Relational Algebra

Learning Maps

Sequence	Title
1	Introduction to databases
2	Relational Databases
3	Relational Algebra
4	Structured Query Language – Part 1
5	Structured Query Language – Part 2
6	Constraints and Triggers
7	Entity Relationship Model
8	Functional Dependency
9	Normalization
10	Storage - Indexing
11	Query Processing
12	Transaction Management – Part 1
13	Transaction Management – Part 2

Intro > Overview

☐ A : Voice and PPT Overview
☐ B : Text-based Overview
■ C : Video and PPT Overview

Opening Message	→ In this lesson, we will study relational algebra Relational algebra, firstly created by Edgar F. Codd, is a family of algebras with a well-founded semantics used for modelling the data stored in relational databases. The main application of relational algebra is providing a theoretical foundation for relational databases, particularly query languages for such databases
Lesson topic	1. Introduction to relational algrebra 2. Set operators 3. Relational operators: Projection, Selection, Rename, Join 4. Common extensions
Learning Goals	Upon completion of this lesson, students will be able to: 1. Understand relational algebra operators 2. Write relational algebraic expressions

Intro > Keywords

Keyword	Description
Relational data model	Is data representation format as a table of values, each row in the table represents a collection of related data values.
Set	is collection of Object
Operator	Is a special token that represent computations such as union, minus , selection, join, etc
Expression	Is a expression built up from operators and operands.

Database Schema

student(**student_id**, first_name,last_name, dob, gender,address,note,*clazz_id*)
clazz(**clazz_id**, name, *lecturer_id*, *monitor_id*)
subject(**subject_id**, name, credit, percentage_final_exam)
enrollment(***student_id, subject_id, semester***, midterm_score, final_score)
lecturer(**lecturer_id**, first_name, last_name, dob, gender, address, email)
teaching(***subject_id, lecturer_id***)
grade(**code**, from_score, to_score)

Database

student

student_id	first_name	last_name	dob	…	clazz_id
20160001	Ngọc An	Bùi	3/18/1987	…	
20160002	Anh	Hoàng	5/20/1987	…	20162101
20160003	Thu Hồng	Trần	6/6/1987	…	20162101
20160004	Minh Anh	Nguyễn	5/20/1987	…	20162101
20170001	Nhật Ánh	Nguyễn	5/15/1988	…	20172201

subject

subject_id	name	credit	percentage_final_exam
IT1110	Tin học đại cương	4	60
IT3080	Mạng máy tính	3	70
IT3090	Cơ sở dữ liệu	3	70
IT4857	Thị giác máy tính	3	60
IT4866	Học máy	2	70

clazz

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001
20172202	CNTT2.02-K62		

enrollment

student_id	subject_id	semester	midterm_score	final_score
20160001	IT1110	20171	9	8.5
20160001	IT3080	20172	8	
20160001	IT3090	20172	6	9
20160001	IT4857	20172	7.5	9
20160001	IT4866	20172	7	9
20160002	IT3080	20172	9	
20160003	IT1110	20171	7	6
20160004	IT1110	20171	6	5

Lesson > Topic 1: Introduction to relational algebra

- Relational algebra providing a theoretical foundation for relational databases, particularly query languages for relational databases
- Relational algebra expression is composed of one or several relational algebraic operators
  - Operator: represent computations
    - Input: one or two relation
    - Output: a relation

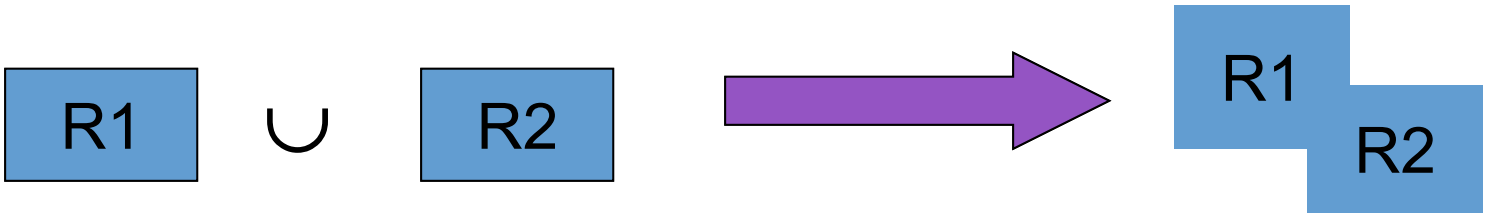
  - Unary operator (one input) vs. binary operator (two inputs)

