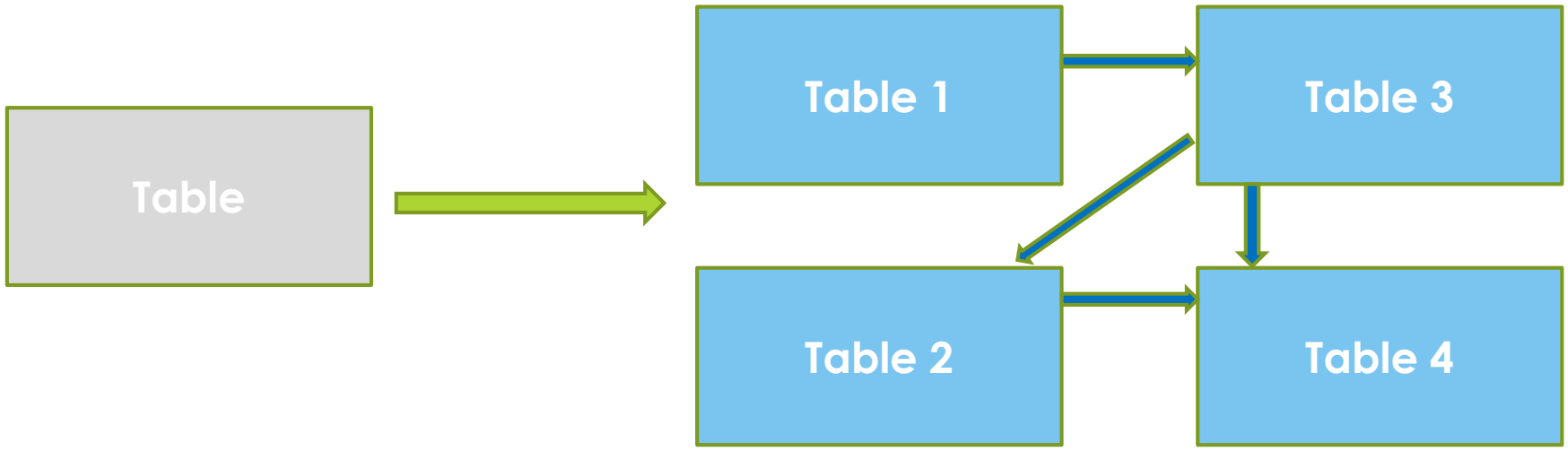
Applied Databases Normalisation

HIGHER DIPLOMA IN DATA ANALYTICS

Normalisation

- Normalization is the process of organizing the columns (attributes) and tables (relations) of a relational database to minimize data redundancy.



Normalisation

Hospital
PPSN varchar(10) Primary Key
First_Name varchar(50)
Surname varchar(50)
Address varchar(200)
Doctor varchar(50)
Doctor_Phone integer

Hospital					
PPSN*	**First_Name**	**Surname**	**Address**	**Doctor**	**Doctor_Phone**
7629913X	John	Smyth	Athlone	Dr. Jones	12345
~~9893333F~~	~~Alan~~	~~Mulligan~~	~~Galway~~	~~Dr. Murphy~~	~~88335~~
9898823W	Fred	Collins	Castlebar	Dr. Jones	12345
NULL	NULL	NULL	NULL	Dr. Rice	88788

Normalisation

Patient_Table
PPSN varchar(10) Primary Key
First_Name varchar(50)
Surname varchar(50)
Address varchar(200)
DoctorID integer

Doctor_Table
DoctorID integer Primary Key
Name varchar(50)
Phone integer

Patient_Table

PPSN*	First_Name	Surname	Address	DoctorID
7629913X	John	Smyth	Athlone	101
9893333F	Alan	Mulligan	Galway	101
9898823W	Fred	Collins	Castlebar	NULL

Doctor_Table

DoctorID*	Name	Phone
~~100~~	~~Dr. Jones~~	~~12345~~
101	Dr. Murphy	88335
102	Dr. Rice	88788

Normalisation

Patient_Table
PPSN varchar(10) Primary Key
First_Name varchar(50)
Surname varchar(50)
Address varchar(200)
DoctorID integer

FOREIGN KEY REFERENCES doctor_table(DoctorID)

Doctor_Table
DoctorID integer Primary Key
Name varchar(50)
Phone integer

Patient_Table				
PPSN*	**First_Name**	**Surname**	**Address**	**DoctorID**
7629913X	John	Smyth	Athlone	100
9893333F	Alan	Mulligan	Galway	101
9898823W	Fred	Collins	Castlebar	100
12345F	Alan	Mulligan	Mayo	102

Doctor_Table		
DoctorID*	**Name**	**Phone**
100	Dr. Jones	12345
101	Dr. Murphy	88335
999	Dr. Rice	9898

SHOW CREATE TABLE

Patient_Table
PPSN varchar(10) Primary Key
First_Name varchar(50)
Surname varchar(50)
Address varchar(200)
DoctorID integer
FOREIGN KEY REFERENCES doctor_table(DoctorID)

Doctor_Table
DoctorID integer Primary Key
Name varchar(50)
Phone integer

```
mysql> SHOW CREATE TABLE patient_table;
| patient_table | CREATE TABLE `patient_table` (
  `ppsn` varchar(10) NOT NULL,
  `first_name` varchar(50) DEFAULT NULL,
  `surname` varchar(50) DEFAULT NULL,
  `address` varchar(200) DEFAULT NULL,
  `doctorID` int(11) DEFAULT NULL,
  PRIMARY KEY (`ppsn`),
  KEY `doctorID` (`doctorID`),
  CONSTRAINT `patient_table_ibfk_1` FOREIGN KEY (`doctorID`) REFERENCES `doctor_table` (`doctorID`)

) ENGINE=InnoDB DEFAULT CHARSET=latin1 |
```

