

Relational Databases

Terminology

Database

An organized collection of related data.

ID	Firstname	Lastname	Address	City	State	Zip
123	John	Doe	981 Main St	Fayetteville	AR	72701
456	Jane	Smith	654 Second St	Bentonville	AR	72712
789	Sarah	Brown	147 Hazel St	Springdale	AR	72762

Database Management System (DBMS)

Software that manages a database.

Allows you to create, add, modify and delete data as well as modify database structure.

Database System

The combination of a database, a database management system and the applications that access the database.

Relational Database

A database whose organization is based on the relational model of data.

Virtually all relational database systems use SQL (Structured Query Language).

Relational Model

An approach to organizing data using a table (relation) structure.

Defined by Edgar F. Codd in 1969/1970 while at IBM.

Relational Database Structure

Tables and Columns

A single table usually represents a particular type of data being stored.

Tables of data have columns and rows.

The structure of what tables the database has, and what attributes are in which table is called the schema.

Columns represent attributes that each instance of data stored in the database will have.

Columns and Rows

The set of allowed values for any particular attribute is called that attribute's domain.

Attributes of a relation can appear in any order.

Rows represent entires of data into the database. These rows are also called records or tuples. A row will have one value per attribute.

Rows and Quantities

Tuples of a relation can appear in any order.

Each tuple in a particular relation must be unique. There can be no duplicate data entries.

The degree of a relation is the number of attributes it contains.

The cardinality of a relation is the number of tuples it contains.

Normalization and Uniqueness

The process of determining the appropriate structure of a database, i.e. the relations in a database should be and what attributes individual relations should have, is determined via a process called normalization.

Each relation (table) must have a unique name relative to its schema.

Each tuple has only one entry per attribute.

Each attribute has a distinct name.

Normalization Demo

Spreadsheets Are Not Databases

Spreadsheets are a front end data recording, calculating and analytics tool. A spreadsheet is a file that is opened and closed.

Databases are a backend information storage and organization tool. A database is persistent and handles requests.

Keys

Key Definitions and Types

Keys are used to identify individual entries in a table, as well as establish relationships between entries in different tables.

- Superkey
- Candidate Key
- Composite Key
- Primary Key
- Alternate Key
- Foreign Key

Superkey

An attribute or set of attributes that uniquely identifies a tuple within a relation.

A relation may contain multiple superkeys by various combinations of attributes.

Any given superkey may contain attributes that are not strictly necessary for unique identification.

Candidate Key

A superkey such that no proper subset is a superkey within the relation.

The minimum viable combination for unique identification.

A relation may contain multiple candidate keys.

Composite Key

A candidate key such that it must contain multiple attributes.

The data in the relation is organized such that multiple attributes are required to ensure a key's uniqueness.

Primary Key

The candidate key that is selected as the preferred means to identify tuples within a relation.

Each relation in the database is required to have a primary key defined.

Candidate keys are named that because they are candidates to become the primary key.

Alternate Key

Candidate keys not chosen to be the primary key.



Foreign Key

An attribute or set of attributes within one relation that matches the candidate key for some relation.

When an attribute appears in more than one relation, that usually indicates a relationship between the tuples of the two relations.

What makes normalization work.

Relational Algebra

Relational Algebra

The eight operations that can be done to relations.

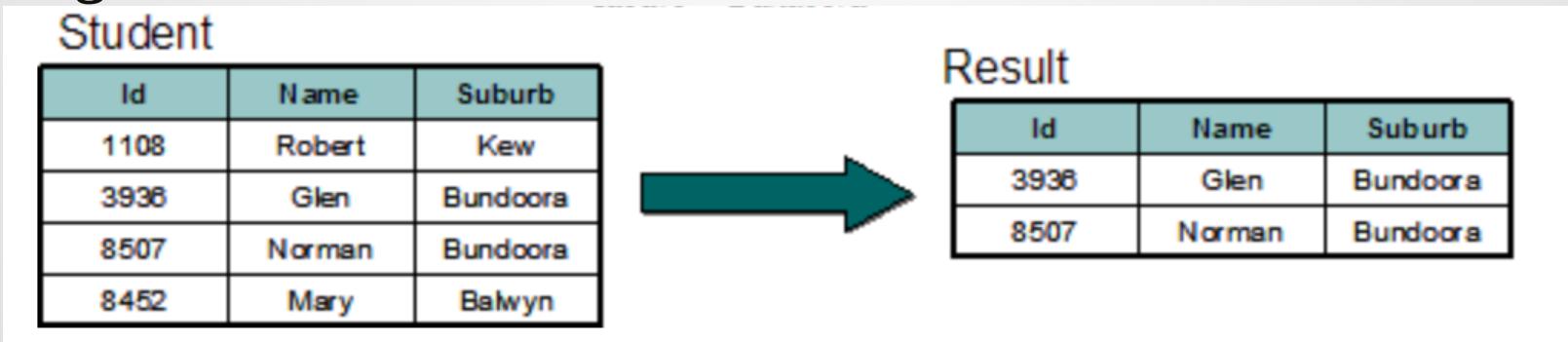
- Select
- Project
- Union
- Intersection
- Difference
- Product
- Join
- Division

