DSE ICT 練習卷：數據庫、SQL 與系統開發

中四至中六（S4 – S6）　資訊及通訊科技　|　時限：75 分鐘　|　總分：100 分　|　共 12 題

甲部 關聯式模型、資料表設計與 SQL（第 1 – 4 題，32 分）

每題 8 分；計算題須列算式及單位，SQL 題須寫出完整可執行的語句。

1. 某學校資料庫的 STUDENT 資料表現有 4,800 筆記錄，每筆記錄平均佔用 180 位元組。(a) 計算該資料表的總容量，以位元組及 KB 表示（KB 取兩位小數）。(b) 若記錄數每年增加 25%，計算兩年後的記錄數目。(c) 說明以 CHAR(60) 與 VARCHAR(60) 儲存姓名的分別，以及對儲存空間的影響。

答案：(a) 總容量 = 4,800 × 180 = 864,000 位元組 = 843.75 KB。(b) 兩年後記錄數目 = 4,800 × 1.25 × 1.25 = 7,500 筆。(c) CHAR(60) 固定佔 60 位元組並以空格補齊，VARCHAR(60) 只佔實際字元長度加長度標記，故以 VARCHAR 儲存短姓名可節省大量空間，但長度不定會令記錄位置浮動。

2. STUDENT 表的班別分佈為：5A 120 筆、5B 118 筆、5C 122 筆、5D 119 筆。(a) 計算全表記錄總數。(b) 計算 5A 班佔全表的百分比，取兩位小數。(c) 寫出可傳回 5A 班所有記錄的 SQL 語句。

答案：(a) 總數 = 120 + 118 + 122 + 119 = 479 筆。(b) 5A 佔 120 ÷ 479 × 100% = 25.05%。(c) `SELECT * FROM STUDENT WHERE class = '5A';`

3. 在 4,800 筆記錄的資料表中搜尋一筆記錄，循序搜尋平均需比較 (n+1)/2 次，而二元搜尋最多只需 ⌈log2(n+1)⌉ 次。(a) 分別計算兩種方法在 n = 4,800 時的比較次數（循序搜尋取一位小數）。(b) 計算兩者次數的比率，取一位小數。(c) 說明為何二元搜尋必須先建立索引。

答案：(a) 循序搜尋 = (4,800 + 1) ÷ 2 = 2,400.5 次；二元搜尋 = ⌈log2(4,801)⌉ = 13 次。(b) 比率 = 2,400.5 ÷ 13 = 184.7 倍。(c) 二元搜尋要先能按鍵值隨機存取已排序的資料，索引提供排序後的鍵值與記錄位址對應，否則資料表只能逐筆掃描而無法跳躍定位。

4. 一個未正規化的關聯包含 6 個屬性，主鍵由 2 個屬性組成，並存在部分相依（非主鍵屬性只依賴主鍵其中一部分）。(a) 指出此關聯違反第幾正規形式。(b) 為消除部分相依，最少要把該關聯拆成幾個資料表？並說明外鍵數目。(c) 說明未消除部分相依會引致的兩項資料異常。

答案：(a) 違反 2NF（同時未達 3NF）：主鍵為複合鍵而存在部分相依。(b) 最少拆成 3 個資料表，並須加設 2 個外鍵把子表連回主表。(c) 插入異常（新增資料必須同時提供無關屬性）與更新異常（同一事實重複儲存，改一處而漏改另一處即產生不一致）。

乙部 正規化、索引與查詢效率（第 5 – 8 題，36 分）

每題 9 分；須引用具體數值或資料表結構作答，並說明推理步驟。

5. STUDENT 表共 4,800 筆記錄，其中女生佔 60%，而 5A 班女生佔全表 15%。(a) 計算女生人數。(b) 計算 5A 班女生人數。(c) 若在 (gender, class) 上建立複合索引，查詢「5A 班女生」時掃描筆數由全表降至多少百分比？另計算減幅。

答案：(a) 女生 = 4,800 × 60% = 2,880 人。(b) 5A 班女生 = 4,800 × 15% = 720 人。(c) 掃描筆數由 4,800 降至 720，即全表的 15.0%，減幅 85.0%。

6. 某系統的資料庫現有 1,200 筆記錄，管理層預計記錄數每 18 個月增加一倍。(a) 計算 54 個月後的記錄數目。(b) 計算由現在至該時的增長百分比。(c) 說明記錄量倍增對查詢效能的兩項影響。

