

## 一、規格要求，違反者以零分計！

- (1) 可編譯與成功執行的 C++ 程式原始碼，特殊編譯需求須在程式碼開頭有清楚註解。
- (2) 任何一部分程式碼都不得被判定為疑似抄襲，程式碼第一列要清楚註解學號姓名。
- (3) 檔名限以「**DS2ex3\_組別\_學號\_學號**」開頭，只限繳交一份**單一檔案**的.cpp 原始碼。
  - 申請核准**分開撰寫各交一份**的組別，檔名改以「**DS2ex3\_組別\_自己學號**」開頭。

## 二、作業內容

### 作業上傳：週一

整合下列任務於同一個程式及單一操作選單下，未整合、無法連續執行或沒有輸入防呆措施，都各扣 5 分。若導致程式檢測系統無法正常運作，該任務以零分計。

### 規範：每個任務違反一項各扣 5 分

- (1) 預先不知道資料筆數，禁止使用固定筆數的靜態陣列，必須採用**動態陣列**或 vector。
- (2) 必須遵循各項任務指定的**雜湊函數**、**步階函數**及其參數設定建立對應的雜湊表。

### （任務零）文字檔轉存二進位檔

輸入：讀入一個文字檔，檔名如 **input301.txt**，沒有標題列，各欄位含意由左而右分別為【學號 sid】、【姓名 sname】、(6 個)【分數 score】、【平均分數 average】，共 **9 個字串欄位**。

步驟：將 6 個分數欄位都改以**整數(unsigned char)**型態儲存，平均分數改以**浮點數(float)**儲存，其他兩個欄位則存成指定大小的字元陣列。

- 【學號 sid】以 10 個字元大小的陣列儲存。
- 【姓名 sname】以 10 個字元大小的陣列儲存。
- (6 個)【分數 score】各自以**整數 unsigned char**型態儲存。
- 【平均分數 average】以**浮點數 float**儲存。

輸出：以 struct 結構暫存上述後的每一筆學生資料，存入另一個**同名但改以.bin 為延伸檔名的二進位檔**，如 **input301.bin**，檔案大小必須和範例程式的輸出一致。

### 注意事項：因違反下列事項而造成程式出錯者以零分計！

- (1) 執行下列任務必須以讀取**.bin** 二進位檔為優先，若只存在.txt 文字檔，就先執行(任務零)。
- (2) 一旦已經存在**.bin** 二進位檔，就必須直接讀取二進位檔以執行任務。

### （任務一）以線性探測 linear probing 建立雜湊表 X

#### 雜湊函數：

- (1) 雜湊表大小 = **大於 1.1 倍資料總筆數的最小質數**。
- (2) 限用函數：hash(key) = (學號每個字母對應的 **ASCII 編碼**相乘) 除以**雜湊表大小**取餘數。

參考網址：<https://simple.wikipedia.org/wiki/ASCII>

輸入：(任務零) 所儲存的**二進位檔**。

步驟：

1. 依照資料存在原始檔案的次序逐筆讀取檔案後，以【學號 sid】為鍵值建立雜湊表 X，並採用線性探測 linear probing 處理碰撞，每個雜湊位址只存放一筆資料，儲存每筆資料的【雜湊值 hvalue】、【學號 sid】、【姓名 sname】及【平均分數 average】。
2. 基於雜湊表 X，依照授課影片 5-10 的方法計算搜尋不存在值（除以雜湊表大小）及搜尋現存值（除以現存資料筆數）的平均比較次數。

輸出：

1. 從位址[0]開始逐筆輸出雜湊表 X 至改以 linear 命名的文字檔，如 linear301.txt，存放每個位址的【雜湊值 hvalue】、【學號 sid】、【姓名 sname】及【平均分數 average】；無資料則在位址後留空白。
2. 輸出搜尋不存在值及搜尋現存值的平均比較次數至螢幕，必須和範例程式的輸出一致。

### （任務二）以雙重雜湊 double hash 建立雜湊表 Y

雜湊函數：同任務一。

步階函數：

- (1)限用函數： $\text{step}(\text{key}) = \text{最高步階} - ((\text{學號對應的 ASCII 編碼相乘}) \text{除以最高步階取餘數})$ 。
- (2)最高步階 = 大於(資料總筆數/5)的最小質數。

輸入：沿用任務一讀檔所暫存在動態陣列的資料。

步驟：

1. 依照資料存在原始檔案的次序逐筆讀取檔案後，以【學號 sid】為鍵值建立雜湊表 Y，並採用雙重雜湊 double hash 處理碰撞，每個雜湊位址只放一筆資料，儲存【雜湊值 hvalue】、【學號 sid】、【姓名 sname】及【平均分數 average】。
2. 基於雜湊表 Y，計算搜尋現存值（除以現存資料筆數）的平均比較次數。

輸出：

1. 同任務一，從位址[0]開始逐筆輸出雜湊表 Y 至以 double 命名的文字檔，如 double301.txt。
2. 只輸出搜尋現存值的平均比較次數至螢幕，必須和範例程式的輸出一致。

程式碼：上機練習前一天上傳原始碼至/評量區/作業/作業 #3 上傳/，隔天可補繳。

流程圖：上機練習前先畫好初稿，以插圖放入貼文，錯過期限不允許更新或補繳。

### 三、評分項目

繳交項目：(A) + (B) + (C) =  $20 \times 3 + 20 + 20 = 100$  分

- (A) 程式碼：每項任務各佔 20 分，一個錯扣 5 分，更多錯以零分計。
- (B) 貼文：簡介文字、流程圖與簡報各佔 5 分，一個錯以零分計。沒有缺交補繳項目佔 5 分。
- (C) 機測：在時限內正確回答關於方法原理或程式寫法的提問，錯一次扣 5 分，共佔 20 分。

### 四、評分流程

- (1) 上機練習兩週前公布作業的題目，上機練習當天才公布上機挑戰的題目。
- (2) 每項任務以公開或隱藏測資評分，程式碼的正確性與效率要自行檢測，並嚴格遵循範例

程式的輸入輸出格式與互動方式！

- (3) 程式碼補繳：截止日期後仍缺交者方可補繳，補繳的每項任務視為一個錯，先扣 5 分。
- (4) 機測時段在機測前一天公告，安排在上課時間進行，若在課守時間會請同學自主填表！
  - 公告的機測若遲到或缺席都視為放棄，該次作業成績以零分計！