DATABASE AS INFORMATION SYSTEMS

By: Joshua Bachuba

INTRODUCTION

- Good decisions require good information derived from raw facts
- Data is managed most efficiently when stored in a database
- Databases evolved from computer file systems
- Understanding file system characteristics is important

WHY DATABASES?

- Databases solve many of the problems encountered in data management
  - Used in almost all modern settings involving data management:
- Business
- Research
- Administration
- Important to understand how databases work and interact with other applications

DATA VS. INFORMATION

- Data are raw facts
- Information is the result of processing raw data to reveal meaning
- Information requires context to reveal meaning
- Raw data must be formatted for storage, processing, and presentation
- Data are the foundation of information, which is the bedrock of knowledge
- Data: building blocks of information
- Information produced by processing data
- Information used to reveal meaning in data
- Accurate, relevant, timely information is the key to good decision making
- Good decision making is the key to organizational survival

DATA PROCESSING

Data processing is generally, "the collection and manipulation of items of data to produce meaningful information."

Data processing is undertaken by any activity which require collection of data. This data collected needs to be stored, sorted, processed and analyzed. This complete process can be divided into 6 basic simple stages which are:

Data collection

Storage of data

Sorting of data

processing of data

Data analysis

Data presentation and conclusions

TYPES OF DATA PROCESSING

Manual Data Processing – data is processed manually without using any machine or tool to get required results. In manual data processing, all the calculations and logical operations are performed manually on the data.

Mechanical Data Processing – data is processed by using different devices like typewriters, mechanical printers or other mechanical devices.

Electronic Data Processing is the modern technique to process data. The data is processed through computer; Data and set of instructions are given to the computer as input and the computer automatically processes the data according to the given set of instructions. The computer is also known as electronic data processing machine.

INTRODUCING THE DATABASE

Database: is a collection of information organized in such a way that it can be accessed easily.

Examples are:

- Tracking customer orders
- Maintaining Employees records
- Maintaining Students Information.

- Database: shared, integrated computer structure that stores a collection of:
  - **End-user data**: raw facts of interest to end user
  - **Metadata**: a set of data that describes and gives information about other data.
- Provides description of data characteristics and relationships in data
- Complements and expands value of data

TYPES OF DATABASES

- Single-user database supports only one user at a time
  - Desktop database: single-user; runs on PC
- Multiuser database supports multiple users at the same time
  - Workgroup and enterprise databases
- Centralized database: data located at a single site
- Distributed database: data distributed across several different sites
- Operational database: supports a company's day-to-day operations

– Transactional or production database

- Data warehouse: stores data used for tactical or strategic decisions

DATABASE DESIGN

Database design is the organization of data according to a database model.

The designer determines what data must be stored and how the data elements interrelate.

It focuses on design of database structure used for end-user data

Designer must identify database's expected use

- Well-designed database:
  - Facilitates data management
  - Generates accurate and valuable information
- Poorly designed database:
  - Causes difficult-to-trace errors

DATABASE MODEL

Database model is a type of data model that determines the logical structure of a database.

It determines in which manner data can be stored, organized and manipulated.

It is also communications tool to facilitate interaction among the designer, the applications programmer, and the end user

Good database design uses an appropriate data model as its foundation

Importance of database model

>> It organizes data for various users

>> End-users have different views and needs for data

DATABASE MODEL BASIC BUILDING BLOCKS

- Entity is anything about which data are to be collected and stored
- Attribute is a characteristic of an entity
- Relationship describes an association among (two or more) entities
  - **One-to-many (1:M) relationship**
  - **Many-to-many (M:N or M:M) relationship**
  - **One-to-one (1:1) relationship**

TYPES OF DATABASE MODELS

There are many kinds of data models. Some of the most common ones include:

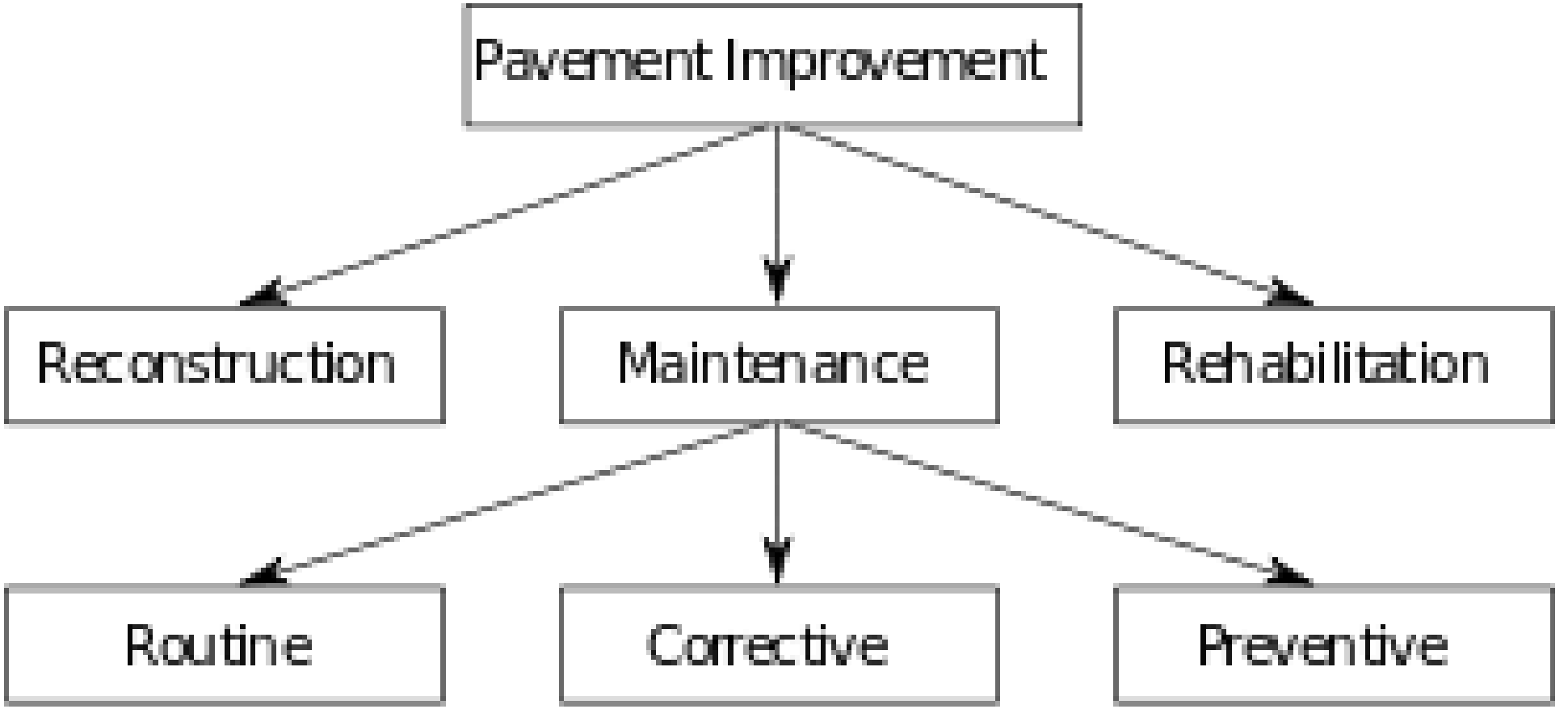
- Hierarchical
- Network
- Relational
- Entity relationship
- Object oriented

TYPES OF DATABASE MODELS.....

HIERARCHICAL MODEL

The hierarchical model organizes data into a tree-like structure, where each record has a single parent or root. Sibling records are sorted in a particular order. That order is used as the physical order for storing the database. This model is good for describing many real-world relationships.

