HW1 is due Wed Sep 13\textsuperscript{th}.

We will release the first project on Mon Sep 11\textsuperscript{th}. It is due Wed Sep 27\textsuperscript{th}.
TODAY’S AGENDA

More Relational Model
Relational Algebra
Relational Calculus
Alternative Data Models
BEFORE THE RELATIONAL MODEL

Database applications were difficult to build and maintain.

Tight coupling between logical and physical layers.

You have to (roughly) know what queries your app would execute before you deployed the database.
BEFORE THE RELATIONAL MODEL

Database applications were difficult to build and maintain.

Tight coupling between logical and physical layers.

You have to (roughly) know what queries your app would execute before you deployed the database.

Edgar F. Codd
A Relational Model of Data for Large Shared Data Banks

E. F. Codd
IBM Research Laboratory, San Jose, California

Abstract

Users need large shared data banks not only to keep track of data but also to analyze the data and use the results. To make the data useful, they must be able to manipulate the information contained in the data. The relational model provides a powerful and general mechanism for doing this. The model also provides a basis for the implementation of the data bank as a system that can be implemented on any computer. The model can be implemented efficiently and can be extended to include other types of data, such as images and text.

1. Relational Model and Normal Form

1.1 Introduction

This paper is concerned with the application of a certain theory to systems which provide shared access to large banks of structured data. This theory, called the relational model, is based on a formalism which is being developed by an IBM research group. The theory is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on the idea that data can be represented in a structured way and that the relationships between the data can be described in terms of the structure.

The relational model is a formalism for representing data and for manipulating the data. It is based on
RELATIONAL MODEL

Proposed in 1970 by Ted Codd.

Database abstraction to avoid this maintenance:
→ Store database in simple data structures.
→ Access data through high-level language.
→ Physical storage left up to implementation.

Edgar F. Codd
A data model is collection of concepts for describing the data in a database.

A schema is a description of a particular collection of data, using a given data model.
DATA MODELS

Relational
Key/Value
Graph
Document
Column-family
Array / Matrix
Hierarchical
Network
RELATIONAL MODEL

**Structure:** The definition of relations and their contents.

**Integrity:** Ensure the database’s contents satisfy constraints.

**Manipulation:** How to access and modify a database’s contents.
RELATIONAL MODEL

A relation is an unordered set that contains the relationship of attributes that represent entities.

A tuple is a set of attribute values (also known as its domain) in the relation.

→ Values are (normally) atomic/scalar.
→ The special value **NULL** is a member of every domain.
A relation is an unordered set that contains the relationship of attributes that represent entities.

A tuple is a set of attribute values (also known as its domain) in the relation.

→ Values are (normally) atomic/scalar.
→ The special value `NULL` is a member of every domain.

### Artist(name, year, country)

<table>
<thead>
<tr>
<th>Name</th>
<th>Year</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wu Tang Clan</td>
<td>1992</td>
<td>USA</td>
</tr>
<tr>
<td>Notorious B.I.G.</td>
<td>1992</td>
<td>USA</td>
</tr>
<tr>
<td>Ice Cube</td>
<td>1989</td>
<td>USA</td>
</tr>
</tbody>
</table>

\[\text{n-ary Relation} = \text{Table with } n \text{ columns}\]
Every relation has at least one candidate key that uniquely identifies a single tuple.

Candidate keys:
→ (name)
→ (name, year)
→ (name, country)
→ (name, year, country)
RELATIONAL MODEL: CANDIDATE KEYS

Every relation has at least one candidate key that uniquely identifies a single tuple.

Candidate keys:

→ (name)
→ (name, year)
→ (name, country)
→ (name, year, country)

<table>
<thead>
<tr>
<th>name</th>
<th>year</th>
<th>country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wu Tang Clan</td>
<td>1992</td>
<td>USA</td>
</tr>
<tr>
<td>Notorious B.I.G.</td>
<td>1992</td>
<td>USA</td>
</tr>
<tr>
<td>Ice Cube</td>
<td>1989</td>
<td>USA</td>
</tr>
<tr>
<td>Ice Cube</td>
<td>2017</td>
<td>India</td>
</tr>
</tbody>
</table>
RELATIONAL MODEL: CANDIDATE KEYS

Every relation has at least one candidate key that uniquely identifies a single tuple.

