**Guidelines:**

1. All problems must be solved by hand. Do not submit any SQL code. It will be ignored.
2. Check your work carefully. I am not very sympathetic to errors that would easily turn up if checked.
3. Do not assume files/tables will not change in size in the future unless stated so.
4. Please state any assumptions you are making. Assumptions are required to be “real-world”, reasonable, and accurate.
5. When primary keys are involved, ***think minimal***.
6. Certainly, if you have questions, you are welcome to call me (908-418-6078) or send an email.

**Submission Requirements:**

1. This assignment is due Wednesday October 25, 2023, at 6pm.
2. The assignment must be submitted via Canvas.
   1. It must be readable, and it is your responsibility to confirm this. If I cannot read it, you run the risk of getting a zero for the problem or the assignment.
3. All work must be your own. The only person you may discuss the assignment with is me (Professor Forman)
   1. You may **NOT** discuss problems with any other student.
   2. You may **NOT** get answers from sites such as Chegg or Homework Hero or any other online site.
   3. Anything not mentioned, that constitutes “***not doing your own work”*** will be considered cheating.
   4. Violation of these requirements will result in a grade of 0.
4. Submit one document only, unzipped.
   1. Handwritten problems will not be accepted unless permission is granted by ME.

**Name: Emiliana Geronimo**

**Problem 1 (15 points)**

Using the Company database attached at the end of the assignment, provide the following information in a neatly designed table. For any project in which an employee worked less than 10 hours, list their First Name, Last Name, their manager’s name, the Project Number, the Project Name, and the hours worked on that project.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Fname** | **Lname** | **ManagerName** | **Pno** | **Pname** | **Hours** |
| John | Smith | Franklin Wong | 2 | ProductY | 7.5 |
| Ahmad | Jabbar | Jennifer Wallace | 30 | Newbenefits | 5 |

In this question, I used the Employee, Works\_On, and Projects table to get my information. The Works\_On table contained information about the hours worked on the project (7.5 and 5), Project number (2 and 30), and SSN that links to the employee who worked on it. The employee table has the Fname and Lname of the employee linked to the SSN of the Works\_on table and contains the Manager’s SSN which links to the Manager’s name. The Projects table contains the Project number and Project name.

**Problem 2 (10 points)**

Determine the cardinality ratio for each relationship shown below based on the attached Company table. (2pts each):

1. Employee \_\_N:1\_\_\_\_ Department
2. Employee \_M:N\_\_ Project
3. Department \_1:N\_ Project
4. Employee\_\_1:N\_ Dependent
5. Department \_M:N\_\_\_Dept\_Locations   
     
   Explain your answers and **list any assumptions** you consider significant. Assumptions are required to be “real-world”, reasonable, and accurate.

**Explanations:**

**A)** The employee to department’s cardinality ratio is N:1 because the Employee table lists the employees and they have one department assigned to the employee.

**B)** The employee to project’s cardinality ratio is M:N because the Works\_On table lists the projects having multiple employees working on them while employees work on more than one project.

**C)** The department to project’s cardinality ratio is 1:N because the Project table has the one department assigned to a project while departments can have multiple projects.

**D)** The employee to dependent’s cardinality ratio is 1:N because the Dependent table lists dependents that are assigned to one employee. This is assuming that none of the employees are related or married.

**E)** The department to department location’s cardinality ratio is M:N because the Dept\_Locations table shows that one department can have multiple locations and one location can have multiple departments.

**Problem 3 (15 points)**

Given the relation below and the associated functional dependencies:

Y = {M, N, O, P, Q, R, S,T, U, V, W}

R 🡪 ST

M 🡪 PQ

P 🡪 UV  
N 🡪 R  
MN 🡪 O

1. **Determine the Primary Key of Y**

R = {MNW , OPQRSTUV}

I believe MNW is the primary key of Y. Using the left-right method in class, I split what attributes were functionally dependent on the right (STPQUVRO) and the attributes that gave function dependency (RMPN). Because R is functionally determined by N and P is functionally determined by M, then RP would not be part of the primary key. Therefore, this leaves MN. However, nothing points to W and W is functionally dependent on itself. Therefore it will be placed in the primary key MNW because the primary functionally determines all attributes.

