

## 2. 關聯資料庫

### 「關聯資料庫」是什麼？

林俊逸  
資訊管理學系  
長庚大學

### 「關聯資料庫」是什麼？

由三個層面說明

- 關聯資料庫是怎麼儲存資料？
  - 是用怎樣的結構來儲存資料？
- 在關聯資料庫如何保護資料的正確性？
  - 什麼資料才是有效的（Valid）資料
- 在關聯資料庫如何維護和運算資料？
  - 維護：新增、修改、刪除
  - 運算：查詢

以上三個答案加在一起稱為「關聯資料模型」

3

### 「資料模型」是什麼？

一個完整的資料模型包括三個部分

- 結構部分 (A Structure Part)
  - 資料結構
- 完整性規則 (Integrity Rules)
  - 確保資料是有效的
- 操作部分 (A Manipulative Part)
  - 允許對資料操作的類型
  - 擷取、新增、修改、刪除

## 關聯資料庫

### 以什麼結構儲存資料？

林俊逸  
資訊管理學系  
長庚大學

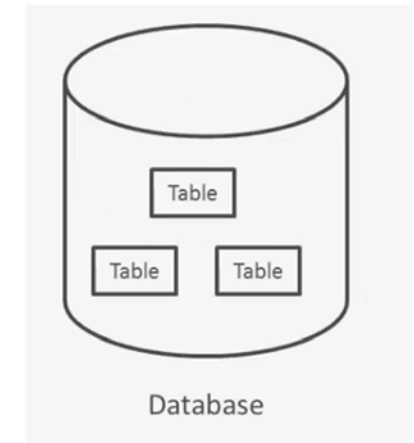
## 「資料模型」是什麼？

一個完整的資料模型包括三個部分

- 結構部分 (A Structure Part)
  - 資料結構
- 完整性規則 (Integrity Rules)
  - 確保資料是有效的
- 操作部分 (A Manipulative Part)
  - 允許對資料操作的類型
  - 擷取、新增、修改、刪除

