**TOPIC: NORMALIZATION**

1. Define the term “Normalization’.

*It is a process, which transforms any complex database table (un-normalized database table) into smaller (normalized) simple or manageable tables*

1. Explain the process involved in the Normalization process.

0NF

**PHASE 1:1ST NORMAL FORM**

Remove all repeating columns from the 0NF (Un-normalized table given)

**PHASE 2:2ND NORMAL FORM**

Remove all partial dependencies from the 1NF (First Normal Form) or from the tables which are already in the 1NF

**PHASE 3:3RD NORMAL FORM**

Remove all transitive dependencies from the 2NF (2ND Normal Form) or from the tables which are already in the 2NF

1. Transform the un-normalized table (UNF) below into first normal form (1NF), second normal form (2NF) and third normal form (3NF).

**0NF or un-normalized table or complex database table**

| *emp-id* | *emp\_name* | *proj\_id* | *project\_name* | *proj\_cost* | *proj\_st\_date* | *proj\_end\_date* | *client\_id* | *client\_name* | *hr\_worked* |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| E1 | JAM | P1 | School Management System | RM 300,000 | Jan-2018 | Dec-2018 | C1 | APU | 10 |
| E1 | JAM | P2 | Hospital Management System | RM 250,000 | Jan-2019 | Feb-2020 | C2 | Pantai | 10 |
| E1 | JAM | P3 | Property Management Systems(PMS) | RM 100,000 | Dec-2019 | Feb-2021 | C3 | Proptiger | 100 |
| E2 | BABA | P4 | Fee Management System | RM 50,000 | Jul-2018 | Aug-2018 | C1 | APU | 50 |
| E2 | BABA | P3 | Property Management Systems(PMS) | RM 100,000 | Dec-2019 | Feb-2021 | C3 | Proptiger | 50 |

**PHASE 1:1ST NORMAL FORM**

Remove all repeating columns from the 0NF (Un-normalized table given)

EMPLYEE\_T(**emp\_id**, emp\_name) - It is a Normalized database table

Primary Key: emp\_id

|  |  |
| --- | --- |
| emp\_id (PK attribute) | emp\_name (Non-PK attribute |
| E1 | Jam |
| E2 | BABA |

Above database table is in 1NF

EMPLOYEE\_PROJECT\_CLINET\_T (**emp\_id, proj\_id**, proj\_name, proj\_cost, proj\_st\_date, proj\_end\_date, cli\_id, cli\_name, hr\_wordked)

Primary Key: (emp\_id+proj\_id) – It is a composite primary PK. A PK with 2 or more attributes

| ***emp\_id*** | ***proj\_id*** | *project\_name* | *proj\_cost* | *proj\_st\_date* | *proj\_end\_date* | *client\_id* | *client\_name* | *hr\_worked* |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| E1 | P1 | School Management System | RM 300,000 | Jan-2018 | Dec-2018 | C1 | APU | 10 |
| E1 | P2 | Hospital Management System | RM 250,000 | Jan-2019 | Feb-2020 | C2 | Pantai | 10 |
| E1 | P3 | Property Management Systems(PMS) | RM 100,000 | Dec-2019 | Feb-2021 | C3 | Proptiger | 100 |
| E2 | P4 | Fee Management System | RM 50,000 | Jul-2018 | Aug-2018 | C1 | APU | 50 |
| E2 | P3 | Property Management Systems(PMS) | RM 100,000 | Dec-2019 | Feb-2021 | C3 | Proptiger | 50 |

Above database table is in 1NF

**PHASE 2:2ND NORMAL FORM**

Remove all partial dependencies from the 1NF (First Normal Form) or from the tables which are already in the 1NF

Partial dependencies – A table may suffer from this partial dependencies, if a non PK attribute does not fully dependent on its PK.

If any non-PK attribute found in 1NF, does not fully dependent on its PK attribute(s), remove that attribute from the table and move it to a new table.