1. **Put the table in Second Normal Form**Assuming that Relation Y is in 1NF,
2. R1 = {N, RST}

This is because N -> R and R-> ST therefore N -> ST by transitivity. All attributes that are functionally determined by N are RST.

1. R2 = {M, PUVQ}

This is because M-> PQ and P -> UV, then by transitivity, M-> UVQ. . All attributes that are functionally determined by M are PUVQ.

1. R3 = {MN, O}

This is because M and N together are the only attributes that functionally determine O.

1. R4 = {W}

Since nothing points to W and W points to nothing, it is safe to assume that W is functionally dependent on itself and can be put in its own table.

1. **Put the table in Third Normal Form**

Assuming 2NF of the table,

1. T1 = {N,R}
2. T2 = {N, ST}

Siince 3NF does not allow transitivity R and ST has to be separated because N -> R and R-> ST.

1. T3 = {M, PQ}
2. T4 = {M, UV}

Since M-> PQ and P -> UV, therefore P and UV are separated to avoid transitivity in the table and made into two tables. Q can go on either table because it doesn’t have transitivity linked to it.

1. T5 = {MN, O}

This stays the same because O is to involved in any transitivity.

1. T6 = {W}

This stays the same because W only points to itself.

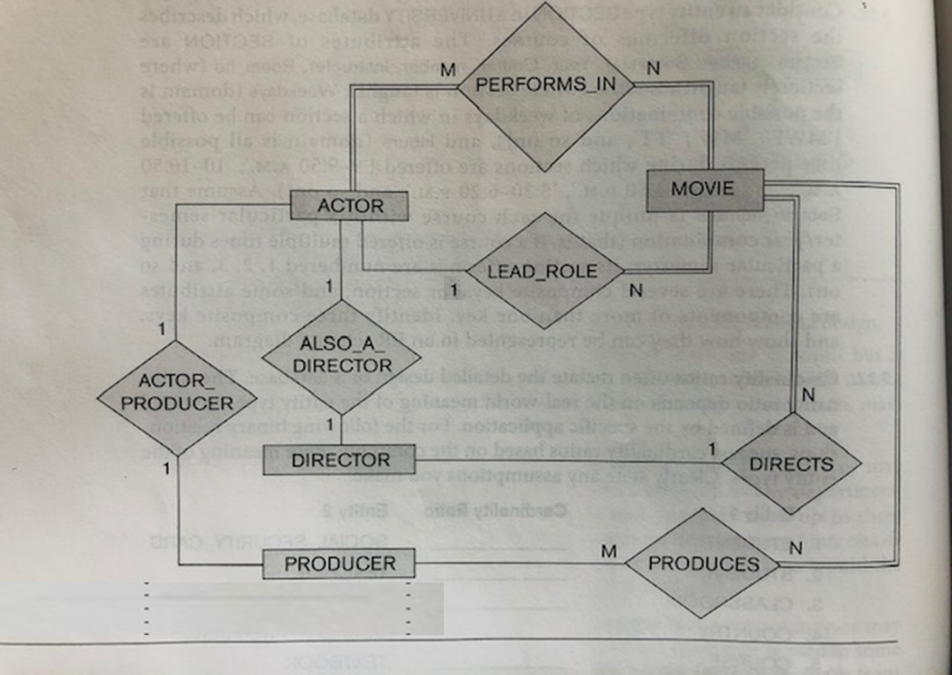
**Problem 4 (10 points)**

Consider the schema below for a database named MOVIES. Using the cardinality constraints shown in the schema below, answer the following questions as True, False, or Maybe. Explain your answer.

Use MAYBE for statements that although not explicitly shown as True cannot be shown to be false based on the schema as it is shown.

Hint: Do not use your knowledge of the movie industry to answer these questions. Your answer should only be based on the diagram below.

