**John Sebastian Ablay – A221 11/15/22**

**1. How many records does the following file contain? How many fields are there per record?**

- The file contains 5 records, and there are 5 fields every record.

**2. What problems would you encounter if you wanted to produce a listing by city? How would you solve this problem by altering the file structure?**

- The problem I encountered would be data redundancies. An example of this is Mati city, while the other cities do not have the stated problems to be encountered. To have a solution for the alteration of the file structure, Foreign Keys should be utilized. This will help me lessen the confusion of repetitive information as such that duplicates will have an assigned Foreign Key for the database to be maintained more effectively.

An example of this would be:

|  |  |
| --- | --- |
| **MANAGER\_ADDRESS** | **FOREIGN KEY** |
| St Michael, Mati | A |
| Dampa, Tagum City | B |
| Maa, Digos | C |
| Obrero, USA | D |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PROJECT\_CODE** | **PROJECT\_MANAGER** | **MANAGER\_PHONE** | **MANAGER\_ADDRESS** | **PROJ\_BID\_PRICE** |
| 1234 | Karen M. Aparece | 221-5265 | A | 2,200,000.00 |
| 6547 | Genelyn Tampos | 225-8696 | B | 5,000,000.00 |
| 1234 | Meljhon V. Aborde | 224-7854 | C | 4,581,257.00 |
| 7896 | Karen M. Aparece | 221-5265 | A | 7,854,129.00 |
| 1236 | Martzel P. Baste | 222-4585 | D | 17,965,823.00 |

**3. How would you alter the file structure if you wanted to produce a listing of the file contents by last name, area code, or city?**

- To alter the file structure of the file contents by last name, I would rearrange it by its alphabetical order. It would now be Aborde, Aparece, then Aparece again, Baste, and then Tampos. To alter the file structure of the area code, I would arrange it according to the ordinality of the first number of their PROJECT\_CODE. Lastly, to alter the file structure of the city, it would be the same idea I had on the previous question, and that is to organize them by alphabetical order, but in accordance with its PROJECT\_CODE.

An example of the suggested alteration of the file structure will be listed below:

(These are not bound altogether)

|  |  |  |
| --- | --- | --- |
| **PROJECT\_MANAGER** | **PROJECT\_CODE** | **MANAGER\_ADDRESS** |
| Aborde, Meljohn V. | 1234 | Maa, Digos |
| Aparece, Karen M. | 1234 | St Michael, Mati |
| Aparece, Karen M. | 1236 | St Michael, Mati |
| Baste, Martzel P. | 6547 | Dampa, Tagum City |
| Tampos, Genelyn | 7896 | Obrero, USA |

**4. What data redundancies do you detect? How could those redundancies lead to anomalies?**

- The data redundancies that I have detected can be found in the PROJECT\_CODE which is code 1234. PROJECT\_MANAGER which has the redundancy of the manager’s name Karen M. Aparece. MANAGER\_PHONE also has repetitive information such as 221-5265, which is Aparece’s telephone number. Data redundancies present on MANAGER\_ADDRESS such as St Michael, and Mati are also visible. Lastly, there are no data redundancies located in PROJ\_BID\_PRICE.

**PROJECT\_CODE**: 1234

**PROJECT\_MANAGER**: Karen M. Aparece

**MANAGER\_PHONE**: 221-5265

**MANAGER\_ADDRESS**: St Michael, Mati

**PROJ\_BID\_PRICE**: None

Whenever data redundancies occur in the presented information, this will lead to the **expanding** of the **size** and **complexity** of its database. This will be more difficult to handle and **maintain** in the future. This will also affect the **time** spent on loading it and can potentially affect the **efficacy** of the employees.

**5. Identify and discuss the serious data redundancy problems exhibited by the file structure shown in Figure 1.2.**

**PROJ\_NUM**: 1, 2, 3

**PROJ\_NAME**: Hurricane, Coast, Satellite

**EMP\_NUM**: 101, 105, 110

**EMP\_NAME**: Aly Punga, Luh Kay, Juny Sia

**JOB\_CODE**: IS, CS

**JOB\_CHG\_HOUR**: 125.00, 130.00

**PROJ\_HOURS**: NONE

**EMP\_PHONE**: 225-2525, 321-5858, 224-5878

Table

Description automatically generated

- The record **PROJ\_NUM** has repeating fields such as 1 which occurred three times, 2 which occurred two times, and 3 which occurred three times. **PROJ\_NAME** has Hurricane appeared 3 times, Coast 2 times, and Satellite 3 times. **EMP\_NUM** has values such that 101, 105, and 110 have 2 duplicates. The same goes for the record **EMP\_NAME**. Aly Punga, Luh Kay, and Juny Sia also have 2 duplicates each. **JOB\_CODE** also has data redundancy, such that IS occurs 2 times, while CS occurs 4 times. **JOB\_CHG\_HOUR** has 125.00 reappearing 5 times, and 130.00 only 2 times. Lastly, the **EMP\_PHONE** record has the following: 225-2525, 321-5858, 224-5878; each repeating 2 times. In contrast, **PROJ\_HOURS** has none.