EMPLOYEE\_PROJECT\_T(**emp\_id, proj\_id**, hr\_wordked) - It is a Normalized database table

| **emp\_id** | ***proj\_id*** | *hr\_worked* |
| --- | --- | --- |
| E1 | P1 | 10 |
| E1 | P2 | 10 |
| E1 | P3 | 100 |
| E2 | P4 | 50 |
| E2 | P3 | 50 |

Above database table is in 2NF

PROJECT\_CLIENT\_T(**proj\_id**, proj\_name, proj\_cost, proj\_st\_date, proj\_end\_date, cli\_id, cli\_name)

| ***proj\_id*** | *project\_name* | *proj\_cost* | *proj\_st\_date* | *proj\_end\_date* | *client\_id* | *client\_name* |
| --- | --- | --- | --- | --- | --- | --- |
| P1 | School Management Syste | RM 300,000 | Jan-2018 | Dec-2018 | C1 | APU |
| P2 | Hospital Management System | RM 250,000 | Jan-2019 | Feb-2020 | C2 | Pantai |
| P3 | Property Management System | RM 100,000 | Dec-2019 | Feb-2021 | C3 | Proptiger |
| P4 | Fee Management System | RM 50,000 | Jul-2018 | Aug-2018 | C1 | APU |

Above database table is in 2NF

**PHASE 3:3RD NORMAL FORM**

Remove all transitive dependencies from the database tables already in 2NF (2ND Normal Form)

If any non-PK attribute found in 2NF, dependents on another non-PK attribute, remove that attributes from the table and transfer to a new table.

Transitive dependencies – A table may suffer from this transitive dependencies, if a non PK attribute dependents on another non PK attribute.

PROJECT\_T

(**proj\_id**, proj\_name, proj\_cost, proj\_st\_date, proj\_end\_date, cli\_id) - It is a Normalized database table

| ***proj\_id*** | *project\_name* | *proj\_cost* | *proj\_st\_date* | *proj\_end\_date* | *client\_id* |
| --- | --- | --- | --- | --- | --- |
| P1 | School Management System | RM 300,000 | Jan-2018 | Dec-2018 | C1 |
| P2 | Hospital Management System | RM 250,000 | Jan-2019 | Feb-2020 | C2 |
| P3 | Property Management Systems(PMS) | RM 100,000 | Dec-2019 | Feb-2021 | C3 |
| P4 | Fee Management System | RM 50,000 | Jul-2018 | Aug-2018 | C1 |

CLIENT\_T

(**cli\_id**, cli\_name) - It is a Normalized database table

| ***client\_id*** | *client\_name* |
| --- | --- |
| C1 | APU |
| C2 | Pantai |
| C3 | Proptiger |

**Below are the normalized tables**

PROPERTY\_T (**prop\_id, prop\_name,prop\_street,prop\_city,prop\_country,prop\_postcode,prop\_area,prop\_uom,prop\_price**)

Primary Key: **prop\_id**

|  |  |  |  |
| --- | --- | --- | --- |
| **prop\_id** | **prop\_name** | **prop\_street** | **prop\_city** |
| E1 | James |  |  |
| E2 | BABA |  |  |

EMPLOYEE\_PROJECT\_T(**emp\_id, proj\_id**, hr\_wordked)

| **emp\_id** | ***proj\_id*** | *hr\_worked* |
| --- | --- | --- |
| E1 | P1 | 10 |
| E1 | P2 | 10 |
| E1 | P3 | 100 |
| E2 | P4 | 50 |
| E2 | P3 | 50 |

EMPLOYEE\_PROJECT\_T(**emp\_id, proj\_id**, hr\_wordked)

| **emp\_id** | ***proj\_id*** | *hr\_worked* |
| --- | --- | --- |
| E1 | P1 | 10 |
| E1 | P2 | 10 |
| E1 | P3 | 100 |
| E2 | P4 | 50 |
| E2 | P3 | 50 |

PROJECT\_T

(**proj\_id**, proj\_name, proj\_cost, proj\_st\_date, proj\_end\_date, cli\_id) ok

| ***proj\_id*** | *project\_name* | *proj\_cost* | *proj\_st\_date* | *proj\_end\_date* | *client\_id* |
| --- | --- | --- | --- | --- | --- |
| P1 | School Management System | RM 300,000 | Jan-2018 | Dec-2018 | C1 |
| P2 | Hospital Management System | RM 250,000 | Jan-2019 | Feb-2020 | C2 |
| P3 | Property Management System | RM 100,000 | Dec-2019 | Feb-2021 | C3 |
| P4 | Fee Management System | RM 50,000 | Jul-2018 | Aug-2018 | C1 |

CLIENT\_T

(**cli\_id**, cli\_name) ok

| ***client\_id*** | *client\_name* |
| --- | --- |
| C1 | APU |
| C2 | Pantai |
| C3 | Proptiger |

A composite key is a primary key composed of multiple columns( 2 or above attributes) used to identify a row/tuple uniquely

A table can have only 1 PK, 1CP and more than 1 FK