

Normalisation of a Class Roll

Class Roll			
Qualification:	ICT50718	Start time:	9:00 AM
Unit:	ICTDB5502	Finish Time:	3:00 PM
Date:	3/03/2020		
TeacherNo.:	1235322	Teacher Name:	James Summer
Student Number	Student Name	Student Email	Attended
1	Steve	SteveSmith@live.com	yes
2	Jennifer	jenniferSmith@live.com	yes
3	Donald	DonalJones@live.com	yes
4	steve 2	SteveSmith@live.com	yes
5	steve 3	SteveSmith@live.com	no
Total Attendance:			4

Scenario image and videos provided by Shaun O'Sullivan

ERD video link:

<https://mediasite.tafeqld.edu.au/Mediasite/Play/f46cca4ddf464811aac81a53c52999f61d>

Normalisation video link:

<https://mediasite.tafeqld.edu.au/Mediasite/Play/b229a7871d3d4e8597c613e2e395ea3d1d>

Via Connect:

ICTICT509 -> Content -> Supplementary Resources -> Videos.

Some Background on Normalisation

Normalisation is a process that complements and reinforces the process of Entity-Relationship diagramming. It is possible to fully identify a set of entities and their relationships and consequently a set of database tables using Entity-Relationship alone. Nevertheless, Normalisation can allow you to supplement or validate these by starting with data from existing forms or reports and breaking them down into individual relations.

Normalisation can also naturally follow from the data elements identified in the Contents column of the data flow and/or data stores information in a data dictionary.

As an example, let us say that we have a class attendance form from a classroom. In the Data Flows section of their data dictionary let us assume we identified the content of the class roll to be:

{ Qualification, Unit, Date, StartTime, FinishTime, TeacherNo, TeacherName, { StudentNumber, StudentName, StudentEmail, Attended }, TotalAttendance }

The process of Normalisation would take the element content of this form and break it down systematically into relations. This process is known to have a number of steps although only the first 3 receive significant attention in general Systems Analysis texts. The output of each step is known as a **normal** form. For example:

Original Un-normalised Data Elements

*Note: Identifying Attributes have been identified in **bold with an underline** within the relations below..*

{ ClassRollNumber, Qualification, Unit, Date, StartTime, FinishTime, TeacherNo, TeacherName, { StudentNumber, StudentName, StudentEmail, Attended }, TotalAttendance }

The TotalAttendance element could be removed in the initial stages, realising that this would likely be a calculated field in the resulting database application and would be displayed when data was presented in forms and reports.

{ ClassRollNumber, Qualification, Unit, Date, StartTime, FinishTime, TeacherNo, TeacherName, { StudentNumber, StudentName, StudentEmail, Attended } }

First Normal Form

In this step, the **Inner Repeating Groups are removed**. The repeating groups are identified by parentheses - { ... }. This step aims to remove or separate the inner repeating group(s) from the outer repeating group. In the process we need to add the Identifying Attribute from the outer group to the inner group so as to retain the link or relationship between the two resulting relations. The Identifying Attribute for the new relation may become a combination of the identifier from the outer group and the identifier for the inner group.

{ ClassRollNumber, Qualification, Unit, Date, StartTime, FinishTime, TeacherNo, TeacherName }

{ ClassRollNumber, StudentNumber, StudentName, StudentEmail, Attended }

Second Normal Form

In this step, the **Partial Dependencies are removed**. This step is only applicable to relations that have a concatenated Identifying Attribute set – ie there are two or more attributes that make up the 'Identifying Attribute'.

It would be a fair statement to say that in this example an element such as StudentName would be better represented by the StudentNumber identifier on its own, rather than by the combination of ClassRollNumber and StudentNumber. IE: the StudentName (and StudentEmail) should be in a relation separate from the 'ClassRollNumber – StudentNumber' relation.

In doing so we also need to copy the StudentNumber to the new relation so as to maintain the link or relationship between the two relations.

{ ClassRollNumber, Qualification, Unit, Date, StartTime, FinishTime, TeacherNo, TeacherName }

{ ClassRollNumber, StudentNumber, Attended }

{ StudentNumber, StudentName, StudentEmail }

Third Normal Form

In this step, the **Transitive Dependencies are removed**. This step is applicable to any section of the current relations where there are attributes that don't link directly to the current Identifying Attribute(s) IE: these attributes are not best represented by the current Identifying Attribute.

Again going back to our example, it can be seen that the Teacher details in the first relation are not best represented by the Identifying Attribute - ClassRollNumber. The Teacher details are more appropriately represented by the TeacherNo (which is not an Identifying Attribute in the first relation). This type of dependency (on a non-key attribute) is often referred to as a transitive dependency. To remove this dependency, the Teacher information can be removed to its own relation.

To maintain the link or relationship, a copy of the TeacherNo needs to be left in the Class Roll relation.

These final set of relations might be named so follows:

CLASS ROLL: { ClassRollNumber, Qualification, Unit, Date, StartTime, FinishTime, TeacherNo }

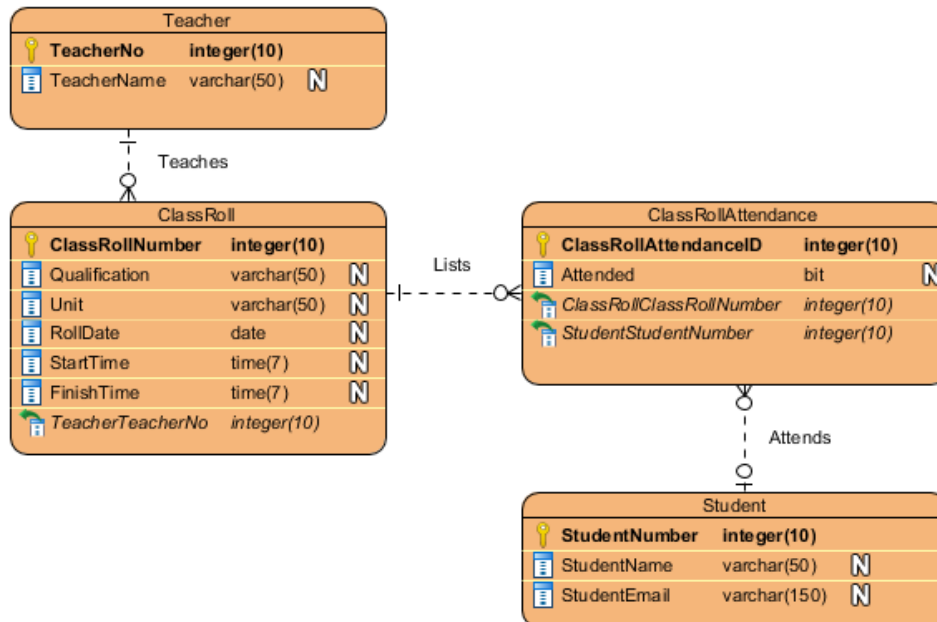
TEACHER: { TeacherNo, TeacherName }

CLASS ROLL ATTENDANCE: { ClassRollNumber, StudentNumber, Attended }

STUDENT: { StudentNumber, StudentName, StudentEmail }

Entity Relationship Diagram of the Class Roll

Based on an ERD analysis, or via the Normalisation detailed on the previous pages:



With Additional Lookup Lists:

