labeled P1 and lead downward to two boxes, which are labeled A and B. Four lines arise from the bottom of the box labeled P2 and leads downward to four boxes, which are labeled C, D, E, and F. 
Three lines arise from the bottom of the second box labeled product line. These lines lead downward to three different boxes. The first box is labeled P3, the second box is labeled P4, and the third box is labeled P5. Two lines arise from the bottom of the box labeled P3 and lead downward to two boxes, which are labeled G and H. Three lines arise from the bottom of the box labeled P4 and lead downward to three boxes, which are labeled I, J, and K. Three lines arise from the bottom of the box labeled P5 and lead downward to three boxes, which are labeled L, M, and N. 
Two lines arise from the bottom of the third box labeled product line. These lines lead downward to two different boxes. The first box is labeled P6, and the second box is labeled P7. Four lines arise from the bottom of the box labeled P6 and lead downward to four boxes, which are labeled O, P, Q, and R. Three lines arise from the bottom of the box labeled P7 and lead downward to three boxes, which are labeled S, T, and U. 

* Network model:
  + Similar to the hierarchical model but records are organized differently.
  + Each record can have multiple parent and child records.

This flowchart depicts the network model in the form of a flowchart. It is divided into three levels. The content against the first level reads customer number. The content against the second level reads invoice number. The content against the third level reads method of payment. 
Three boxes have been positioned horizontally at the top of the flowchart. These boxes are labeled 2000, 3000, and 9000.  A line arises from the box labeled 2000 and leads downward to a box labeled 111. Two lines arise from the box labeled 3000 and lead downward to two different boxes, which are labeled 222 and 333. Two lines arise from the box labeled 9000 and lead downward to two different boxes, which are labeled 444 and 555. 
Three lines arise from each of the boxes labeled 111, 333, and 444 and lead downward to a box that is labeled cash. Two lines arise from each of the boxes labeled 222 and 555 and lead downward to a box that is labeled credit. 

**The Relational Model**

* Uses a two-dimensional table of rows and columns of data:
  + Rows are records (i.e., tuples)
  + Columns are fields (i.e., attributes)
* Data dictionary:
  + Stores definitions, such as data types for fields, default values, and validation rules for data in each field.
* Primary key:
  + Uniquely identifies every record in a relational database.
* Foreign key:
  + Field in a relational table that matches the primary key column of another table.
  + Used to cross-reference tables.
* Normalization:
  + Used to improve database efficiency.
    - Eliminates redundant data.
    - Ensures only related data is stored in a table.
  + Goes through different stages, from the first normal form (1NF) to the fifth normal form (5NF).
* Operations:
  + Help retrieve data from tables.
  + Common operations: select, project, join, intersect, union, and difference.

**Components of a DBMS**

* DBMS software components:
  + Database engine
  + Data definition
  + Data manipulation
  + Application generation
  + Data administration
* Database engine:
  + Responsible for data storage, manipulation and retrieval.
  + Interacts with other components of the DBMS to convert logical requests from the users into their physical equivalents.
* Data definition:
  + Used to create and maintain the data dictionary and define database file structure.
  + Makes changes to a database’s structure.
* Data manipulation:
  + Used to add, delete, modify, and retrieve records from a database.
  + Uses a query language, such as Structured Query Language (SQL).
* Application Generation:
  + Designs elements of an application using a database.
  + Used by IT professionals and database administrators.
* Data Administration:
  + Used for tasks such as backup and recovery, security, and change management.
  + Used to determine who has permission to perform certain functions, summarized as create, read, update, and delete (CRUD)
* Database Administrators (DBAs)
  + Handle database design and management.

**Recent Trends in Database Design and Use**

* Data-driven web sites.
* Natural language processing.
* Distributed databases.
* Object-oriented databases.
* Advances in artificial intelligence.

**Data-Driven Web Sites**

* Acts as an interface to a database
  + Retrieves data and allows users to enter data in the database.
* Improves access to information:
  + Reduces support and overhead needed to maintain static web sites.
  + Gives users more current information from a variety of data sources.

**Distributed Databases**

* Distributed Database Management System (DDBMS)
  + Stores data on multiple servers throughout an organization.
  + Several advantages:
    - Design better reflects the firm’s structure.
    - Local data storage reduces response time.
    - Minimizes effects of computer failure.
    - Cost advantage.
    - Not limited by physical location of the data.
* Approaches to setting up a DDBMS:
  + Fragmentation: addresses how tables are divided among multiple locations.
  + Replication: each site stores a copy of the data in the organization’s database.
  + Allocation: combines fragmentation and replication.
* Object-Oriented Databases:
  + Data and their relationships are contained in a single object.
    - Object consists of attributes and methods that can be performed on the object’s data.
      * Encapsulation: grouping objects along with their attributes and methods into a class.
      * Inheritance: new objects can be created faster and more easily by entering new data in attributes.
  + Advantages of object-oriented database:
    - Supports more complex data management.
    - Handles storing and manipulating all types of multimedia as well as numbers and characters.

**Data Warehouses**

* Collection of data from a variety of sources.
  + Support decision-making applications.
  + Generate business intelligence.
* Called hypercubes because they can store multidimensional data.
* Characteristics of data in a warehouse
  + Subject oriented.
  + Comes from a variety of sources.
  + Categorized based on time.
  + Captures aggregated data.
  + Used for analytical purposes.