Hierarchical Model

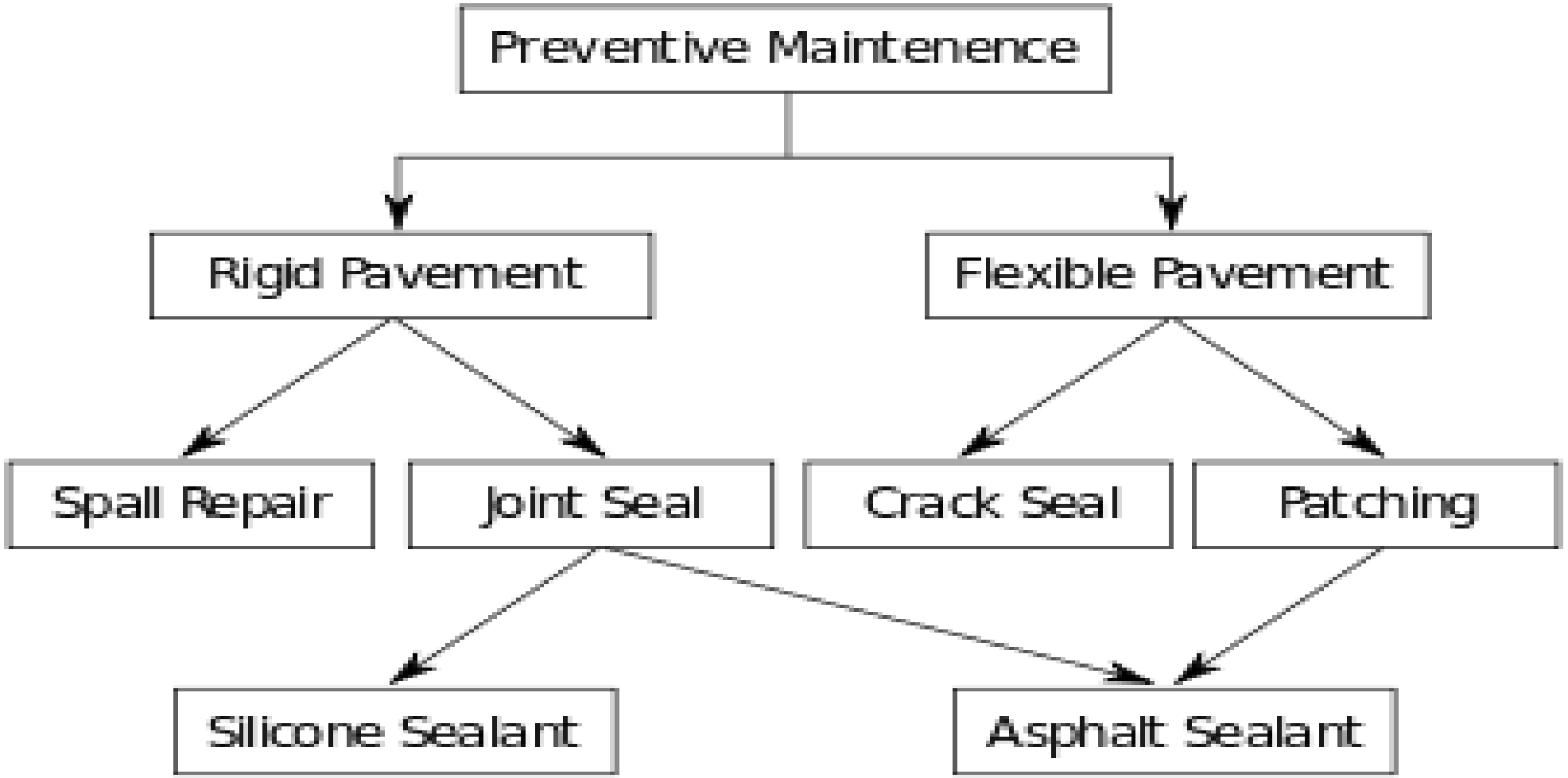


TYPES OF DATABASE MODELS.....

NETWORK MODEL

The network model builds on the hierarchical model by allowing many-to-many relationships between linked records, implying multiple parent records. Based on mathematical set theory, the model is constructed with sets of related records. Each set consists of one owner or parent record and one or more member or child records. A record can be a member or child in multiple sets, allowing this model to convey complex relationships.

Network Model



TYPES OF DATABASE MODELS.....

RELATIONAL MODEL

The most common model, the relational model sorts data into tables, also known as relations, each of which consists of columns and rows. Each column lists an attribute of the entity in question, such as price, zip code, or birth date.

A particular attribute or combination of attributes is chosen as a primary key that can be referred to in other tables, when it's called a foreign key.

Each row, also called a tuple, includes data about a specific instance of the entity in question, such as a particular employee.

Student ID	First name	Last name
52-743965	Charles	Peters
48-209689	Anthony	Sondrup
14-204968	Rebecca	Phillips

ProviderID	Provider name
156-983	UnitedHealth
146-823	Blue Shield
447-784	Carefirst Inc.

