

經濟部所屬事業機構 108 年新進職員甄試試題

類別：統計資訊

節次：第三節

科目：1. 資料庫及資料探勘 2. 程式設計

注意事項

1. 本試題共 3 頁(A3 紙 1 張)。
2. 可使用本甄試簡章規定之電子計算器。
3. 本試題分 6 大題，每題配分於題目後標明，共 100 分。須用藍、黑色鋼筆或原子筆在答案卷指定範圍內作答，不提供額外之答案卷，作答時須詳列解答過程，於本試題或其他紙張作答者不予計分。
4. 本試題採雙面印刷，請注意正、背面試題。
5. 考試結束前離場者，試題須隨答案卷繳回，俟本節考試結束後，始得至原試場或適當處所索取。
6. 考試時間：120 分鐘。

一、下列距離矩陣呈現了 A ~ F 這 6 個物件兩兩間之距離，另群集為物件之集合，試回答下列問題：(15 分)

B	16				
C	47	37			
D	72	57	40		
E	77	65	30	31	
F	79	66	35	23	10
	A	B	C	D	E

(一)若定義群集間之距離，為各自包含物件間之最小距離(單一鏈結 Single-linkage)，請以聚合式階層分群法，由下而上建立 A ~ F 這 6 個物件之群集樹狀圖(Dendrogram)並簡述各步驟。(10 分)

(二)若將這 6 個物件分為 3 個群集，試問其做法為何？另各群集中含有哪些物件？(5 分)

二、在設計資料庫的過程中，正規化(Normalization)是建構資料模式所運用的一種技術，其目的是為了降低資料的重覆性與避免更新異常的情況發生。故於正規化的過程中，常將原先關聯表格的所有資料進行分解(Decomposition)，試回答下列問題：(15 分)

(一)請比較有損分解(Lossy Decomposition)與無損分解(Lossless Decomposition)之差異。(2 分)

(二)請針對下方表格，列出無損分解(Lossless Decomposition)的 BCNF (Boyce-Codd Normal Form)型式：(13 分)

cID	cName	tID	tName	sID	sName	score
CS101	Database	1001	Scott	96010	Chen	70
				96013	Wang	90
				96020	Lee	88
CS103	Probability	1015	Andy	96033	Ho	58
				96045	Tan	92

三、有一船員資料表 Sailors 如下，試回答下列問題：（20 分）

編號(sid)	姓名(sname)	評價等第(rating)	年齡(age)
22	Dustin	7	45.0
31	Lubber	8	55.5
71	Zorba	10	16.0
64	Horatio	7	35.0
29	Brutus	1	33.0
58	Rusty	10	35.0

(一)下列 SQL 指令之查詢結果為何？（5 分）

```

SELECT      S.rating, MIN(S.age) AS min_age
FROM        Sailors S
WHERE       S.age >= 18
GROUP BY   S.rating
HAVING     COUNT (*) > 1 ;
    
```

(二)請寫出 SQL 指令以查詢：年滿 20 歲船員之平均評價等第與平均年齡。（5 分）

(三)請寫出 SQL 指令以查詢：最年長船員的姓名與年齡。（5 分）

(四)請寫出 SQL 指令以查詢：各評價等第中，平均年齡最低者及其平均年齡。（5 分）

四、有一 Web 程式系統，由使用者輸入 ID 及地址，並由後端接收處理程式進行資料庫相關處理，其前端畫面輸入欄位、對應之 Html 碼及後端接收處理程式如下，試回答下列問題：（14 分）

前端畫面輸入欄位：

請輸入 ID：

請輸入地址：

對應之輸入欄位 Html 碼：

```

< input type= " text " name= " id " >
< input type= " text " name= " address " >
    
```

後端接收處理程式及註解(雙斜線//後為註解)：

```

< %
.....//以上省略

$Id = Request( " id " ) ;           //取得輸入 ID 欄位資料
$address = Request( " address " ) ; //取得輸入地址欄位資料

$sqlstring = " Select * from idtable where id = '"+$Id+"' " ; //id 是否存在於 idtable 資料表
$result = sql_query( $sqlstring ) ; //進行資料庫查詢及結果
.....//以下省略

% >
    
```

(一)請問上述設計會產生何種資料庫資安風險？(2分)

(二)請列舉上述資安風險對系統及資料所產生之危害。(4分)

(三)請在不變動系統任何設定及前端輸入下，就你所知道之任一程式語言(含虛擬碼)，在後端接收處理程式新增字串處理函數 `checkdata(XXXX)`，傳入參數為使用者輸入之資料，經處理後傳回無風險之資料(除宣告函數外，每行程式碼皆須註解處理內容，未加註解者不予計分)。(8分)

五、請就二元搜尋樹回答下列問題：(20分)

(一)請說明欲建立二元搜尋樹，必須滿足哪些條件？(6分)

(二)數列 27、35、17、33、20、3、38，試以第 1 個數字為根，寫出其二元搜尋樹及建立的步驟。(10分)

(三)在上述二元搜尋樹中，若欲刪除元素 27，請寫出 2 種做法。(4分)

六、在常見的程式設計語言中，變數常區分為全域變數(Global Variable)與區域變數(Local Variable)，並在某些情況下使用靜態變數(Static Variable)，試回答下列問題：(16分)

(一)何謂靜態變數與區域變數？並說明兩者的特性、差別及個別的生命週期。(6分)

(二)在一般程式設計中，若區域變數與全域變數同名，試問何者優先使用？(2分)

(三)有一打彈珠機函數 `balls`，接收引數 `input1`、`input2` 分別代表輸入值與輸出倍數，其功能要求：80%機率失敗，回傳值為 0；20%機率成功，回傳值 = 輸入值 * 輸出倍數，宣告一變數 `count` 初始值 1,000，並記錄每次呼叫後之剩餘值。請在下列底線部分填入適當程式碼：(每項 2 分，共 8 分)

```
int balls ( int input1, int input2 ) { // input1 : 輸入值 ; input2 : 輸出倍數
    _____ (1)
    int output = 0 ;
    if ( rand() >= _____ (2) ) // rand() 為亂數函數 , 0 <= rand() < 1
        {
            _____ (3)
        }
    count = _____ (4)
    return output ;
}
```