**Data Warehouses Configuration**

* Input:
  + Different data sources provide the input for a data warehouse to perform analyses and generate reports
    - External data sources, databases, and transaction files.
    - Enterprise resources planning (ERP) systems.
    - Customer relationship management (CRM) systems.

This image depicts the configuration of a data warehouse and highlights its components. Four cylinders have been positioned vertically, one below the other, on the left side of the image. From the top to the bottom, these boxes are labeled databases, transaction files, enterprise resource planning systems, and customer relationship management systems. Another cylinder has been positioned at the top of the image. This cylinder is labeled external data sources. 
Arrows arise from the right side of each of the four cylinders positioned on the left side and from the bottom of the cylinder positioned at the top of the image. These arrows point at a rectangular box positioned vertically at the center of the image. This box is labeled extraction, transforming, and loading. An arrow arises from the right side of the box at the center and leads to another long, vertically-positioned cylinder. This cylinder is labeled raw data, summary data, and metadata. 
Three arrows arise from the right side of the vertically-positioned, long cylinder. The first arrow points at content that reads O L A P analysis. The second arrow points at content that reads data-mining analysis. The third arrow points at content that reads decision-making reports.

* Extraction, Transformation, and Loading (ETL):
  + Processes used in a data warehouse.
    - Collecting data from a variety of sources.
    - Converting data into a format that can be used in transformation processing.
    - Loading data into the data warehouse.
* Storage:
  + Collected information is organized in a date warehouse as:
    - Raw data: information in its original form.
    - Summary data: gives users subtotals of various categories.
    - Metadata: information about data’s content, quality, condition, origin, and other characteristics.
* Output:
  + Data warehouses use the following to generate reports:
    - Online analytical processing (OLAP)
      * Uses multiple sources of information and provides multidimensional analysis.
      * Generates business intelligence.
    - Data mining analysis
      * Used to discover patterns and relationships.
  + Benefits of data warehouses:
    - Cross-reference segments of an organization’s operations for comparison.
    - Generate complex queries and reports faster than when using databases.
    - Generate reports efficiently using data from a variety of sources.
    - Find patterns and trends that cannot be found with databases.
    - Analyze large amounts of historical data quickly.
    - Assist management in a making well-informed business decisions.
    - Manage a high demand for information from many users with different needs and decision-making styles.

**Data Marts**

* Smaller version of a data warehouse used by a single department or function.
  + Advantages over warehouses:
    - Faster access to data owing to its smaller size.
    - Improved response time for users.
    - Easier to create because of its size and simplicity
    - Less expensive.
    - Effective targeting of users.
  + Disadvantages:
    - Limited scope.
    - Difficulty in consolidating information from different departments or functional areas.

**Business Analytics**

* Uses data and statistical methods.
  + Gains insight into the data.
  + Provides decision makers with information to act on.
* Methods:
  + Descriptive
  + Predictive
  + Prescriptive
  + Descriptive Analytics:
    - Reactive strategy
    - Reviews past events, analyzes the data, and provides a report indicating:
      * What happened in a given period of time?
      * How to prepare for the future.
  + Predictive analytics:
    - Proactive strategy
    - Prepares decision makers for future events
  + Prescriptive analytics
    - Recommends a course of action that decision makers should follow.
    - Shows the likely outcome of each decision.

**The Big Data Era**

* Voluminous data
  + Conventional computing methods are unable to efficiently process and manage it.
* Involves five dimensions:
  + Volume
  + Variety
  + Velocity
  + Veracity
  + Value
* Provides competitive advantage in many areas
  + Retail, financial services, advertising and public relations, government, manufacturing, healthcare, etc.
* Many technologies and applications have contributed to growth and popularity
  + Mobile and wireless technology, the popularity of social networks, etc.
* Executives should guard against privacy risks
  + Discrimination, privacy breaches and embarrassments, unethical actions based on interpretations, loss of anonymity, etc.

**Database Marketing**

* Uses an organization’s database of customers and potential customers to promote products or services
  + Main goal: use information within the database to implement marketing strategies.
    - Increase profits.
    - Enhance competitiveness.
  + Transforms marketing from a reactive to a proactive process
    - Multivariate analysis
    - Data segmentation
    - Automated tools
  + Tasks performed by successful database marketing campaigns
    - Calculating customer lifetime value (CLTV)
    - Conducting recency, frequency, and monetary analysis (RFM)
    - Using different techniques to communicate effectively with customers
    - Using different techniques to monitor customer behavior across a number of retail channels.

**Relationship Diagrams Review**

Table is in 1NF if the table satisfies the following five conditions:

1. There’s not top-to-bottom ordering to the rows.
2. There’s no left-to-right ordering to the columns.
3. There are no duplicate rows.
4. Every row-and-column intersection contains exactly one value from the applicable domain (and nothing else).
5. All columns are regular [i.e. rows have no hidden components such as row IDs, object IDs, or hidden timestamps.]

Table

Description automatically generated

This table is not on the 1NF. Because you can observe the same information being repeated in the same row.

**One to Many Relationship**

One-to-many (1:M) – is used to relate one record in the table A with many records in the table B. A record in table A can have many matching in table B, and a record in table B can have one matching record in table A. In a one-to-many relationship, each row in one table can be related to many rows in the others table.

Graphical user interface

Description automatically generated with medium confidence

**Modelling a one-to-many relationship**

Step 01: Create individual tables with associated primary keys.

Step 02: Take the primary key from the table whose multiplicity is 1 and add it in table with multiplicity M.

Step 03: Create the relationship.