1. There are some actors who have acted in more than 5 movies. **MAYBE;** This is a maybe because there are multiple actors who work in multiple movies based on the M:N relationship shown in the schema below, but we cannot know for sure if they acted in more than 5 movies specifically.
2. An actor can only be the lead in multiple movies. **TRUE;** Based on the schema, the relationship between a lead actor and movies is 1:N, therefore an actor can only be the lead for multiple movies.
3. A movie can have multiple lead actors. **FALSE;** Based on the schema, the relationship between a lead actor and movies is 1:N, which means that there is only one lead actor to a multiple movies and not the other way around.
4. Leonardo DiCaprio has acted in multiple movies. **TRUE;** The actor to movie’s cardinality ratio shows M:N. Leonardo is an actor and, based on this ratio, he acted in multiple movies.
5. A movie can have multiple producers. **TRUE;** The movie to producer’s cardinality ratio is N:M, therefore a movie can have multiple producers while a producer can produce multiple movies.



**Problem 5 (20 points):**

Given Relation Y = {A, B, C, D, E, F, G, H, J, K} and given these functional dependencies:

G 🡪 AB

D 🡪 F

C 🡪 H

EG 🡪 K

EJ 🡪 ACD

1. **Determine the Primary Key of Y.**Y = {EGJ , ABCDFHK)

I believe that EGJ is the primary key of Y. I attacked this problem by using the left right method and seperating what attributes are functionally dependent on other attributes which are ABFHKCD, therefore they cannot be in the primary key. CD are on both sides of left right method, therefore they are linked to transivity and cannot be in the primary key. This leaves EGJ being the only attributes that functionally determines other attributes and not vice versa.

1. **Is this table in first normal form?**   
   Maybe, we do not know if the table has multivalued attributes, composite attributes, or nested values because you cannot see the data, therefore we can not say yes or no.
2. **Put the table in Second Normal Form**

Assuming that Relation Y is in 1NF,

1. R1 = {EG,K}

This is because EG -> K, K does not functionally determine anything else and only EG determines K.

1. R2 = {EJ, CDFH}

Because C-> H, D-> F, and EJ-> CD, then by pseudo transitivity, EJ -> CDFH. I chose to remove A because G also points to A. For simplicity, I kept A with G.

1. R3 = {G, AB}Because A is functionally dependent on G as well and G only functionally determines B, then this would be the final relation.
2. **Put the table in Third Normal Form**
3. T1 = {EG,K}

This is kept the same because EG points to only K and K has no transitivity.

1. T2 = {EJ, FH}
2. T3 = {EJ, CD}

Since C -> H and D -> F, we need to separate CH and DF because of transitivity therefore I created one table with FH and one with CD. This separation creates no transitivity in the two tables.

1. T4 = {G, AB}

I kept this the same because there is no transivity with A and B.

**Problem 6: (10 points)**

Given the Relation Y = {P, Q, R, S, T} with the following dependencies:

{P, Q}🡪 S

{R, S}🡪 T

{S, T}🡪 P

1. **Is {P, Q} a candidate key of Y? Explain (4 pts)**

No, PQ is not a candidate key because P is functionaly dependent on ST so it cannot be apart of the candidate key.

1. **Is {P,Q,R} a candidate key of Y? Explain (3pts)**No, PQR is not a candiate key. I believe that the primary key is QR, therefore PQR is a superkey and not considered minimal enough to be a primary key.
2. **Is {D,E,F,H} a candidate key of Y? Explain (3 pts)**No, DEFH is not a candidate key of Y because DEFH contains all attributes that are not apart of Relation Y therefore it cannot functionally determine anything and be considered a candidate key.

**Problem 7 (10 points):**

***You are given a single-table database of the following manager-employee relations.***

|  |  |
| --- | --- |
| **Column Heading** | **Explanation** |
| MID | Manager unique identifier |
| MFNAME | Manager first name |
| MLNAME | Manager last name |
| EEFNAME | Employee first name |
| EELNAME | Employee last name |
| EEPHONE# | Employee phone number |
| EEADDRESS | Employee address |
| EESEX | Employee sex |

***Assume the company has one level of manager and that managers are considered separate from employees.*** Assume each manager has a unique identifier (MID). Also, assume each manager can have multiple employees, but those employees’ name **will always be** unique per manager (but not necessarily unique across the whole table, i.e., no manager will have two employees with the same name, but two managers may have employees named John Smith).