Student ID	ProviderID	Type of plan	Start date
52-743965	156-983	HSA	04/01/2016
48-209689	146-823	HMO	12/01/2015
14-204968	447-784	HSA	03/14/2016

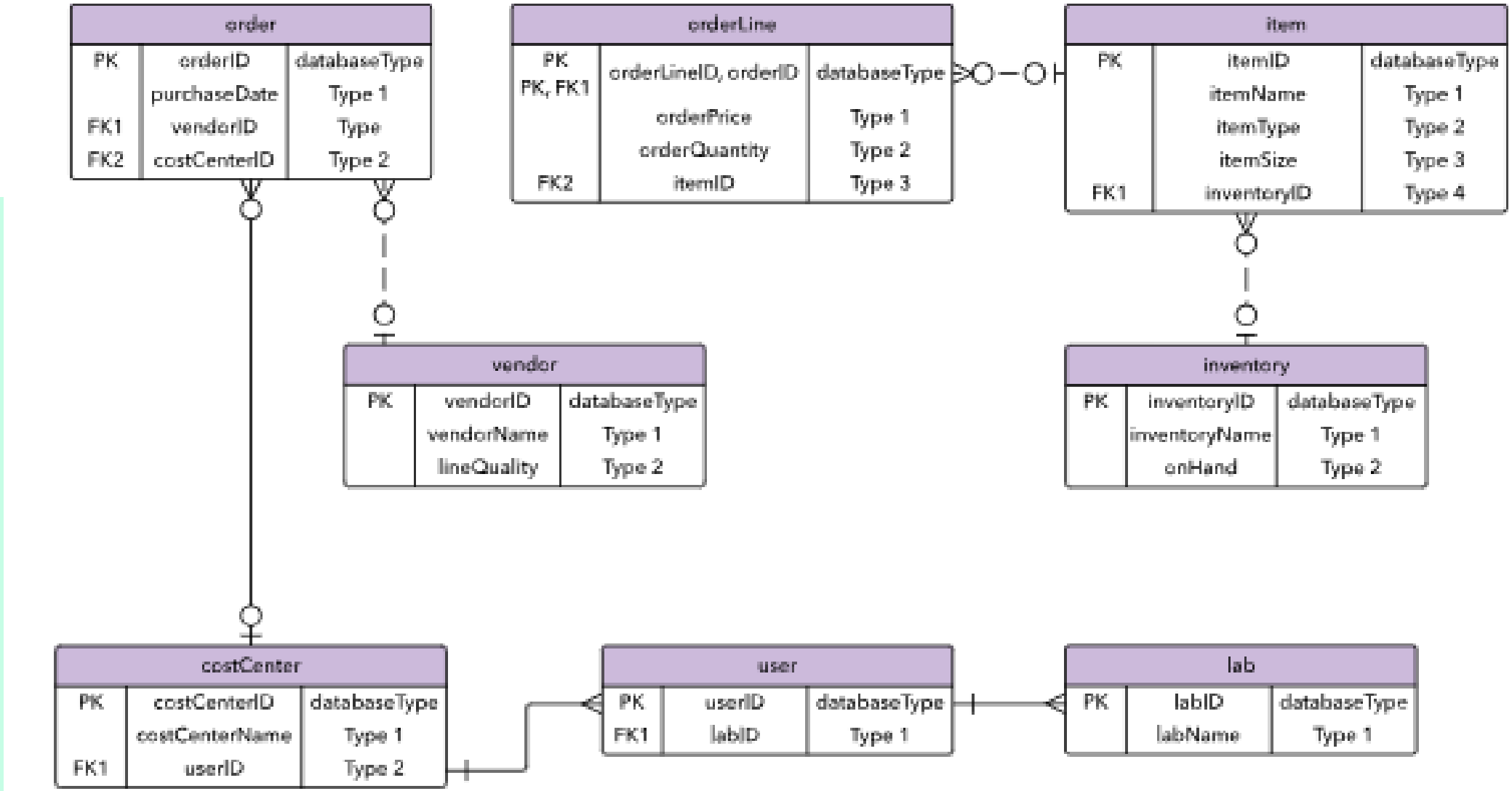
TYPES OF DATABASE MODELS……

ENTITY-RELATIONSHIP MODEL

This model captures the relationships between real-world entities much like the network model, but it isn't as directly tied to the physical structure of the database. Instead, it's often used for designing a database conceptually.

A common form of the ER diagram is the star schema, in which a central fact table connects to multiple dimensional tables.