Lesson > Topic 2: Set operators

- Union
- Intersection
- Difference
- Cartesian product

Lesson > Topic 2: Set operators

- **Union**: combining the tuples from two input *union-compat ible relations* (having the same set of attributes)



clazz

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003

∪

clazz 2

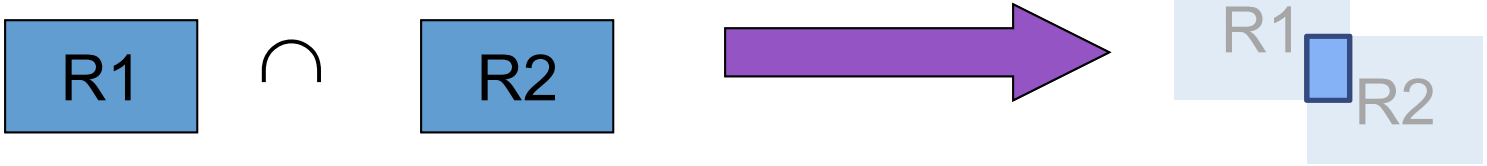
clazz_id	name	lecturer_id	monitor_id
20172201	CNTT2.01-K62	02002	20170001

result

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20172201	CNTT2.01-K62	02002	20170001

Lesson > Topic 2: Set operators

- **Intersection**: Keeping only common tuples from 2 input u nion-compatible relation



clazz

clazz_id	name	lecturer_id	monitor_id
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001
20172202	CNTT2.02-K62		

∪

clazz 2

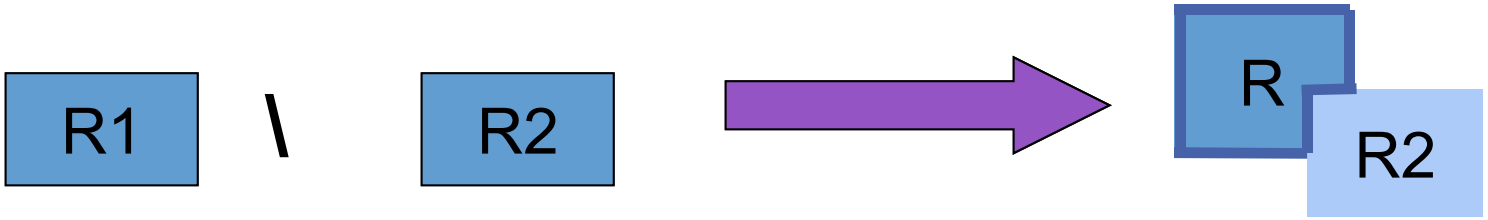
clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001

result

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001

Lesson > Topic 2: Set operators

- **Difference**: containing tuples occurred in the first relation but not in the second



clazz

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001
20172202	CNTT2.02-K62		

\

clazz 2

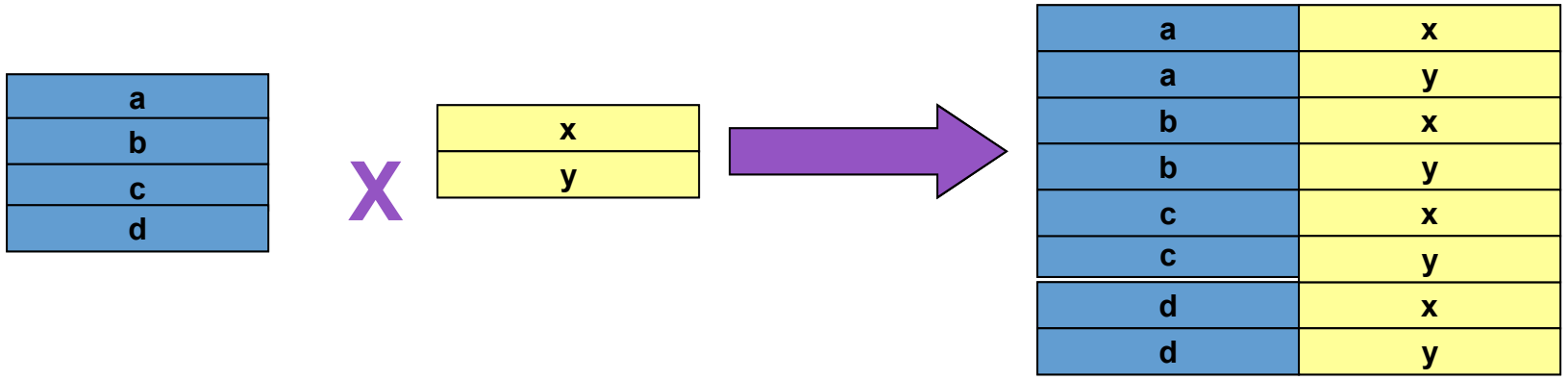
clazz_id	name	lecturer_id	monitor_id
20172202	CNTT2.02-K62		
20162102	CNTT1.02-K61		

result

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20172201	CNTT2.01-K62	02002	20170001

Lesson > Topic 2: Set operators

- **Cartesian Product**: the concatenation of every tuple of o ne relation with every tuple of the other relation.



