

《資料庫應用》

一、下列為包含五筆交易之交易資料庫，在購買的商品欄位中，每一個英文字母代表一種商品，例如編號T100的交易購買M、O、N、K、E、Y六種商品。

交易編號	購買的商品
T100	{M, O, N, K, E, Y}
T200	{D, O, N, K, E, Y}
T300	{M, A, K, E}
T400	{M, U, C, K, Y}
T500	{C, O, K, I, E}

關聯規則的形式為 $A \Rightarrow B$ ，其中A與B皆為商品的集合（例如{M, K}），表示如果顧客買了A集合的商品就會買B集合的商品。假設最小支持度與最小信賴度分別為60%與80%，請找出所有滿足最小支持度與最小信賴度的關聯規則。（20分）

試題評析	Association Rule的計算題型已經很多年沒出現在資料庫應用的考試中了。資料探勘是資料庫大量資料應用的一種常見方式，因此偶爾資料庫應用便會出現資料探勘的考題。其中Association Rule在考試上是資料探勘最常被考到的技術。考生應對其Support, Confidence的定義與計算都要稍加熟悉。此題關聯規則資料較多，需要多花時間仔細作答。
考點命中	《高點·高上資料庫應用講義》第三回，唐筆編撰，頁16-17。

答：

(一)各商品在五個交易中出現次數分別為M=3,O=3,N=2,K=5,E=4,Y=3,D=1,A=1,U=1,C=2,I=1。

故大於等於minimum support (60%)的large itemset為{M,O,K,E,Y}，即五筆交易至少有三筆購買該商品。

{M,O,K,E,Y}組合成的 $A \Rightarrow B$ 形式的關聯規則如下，並推算各規則的Support與Confidence:

Association Rule	Support	Confidence	Association Rule	Support	Confidence
M→O	0.2	0.33	O→M	0.2	0.33
M→K	0.6	1	K→M	0.6	0.6
M→E	0.4	0.67	E→M	0.4	0.5
M→Y	0.4	0.67	Y→M	0.4	0.67
O→K	0.6	1	K→O	0.6	0.6
O→E	0.6	1	E→O	0.6	0.75
O→Y	0.4	0.67	Y→O	0.4	0.67
K→E	0.8	0.8	E→K	0.8	1
K→Y	0.6	0.6	Y→K	0.6	1
E→Y	0.4	0.5	Y→E	0.4	0.67

上表超過最小支持度(60%)與最小信賴度(80%)之關聯規則為：M→K, O→K, O→E, K→E, E→K, Y→K。

承上， $A \Rightarrow B$ 的A或B為兩樣商品集合，即 $A \Rightarrow B$ 供包含三樣商品的可能組合為(表三樣商品至少一起出現於三筆交易)：

OKE：共同出現次數為3

可能組合為

Association Rule	Support	Confidence	Association Rule	Support	Confidence
O→KE	0.6	1	KE→O	0.6	0.75
K→OE	0.6	0.6	OE→K	0.6	1
E→OK	0.6	0.75	OK→E	0.6	1

上表超過最小支持度(60%)與最小信賴度(80%)之關聯規則為：O→KE, OE→K, OK→E

結論：所有滿足之關聯規則為 $M \rightarrow K, O \rightarrow K, O \rightarrow E, K \rightarrow E, E \rightarrow K, Y \rightarrow K, O \rightarrow KE, OE \rightarrow K, OK \rightarrow E$

二、檜木傢俱公司製造與販售傢俱，其關聯式資料庫（Relational Database）包含了顧客、訂單、傢俱及訂購四個表格（Table）。顧客表格記錄顧客編號、顧客名稱和顧客地址；訂單表格記錄訂單編號、訂單日期和下訂單的顧客編號，每一個訂單只有一位下訂單的顧客；傢俱表格記錄傢俱編號、傢俱名稱和傢俱定價；訂購表格記錄某一個訂單訂購某一種傢俱以及訂購傢俱的數量。這四個表格的關聯網要（Relational Schema）如下所示：

顧客（顧客編號、顧客名稱、顧客地址）

訂單（訂單編號、訂單日期、顧客編號）

傢俱（傢俱編號、傢俱名稱、傢俱定價）

訂購（訂單編號、傢俱編號、訂購數量）

有底線的屬性（Attribute）為該表格之主鍵（Primary Key）；例如，訂購表格之主鍵為訂單編號與傢俱編號兩個屬性的組合。訂單表格有一個外來鍵（Foreign Key）顧客編號，參照顧客表格的主鍵。訂購表格有訂單編號與傢俱編號兩個外來鍵，分別參照訂單表格與傢俱表格的主鍵。

針對下列三個查詢問題，請各寫出一個SQL指令來進行查詢。

（一）列出售價（定價打八折）小於10000的每一種傢俱之傢俱編號與傢俱名稱。（10分）

（二）列出訂購各種傢俱的總數量大於12的每一個訂單之訂單編號及其訂購各種傢俱的總數量。（15分）

（三）列出所有訂單都有訂購的傢俱之傢俱編號與傢俱名稱。（15分）

試題評析	SQL是每年必考題型。今年的語法小小變化但不算複雜，仔細閱讀題目、小心作答，有機會完整取分。
考點命中	《高點·高上資料庫應用講義》第二回，唐箏編撰，頁100-103。