答案：(a) 54 ÷ 18 = 3 個倍增期，記錄數 = 1,200 × 2^3 = 9,600 筆。(b) 增長 = (9,600 − 1,200) ÷ 1,200 × 100% = 700%。(c) 全表掃描的 I/O 次數隨記錄數同步上升，索引層數亦增加，令每次查詢的磁碟讀取與比較次數上升。

7. 成績表 MARKS 有 5 筆記錄：62、78、85、91、54。(a) 寫出計算平均值與最高分的 SQL 聚合函數語句。(b) 計算總和與平均值。(c) 若要同時按班別分組計算平均分，須加入哪一個子句？另可用甚麼子句篩選分組結果？

答案：(a) SELECT AVG(score), MAX(score) FROM MARKS;(b) 總和 = 62 + 78 + 85 + 91 + 54 = 370，平均 = 370 ÷ 5 = 74.0，最高分為 91。(c) 須加入 GROUP BY class 子句；可用 HAVING 篩選分組結果。

8. 對 1,024 筆記錄排序，合併排序約需 $n \cdot \log_2(n)$ 次比較，氣泡排序平均約需 $n^2/2$ 次比較。(a) 分別計算兩者的比較次數。(b) 計算兩者的比率，取一位小數。(c) 說明為何在資料量大時合併排序明顯較優。

答案：(a) 合併排序 = 1,024 × 10 = 10,240 次；氣泡排序 = $1{,}024^2 \div 2 = 524{,}288$ 次。(b) 比率 = 524,288 ÷ 10,240 = 51.2 倍。(c) 合併排序時間複雜度為 $O(n \log n)$，比較次數隨資料量以對數增長；氣泡排序為 $O(n^2)$，資料量倍增時工作量升四倍，故差距隨規模擴大。

丙部 系統開發流程、測試與備份管理（第 9 – 12 題，32 分）

每題 8 分；原因與風險須分點說明，並以本題數據支持結論。

9. 某系統開發專案的時程為：需求分析 15 日、系統設計 20 日、程式實作 40 日、系統測試 25 日。(a) 計算專案總日數。(b) 計算實作階段佔總日數的百分比。(c) 說明為何測試階段不宜壓縮至 5 日以內。

答案：(a) 總日數 = 15 + 20 + 40 + 25 = 100 日。(b) 實作階段佔 40 ÷ 100 × 100% = 40.0%。(c) 測試需時執行測試個案、記錄缺漏、修正後再做回歸測試；壓縮至 5 日會令缺漏未被發現便上線，後期修復成本遠高於前期。

10. 系統測試共設計 240 個測試個案，其中 222 個通過。(a) 計算通過率，取一位小數。(b) 計算未通過的個案數目。(c) 指出未通過個案在系統推出前必須完成的兩項工作。

答案：(a) 通過率 = 222 ÷ 240 × 100% = 92.5%。(b) 未通過個案 = 240 − 222 = 18 個。(c) 須修正相關缺漏並重新執行該個案（回歸測試），同時由測試人員或使用者簽署確認修正結果。

11. 驗收測試（UAT）在 12.5 千行原始碼的系統中找到 36 個缺漏，其中 9 個屬嚴重缺漏。(a) 計算嚴重缺漏佔全部缺漏的百分比。(b) 計算每千行原始碼的缺漏密度，取兩位小數。(c) 說明缺漏密度可用作哪一項管理判斷。

答案：(a) 嚴重缺漏佔 9 ÷ 36 × 100% = 25.0%。(b) 缺漏密度 = 36 ÷ 12.5 = 2.88 個 / 千行。(c) 可用以比較不同模組或版本的品質，並決定是否適合推出：密度偏高的模組須加測或重構後才上線。

12. 備份策略為每週一次全備份 12.0 GB，其餘 6 日每日一次增量備份 1.6 GB。(a) 計算一週的備份總量。(b) 若壓縮後體積為原來的 60%，計算壓縮後的每週備份量，取兩位小數。(c) 說明只保留全備份而取消增量備份的兩項風險。

答案：(a) 每週總量 = 12.0 + 6 × 1.6 = 21.6 GB。(b) 壓縮後 = 21.6 × 60% = 12.96 GB。(c) 只做全備份會令備份窗口過長並佔用大量儲存；資料一旦損壞只能回到上一週的狀態，復原點目標（RPO）由一日退化為七日。