Lesson > Topic 3: Relational algebraic operators

- Selection
- Projection
- Rename
- Join
- Division

Lesson > Topic 3: Relational algebraic operators

- **Selection**: choose from R each tuple where the conditio n holds.

$$\sigma_{<condition>}(R)$$

clazz

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001
20172202	CNTT2.02-K62		

$$\sigma_{\text{lecturer\_id != null}}(\text{clazz})$$

result

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20172201	CNTT2.01-K62	02002	20170001

Lesson > Topic 3: Relational algebraic operators

- **Projection**: Choose some attributes

$$\prod_{A1, A2, \ldots}(R)$$

clazz

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001
20172202	CNTT2.02-K62		

$$\prod_{\text{clazz\_id, name}}(\text{clazz})$$

result

clazz_id	name
20162101	CNTT1.01-K61
20172201	CNTT2.01-K62

Lesson > Topic 3: Relational algebraic operators

- Rename: result is identical to *R* except that the *b* attrib ute in all tuples is renamed to an *a* attribute

$$\rho_{a|b}(R)$$

clazz

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001
20172202	CNTT2.02-K62		

$$\rho_{\text{name of class} \mid \text{name}}(\text{clazz})$$

clazz

clazz_id	Name of class	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001
20172202	CNTT2.02-K62		

Lesson > Topic 3: Relational algebraic operators

- **Join**: Combine attributes from 2 tables

$$R_1 \rhd\lhd_{<join\_condition>} R_2$$

Lesson > Topic 3: Relational algebraic operators

Student ⋈ clazz

student

student_id	first_name	last_name	…	clazz_id
20160001	Ngọc An	Bùi		
20160002	Anh	Hoàng		20162101
20160003	Thu Hồng	Trần		20162101
20160004	Minh Anh	Nguyễn		20162101
20170001	Nhật Ánh	Nguyễn		20172201

clazz

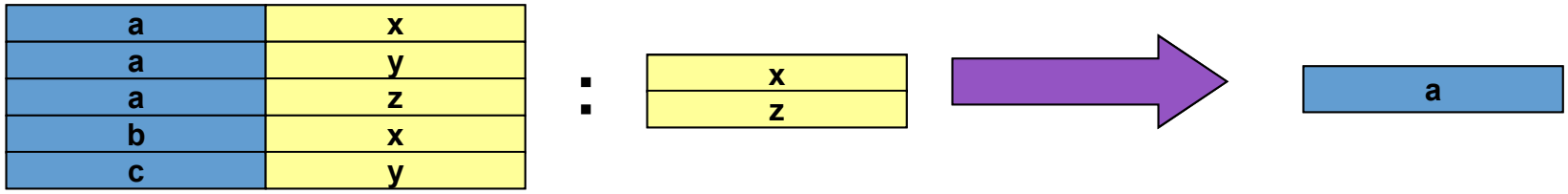
clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001
20172202	CNTT2.02-K62		

result

student_id	first_name	last_name	name
20160002	Anh	Hoàng	CNTT1.01-K61
20160003	Thu Hồng	Trần	CNTT1.01-K61
20160004	Minh Anh	Nguyễn	CNTT1.01-K61
20170001	Nhật Ánh	Nguyễn	CNTT2.01-K62

Lesson > Topic 3: Relational algebraic operators

- **Division**: divides a dividend relation R1 or degree m+n by a divisor relation R2 of degree n, and produces a qu otient relation of degree m.



Lesson > Topic 3: Relational algebraic operators

- **List student_id who enroll in all subjects**

enrollment

student_id	subject_id	semester	midterm_ score	final_ score
20160001	IT1110	20171	9	8.5
20160001	IT3080	20172	8	
20160001	IT3090	20172	6	9
20160001	IT4857	20172	7.5	9
20160001	IT4866	20172	7	9
20160002	IT3080	20172	9	
20160003	IT1110	20171	7	6
20160004	IT1110	20171	6	5

subject

subject_id	name	credit	percentage_ final_exam
IT1110	Tin học đại cương	4	60
IT3080	Mạng máy tính	3	70
IT3090	Cơ sở dữ liệu	3	70
IT4857	Thị giác máy tính	3	60
IT4866	Học máy	2	70