See the diagram below


order
PK
orderID
databaseType
purchaseDate
Type 1
FK1
vendorID
Type
FK2
costCenterID
Type 2
orderLine
PK
PK, FK1
orderLineID, orderID
databaseType
orderPrice
Type 1
orderQuantity
Type 2
FK2
itemID
Type 3
item
PK
itemID
databaseType
itemName
Type 1
itemType
Type 2
itemSize
Type 3
FK1
inventoryID
Type 4
vendor
PK
vendorID
databaseType
vendorName
Type 1
lineQuality
Type 2
inventory
PK
inventoryID
databaseType
inventoryName
Type 1
onHand
Type 2
costCenter
PK
costCenterID
databaseType
costCenterName
Type 1
FK1
userID
Type 2
user
PK
userID
databaseType
FK1
labID
Type 1
lab
PK
labID
databaseType
labName
Type 1

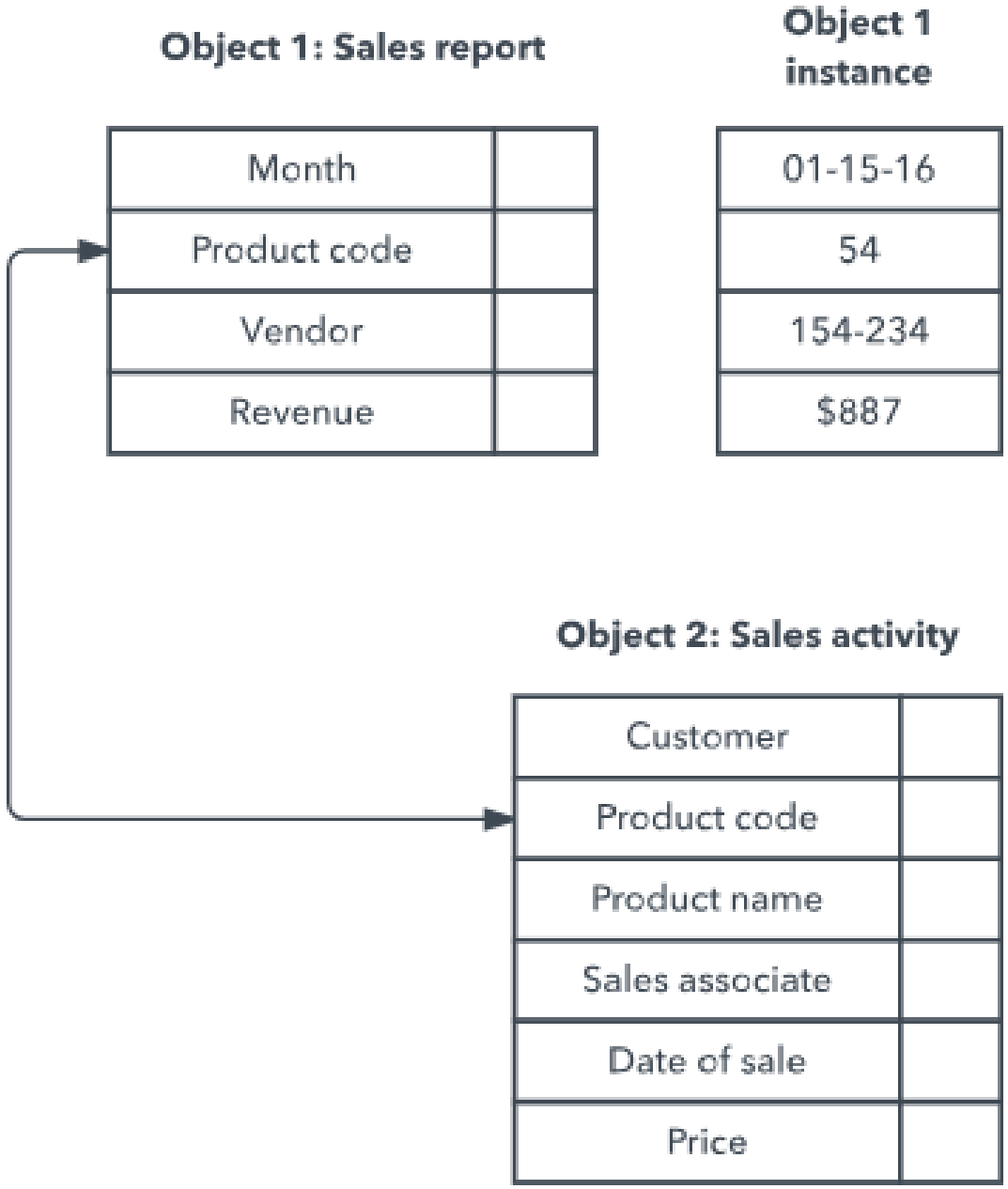
TYPES OF DATABASE MODELS……

OBJECT-ORIENTED DATABASE MODEL

This model defines a database as a collection of objects, or reusable software elements, with associated features and methods. There are several kinds of object-oriented databases:

A multimedia database incorporates media, such as images, that could not be stored in a relational database.

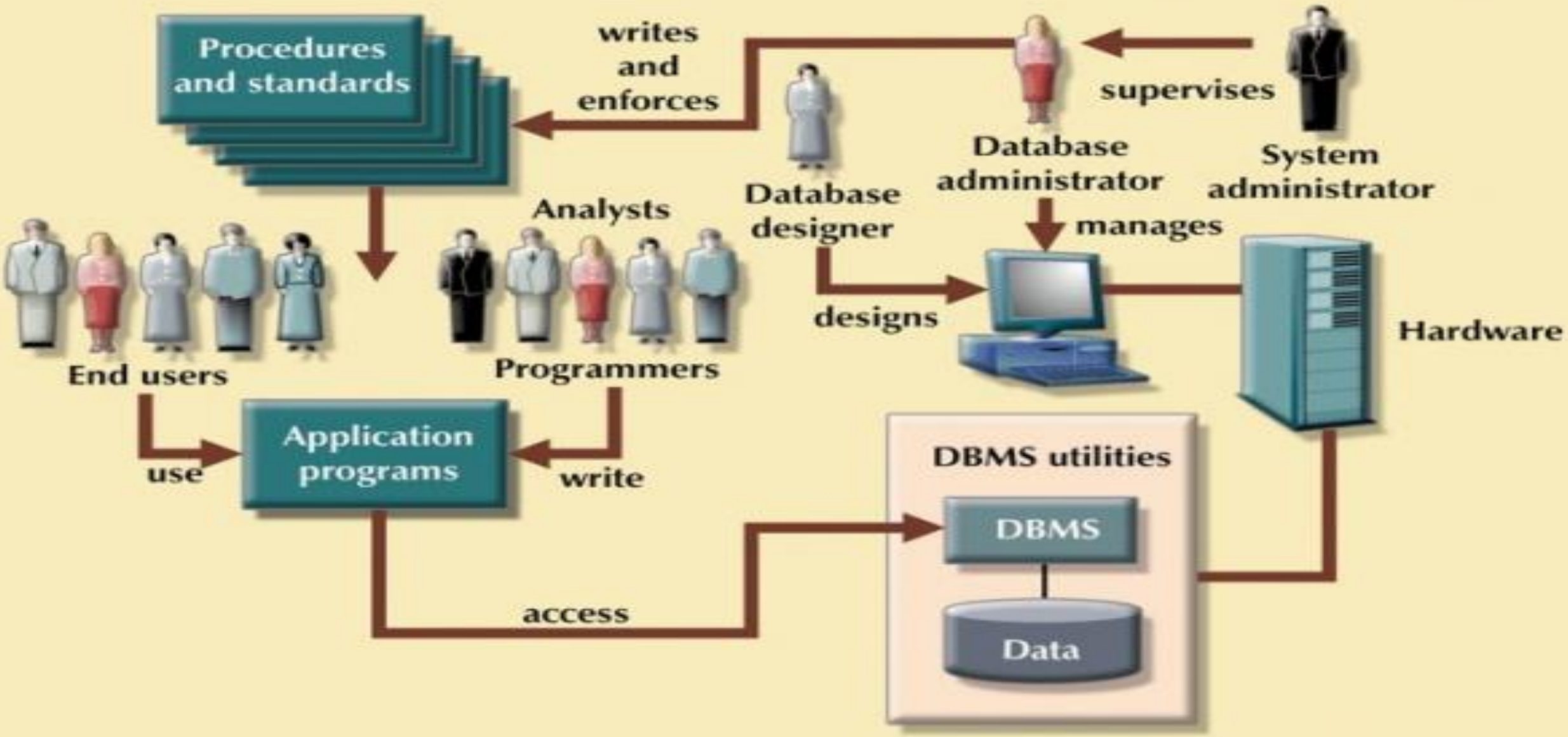
A hypertext database allows any object to link to any other object. It’s useful for organizing lots of disparate data, but it’s not ideal for numerical analysis.