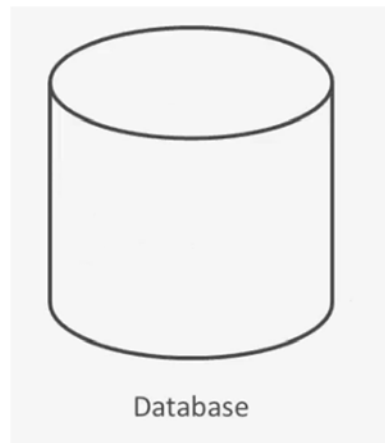
## 關聯資料庫

- 一個關聯資料庫裡有多個資料表
- 資料表是一個資料庫的最根本的構建塊
- 所有的數據都存進某一個資料表中



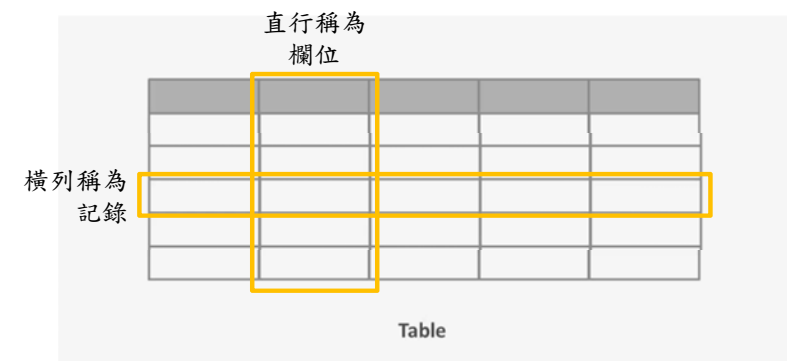
## 一個沒有資料表的資料庫

- 只是一個毫無意義的空殼
- 不只是裡面沒有任何資料
- 且沒有任何地方可以儲存資料



## 資料表

- 資料表結構看起來像是一張試算表
- 資料表包含欄位跟記錄
  - 欄位：由上而下
  - 記錄：由左而右



## 多個資料表

- 一個關聯資料庫裡有多個資料表
- 每個不同的資料表含有多筆記錄
  - 但都是屬於相同種類的資料

例如：

- 一個存著所有客戶的資料表
- 另一個存著所有訂單的資料表
- 另一個存著所有員工的資料表

## 多個資料表

- 每一筆記錄代表真實世界中的一個個體
  - 一個顧客，或是一筆訂單，或是一位員工
- 但是這些記錄並不是以自由形式存在
- 你必須應用結構在這些記錄上面
- 你必須規定一筆記錄是如何建構的

## 資料表 (Tables)

- 每個資料表都有自己的名子
- 資料表是一個有兩個維度的表格
  - 欄位：有自己的名子和可以儲存的資料種類（結構）
  - 記錄：存放在資料表裡的數據（內容物）
- 資料表的結構
  - 像是在描述一類擁有有多個儲存槽（欄位），且結構、大小一模一樣的容器
- 一筆一筆的記錄
  - 像是裝在一個這種容器內的內容物
- 理論上，一個資料表的欄位個數和擁有的記錄筆數是沒有上限的

## 「員工」資料表

		columns			
ROWS	FirstName (text)	LastName (text)	HireDate (date)	Grade (numeric)	Salary (currency)
	Alice	Mann	4/4/2009	4	75000
	James	Black	3/1/2010	4	75000
	Colista	Guerra	10/1/2006	6	80000
	Fay	Fitzgerald	7/21/2002	7	100000
	John	Bowen	11/11/2011	3	45000

- 每一筆記錄就是一名員工
- 每個欄位描述關於這個員工一個非常特定的資料段

## 「員工」資料表

		columns			
rows	FirstName (text)	LastName (text)	HireDate (date)	Grade (numeric)	Salary (currency)
	Alice	Mann	4/4/2009	4	75000
	James	Black	3/1/2010	4	75000
	Calista	Guerra	10/1/2006	6	80000
	Fay	Fitzgerald	7/21/2002	7	100000
	John	Bowen	11/11/2011	3	45000

- 結構：
  - 欄位名稱
  - 每個欄位的資料類型
- 當你在描述這些細節時，其實你是在制定這些欄位的必須遵守的規定
- 每一筆記錄必須遵循相同的結構（格式）

## 「Employees」資料表

		columns			
rows	FirstName (text)	LastName (text)	HireDate (date)	Grade (numeric)	Salary (currency)
	Alice	Mann	4/4/2009	4	75000
	James	Black	3/1/2010	4	75000
	Calista	Guerra	10/1/2006	6	80000
	Fay	Fitzgerald	7/21/2002	7	100000
	John	Bowen	11/11/2011	3	45000

- 資料庫管理系統不允許違反這些規定
- 例如：HireDate 欄位的資料類型是日期，就只能存日期，不可以存其它不是日期的資料
- 例如：如果規定這個欄位一定要有資料，就不允許建立沒有這個欄位資料的一筆記錄

## 「Students」資料表

std_no	name	sex	dept_no	DOB	height	phone_no	photo_id
B0341001	周潔倫	男	醫管	1996/12/15	176	0932333444	N459852364
B0341002	蔡依林	女	醫管	1997/1/1	155	0927555666	P987456321
B0344103	張惠妹	女	資管	1997/4/6	150	0968888666	K785426542
B0344204	李宗盛	男	資管	1996/10/12	174	0912111222	A457863256
B0244201	史彥文	男	資管	1996/9/21	168	0936123456	B784521458
B0244103	張三豐	男	資管	1997/6/6	178	0928999000	F457862534

- 一筆記錄→一位學生
- 一個欄位→學生的一個資料項
- 如果需要一個資料表來儲存「課程」資料，應該會有哪些欄位？

## 資料類型 (Data Type)

- 定義一種資料種類和在這些資料允許的操作
  - 「文字」和「數字」有何不同
- 每個資料類型都有一個名稱
- 有些資料類型可以宣告長度或大小
  - 八個字元的文字
  - 最大到百萬的整數
- 每個欄位必須宣告屬於哪一種資料類型和大小