Candidate keys:
→ (name)
→ (name, year)
→ (name, country)
→ (name, year, country)

<table>
<thead>
<tr>
<th>Artist (name, year, country)</th>
</tr>
</thead>
<tbody>
<tr>
<td>name</td>
</tr>
<tr>
<td>Wu Tang Clan</td>
</tr>
<tr>
<td>Notorious B.I.G.</td>
</tr>
<tr>
<td>Ice Cube</td>
</tr>
<tr>
<td>Ice Cube</td>
</tr>
<tr>
<td>Ice Cube</td>
</tr>
</tbody>
</table>
RELATIONAL MODEL: CANDIDATE KEYS

Every relation has at least one candidate key that uniquely identifies a single tuple.

Candidate keys:
→ (name)
→ (name, year)
→ (name, country)
→ (name, year, country)
### RELATIONAL MODEL: CANDIDATE KEYS

Every relation has at least one candidate key that uniquely identifies a single tuple.

**Candidate keys:**
- \((id, name, year, country)\)
- \((id)\)
- \((id, name)\)
- \((id, year)\)
- \((id, country)\)
- *All other combinations...*

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
<th>year</th>
<th>country</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>Wu Tang Clan</td>
<td>1992</td>
<td>USA</td>
</tr>
<tr>
<td>102</td>
<td>Notorious B.I.G.</td>
<td>1992</td>
<td>USA</td>
</tr>
<tr>
<td>103</td>
<td>Ice Cube</td>
<td>1989</td>
<td>USA</td>
</tr>
<tr>
<td>104</td>
<td>Ice Cube</td>
<td>2017</td>
<td>India</td>
</tr>
<tr>
<td>105</td>
<td>Ice Cube</td>
<td>2017</td>
<td>USA</td>
</tr>
<tr>
<td>106</td>
<td>Ice Cube</td>
<td>2017</td>
<td>USA</td>
</tr>
</tbody>
</table>
RELATIONAL MODEL: PRIMARY KEY

A relation’s primary key is a candidate key that is deemed more “important” than other candidate keys.

Some DBMSs automatically create an internal primary key for a table if you do not define one.
A foreign key specifies that an attribute from one relation has to map to a tuple in another relation.
RELATIONAL MODEL: FOREIGN KEYS

A foreign key specifies that an attribute from one relation has to map to a tuple in another relation.

**Artist(id, name, year, country)**

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
<th>year</th>
<th>country</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>Wu Tang Clan</td>
<td>1992</td>
<td>USA</td>
</tr>
<tr>
<td>102</td>
<td>Notorious B.I.G.</td>
<td>1992</td>
<td>USA</td>
</tr>
<tr>
<td>103</td>
<td>Ice Cube</td>
<td>1989</td>
<td>USA</td>
</tr>
</tbody>
</table>

**ArtistAlbum(artist_id, album_id)**

<table>
<thead>
<tr>
<th>artist_id</th>
<th>album_id</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>11</td>
</tr>
<tr>
<td>101</td>
<td>22</td>
</tr>
<tr>
<td>103</td>
<td>22</td>
</tr>
</tbody>
</table>

**Album(id, name, year)**

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
<th>year</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Enter the Wu Tang</td>
<td>1993</td>
</tr>
<tr>
<td>22</td>
<td>St.Ides Mix Tape</td>
<td>1994</td>
</tr>
</tbody>
</table>
DATA MANIPULATION LANGUAGES (DML)

How to store and retrieve information from a database.

**Procedural:**
→ The query specifies the (high-level) strategy the DBMS should use to find the desired result.

**Non-Procedural:**
→ The query specifies only what data is wanted and not how to find it.

← Relational Algebra

← Relational Calculus
RELATIONAL ALGEBRA

Fundamental operations to retrieve and manipulate tuples in a relation.
→ Based on set algebra.

Each operator takes one or more relations as its inputs and outputs a new relation.
→ We can “chain” operators together to create more complex operations.
**RELATIONAL ALGEBRA:**

**SELECT**

Choose a subset of the tuples from a relation that satisfies a selection predicate.

→ Predicate acts as a filter to retain only tuples that fulfill its qualifying requirement.

→ Can combine multiple predicates using conjunctions / disjunctions.

**Syntax:** $\sigma_{\text{predicate}}(R)$

```
SELECT * FROM R
WHERE a_id='a2' AND b_id>102;
```
RELATIONAL ALGEBRA: PROJECTION

Generate a relation with tuples that contains only the specified attributes.
→ Can rearrange attributes’ ordering.
→ Can manipulate the values.

Syntax: $\pi_{A_1,A_2,...,A_n}(R)$
RELATIONAL ALGEBRA: UNION

Generate a relation that contains all tuples that appear in either only one or both of the input relations.

Syntax: \((R \cup S)\)
RELATIONAL ALGEBRA: INTERSECTION

Generate a relation that contains only the tuples that appear in both of the input relations.

Syntax: \((R \cap S)\)
RELATIONAL ALGEBRA: DIFFERENCE

Generate a relation that contains only the tuples that appear in the first and not the second of the input relations.

