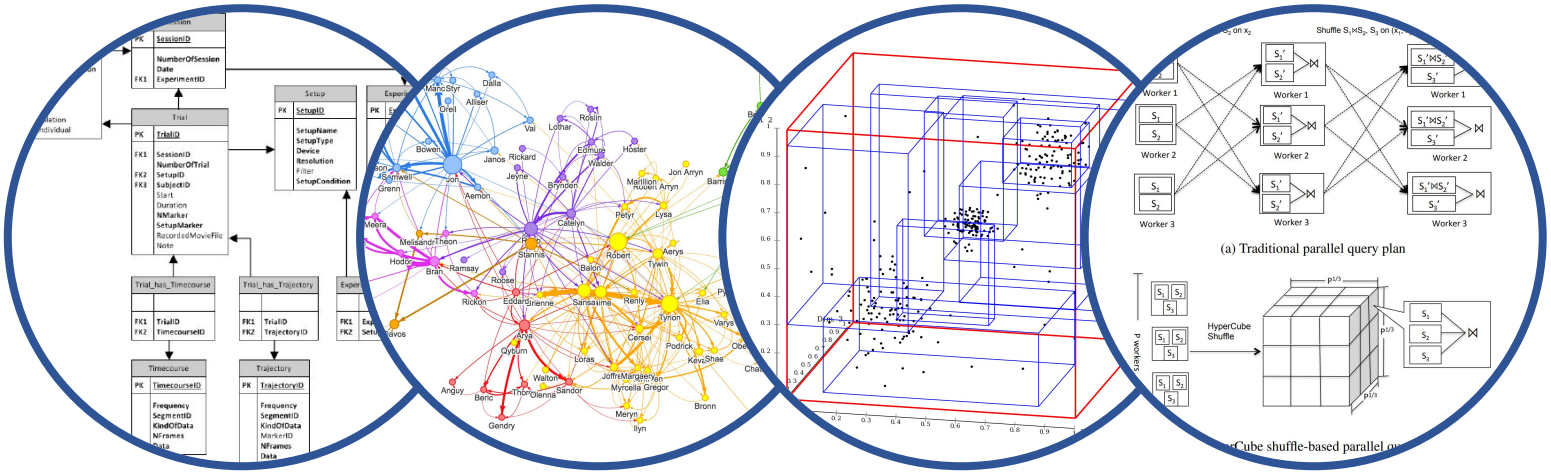


Introduction to Data Management

SQL Basics, Keys, and Joins

Rob Thompson
Based on slides by Alyssa Pittman, Jonathan Leang, Dan Suciu, et al

Paul G. Allen School of Computer Science and Engineering
University of Washington, Seattle

Recap – The Relational Model

- Flat tables, static and typed attributes, etc.
  - "It's a spreadsheet with rules"

Table/ Relation

Columns/Attributes/Fields

Rows/ Tuples/ Records

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Recap - The Relational Model

- **Set semantics**
  - No duplicate tuples
- Attributes are **typed** and **static**
  - INTEGER, FLOAT, VARCHAR(n), DATETIME, …
- Tables are **flat**

Outline

- Semantics
- Keys Identification
- Foreign Keys Relationships
- Joins in SQL
  - Inner joins
  - Outer joins
  - Self joins

Hello World

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

SELECT
What kind of data I want

FROM
Where the data coming from

WHERE
Filter the data

Hello World

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

?

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

Hello World

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Name	UserID
Jack	123
Allison	345

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

SQL semantics

- What is the meaning (semantics) of this query?
- What results can we expect to get?

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

For-each semantics

```
for each row in P:
  if (row.Job == 'TA'):
    output (row.Name, row.UserID)
```

Hello World

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Name	UserID

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

Hello World

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Name	UserID
Jack	123

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

Hello World

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Name	UserID
Jack	123

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

Hello World

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Name	UserID
Jack	123
Allison	345

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

Hello World

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Name	UserID
Jack	123
Allison	345

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

Hello World

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Name	UserID
Jack	123
Allison	345

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

Hello World

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Name	UserID
Jack	123
Allison	345

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

SQL Common Keywords

- **`SELECT`** col1, col2, ...
- **`FROM`** rel1, rel2, ...
- **`WHERE`** cond1 AND cond2 OR cond3 ...
- **`GROUP BY`** aggregate1, aggregate2, ...
- **`HAVING`** cond1 AND cond2 OR cond3 ...