Lesson > Topic 3: Relational algebraic operators

- **List student_id who enroll in all subjects**

enrollment

student_id	subject_id	semester	midterm_ score	final_ score
20160001	IT1110	20171	9	8.5
20160001	IT3080	20172	8	
20160001	IT3090	20172	6	9
20160001	IT4857	20172	7.5	9
20160001	IT4866	20172	7	9
20160002	IT3080	20172	9	
20160003	IT1110	20171	7	6
20160004	IT1110	20171	6	5

subject

subject_id	name	credit	percentage_ final_exam
IT1110	Tin học đại cương	4	60
IT3080	Mạng máy tính	3	70
IT3090	Cơ sở dữ liệu	3	70
IT4857	Thị giác máy tính	3	60
IT4866	Học máy	2	70

$\prod_{student\_id,\ subject\_id}(enrollement)$ $\prod_{subject\_id}(subject)$

Lesson > Topic 3: Relational algebraic operators

- **List student_id who enroll in all subjects**

enrollment

student_id	subject_i
20160001	IT1110
20160001	IT3080
20160001	IT3090
20160001	IT4857
20160001	IT4866
20160002	IT3080
20160003	IT1110
20160004	IT1110

:

subject

subject_id
IT1110
IT3080
IT3090
IT4857
IT4866

Result

student_id
20160001

$(\Pi_{\text{student\_id, subject\_id}}(\text{enrollement})) : (\Pi_{\text{subject\_id}}(\text{subject}))$

Lesson > Topic 4: Common extension

- Natural join
- Outer join
- Aggregation

Lesson > Topic 4: Common extension

- **Natural join**: Special join operation with equal join co ndition on their common attributes, noted *

student

student_id	first_name	last_name	dob	…	clazz_id
20160001	Ngọc An	Bùi	3/18/1987	…	
20160002	Anh	Hoàng	5/20/1987	…	20162101
20160003	Thu Hồng	Trần	6/6/1987	…	20162101
20160004	Minh Anh	Nguyễn	5/20/1987	…	20162101
20170001	Nhật Ánh	Nguyễn	5/15/1988	…	20172201

clazz

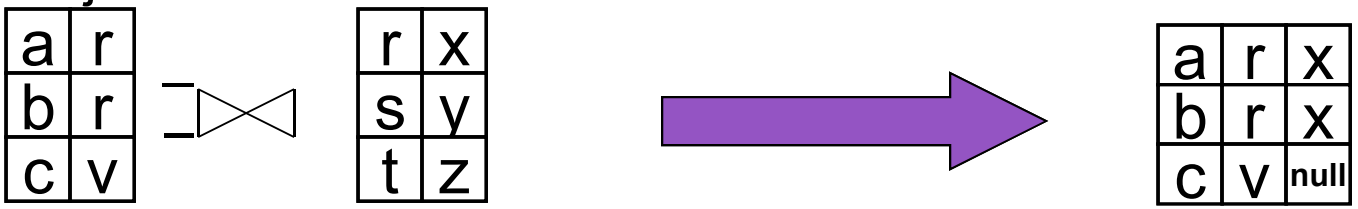
clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001
20172202	CNTT2.02-K62		

results

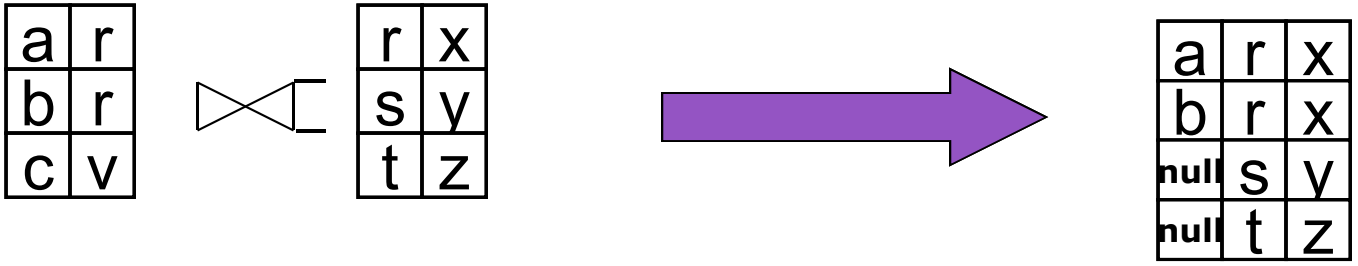
student_id	first_name	last_name	dob	…	clazz_id	Name	lecturer_id	Monitor_id
20160002	Anh	Hoàng	5/20/1987	…	20162101	CNTT1.01-K61	02001	20160003
20160003	Thu Hồng	Trần	6/6/1987	…	20162101	CNTT1.01-K61	02001	20160003
20160004	Minh Anh	Nguyễn	5/20/1987	…	20162101	CNTT1.01-K61	02001	20160003
20170001	Nhật Ánh	Nguyễn	5/15/1988	…	20172201	CNTT2.01-K62	02002	20170001

