ITC-6000

Assignment 1

July 13, 2020

**[Sunil Raj Thota]**

# Question 1

**Use the table from the data file below to answer these questions regarding file structure:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| First | Last Name | ID | Hours | GrossPay | Withholding | NetPay | PayDate |
| Roger | McHaney | 1 | 38.5 | $561.56 | $112.31 | $449.25 | 10/23/2000 |
| Roger | McHaney | 1 | 38.5 | $561.56 | $112.31 | $449.25 | 1/17/2000 |
| Telly | O'Grady | 2 | 30 | $489.06 | $97.81 | $391.25 | 3/1/2000 |
| Telly | O'Grady | 2 | 30 | $489.06 | $97.81 | $391.25 | 2/22/2000 |
| Telly | O'Grady | 2 | 20 | $326.04 | $65.21 | $260.83 | 2/28/2000 |
| Telly | O'Grady | 2 | 20 | $326.04 | $65.21 | $260.83 | 3/6/2000 |
| Telly | O'Grady | 2 | 44.5 | $725.44 | $145.09 | $580.35 | 1/21/2000 |
| Telly | O'Grady | 2 | 44.5 | $725.44 | $145.09 | $580.35 | 1/17/1998 |
| Grant | Brikestein | 3 | 18 | $408.60 | $81.72 | $326.88 | 10/23/2000 |
| Grant | Brikestein | 3 | 18 | $408.60 | $81.72 | $326.88 | 1/17/2000 |
| Alarry | Abu | 4 | 40 | $728.80 | $145.76 | $583.04 | 1/17/2000 |
| Betty | Marygrade | 5 | 35 | $792.40 | $158.48 | $633.92 | 1/17/2000 |
| Samuel | Timod | 6 | 40 | $927.20 | $185.44 | $741.76 | 1/17/2000 |

1. **How many records does this file contain?**

The above file contains **13** records

1. **How many fields are there per record?**

There are **8** fields per record in the above file shown

1. **What problem would you encounter if you tried to create a listing by last name?**

The **Last Name** shown in the table occurs multiple times and due to this we might face **Data Redundancy** problems if we tried to create a listing by **Last Name.** The Data Redundant list may occur something similar to the below shown table**:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| First | Last Name | ID | Hours | GrossPay | Withholding | NetPay | PayDate |
| Roger | McHaney | 1 | 38.5 | $561.56 | $112.31 | $449.25 | 10/23/2000 |
| Telly | O'Grady | 2 | 30 | $489.06 | $97.81 | $391.25 | 3/1/2000 |
| Grant | Brikestein | 3 | 18 | $408.60 | $81.72 | $326.88 | 10/23/2000 |
| Alarry | Abu | 4 | 40 | $728.80 | $145.76 | $583.04 | 1/17/2000 |
| Betty | Marygrade | 5 | 35 | $792.40 | $158.48 | $633.92 | 1/17/2000 |
| Samuel | Timod | 6 | 40 | $927.20 | $185.44 | $741.76 | 1/17/2000 |

1. **What are the field names? Separate by semicolon.**

The field names are as follows:

**First; Last Name; ID; Hours; GrossPay; Withholding; NetPay; PayDate**

1. **What other problems do you see?**

Based on my observation, I see following points as other problems:

* First, Same data of **ID 1** and **ID 3** have been repeated two times with separate **PayDate(s)**
* Second, **ID 2** holds **6** total records with **3** records each of **2** pairs but by separate **PayDate(s)**
* These records might probably make the **dataset inconsistent** and chances of occurring data redundancy are **high**

# Question 2

**Answer the following two questions in paragraph form:**

1. **A database system is made up of five major parts. Identify these parts. Provide specifics of what make up these components.**

The Database consists of 5 main components. They are:

* **Hardware:** The systems which are made up of physical devices are called Hardware. These provide an interface between computers and real-world systems.  **Examples:** Monitor, I/ O Devices, Keyboard, Mouse, Storage Devices, CPU, Printer, Hard drives, routers are few examples of a hardware
* **Software:** It is important as to define what to do for a hardware.A program or set of commands that run on computer is known as Software. It controls and manages everything in a database and acts as the core element. This includes any software used for OS, Networks, Data Accessing, Data sharing or to run applications
* **People:** People are required to run the ecosystem, to monitor the data, to relate with data, to generate use cases and implement for better predictions to help the businesses. **Examples:** End Users, Systems Analysts, Systems Admins, Database Admins, and System Programmers
* **Data:** The Database Management Systems collects, stores, processes and access the data. The Databases holds both the operational or actual data and the metadata. The Data has to be stored in the system initially. It is processed further to reap the information from it which is highly important, secure and beneficial to the end-user
* **Procedures:** They are the instructions and rules which govern the design and the use of the database systems. It acts as a guide to follow proper guidelines to manage or to operate