## 常用資料類型

Data Type	Description
CHAR [(size)]	固定長度的文字資料 Example: 儲存「學號」資料 CHAR(8)
VARCHAR2(size)	變長的文字資料 Example: 儲存「地址」資料 VARCHAR2(50)
NUMBER [(p,s)]	數字 Example: NUMBER(3,2) → -9.99~9.99 NUMBER(2) → -99~99
DATE	日期和時間 Example: 儲存「加選時間」資料 20-2月-2016 上午08:05:46

## 「Students」資料表

std_no	name	sex	dept_no	DOB	height	phone_no	photo_id
B0341001	周潔倫	男	醫管	1996/12/15	176	0932333444	N459852364
B0341002	蔡依林	女	醫管	1997/1/1	155	0927555666	P987456321
B0344103	張惠妹	女	資管	1997/4/6	150	0968888666	K785426542
B0344204	李宗盛	男	資管	1996/10/12	174	0912111222	A457863256
B0244201	史彥文	男	資管	1996/9/21	168	0936123456	B784521458
B0244103	張三豐	男	資管	1997/6/6	178	0928999000	F457862534

幫每個欄位找一個適合的資料類型與大小

### 選項

- 固定寬度的文字
- 數字
- 日期
- 變動寬度文字

## 試看看

■ 下列的欄位，你覺得哪一種資料類型和大小比較適合？

- 地址
- 修課成績
- 性別
- 月薪
- 折扣
- 流水號

### 選項

- 固定寬度的文字
- 數字
- 日期
- 變動寬度文字

## 關聯資料表的 7 個特性

1. 資料表名稱具 唯一性 (與所有其它資料表名稱不同)
2. 在資料表內每個欄位都有一個 獨特 的名字
3. 同一個欄位的值都來自同一個 值域 ( 資料類型 和 大小 )
4. 資料表中的每個儲存格最多只可以包含 一個 值
  - 因此將多個電話號碼存放在單個儲存格的作法是不對的
  - 稱為：沒有重複的組 (No Repeating Groups)
5. 每筆記錄都是 唯一的
  - 沒有任何 重複的 記錄
6. 欄位出現的順序是 不具意義的
7. 記錄出現的順序是 不具意義的

## 替代用語

Table-oriented	Set-oriented	Record-oriented
Table	Relation (關聯)	Record-type, file
Row	Tuple (值組)	Record
Column	Attribute (屬性)	Field
End Users	Academic Researchers	Information System Professionals

## 替代用語

Employee Data Table					
EMPLOYEE_ID	LAST_NAME	FIRST_NAME	SALARY	COMMISSION_PCT	DEPARTMENT_ID
100	King	Steven	24000		90
101	Kochhar	Neena	17000		90
102	De Haan	Lex	17000		90
103	Hunold	Alexander	9000		60
104	Ernst	Bruce	6000		60
107	Lorentz	Diana	4200		60
124	Mourgos	Kevin	5800		50
141	Rajs	Trenna	3500		50
142	Davies	Curtis	3100		50
143	Matos	Randall	2600		50
144	Vargas	Peter	2500		50
149	Zlotkey	Eleni	10500	.2	80
174	Abel	Ellen	11000	.3	80
176	Taylor	Jonathon	8600	.2	80
178	Grant	Kimberely	7000	.15	
200	Whalen	Jennifer	4400		10
201	Hartstein	Michael	13000		20
202	Fay	Pat	6000		20
205	Higgins	Shelley	12000		110
206	Gietz	William	8300		110

Diagram labels: Column Attribute points to DEPARTMENT\_ID; Field points to the DEPARTMENT\_ID cell for employee 102; Row Tuple Record points to the entire row for employee 201.