答：

- （一）
- ```
SELECT 傢俱編號, 傢俱名稱
FROM 傢俱
WHERE 傢俱定價*0.8<10000
```
- （二）
- ```
SELECT  訂單編號, SUM(訂購數量) AS 總數量
FROM    訂購
GROUP BY  訂單編號
HAVING  SUM(訂購數量)>12
```
- （三）
- ```
SELECT 傢俱編號, 傢俱名稱
FROM 傢俱
WHERE NOT EXISTS
 (SELECT *
 FROM 訂購
 WHERE NOT EXISTS
 (SELECT *
 FROM 訂購
 WHERE 傢俱.傢俱編號=訂購.傢俱編號 AND 訂購.訂單編號=訂購.訂單編號))
```

三、「資料獨立」（Data Independence）是使用資料庫的方法來管理資料的優點之一，請說明資料獨立的做法及資料獨立的優點。（20分）

|      |                                                 |
|------|-------------------------------------------------|
| 試題評析 | 此題為資料庫基本觀念題，資料獨立性是第一回講義在看資料庫架構時的重點之一，考生們應可輕鬆取分。 |
| 考點命中 | 《高點·高上資料庫應用講義》第一回，唐箏編撰，頁8。                      |

**答：**

以ANSI/SPARC三層式架構的觀點，資料獨立性指資料庫的內部層(Internal level)資料儲存結構改變，或概念層(Conceptual level)資料格式的改變，皆不會影響到外部層(External level)的應用程式。一般分為邏輯資料獨立(Logical Data Independence)，以及實體資料獨立(Physical Data Independence)。

(一)做法：資料庫結構與上層應用程式互相獨立，資料庫的結構、存取方式、資料格式的改變，只需透過資料庫管理系統，從事適當的管理或修改，而不必更改資料庫的應用程式。

(二)定義與優點：

- 1.邏輯資料獨立(Logical Data Independence)：改變概念層次時，不必改變外部層次或應用程式。優點例如：在資料庫中新增或修改一個欄位或資料項目，甚至新增修改一個table時，外部層次不需作任何更動。
- 2.實體資料獨立(Physical Data Independence)：改變內部層次，如資料重組時，不必改變概念層次或外部層次。優點例如：因應資料量的增加，將循序檔更改為索引檔時，概念層次與外部層次不必變動。

四、供應商供應（供應商編號，供應商城市，供應商等級，零件編號，零件供應數量）是一個關聯網要，其中供應商供應為表格名稱，供應商編號、供應商城市、供應商等級、零件編號和零件供應數量為表格的五個屬性，供應商編號與零件編號兩個屬性的組合為表格的主鍵，同時也是該表格唯一的候選鍵（Candidate Key）。另外，供應商供應表格有下列四個功能相依（Functional Dependency）：（每小題10分，共20分）

{ 供應商編號，零件編號 } → { 零件供應數量 }

{ 供應商編號 } → { 供應商城市 }

{ 供應商編號 } → { 供應商等級 }

{ 供應商城市 } → { 供應商等級 }

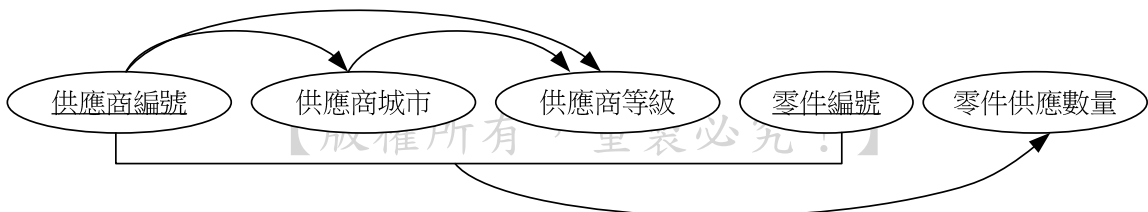
(一)請問供應商供應表格符合第幾正規型式（Normal Form）？請說明理由。

(二)供應商供應表格應該要正規化到第幾正規型式才適當？請將供應商供應表格正規化到適當的正規型式，寫出正規化之後每一個表格的網要，包括表格名稱、屬性、主鍵及外來鍵（如果有外來鍵的話）。

|      |                                                             |
|------|-------------------------------------------------------------|
| 試題評析 | 正規化是第二題SQL語法外，資料庫的第二大命題重點。此題型為基本的第二與第三正規化，稍有準備的考生，相信應可輕鬆應付。 |
| 考點命中 | 《高點·高上資料庫應用講義》第二回，唐箏編撰，頁12-14。                              |

**答：**

(一)依題目各屬性功能相依關係圖示如下：



其中，供應商城市、供應商等級，僅部分功能相依於主鍵中的供應商編號，不滿足第二正規(2NF)定義。因此，假設各欄位值皆為基元值，此供應商供應表格，僅符合第一正規形式(1NF)。

(二)1.供應商供應表格，正規化至2NF後，去除部分功能相依屬性如下：

供應商供應(供應商編號, 零件編號, 零件供應數量)

供應商(供應商編號, 供應商城市, 供應商等級)

其中, 供應商供應的供應商編號欄位為外鍵, 參考至供應商表格的供應商編號欄位。

2. 以上正規化之後, 供應商供應已滿足3NF; 但供應商表格仍存在遞移相依, 即供應商等級遞移相依於主鍵。故將供應商表格3NF之後, 各表格綱要如下:

供應商供應(供應商編號, 零件編號, 零件供應數量)

供應商(供應商編號, 供應商城市)

供應商城市等級(供應商城市, 供應商等級)

其中,

供應商供應的供應商編號欄位為外鍵, 參考至供應商表格的供應商編號欄位。

供應商的供應商城市欄位為外鍵, 參考至供應商城市等級表格的供應商城市欄位。

高點  
·  
高上

【版權所有，重製必究！】