Syntax: \((R - S)\)
RELATIONAL ALGEBRA: PRODUCT

Generate a relation that contains all possible combinations of tuples from the input relations.

Syntax: \((R \times S)\)

\[
\text{SELECT * FROM R CROSS JOIN S;}
\]

\[
\text{SELECT * FROM R, S;}
\]
RELATIONAL ALGEBRA: JOIN

Generate a relation that contains all tuples that are a combination of two tuples (one from each input relation) with a common value(s) for one or more attributes.

Syntax: \((R \bowtie S)\)

```
SELECT * FROM R NATURAL JOIN S;
```
RELATIONAL ALGEBRA: JOIN TYPES

Cross Join
→ Same thing as Cartesian product

Inner Join
→ Each tuple in the first relation must have a corresponding match in the second relation.

Outer Join
→ Each tuple in one relation does not need to have a corresponding match in the other relation.
RELATIONAL ALGEBRA: INNER JOINS (1)

**Natural Join** \((R \bowtie S)\)

→ Match tuples in \(R\) and \(S\) where the shared attributes are equivalent.

**Theta / Equi Join** \((R \bowtie_\theta S)\)

→ Match tuples using some arbitrary join predicate defined by \(\theta\).
→ If \(\theta\) is just an equality predicate, then it is called an "Equi Join".
**RELATIONAL ALGEBRA: INNER JOINS (1)**

**Natural Join (R⋈S)**

→ Match tuples in R and S where the shared attributes are equivalent.

** Theta / Equi Join (R⋈θS)**

→ Match tuples using some arbitrary join predicate defined by θ.
→ If θ is just an equality predicate, then it is called an "Equi Join".

```
SELECT * FROM R INNER JOIN S 
ON R.a_id = S.a_id 
AND R.b_id = S.b_id;
```
RELATIONAL ALGEBRA: INNER JOINS (1)

Natural Join \((R \bowtie S)\)

→ Match tuples in \(R\) and \(S\) where the shared attributes are equivalent.

Theta / Equi Join \((R \bowtie_\theta S)\)

→ Match tuples using some arbitrary join predicate defined by \(\theta\).

→ If \(\theta\) is just an equality predicate, then it is called an "Equi Join".

\[
\begin{array}{|c|c|}
\hline
a_id & b_id \\
\hline
a3 & 103 \\
\hline
\end{array}
\]

\[
\begin{array}{|c|c|c|c|c|}
\hline
R.a_id & R.b_id & S.a_id & S.b_id \\
\hline
a3 & 103 & a3 & 103 \\
\hline
\end{array}
\]

```sql
SELECT * FROM R INNER JOIN S
ON R.a_id = S.a_id
AND R.b_id = S.b_id;
```
RELATIONAL ALGEBRA: INNER JOINS (1)

Natural Join \((R \bowtie S)\)
→ Match tuples in \(R\) and \(S\) where the shared attributes are equivalent.

Theta / Equi Join \((R \bowtie_\theta S)\)
→ Match tuples using some arbitrary join predicate defined by \(\theta\).
→ If \(\theta\) is just an equality predicate, then it is called an "Equi Join".

\[
\begin{array}{|c|c|}
\hline
a_id & b_id \\
\hline
a3 & 103 \\
\hline
\end{array}
\]

\[
\begin{array}{|c|c|c|c|}
\hline
R.a_id & R.b_id & S.a_id & S.b_id \\
\hline
a3 & 103 & a3 & 103 \\
\hline
\end{array}
\]

\[
\begin{align*}
\text{SELECT} & \quad \text{\textast} \\ & \quad \text{FROM} \\ R & \quad \text{INNER JOIN} \\ S
\end{align*}
\]

\[
\begin{align*}
\text{SELECT} & \quad \text{\textast} \\ & \quad \text{FROM} \\ R & \quad \text{JOIN} \\ S
\end{align*}
\]

\[
\begin{align*}
\text{SELECT} & \quad \text{\textast} \\ & \quad \text{FROM} \\ R, S \\ \text{WHERE} & \quad R.a_id = S.a_id \\ & \quad \text{AND} \\ & \quad R.b_id = S.b_id;
\end{align*}
\]
RELATIONAL ALGEBRA: INNER JOINS (2)

Semi Join \((R \bowtie S)\)

→ Generate relation that contains tuples of \(R\) that match with tuples in \(S\).

→ Same as natural join except that output relation only contains tuples from the first relation.