SQL Common Keywords

- **SELECT** col1, col2, ...
- **FROM** rel1, rel2, ...
- **WHERE** cond1 AND cond2 OR cond3 ...
- **GROUP BY** aggregate1, aggregate2, ...
- **HAVING** cond1 AND cond2 OR cond3 ...

Aggregates next week

A little extra SQL

- **`ORDER BY`** – Orders result tuples by specified attributes (default ascending)

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA'
 ORDER BY P.Salary, P.Name;
```

- **`DISTINCT`** – Deduplicates result tuples

```
SELECT DISTINCT P.Job
  FROM Payroll AS P
 WHERE P.Salary > 70000;
```

Create Table Statement

Payroll(UserId, Name, Job, Salary)

```
CREATE TABLE Payroll (
 UserId INT,
 Name VARCHAR(100),
 Job VARCHAR(100),
 Salary INT
);
```

* Case insensitive, but useful for readability

Insert Statement

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
789	Dan	Prof	100000

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
789	Dan	Prof	100000
912	Alyssa	Lecturer	45000

```
INSERT INTO
   Payroll(UserId, Name, Job, Salary)
VALUES(912, 'Alyssa', 'Lecturer', 45000);
```

Insert Statement

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
789	Dan	Prof	100000

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
789	Dan	Prof	100000
912	Alyssa	Lecturer	45000

Can omit column names if inserting to all columns in order

```
INSERT INTO
   Payroll(UserId, Name, Job, Salary)
VALUES(912, 'Alyssa', 'Lecturer', 45000);
```

Insert Statement

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
789	Dan	Prof	100000

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
789	Dan	Prof	100000
912	Alyssa	Lecturer	45000

Can omit column names if inserting to all columns in order

```
INSERT INTO
   Payroll
VALUES(912, 'Alyssa', 'Lecturer', 45000);
```

Keys

Key

A **Key** is one or more attributes that uniquely identify a row.

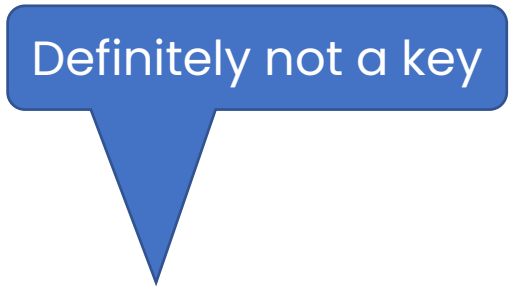
Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Keys

Key

A **Key** is one or more attributes that uniquely identify a row.



UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Keys

Key

A **Key** is one or more attributes that uniquely identify a row.

Good candidate for a key

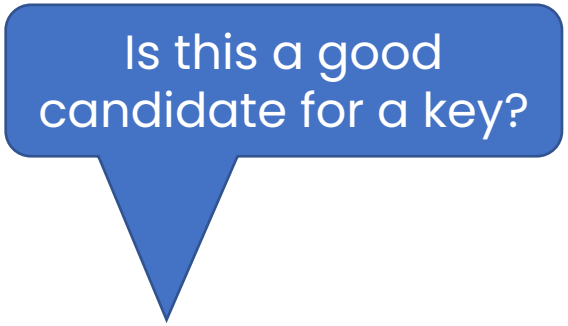
Definitely not a key

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Keys

Key

A **Key** is one or more attributes that uniquely identify a row.



UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Keys

Key

A **Key** is one or more attributes that uniquely identify a row.

Is this a good candidate for a key?

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Keys

Key

A **Key** is one or more attributes that uniquely identify a row.

Is this a good candidate for a key?

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000
913	Peter	TA	60000

Keys

Key

A **Key** is one or more attributes that uniquely identify a row.

Data comes from the real world so models ought to reflect that

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000
913	Peter	TA	60000

Keys