Lesson > Topic 4: Common extension

- Outer join
  - Outer join left



  - Outer join right



Lesson > Topic 4: Common extension

- **Example of left-outer join**: List all students and class i nformation if any

student

student_id	first_name	last_name	dob	…	clazz_id
20160001	Ngọc An	Bùi	3/18/1987	…	
20160002	Anh	Hoàng	5/20/1987	…	20162101
20160003	Thu Hồng	Trần	6/6/1987	…	20162101
20160004	Minh Anh	Nguyễn	5/20/1987	…	20162101
20170001	Nhật Ánh	Nguyễn	5/15/1988	…	20172201

clazz

clazz_id	name	lecturer_id	monitor_id
20162101	CNTT1.01-K61	02001	20160003
20162102	CNTT1.02-K61		
20172201	CNTT2.01-K62	02002	20170001
20172202	CNTT2.02-K62		

results

student_id	first_name	last_name	dob	…	clazz_id	Name	lecturer_id	Monitor_id
20160001	Ngọc An	Bùi	3/18/1987					
20160002	Anh	Hoàng	5/20/1987	…	20162101	CNTT1.01-K61	02001	20160003
20160003	Thu Hồng	Trần	6/6/1987	…	20162101	CNTT1.01-K61	02001	20160003
20160004	Minh Anh	Nguyễn	5/20/1987	…	20162101	CNTT1.01-K61	02001	20160003
20170001	Nhật Ánh	Nguyễn	5/15/1988	…	20172201	CNTT2.01-K62	02002	20170001

Lesson > Topic 4: Common extension

- Aggregation

$$G_1, G_2, \ldots G_n \; \mathcal{G} \; F_1(A_1), F_2(A_2), \ldots F_n(A_n) \; (\mathcal{R})$$

$G_1, G_2, \ldots G_n$ is a list of attributes on which to group
$F_1(A_1), F_2(A_2), \ldots F_n(A_n)$ is a list of aggregation function on a ttribute $A_1, A_2, .. A_n$

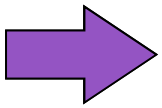
Lesson > Topic 4: Common extension

- Example of Aggregation

$$G_{student\_id}\ \mathcal{G}\ count(subject\_id)\ (enrollment\ )$$

enrollment

student_id	subject_id	semester	midterm_ score	final_ score
20160001	IT1110	20171	9	8.5
20160001	IT3080	20172	8	
20160001	IT3090	20172	6	9
20160001	IT4857	20172	7.5	9
20160001	IT4866	20172	7	9
20160002	IT3080	20172	9	
20160003	IT1110	20171	7	6
20160004	IT1110	20171	6	5



results

student_id	count(subject_id)
20160001	5
20160002	1
20160003	1
20160004	1

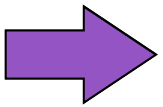
Lesson > Topic 4: Common extension

- Example of Aggregation

$$\mathcal{G}_{\text{count(student\_id)}}(\text{student})$$

student

student_id	first_name	last_name	dob	…	clazz_id
20160001	Ngọc An	Bùi	3/18/1987	…	
20160002	Anh	Hoàng	5/20/1987	…	20162101
20160003	Thu Hồng	Trần	6/6/1987	…	20162101
20160004	Minh Anh	Nguyễn	5/20/1987	…	20162101
20170001	Nhật Ánh	Nguyễn	5/15/1988	…	20172201



value not a relation

5

Remarks

- Relational algebra
  - operators
  - expressions
- Set operators
- Relational algebraic operators
- Common extension : not standard

Quiz

No	Question (Multiple Choice)	Answer (1,2,3,4)	Commentary
1	Unary operator has 2 input relations A. True B. False	B	
2	The result of a selection is A. A relation B. A value	A	
3	The result of an aggregation can be A. A relation B. A value C. Both	C	
4	The result of an algebraic expression can be A. A relation B. A value C. Both	C	

Outro > Summary

No	Topic	Summary
1	Introduction to relationa l algrebra	- Procedural langue
2	Set operators	- Union, intersection, difference,
3	Relational operators:	- Projection, Selection, Rename, Join
4	Common extensions	Natural join, Outer join, Aggregation

You've just have an overview of databases

Next lesson:

Structured Query Language

1. History
2. Syntax