PROJECT_CODE	PROJECT_MANAGER	MANAGER_PHONE	MANAGER_ADDRESS	PROJECT_BID_PRICE
21-5Z	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	16833460.00
25-2D	Jane D. Grant	615-898-9909	218 Clark Blvd., Nashville, TN 36362	12500000.00
25-5A	George F. Dorts	615-227-1245	124 River Dr., Franklin, TN 29185	32512420.00
25-9T	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	21563234.00
27-4Q	George F. Dorts	615-227-1245	124 River Dr., Franklin, TN 29185	10314545.00
29-2D	Holly B. Parker	904-338-3416	3334 Lee Rd., Gainesville, FL 37123	25559999.00
31-7P	William K. Moor	904-445-2719	216 Morton Rd., Stetson, FL 30155	56850000.00

1. How many records does the file contain, and how many fields are there per record?
2. Using the table above, what problem would you encounter if you wanted to produce a listing by city? How would you solve this problem by altering the table structure?

## 定義資料表、欄位之後

- 在你定義資料表、欄位之後，接下來的日常工作所接觸到的幾乎都是記錄
  - 新增
  - 修改
  - 刪除
  - 讀取
- 要完成這些工作，你必須要有能力來明確的指定特定的一筆記錄
  - 某一顧客或是某一筆訂單
  - 要完成這些任務，我們需要「主鍵」 (Primary Key) 的概念

## 關聯資料庫

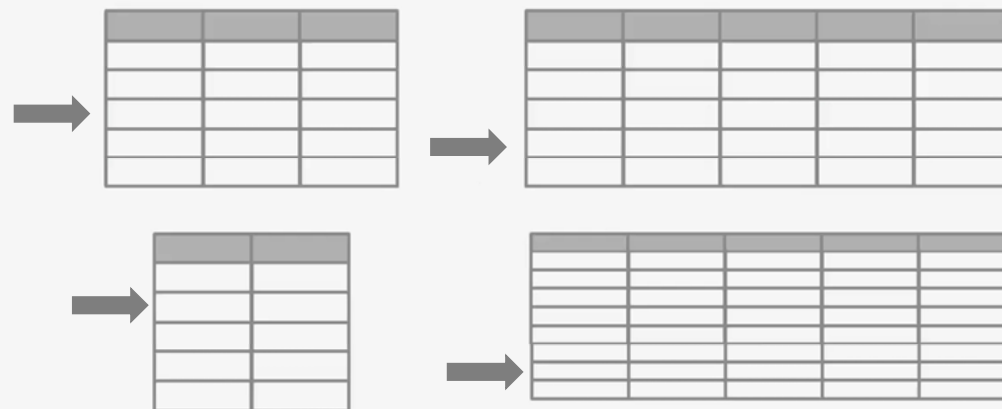
如何在眾多記錄中，  
明確的指定特定的一筆記錄？

林俊逸  
資訊管理學系  
長庚大學

### 關聯資料表的 7 個特性

1. 資料表名稱具唯一性（與所有其它資料表名稱不同）
2. 在資料表內每個欄位都有一個獨特的名字
3. 同一個欄位的值都來自同一個值域（資料類型和大小）
4. 資料表中的每個儲存格最多只可以包含一個值
  - 因此將多個電話號碼存放在單個儲存格的作法是不對的
5. 每筆記錄都是唯一的
  - 沒有任何重複的記錄
6. 欄位出現的順序是不具意義的
7. 記錄出現的順序是不具意義的

### 如何明確的指定特定的一筆記錄？



### 「Students」資料表

std_no	name	sex	dept_no	DOB	height	phone_no	photo_id
B0341001	周潔倫	男	醫管	1996/12/15	176	0932333444	N459852364
B0341002	蔡依林	女	醫管	1997/1/1	155	0927555666	P987456321
B0344103	張惠妹	女	資管	1997/4/6	150	0968888666	K785426542
B0344204	李宗盛	男	資管	1996/10/12	174	0912111222	A457863256
B0244201	史彥文	男	資管	1996/9/21	168	0936123456	B784521458
B0244103	張三豐	男	資管	1997/6/6	178	0928999000	F457862534