```
CREATE TABLE Payroll (
  UserID INT,
  Name VARCHAR(100),
  Job VARCHAR(100),
  Salary INT);
```

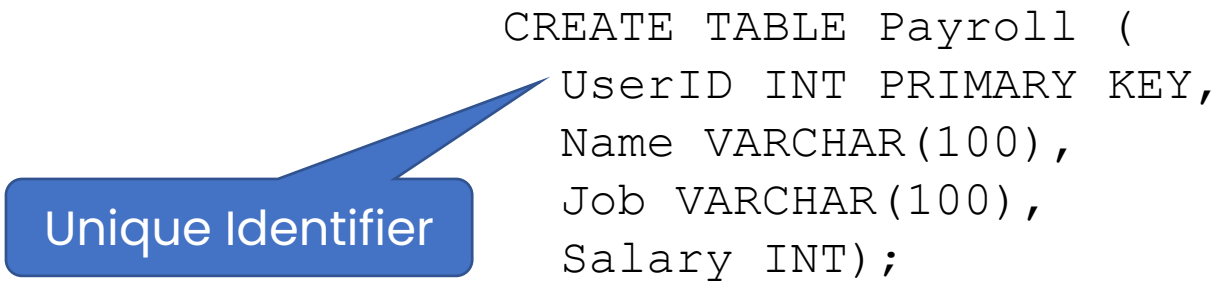
Payroll(UserId, Name, Job, Salary)

Keys


```
CREATE TABLE Payroll (
  UserID INT,
  Name VARCHAR(100),
  Job VARCHAR(100),
  Salary INT);
```

Payroll(UserId, Name, Job, Salary)

Keys




```
CREATE TABLE Payroll (
  UserID INT PRIMARY KEY,
  Name VARCHAR(100),
  Job VARCHAR(100),
  Salary INT);
```

Payroll(<u>UserId</u>, Name, Job, Salary)

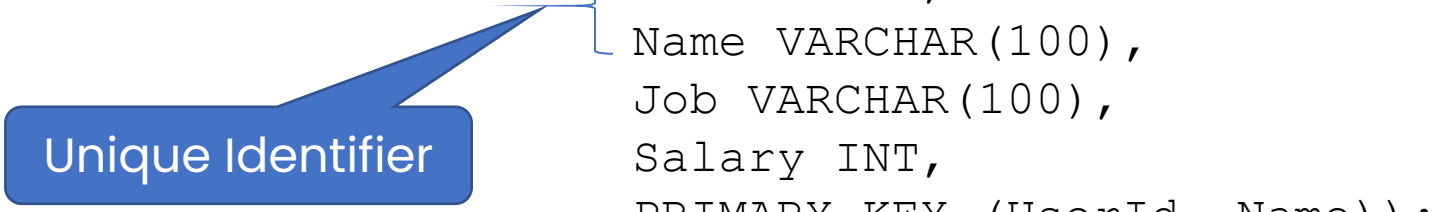
Keys

```
CREATE TABLE Payroll (
  UserID INT,
  Name VARCHAR(100),
  Job VARCHAR(100),
  Salary INT);
```

Unique Identifier

Payroll(UserId, Name, Job, Salary)

Keys