Select

A unary operation (only works with one relation).

Accepts a condition, then applies the condition and returns the tuples that satisfy the condition.

The degree remains unchanged, cardinality may or may not change.

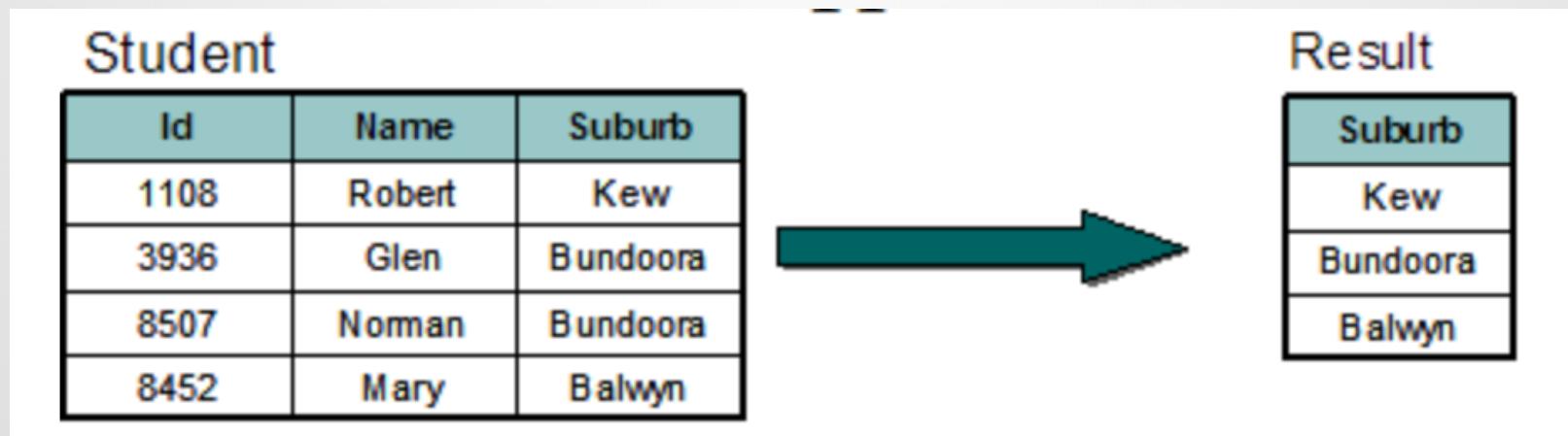


Project

A unary operation (only works with one relation).

Accepts a list of attributes and returns only the values for those attributes for all tuples.

The degree will change, cardinality may or may not.



Union

A binary operation (only works with two relations).

Combines tuples from both relations, removing any duplicates.

