

類科：資訊處理

科目：資料庫應用

考試時間：2小時

座號：_____

※注意：(一)禁止使用電子計算器。

(二)不必抄題，作答時請將試題題號及答案依照順序寫在試卷上，於本試題上作答者，不予計分。

一、某市政府希望透過網路來處理市民的反映問題，經過訪談和分析後，獲得以下需求：

- 1.陳情案 (Case)：必須記載陳情人的電子郵件信箱 (email) 和姓名 (name)、陳情時間 (dateTime)、主旨 (title)、內容 (content) 和附檔 (files)。其中陳情人的電子郵件信箱、姓名和陳情時間合起來為唯一，附檔可有多個。
- 2.局處 (Division)：必須記載局處名稱 (dName) 和首長姓名 (director)。其中局處名稱為唯一，且每一陳情案可能會分配給數個相關局處來處理。
- 3.員工 (Employee)：必須記載員工編號 (eId)、姓名 (eName) 和身分證字號 (pId)。其中員工編號和身分證字號均為唯一，且一位員工剛好屬於一個局處。員工一定是以下兩種身分中的一種：行政員工 (Admin) 和技術員工 (Tech)，行政員工要記載其職級 (level)，技術員工則要記載其技術能力 (expertise)。此外，一個陳情案會由一位行政員工來總負責 (Handles)。

(一)請依以上需求敘述，畫出其實體關係圖 (Entity Relationship Diagram)，員工的分類請使用繼承 (Inheritance) 來表示。(20分)

(二)請將(一)的實體關係圖轉成關聯綱目並指出主鍵、外鍵，以及外鍵的參考關係。(10分)

二、考慮以下記載討論版文章的關聯：

$R(\text{發表人 id}, \text{發表人姓名}, \text{發表時間}, \text{主旨}, \text{內容}, \text{字數}, \text{回應發表人 id}, \text{回應發表時間})$

函數相依 (functional dependency) 以 \rightarrow 表示，包含於 (is the subset of) 以 \sqsubseteq 表示：
 $\text{發表人 id} \rightarrow \text{發表人姓名}$

$\{\text{發表人 id}, \text{發表時間}\} \rightarrow \{\text{主旨}, \text{內容}, \text{回應發表人 id}, \text{回應發表時間}\}$

$\text{內容} \rightarrow \text{字數}$

$\{(\text{回應發表人 id}, \text{回應發表時間})\} \sqsubseteq \{(\text{發表人 id}, \text{發表時間})\}$

(一)請推論出以上關聯 R 的主鍵和外鍵，推導過程必須列出。(6分)

(二)請將關聯分解成第二正規式 (Second Normal Form, 2NF) 但不滿足第三正規式 (Third Normal Form, 3NF)，請指明主鍵、外鍵和其參考關係。(7分)

(三)請將關聯分解成第三正規式，請指明主鍵、外鍵和其參考關係。(7分)

三、考慮以下記載討論版文章的關聯：

$\text{Member}(\underline{\text{mId}}, \text{mName}, \text{email})$

// 會員(編號, 姓名, 電子郵件信箱)

$\text{Post}(\underline{\text{pId}}, \text{mId}, \text{dateTime}, \text{title}, \text{pContent})$

// 文章(文章 id, 作者編號, 日期時間, 主旨, 內容)

(請接背面)

類科：資訊處理
科目：資料庫應用

Reply(pId, rId, rmId, dateTime, rContent)
// 回文(文章 id, 回文 id, 回文作者編號, 日期時間, 內容)

其中主鍵以底線表示，外鍵如下：

Post(mId)參考到 Member(mId)
Reply(pId)參考到 Post(pId)
Reply(rmId)參考到 Member(mId)

請用一個 SQL 查詢句(Queries)回答以下每一個查詢：

- (一)列出 2016 年有 PO 過文章（即出現在 Post 關聯裡）的會員之會員編號和姓名。(5 分)
- (二)對於每一篇 PO 文，列出其 pId, PO 文時間，PO 文作者姓名，以及在 2016 年被回應的次數。(8 分)
- (三)列出有原作者回應的 PO 文之 pId, title 和作者姓名。(7 分)
- (四)找出 PO 文中(即出現在 Post 關聯裡)的內容有疑似西元日期的文章之 pId, title, PO 文作者姓名。所謂疑似西元日期是指出現連續四個數字，且第一個數字為 1 或 2，或是(連續)數字前有西元或公元的字眼。(10 分)

四、MongoDB 是目前使用愈來愈普遍的 NoSQL 資料庫管理系統，NoSQL 的主要特色包括 Schemaless, eventual consistency 和 horizontal scaling

- (一)請說明 Schemaless, eventual consistency 和 horizontal scaling 的意義。(10 分)
- (二)假設你有一個 order(訂單)的 collection，包含四個欄位：cust_id(顧客 id), product(產品), amount(金額), type(產品型態)，以下 Mongo 查詢句可用來找出衣服(clothing)類產品總購買金額超過 10,000 元的顧客之顧客 id 和衣服類產品購買總金額：

```
db.orders.aggregate([
  { $match: {type: "clothing"}}
  {
    $group: {
      _id: "cust_id",
      total: { $sum: "amount"}
    }
  }
  { $match: {total: { $gt: 10000}}}
])
```

請仿以上 MongoDB 語法，找出購買金額超過 1,000 元的訂單中，各類產品的總購買金額。(10 分)