1. **A DBMS provides at least nine functions. Identify these functions and provide a brief explanation of each**

The functions of DBMS are as follows:

* **Data Dictionary Management:** The DBMS stores the definitions of the data elements and their relationships **(meta data)**.Programs go through the database system in order to access data from the database. Any change made in the structure of the data is recorded in it. DBMS provides data abstraction and removes structural data and data dependency
* **Data Storage Management:** DBMS creates and manages complex structures required for data storage. Also stores necessary data entry forms, definitions, data validation rules, procedural codes etc., It also performs Performance tuning activities to make the database perform efficiently and effectively. The internal schema defines how the data should be stored by the storage management mechanism and the storage manager interfaces with the OS to access the physical storage. DBMS stores the database in multiple physical data files
* **Data Transformation and Presentation:** The DBMS transfers the entered data to conform to the data structures that are required to store the data. It translates Logical request for the data into physically locate and retrieve the requested data. DBMS formats the physically retrieved data to make it confirm to the user’s logical expectation. **Data Independence –** A condition that exists when data access is unaffected by changes in physical data storage characteristics. The DBMS formats the physically retrieved data to make it conform to the user’s logical expectations
* **Security Management:** It creates a security system that enforces user security and data privacy within the database. Security rules determine which users can access the database, which data items each user may access and which data operations the user may perform like **CRUD (Create, Read, Update, and Delete)**. It is important in a multi-user environment as many users access the database simultaneously. Uses multiple tools to manage
* **Multi-User Access Control:** The DBMS creates the complex structures that allow multiple users access to the data. In order to provide data integrity and data consistency the DBMS uses sophisticated algorithms to ensure that multiple users can access the database concurrently without compromising the integrity of the database. Enforces transaction management and concurrency control
* **Backup and Recovery Management:** DBMS provides mechanisms for backing up the data periodically and recovering from different types of failures. This prevents loss of data and ensures data safety and integrity
* **Data Integrity Management:** The DBMS promotes and enforces integrity rules to eliminate data integrity problems. Thus, minimizing data redundancy and maximizing data consistency. The data relationships stored in the data dictionary are used to enforce data integrity. It is important in transaction-oriented database systems
* **Database Access Languages and Application Programming Interfaces:** The DBMS provides data access through a query language. SQL is the de facto query language. It’s a Non-procedural language. This contains two components mainly and they are: **Data Definition Language (DDL) –** It allows a database admin to define the database structure, schema, and sub-schema. **Data Manipulation Language (DML) –** These are set of commands allows end users to manipulate the data **(SELECT, INSERT, UPDATE, DELETE)**. The DBMS provides data access to programmers via procedural languages like COBOL, PASCAL, C, VB, etc.,
* **Database Communication Interface:** Current DBMS’s provide communication interfaces designed to allow the database to accept end-user request within a computer network environment. Communications are accomplished in several ways. End-users generate answers to the queries by filling in screen forms though Web browser. DBMS automatically publishes predefines reports on a website. DBMS connects to 3rd-party systems to distribute information via e-mail

# Question 3

# A conceptual model provides a bird's-eye view of a database. Conceptual models are derived from the business rules that make up an application.

From the following business rules, create a conceptual model.

Business Rules:

* An author writes books.
* A book is written by one or more authors.
* A book has one subject.
* A subject has at least 20 books.
* A book is published in a year

1. **Identify the entities involved in the problem and determine the relationships that exist between entities.**

**Entities:**

* Author - **A**
* Book - **B**
* Subject - **S**
* Year - **Y**

**Relationships:**

Based on the Business Rules we can determine the following relationships that exist between entities:

* **A: B = M: N** – An Author writes **many** books, and a book can be written by **one or more** authors
* **B: S = 1: N** – A book has **one** subject, and one subject has **many (at least 20)** books
* **B: Y = 1: 1** – A book is published in **a** year

1. **Create and paste below your conceptual model.**