```
CREATE TABLE Payroll (
  UserID INT,
  Name VARCHAR(100),
  Job VARCHAR(100),
  Salary INT,
  PRIMARY KEY (UserId, Name));
```

Payroll(<u>UserId</u>, <u>Name</u>, Job, Salary)

Foreign Keys

- Databases can hold multiple tables
- How do we capture relationships *between* tables?

Payroll

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Regist

UserID	Car
123	Charger
567	Civic
567	Pinto

Foreign Keys

- Databases can hold multiple tables
- How do we capture relationships *between* tables?

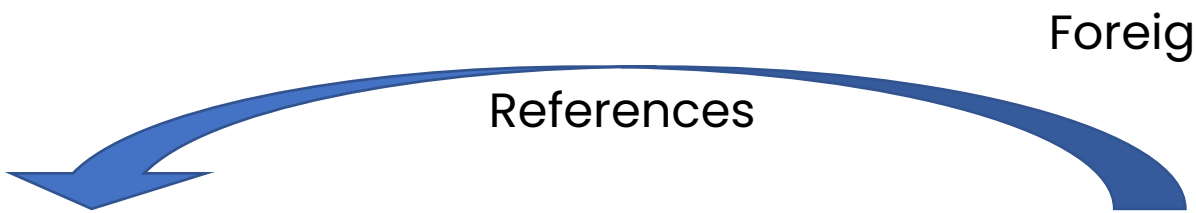
UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

Foreign Key

UserID	Car
123	Charger
567	Civic
567	Pinto

Foreign Keys

- Databases can hold multiple tables
- How do we capture relationships *between* tables?



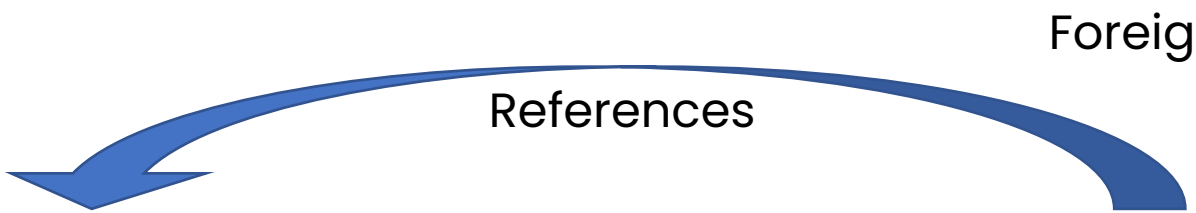
UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Foreign Keys

Foreign Key

A **Foreign Key** is one or more attributes that uniquely identify a row in *another table*.



UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Foreign Keys

Foreign Key

A **Foreign Key** is one or more attributes that uniquely identify a row in *another table*.

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Foreign Keys

Foreign Key

A **Foreign Key** is one or more attributes that uniquely identify a row in *another table*.

Is this valid?
Nope

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Foreign Keys

```
CREATE TABLE Payroll (
  UserID INT PRIMARY KEY,
  Name VARCHAR(100),
  Job VARCHAR(100),
  Salary INT);
```

```
CREATE TABLE Regist (
  UserID INT,
  Car VARCHAR(100));
```

Payroll(<u>UserId</u>, Name, Job, Salary)

Regist(UserId, Car)

Foreign Keys

```
CREATE TABLE Payroll (
  UserID INT PRIMARY KEY,
  Name VARCHAR(100),
  Job VARCHAR(100),
  Salary INT);
```

```
CREATE TABLE Regist (
  UserID INT
    REFERENCES Payroll(UserID),
  Car VARCHAR(100));
```

Payroll(<u>UserId</u>, Name, Job, Salary)

Regist(UserId, Car)

Foreign Keys

```
CREATE TABLE Payroll (
  UserID INT PRIMARY KEY,
  Name VARCHAR(100),
  Job VARCHAR(100),
  Salary INT);
```

```
CREATE TABLE Regist (
  UserID INT
    REFERENCES Payroll(UserID),
  Car VARCHAR(100));
```

Payroll(<u>UserId</u>, Name, Job, Salary)

Regist(UserId, Car)

The Relational Model Revisited

- More complete overview of the Relational Model:
  - Database □ collection of tables
  - All tables are flat
  - Keys uniquely ID rows
  - Foreign keys act as a “semantic pointer”
  - **Physical data independence**

Joins

- Foreign keys are able to *describe* a relationship between tables
- Joins are able to *realize* combinations of data

Inner Joins

- Bread and butter of SQL queries
  - “Inner join” is often interchangeable with just “join”

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

```
SELECT P.Name, R.Car
  FROM Payroll AS P JOIN Regist AS R
       ON P.UserID = R.UserID;
```

How do we algorithmically get our results?

Name	Car
Jack	Charger
Magda	Civic
Magda	Pinto

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

```
SELECT P.Name, R.Car
  FROM Payroll AS P JOIN Regist AS R
       ON P.UserID = R.UserID;
```

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger
Magda	Civic

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger
Magda	Civic

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger
Magda	Civic
Magda	Pinto

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger
Magda	Civic
Magda	Pinto

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger
Magda	Civic
Magda	Pinto

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger
Magda	Civic
Magda	Pinto

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Nested-Loop Semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger
Magda	Civic
Magda	Pinto