Requires the relations be union compatible: the degrees of the two relations must be equal

Subject

Name	Course
Systems	BCS
Database	BCS
Database	MCS
Algebra	MCS



Subject2

Name	Course
DataMining	MCS
Writing	BCS

Result

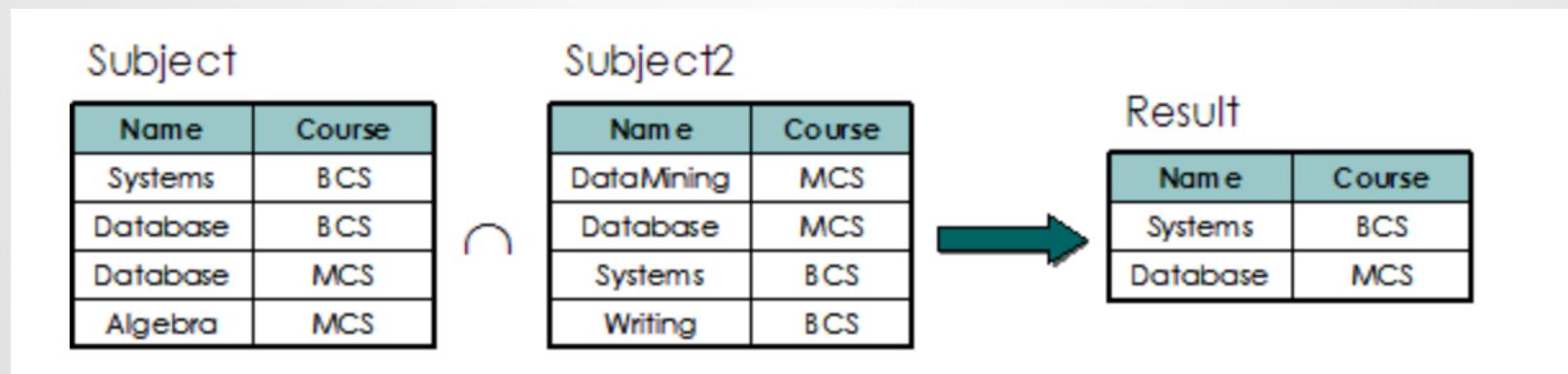
Name	Course
Systems	BCS
Database	BCS
Database	MCS
Algebra	MCS
DataMining	MCS
Writing	BCS

Intersection

A binary operation (only works with two relations).

Finds the tuples that exist in both relations.

Requires the relations to be union compatible.



Difference

A binary operation (only works with two relations).

Finds the tuples that exist in one relation but not the other.

Requires the relations to be union compatible.

Subject	
Name	Course
Systems	BCS
Database	BCS
Database	MCS
Algebra	MCS

Subject2	
Name	Course
DataMning	MCS
Database	MCS
Systems	BCS
Writing	BCS

Result



Name	Course
Database	BCS
Algebra	MCS

Product

A binary operation (only works with two relations).

Takes every tuple in one relation and appends every tuple in the other relation.

Student		
Id	Name	Suburb
1108	Robert	Kew
2936	Glen	Bundoora
8307	Norman	Bundoora
8452	Mary	Balwyn



Sport	
SportID	Sport
01	Swimming
02	Dancing



Student_Sport				
Id	Name	Suburb	SportID	Sport
1108	Robert	Kew	01	Swimming
2936	Glen	Bundoora	01	Swimming
8307	Norman	Bundoora	01	Swimming
8452	Mary	Balwyn	01	Swimming
1108	Robert	Kew	02	Dancing
2936	Glen	Bundoora	02	Dancing
8307	Norman	Bundoora	02	Dancing
8452	Mary	Balwyn	02	Dancing

Join

A binary operation (only works with two relations).

Will combine the tuples of the different relations in different ways depending on the type of join.

<i>Employee</i>		
Name	Empld	DeptName
Harry	3415	Finance
Sally	2241	Sales
George	3401	Finance
Harriet	2202	Sales

<i>Dept</i>	
DeptName	Manager
Finance	George
Sales	Harriet
Production	Charles

<i>Employee</i> \bowtie <i>Dept</i>			
Name	Empld	DeptName	Manager
Harry	3415	Finance	George
Sally	2241	Sales	Harriet
George	3401	Finance	George
Harriet	2202	Sales	Harriet

Division

A binary operation (only works with two relations).

<i>Completed</i>	
Student	Task
Fred	Database1
Fred	Database2
Fred	Compiler1
Eugene	Database1
Eugene	Compiler1
Sarah	Database1
Sarah	Database2

<i>DBProject</i>
Task
Database1
Database2

$$\begin{array}{c} \textit{Completed} \\ \div \\ \textit{DBProject} \\ \hline \textit{Student} \\ \hline \text{Fred} \\ \hline \text{Sarah} \end{array}$$

SQL