**What would the primary key be? Design the table accordingly, but you *may not create any new columns*. Show the attributes and a few sample records (enough to make your thinking clear).**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **MID** | **MFNAME** | **MLNAME** | **EEFNAME** | **EELNAME** | **EEPHONE#** | **EEADDRESS** | **EESEX** |
| 253 | Justin | Bieber | John | Smith | 843-123-4569 | 123 Sesame St | M |
| 253 | Justin | Bieber | Harold | Velez | 324-345-6788 | 56 Apple Ave | M |
| 111 | Selena | Gomez | John | Smith | 790-456-7986 | 621 Ridge Rd | M |
| 323 | Hailey | Baldwin | Kylie | Jenner | 583-256-6597 | 999 Brook Blvd | F |

The primary key for this table would be MID, EEFNAME, and EELNAME because MID is the unique identifier to the manager and the employee’s name (assuming full name) will always be unique to the manager. EEFNAME and EELNAME would both be in the primary key because people can have the first name or last name so together, the name is more unique. In my table that I made, MID of 253 is uniquely linked to Justin Bieber, however it is not unique in the table and is tied to different employees like John Smith and Harold Velez. It is the same case for the employee’s name because it can show up multiple times in the table, ex: John Smith. If we know that no manager wil have two employees that share the same name, then it safe to assume that MID with the employee’s name (EEFNAME, EELNAME) will uniquely determine each record (row).

**Problem 8 (10 points):**

***Similar to Problem 7 but not identical.***

***You are given a single-table database of the following manager-employee relations.***

|  |  |
| --- | --- |
| **Column Heading** | **Explanation** |
| MID | Manager unique identifier |
| MFNAME | Manager first name |
| MLNAME | Manager last name |
| EEFNAME | Employee first name |
| EELNAME | Employee last name |
| EEPHONE# | Employee phone number |
| EEADDRESS | Employee address |
| EESEX | Employee sex |

***Assume the company has one level of manager and that managers are considered separate from employees.*** Assume each manager has a unique identifier (MID). Also, assume each manager can have multiple employees, but those employee’s name may not be unique per manager.

**What would the primary key be? Design the table accordingly, but in this problem, you *may create new columns*. You may not use SSN or a surrogate key as a primary key. Show the attributes and a few sample records (enough to make your thinking clear).**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **MID** | **MFNAME** | **MLNAME** | **EEFNAME** | **EELNAME** | **EEPHONE#** | **EEADDRESS** | **EESEX** | **EEID** |
| 253 | Justin | Bieber | John | Smith | 843-123-4569 | 123 Sesame St | M | 32245601 |
| 253 | Justin | Bieber | Harold | Velez | 324-345-6788 | 56 Apple Ave | M | 43401323 |
| 111 | Selena | Gomez | John | Smith | 790-456-7986 | 621 Ridge Rd | M | 32245601 |
| 323 | Hailey | Baldwin | Kylie | Jenner | 583-256-6597 | 999 Brook Blvd | F | 54421049 |
| 253 | Justin | Bieber | Harold | Velez | 657-975-9075 | 823 Lipton Lane | M | 29485739 |

I created EEID column which is an unique employee number assigned to each employee. With this new column, the new primary key would be EEID and MID. The EEID will help distinguish people with same name, therefore a combination of EEID and MID will uniquely identify each row. This example is seen on the table where Justin Bieber and Harold Velez relationship occurs twice, however the two Harold Velez are completely different people based on the rest of the information and EEID number. Using just the name is not enough to be in the primary key. To point out, one assumption of this relation is that there are no duplicate manager-employee relationship records on this table. For instance, there is no second record of MID of 253 and EEID of 32245601. The duplication of records of the relationships may happen if the table wanted to document management over time, however if it was, this would not be the optimal way to do it because there will be a ton of information repeated and space wasted. Therefore, I will support the assumption of there being no manager-employee duplicates, so the primary key would still be MID and EEID.

**Text

Description automatically generated**