Anti Join \((R \triangleright S)\)

→ Generate relation that contains tuples of \(R\) that do not match with any tuple in \(S\).
RELATIONAL ALGEBRA: OUTER JOINS

Left Outer Join \((R \bowtie S)\)
→ Generate all combinations of tuples in \(R\) and \(S\) that are equal on their shared attributes, in addition to tuples in \(R\) that have no matching tuples in \(S\).

Right Outer Join \((R \bowtie S)\)
→ Same as LEFT OUTER but with the input relations reversed.

Full Outer Join \((R \bowtie S)\)
→ Union of LEFT OUTER and RIGHT OUTER.

<table>
<thead>
<tr>
<th>R.a_id</th>
<th>R.b_id</th>
<th>S.a_id</th>
<th>S.b_id</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>101</td>
<td>NULL</td>
<td>NULL</td>
</tr>
<tr>
<td>a2</td>
<td>102</td>
<td>NULL</td>
<td>NULL</td>
</tr>
<tr>
<td>a3</td>
<td>103</td>
<td>a3</td>
<td>103</td>
</tr>
</tbody>
</table>

```
SELECT *
FROM R LEFT OUTER JOIN S
ON R.a_id = S.a_id
AND R.b_id = S.b_id;
```
RELATIONAL ALGEBRA: EXTRA OPERATORS

Rename ($\rho$)
Assignment ($R\leftarrow S$)
Duplicate Elimination ($\delta$)
Aggregation ($\gamma$)
Sorting ($\tau$)
Division ($R \div S$)
OBSERVATION

Relational algebra still defines the high-level steps of how to compute a query.

\[ \sigma_{b\_id=102}(R \bowtie S) \text{ vs. } (R \bowtie (\sigma_{b\_id=102}(S))) \]

A better approach is just state the high-level query you want

\[ \text{Retrieve the joined tuples from } R \text{ and } S \text{ where } b\_id \text{ equals } 102\].
RELATIONAL CALCULUS

Tuple Relational Calculus
→ Uses variables whose permitted values are tuples of that relation.

Domain Relational Calculus
→ Uses domain variables to select attributes instead of whole tuples.

Both are equivalent to relational algebra.

Relational Algebra
\[ \sigma_{a_id='a2' \land b_id>102}(R) \]

Tuple Relational Calculus
\[ \{ t \mid t \in R \land t.a_id='a2' \land t.b_id>102 \} \]

Domain Relational Calculus
\[ \{ <a_id,b_id> \mid <a_id,b_id> \in R \land a_id='a2' \land b_id>102 \} \]
DATA MODELS

Relational
Key/Value
Graph
Document
Column-family
Array / Matrix
Hierarchical
Network

← NoSQL
NON-RELATIONAL DATA MODELS: KEY-VALUE

Store records in an **associative array** that maps a **key** to a **value**.
→ Sometimes called a **dictionary** or **hash table**.

The contents of a record’s **value** is often opaque to DBMS.
→ It is up to the application to interpret its contents.
→ Some systems allow for compound values.

**Artists**: name→(year,country)

<table>
<thead>
<tr>
<th>Key</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wu Tang Clan</td>
<td>1992,&quot;USA&quot;</td>
</tr>
<tr>
<td>Notorious BIG</td>
<td>1992,&quot;USA&quot;</td>
</tr>
<tr>
<td>Ice Cube</td>
<td>1989,&quot;USA&quot;</td>
</tr>
</tbody>
</table>
NON-RELATIONAL DATA MODELS: GRAPH

Represent the database as a collection of nodes and edges.
→ Each node/edge can also be annotated with additional properties.

No standard query language:
→ SPARQL, Gremlin, Cypher
A document is a self-contained record that contains the description of its data attributes and their corresponding values.
NON-RELATIONAL DATA MODELS: DOCUMENT

A document is a self-contained record that contains the description of its data attributes and their corresponding values.

Support nesting together all of the documents for a single entity. → This is called denormalization in the relational model.
NON-RELATIONAL DATA MODELS: COLUMN-FAMILY

Hybrid data model that maps a key to a column family.

Each column family contains any number of rows that each has one or more column names and values.

<table>
<thead>
<tr>
<th>Row Key</th>
<th>year</th>
<th>country</th>
<th>city</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wu Tang Clan</td>
<td>1992</td>
<td>USA</td>
<td></td>
</tr>
<tr>
<td>Ice Cube</td>
<td>1992</td>
<td>USA</td>
<td>Los Angeles</td>
</tr>
</tbody>
</table>
CONCLUSION

Relational algebra defines the primitives for processing queries on a relational database.

We will see relational algebra again when we talk about query optimization + execution.
NEXT CLASS

Aggregations + Group By
Output Control / Redirection
Nested Queries
Common Table Expressions
Window Functions