

### 三．操作實踐部分：（共30分）

根據題目要求使用SPSS讀入相關的資料，進行適當的資料處理，然後選擇恰當的方法進行相應的分析和檢驗，製作出要求的圖表。進行恰當的編輯處理後形成可印刷的簡要報告，並概括得到的結論。

提交電子版報告。使用Word格式（.doc），檔案名為姓名+學號，如“李丁20110102”

相應的操作命令附在報告的最後。命令要求結構清晰，注釋得當，不要出現多餘不必要的命令。

#### 題目如下：

考試提供的資料來源於一家美國信用卡公司的 5000 名客戶調查資料。資料包括受訪者所住城市規模，性別，年齡，教育水準，職業類型，配偶教育水準，家庭收入等資訊（具體見資料檔案）。

由於調查進度的影響，公司是分兩次得到資料的。第一次得到的資料為 SPSS 格式的 customer01.sav，共有 2500 條記錄。已經包含了完整的變數資訊。第二次得到資料同樣有 2500 條記錄，但剛剛從一種調查軟體中匯出來，是固定格式的 ASCII 資料檔案，檔案名為 customer02.dat。該資料包括如下變數。與 customer01.sav 相比，少了配偶受教育水準變數 spousedcat——它是根據受教育年數重新分組得到的一個變數。

customer02.dat 中各個變數的起止位置與格式如下：

變數名	起始位置	結束位置	格式
townsize	1	4	F4.0
gender	5	8	F4.0
age	9	12	F4.0
agecat	13	16	F4.0
ed	17	20	F4.0
edcat	21	24	F4.0
jobcat	25	28	F4.0
employ	29	32	F4.0
retire	33	36	F4.0
income	37	44	F8.2
inccat	45	48	F4.0
marital	49	52	F4.0
spoused	53	56	F4.0
ID	57	60	F4.0

### **任務一、資料讀入與合併 ( 5 分 ):**

使用 SPSS 讀入上述 customer02.dat 資料，另存為檔 customer02.sav。  
將文件 customer01.sav 與 customer02.sav 兩個資料合併成為檔 customerall.sav，為下一步的分析做準備。

如果你無法完成任務一，你可以使用 customer.sav 直接進行下面的任務二，但你將無法獲得任務一的分數。customer.sav 為已經合併好的上述兩個資料。

### **任務二、變數的修改與創建 ( 5 分 ):**

( 1 ) 由於 customer02.dat 中的資料缺少 spousedcat 變數資訊，你需要將這一變數資訊在合併得到的資料中補全，如果你直接使用 customer.sav，同樣需要補全這一變數。

( 2 )，原資料中受訪者的教育水準 edcat 的變數標籤和值標籤還是英文，請修改成中文。

如果你不知道怎麼補全原有變數，你也可以選擇新生成一個變數 ( 如 spousedcat2 ) 使得所有記錄都含有配偶受教育水準資訊 ( 注意加上變數標籤和取值標籤 )。

### **任務三、統計描述 ( 5 分 )**

調查物件的家庭收入分佈情況如何？新產生一個變數  $\ln inc$ ，使其等於  $income$  的自然對數。這個新變數的分佈如何？

#### **任務四、均值比較與檢驗 ( 10 分 )**

對比已婚者的家庭收入與未婚者的家庭收入 (  $\ln inc$  )，他們之間存在顯著的差異嗎？不同年齡組的人們之間呢？你可以選擇合適的圖表以及統計檢驗來表述表達自己的觀點。

#### **任務五、以統計資料來支援自己的論述 ( 5 分 )**

根據資料中的資訊，受訪者的家庭收入主要受什麼因素影響呢？依據現有的資料做出你的判斷，並使用恰當的圖表來論述自己的想法。論述可以簡潔表述。