6. **Looking at the EMP\_NAME and EMP\_PHONE contents in Figure 1.2, what change/s would you recommend?**

- Arrange EMP\_NAME with a use of Foreign Key, while the EMP\_PHONE will be arranged in accordance with the newly arranged EMP\_NAME.

|  |  |  |  |
| --- | --- | --- | --- |
| **EMP\_NAME** | **FOREIGN KEY** | **EMP\_PHONE** | **FOREIGN KEY** |
| Aly Punga | 1 | 225-2525 | A |
| Luh Kay | 2 | 321-5858 | B |
| Juny Sia | 3 | 224-5878 | C |
| Cory Khong | 4 | 321-9696 | D |
| Mely Cirus | 5 | 221-8574 | E |

|  |  |
| --- | --- |
| **EMP\_NAME** | **EMP\_PHONE** |
| 1 | A |
| 2 | B |
| 3 | C |
| 1 | A |
| 4 | D |
| 3 | C |
| 2 | B |
| 5 | E |

**7. Identify the various data sources in the file you examined in Problem 5.**

- The data sources in the file that I have examined in Problem 5 are the following:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **PROJ\_NUM** | **PROJ\_NAME** | **EMP\_NUM** | **EMP\_NAME** | **JOB\_CODE** | **JOB\_CHG\_HOUR** | **PROJ\_HOURS** | **EMP\_PHONE** |

**8. Given your answer to Problem 7, what new files should you create to help eliminate the data redundancies found in the file shown in Figure 1.2?**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **PROJ\_NUM** | **PROJ\_NAME** | **EMP\_NUM** | **EMP\_NAME** | **JOB\_CODE** | **JOB\_CHG\_HOUR** | **PROJ\_HOURS** | **EMP\_PHONE** |
| O | R | A | 1 | K | X | 15.0 | F |
| O | R | B | 2 | L | X | 12.5 | G |
| O | R | C | 3 | L | Y | 15.4 | H |
| P | S | A | 1 | K | X | 14.0 | F |
| P | S | D | 4 | M | X | 25.0 | I |
| Q | T | C | 3 | L | Y | 16.0 | H |
| Q | T | B | 2 | L | X | 18.5 | G |
| Q | T | E | 5 | N | Z | 12.0 | J |

**Assignation of Foreign Keys:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **EMP\_NAME** | **FOREIGN KEY** | **EMP\_PHONE** | **FOREIGN KEY** | **EMP\_NUM** | **FOREIGN KEY** |
| Aly Punga | 1 | 225-2525 | F | 101 | A |
| Luh Kay | 2 | 321-5858 | G | 105 | B |
| Juny Sia | 3 | 224-5878 | H | 110 | C |
| Cory Khong | 4 | 321-9696 | I | 108 | D |
| Mely Cirus | 5 | 221-8574 | J | 123 | E |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **JOB\_CHG\_HOUR** | **FOREIGN KEY** | **JOB\_CODE** | **FOREIGN KEY** | **PROJ\_NUM** | **FOREIGN KEY** |
| 125.00 | X | IS | K | 1 | O |
| 130.00 | Y | CS | L | 2 | P |
| 140.00 | Z | IT | M | 3 | Q |
| - | - | EMC | N | - | - |

|  |  |
| --- | --- |
| **PROJ\_NAME** | **FOREIGN KEY** |
| Hurricane | R |
| Coast | S |
| Satellite | T |

|  |
| --- |
| **Insights:**  I have learned that data redundancies are pieces of data that occur in multiple places. This can cause data inconsistency and can be linked to the unnecessary expansion of the size and complexity of the user database. This will in turn provide hindrances to computer scientists in handling information. Ideally, this must not be present within the database to avoid confusion and many problems. That is why to counter this possibility, the utility of Foreign Keys is used. These are implemented to help build referential integrity within the database. This eradicates the errors and data redundancies present by avoiding the existence of repeating and duplicating information within records and fields. This will also help in having presentable pieces of information to avoid confusion. Thus, intricate and thorough processes are always mandatory in processing databases. |