Object 1: Sales report
Object 1 instance
Month
Product code
Vendor
Revenue
01-15-16
54
154-234
$887
Object 2: Sales activity
Customer
Product code
Product name
Sales associate
Date of sale
Price

THE DATABASE SYSTEM ENVIRONMENT

- **System:** a set of things working together as parts of a mechanism or an interconnecting network; a complex whole.
- Database system: defines and regulates the collection, storage, management, use of data
- Five major parts of a database system:
  - Hardware
  - Software
  - People
  - Procedures
  - Data

The database system environment



DATABASE TERMINOLOGIES

In studying Databases the following Terminologies must be clear

Data

- This is the fact or facts about specific entity (person, place or thing).

Information

- This is the processed facts perceived to be useful by the user.

Field

- This is a single item of information or data in entity. E.g. employee Name

Record

- This can be defined as data or group of fields about an entity. E.g. employee particulars

Subject (relation or table)

- This can be defined as a collection of records that are related to a particular entity. For example, Records for all employees will make up one subject.

DATABASE TERMINOLOGIES

Table is a set of data elements (values) using a model of vertical columns (identifiable by name) and horizontal rows, the cell being the unit where a row and column intersect.

Form is a window or screen that contains numerous fields, or spaces to enter data. Each field holds a field label so that any user who views the **form** gets an idea of its contents. A **form** is more user friendly than generating queries to create tables and insert data into fields.

Database report is the formatted result of database queries and contains useful data for decision-making and analysis. Most good business applications contain a built-in reporting tool; this is simply a front-end interface that calls or runs back-end database queries that are formatted for easy application usage.

RELATIONAL TERMINOLOGIES

- Table or Relation
- Null values
- Duplicate Values
- Changeable Values
- Primary Keys
- Foreign Keys

TABLE OR RELATION

- Table will store information for a particular entity
- *Table name must be unique*
- *The Table name should be descriptive*
- *Column Name must be unique within the Table*
- *Rows must be unique*

NULL VALUES

- missing or unknown value in a column of a table
- Nulls are not the same as zeros
- Most arithmetic operations can be performed on zero values
- nulls must be excluded from mathematical manipulations

DUPLICATE VALUES

- A duplicate value is a value in a column of a table that exactly matches some other values within the same column.

CHANGEABLE VALUES

- value in a table that may vary over time.
- Most values in most tables are Changeable
- You can prevent changes when it is desirable to prevent changes in a given column of a table

PRIMARY KEYS

- Uniquely identify each row of that table.
- Every table must have only one Primary key

RULES FOR PRIMARY KEY

- Must always have a value (null values are not allowed)
- Value should be unique (duplicate values are not allowed)
- Value should not change over time

FOREIGN KEYS

- key which relates the rows of the Table to other Tables
- value can be null,
- value can change
- value can be duplicated