唯一的 unique	不唯一的 not unique
245	100
396	200
124	100
969	300
246	200

## 候選鍵 (Candidate Key)

- 一個候選鍵可以只包含一個欄位，或是多個欄位
- 候選鍵具有下列兩個特質
  - 唯一性 (Uniqueness)
    - 候選鍵的值能夠唯一識別每一筆記錄
  - 不可減少性 (Irreducibility)
    - 如果從中刪除任何欄位，將使它不再具有唯一識別特性

## 唯一值欄位 (Unique Column)

std_no	name	sex	dept_no	DOB	height	phone_no	photo_id
B0341001	周潔倫	男	醫管	1996/12/15	176	0932333444	N459852364
B0341002	蔡依林	女	醫管	1997/1/1	155	0927555666	P987456321
B0344103	張惠妹	女	資管	1997/4/6	150	0968888666	K785426542
B0344204	李宗盛	男	資管	1996/10/12	174	0912111222	A457863256
B0244201	史彥文	男	資管	1996/9/21	168	0936123456	B784521458
B0244103	張三豐	男	資管	1997/6/6	178	0928999000	F457862534

## 候選鍵 (Candidate Key)

- 一個資料表必須至少擁有一個候選鍵
- 一個資料表可能具有多個候選鍵
- 當一個候選鍵由多個欄位組成，我們把它叫做複合鍵 (Composite Key)



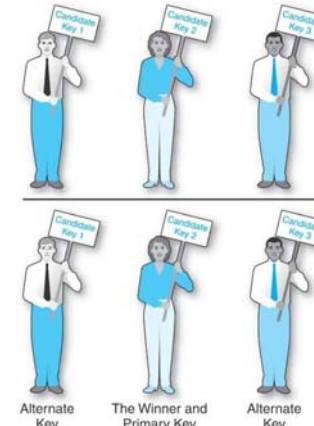
## 候選鍵

### Students 資料表

std_no	name	sex	dept_no	DOB	height	phone_no	photo_id
--------	------	-----	---------	-----	--------	----------	----------

欄位組合	唯一性?	不可減少性?	候選鍵?
std_no			
std_no+dept_no			
std_no+name			
name			
dept_no+name			
dept_no			
photo_id			
photo_id+std_no			
phone_no			

## 候選鍵、主鍵 & 替代鍵



- **主鍵 (Primary Key, PK)**
  - 在所有候選鍵當中，選擇其中一個充當主鍵
  - 一個資料表只會有一個主鍵
  - 在資料運算時，需要唯一識別記錄時，便使用主鍵的值來識別記錄
- **替代鍵 (Alternate Key, AK)**
  - 其它未被選來當主鍵的候選鍵
  - 一個資料表可以有多個替代鍵

## 選擇主鍵的原則

- 在可見的未來應該不會喪失唯一識別特性
- 任何時間點皆擁有資料
- 內容改變的可能性越小越好
- 欄位個數越少越好
  - 最少：1 個
- 內容越小（短）越好
- 使用者在資料使用上越方便越好

## 選擇主鍵

### Students 資料表

std_no	name	sex	dept_no	DOB	height	phone_no	photo_id
--------	------	-----	---------	-----	--------	----------	----------

欄位組合	候選鍵?	主鍵?	替代鍵?
std_no			
photo_id			
phone_no			

## 主鍵 (Primary Key)

- 在資料表中，將某個具有唯一性的欄位定為『主鍵』
  - 通常只包含一個欄位
  - 有時會包含多個欄位
- 每一筆記錄的主鍵含有一個不會重複（唯一）的值
- 一個資料庫中，幾乎所有資料表都需要一個主鍵

## 主鍵有時會包含多個欄位

修課

開課序號	學號	成績
14980	B8844001	82
14980	B8844002	88
14985	B8844002	80
14989	B8844001	88
14990	B8844001	96
14990	B8844002	90
15005	B8844001	88

## 作業

HW 第三章 69頁 Questions : 3