Introduction

Data v.s. Information

Data

known facts stored and recorded (text, numbers, dates, images, sound, videp...)

Information

Data presented in context (summarised data), processed data that can increase users knowledge. (tabled data; data in charts)

Data v.s. Informatin

1. data: known and available
2. infromation: processed and more useful

Metadata

A data dictionary that descrips about data in dataset (structure, rules, constraints...); required in systems

make dataset more consistency and meaningful

Database

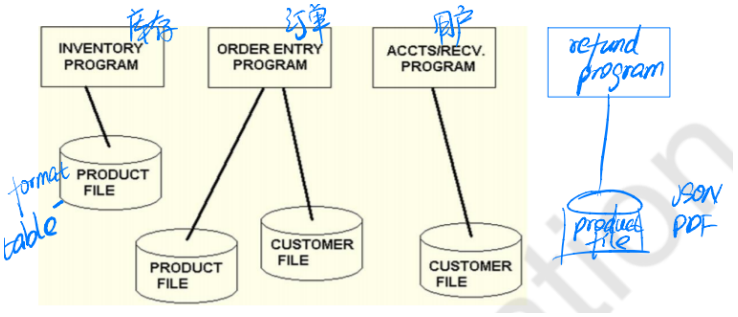
A large, intefrated, structured collection of data

usually intended to model some real-world entreprise; e.g. a university (Entities: courses, students, professors; Relationships: enrollment, teaching...)

Database Management System (DBMS)

software system to store, manage, and facilitate access to databases

File Processing Systems V.S. DBMS



cons of fps

1. Program-data dependence (program depends on the file structure)
2. Duplication of data (redundancy, lack of integrity: accuracy and consistency)
3. limited data sharing
4. lengthy development times (figure out file format)

5. Excessive program maintenance

pros of dbms

Manage sata in a structured way; relational model dominant since 1980 (table-rows & columns forming relations)

1. Data independence
2. Minimal data redundancy
3. Improved data consistency
4. Improved data sharing
5. Reduced program maintenance
6. Novel ad hoc data access without programming

Database development lifecycle

1. Database Planning (out of course)

based on data model; top level perspective on data requirements

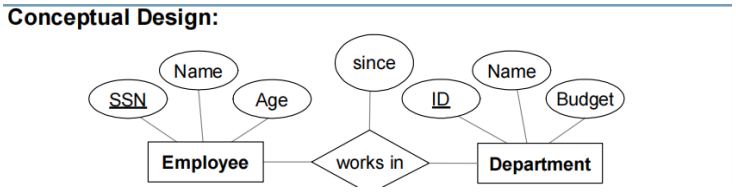
1. **Systems Definition** (How the database interacts with other information systems in the organisation; specify scope and boundaries; how to interfere with other systems)
2. **Requirements Definition and Analysis** (collection and analysis of requirements; a written documentation)

会给文字形式的题目，从中抽取对设计者的要求 (Business rule); 常见词汇：keep track of/store/need to know

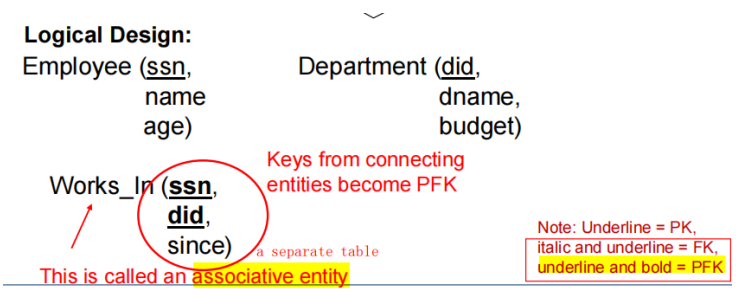
2. Design

A description of implementation of logical design.

1. **Conceptual Design** (construction of a model of the data: entities & attributes, relationships; no 'Database Model' has been applied; ER diagrams describe data model; independent of Specific DBMS)



1. **Logical Design** (Construction of a (relational or hierachical, ...) model (table) of the data based on the conceptual design; ER + key; independent of Specific DBMS)



1. **Physical Design** (data types selection, file organisation, indexed (based on mySQL/NoSQL); for a specific DBMS)

Physical Design:

Employee	Department	Works_In(
(ssn CHAR(11),	(did INTEGER,	ssn CHAR(11),
name VARCHAR(20),	dname VARCHAR(20),	did INTEGER,
age INTEGER)	budget FLOAT)	since DATE)

e.g.1 Mysql data types

1. Character Types CHAR(0~255)[5], VARCHAR(1~65535)[0~45], BIT, BOOL, CHAR(1), BLOB, TEXT(up to 65535 bytes), ENUM('value1','value2',...; up to 65535 mems)[多选一], SET('value1','value2',...; up to 64 items), BLOB[视频图片]
2. Integer Types TINYINT(-128~127), SMALLINT(0~65535/-32768~32767), MEDIUMINT(0~16777215),INT(0~4294967295),BIGINT...
3. Real Types FLOAT, DOUBLE/REAL, DECIMAL
4. Time and Date Types DATE, TIMES, DATETIME, TIMESTAMP, TEAR

e.g.2 Loof-up table (ID or name?)