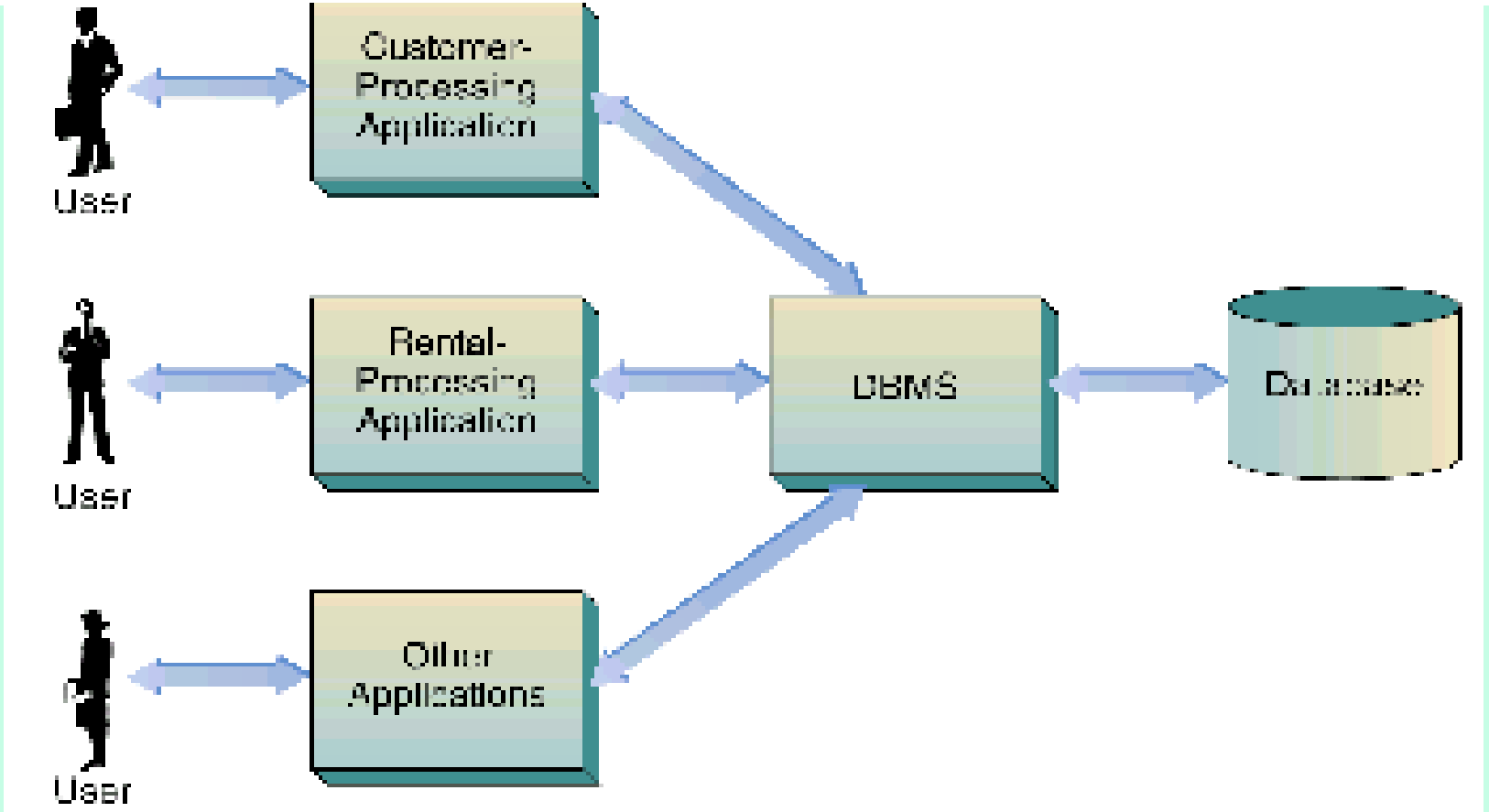
RULES FOR FOREIGN KEYS

- Value must refer to the existing primary key.

DATABASE MANAGEMENT SYSTEMS (DBMS)

- A program that allows users to define, create, manipulate, store, maintain, retrieve, and process the data in a database in order to produce meaningful information.
- Focus on information representation
- Data stored as records in various database files that can be combined to produce meaningful information for users
- DBMS controls all functions of capturing, processing, storing, retrieving data and generates various forms of data output
- Manages access by multiple users and multiple programs to a common store of data

DBMS Relationships



DBMS OVERCOMES ALL LIMITATIONS OF FPS.

- Eliminates separation and isolation of data
- Reduces data redundancy
- Eliminates dependence between programs and data
- Allows for representation of data from user's view
- Increases data flexibility
- Superior flexibility and security over spreadsheet applications

CHARACTERISTICS OF A DBMS

- Computerized record-keeping system
- Contains facilities that allow the user to:
  - Add, delete files
  - Insert, retrieve, update, delete data
- Collection of databases; each can be used for separate purposes or combined

FUNCTIONS AND USES OF A DBMS

- To store data
- To organize data
- To control access to data
- To protect data
- To provide decision support
- To provide transaction processing

ADVANTAGES AND DISADVANTAGES OF A DBMS

- Advantages:
  - Centralized data reduces management problems
  - Data redundancy and consistency are controllable
  - Program - data interdependency is diminished
  - Flexibility of data is increased

- Disadvantages:
  - Reduction in speed of data access time
  - Requires special knowledge
  - Possible dependency of application programs to specific DBMS versions