Getting information from multiple tables

Patient_Table				
PPSN*	**First_Name**	**Surname**	**Address**	**DoctorID**
7629913X	John	Smyth	Athlone	100
9893333F	Alan	Mulligan	Galway	101
9898823W	Fred	Collins	Castlebar	100

Doctor_Table		
DoctorID*	**Name**	**Phone**
100	Dr. Jones	12345
101	Dr. Murphy	88335

Patient & Doctor Information			
PPSN*	**First_Name**	**Name**	**Phone**
7629913X	John	Dr. Jones	12345
9893333F	Alan	Dr. Murphy	88335
9898823W	Fred	Dr. Jones	12345

Getting information from multiple tables

Patient_Table				
PPSN*	**First_Name**	**Surname**	**Address**	**DoctorID**
7629913X	John	Smyth	Athlone	100
9893333F	Alan	Mulligan	Galway	101
9898823W	Fred	Collins	Castlebar	100

Doctor_Table		
DoctorID*	**Name**	**Phone**
100	Dr. Jones	12345
101	Dr. Murphy	88335

```
SELECT patient_table.ppsn, patient_table.first_name,
       doctor_table.name, doctor_table.phone
FROM patient_table
INNER JOIN doctor_table
     on patient_table.doctorID = doctor_table.doctorID;
```

Getting information from multiple tables

Patient_Table

PPSN*	First_Name	Surname	Address	DoctorID
7629913X	John	Smyth	Athlone	100
9893333F	Alan	Mulligan	Galway	101
9898823W	Fred	Collins	Castlebar	100

Doctor_Table

DoctorID*	Name	Phone
100	Dr. Jones	12345
101	Dr. Murphy	88335

```
SELECT pt.ppsn, pt.first_name, pt.surname, patient_table.first_name,
       dt.name AS doc, doctor_table.phone
FROM patient_table pt
INNER JOIN doctor_table dt
      on pt.doctorID = dt.doctorID;
```

```
mysql> SELECT name as doc
    -> FROM doctor_table;
+------------+
| doc        |
+------------+
| Dr. Jones  |
| Dr. Murphy |
+------------+
2 rows in set (0.00 sec)
```

Getting information from multiple tables

Patient_Table

PPSN*	First_Name	Surname	Address	DoctorID
7629913X	John	Smyth	Athlone	100
9893333F	Alan	Mulligan	Galway	101
9898823W	Fred	Collins	Castlebar	100
2344234S	Mary	Burke	Galway	NULL

Doctor_Table

DoctorID*	Name	Phone
100	Dr. Jones	12345
101	Dr. Murphy	88335

```
mysql> SELECT pt.ppsn, pt.first_name, dt.name
    -> FROM patient_table pt
    -> INNER JOIN doctor_table dt
    ->   on pt.doctorID = dt.doctorID;
+----------+------------+------------+
| ppsn     | first_name | name       |
+----------+------------+------------+
| 7629913X | John       | Dr. Jones  |
| 9898823W | Fred       | Dr. Jones  |
| 989333F  | Alan       | Dr. Murphy |
+----------+------------+------------+
3 rows in set (0.00 sec)
```

INNER JOIN v LEFT JOIN

INNER JOIN	LEFT JOIN
Returns rows from two tables only when the JOIN condition is met. If JOIN condition is NOT met, nothing is returned from either table	Returns rows from two tables when the JOIN condition is met. If JOIN condition is NOT met, rows from the first (LEFT) table are returned and NULL is returned instead of rows from the second table.

INNER JOIN v LEFT JOIN

Patient_Table

PPSN*	First_Name	Surname	Address	DoctorID
7629913X	John	Smyth	Athlone	100
9893333F	Alan	Mulligan	Galway	101
9898823W	Fred	Collins	Castlebar	100
2344234S	Mary	Burke	Galway	NULL

Doctor_Table

DoctorID*	Name	Phone
100	Dr. Jones	12345
101	Dr. Murphy	88335

```
mysql> SELECT pt.ppsn, pt.first_name, dt.name
    -> FROM patient_table pt
    -> INNER JOIN doctor_table dt
    ->   on pt.doctorID = dt.doctorID;
+----------+------------+------------+
| ppsn     | first_name | name       |
+----------+------------+------------+
| 7629913X | John       | Dr. Jones  |
| 9898823W | Fred       | Dr. Jones  |
| 989333F  | Alan       | Dr. Murphy |
+----------+------------+------------+
3 rows in set (0.00 sec)
```

```
mysql> SELECT pt.ppsn, pt.first_name, dt.name
    -> FROM patient_table pt
    -> LEFT JOIN doctor_table dt
    ->   on pt.doctorID = dt.doctorID;
+----------+------------+------------+
| ppsn     | first_name | name       |
+----------+------------+------------+
| 7629913X | John       | Dr. Jones  |
| 9898823W | Fred       | Dr. Jones  |
| 989333F  | Alan       | Dr. Murphy |
| 2344234S | Mary       | NULL       |
+----------+------------+------------+
4 rows in set (0.02 sec)
```

```
mysql> SELECT pt.ppsn, pt.first_name, dt.name
    -> FROM doctor_table dt
    -> LEFT JOIN patient_table pt
    ->   on dt.doctorID = pt.doctorID;
+----------+------------+------------+
| ppsn     | first_name | name       |
+----------+------------+------------+
| 7629913X | John       | Dr. Jones  |
| 9898823W | Fred       | Dr. Jones  |
| 989333F  | Alan       | Dr. Murphy |
+----------+------------+------------+
3 rows in set (0.00 sec)
```