ensure data field integrity (accuracy and consistency); handling missing data (NULL data); cons: speed down

e.g.3 De-Normalise or Not

Normalisation(大表拆小表): speed down; data integrity up De-Normalsation(小表并大表): speed up; wasted storage space + data integrity down

3. Application Design

design interface and application programs that use and process the database（与应用的交互设计）

4. Implementation

physical realisation of the db (create database tables)

5. Data Conversion and Loading

Transfer existing data into the database; Conversion from old systems; huge work!（新老data转换）

6. Testing (ooc)

Run and find errors; Performance, Robustness, Recoverability, Adaptability

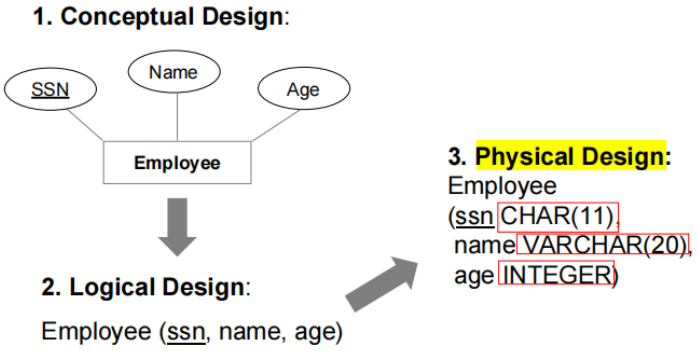
7. Operational Maintenance (ooc)

moniroting and maintaining the database system following its commissioning; handle new requirements or changes

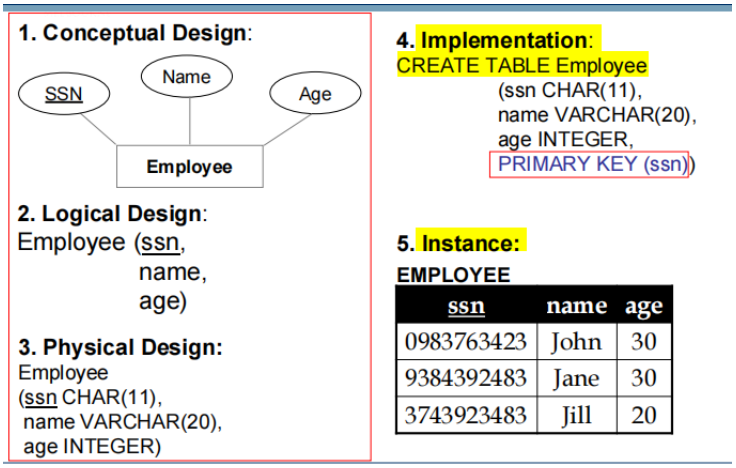
Logical & Physical Design: ER to Relational Model & Relational to Physical

logical design: entity set -> relation

physical design: choose data types



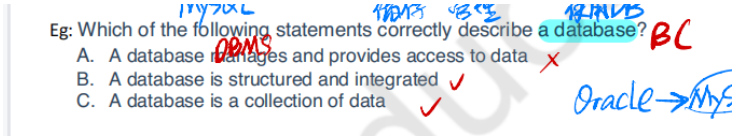
the entire cycle includes

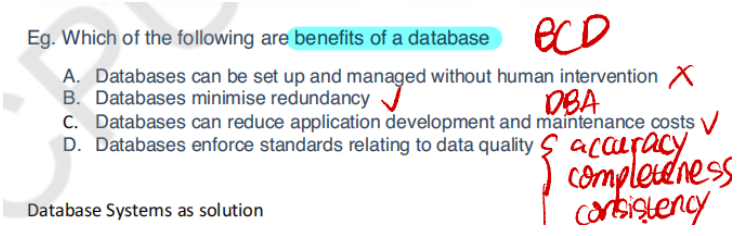


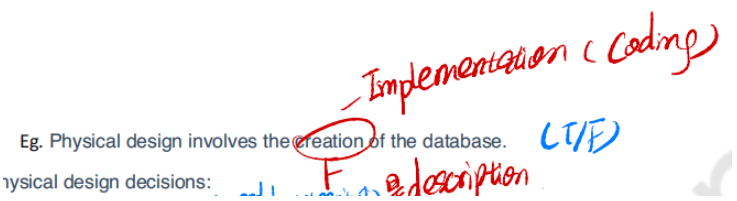
Implementation: choose the primary key

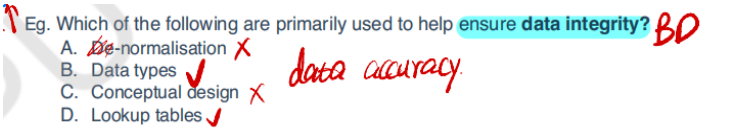
Instance: add data

Quizs









Eg. Conceptual design provides a model of data. (T/F)

T

Logical design

Construction of a model of how data will be stored

延续了 conceptual design 形成表格

Eg. Logical design is independent of a specific database management system. (T/F)

T

Physical design (datatype selection)

延续了 logical design，进行改良（选择 datatype）

A description of implementation of logical design

Implementation (coding)

Eg. Physical design involves the creation of the database. (T/F)

F 是description

ysical design decisions:

Eg.

Q10A. Which of the following is **NOT** part of the database development lifecycle? D

A) Implementation
B) Maintenance
C) Requirement analysis
D) First-level support