```
for each row1 in Payroll:
   for each row2 in Regist:
      if (row1.UserID = row2.UserID):
         output (row1.Name, row2.Car)
```

Inner Joins

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Explicit

```
SELECT P.Name, R.Car
  FROM Payroll AS P JOIN Regist AS R
       ON P.UserID = R.UserID;
```

Implicit

```
SELECT P.Name, R.Car
  FROM Payroll AS P, Regist AS R
 WHERE P.UserID = R.UserID;
```

Outer Joins

- LEFT OUTER JOIN
  - All rows in left table are preserved
- RIGHT OUTER JOIN
  - All rows in right table are preserved
- FULL OUTER JOIN
  - All rows are preserved

Outer Joins

Now I want to include everyone, even if they don't drive.

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Outer Joins

Now I want to include everyone, even if they don't drive.

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

```
SELECT P.Name, R.Car
  FROM Payroll AS P LEFT OUTER JOIN Regist AS R
       ON P.UserID = R.UserID;
```

Outer Joins

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger
Allison	NULL
Magda	Civic
Magda	Pinto
Dan	NULL

Outer Joins

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

Name	Car
Jack	Charger
Allison	NULL
Magda	Civic
Magda	Pinto
Dan	NULL

NULL is a value placeholder. Depending on context, it may mean unknown, not applicable, etc.

Self Joins

Find all people who drive a Civic and Pinto

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

```
SELECT P.Name, R.Car
  FROM Payroll AS P, Regist AS R
 WHERE P.UserID = R.UserID AND
       R.Car = 'Civic';
```

Self Joins

Find all people who drive a Civic and Pinto

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

```
SELECT P.Name, R.Car
  FROM Payroll AS P, Regist AS R
 WHERE P.UserID = R.UserID AND
       R.Car = 'Civic' AND
       R.Car = 'Pinto';
```

Will this work?

Self Joins

Find all people who drive a Civic and Pinto

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

```
SELECT P.Name, R.Car
  FROM Payroll AS P, Regist AS R
 WHERE P.UserID = R.UserID AND
       R.Car = 'Civic' AND
       R.Car = 'Pinto';
```

Will this work?

Nope, empty set is returned

Self Joins

Find all people who drive a Civic and Pinto

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

```
SELECT P.Name, R.Car
  FROM Payroll AS P, Regist AS R
 WHERE P.UserID = R.UserID AND
       (R.Car = 'Civic' OR
       R.Car = 'Pinto');
```

Will this work?

Self Joins

Find all people who drive a Civic and Pinto

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto
789	Civic

```
SELECT P.Name, R.Car
  FROM Payroll AS P, Regist AS R
 WHERE P.UserID = R.UserID AND
       (R.Car = 'Civic' OR
       R.Car = 'Pinto');
```

Will this work?
Nope, returns people who had just one type of car

Self Joins

Find all people who drive a Civic and Pinto

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

```
SELECT P.Name, R1.Car, R2.Car
  FROM Payroll AS P, Regist AS R1, Regist AS R2
 WHERE P.UserID = R1.UserID AND
       P.UserID = R2.UserID AND
       R1.Car = 'Civic' AND
       R2.Car = 'Pinto';
```

Self Joins

Find all people who drive a Civic and Pinto

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

UserID	Car
123	Charger
567	Civic
567	Pinto

All pairs of cars a person can drive

```
SELECT P.Name, R1.Car, R2.Car
  FROM Payroll AS P, Regist AS R1, Regist AS R2
 WHERE P.UserID = R1.UserID AND
       P.UserID = R2.UserID AND
       R1.Car = 'Civic' AND
       R2.Car = 'Pinto';
```

Today's Takeaways!

- The **Relational Model** concept
- How a basic **SELECT-FROM-WHERE** query works
- Basic for-each semantics

UserID	Name	Job	Salary
123	Jack	TA	50000
345	Allison	TA	60000
567	Magda	Prof	90000
789	Dan	Prof	100000

```
SELECT P.Name, P.UserID
  FROM Payroll AS P
 WHERE P.Job = 'TA';
```

Name	UserID
Jack	123
Allison	345

Takeaways

- We can describe relationships between tables with keys and foreign keys
- Different joining techniques can be used to achieve particular goals

- Our SQL toolbox is growing!
  - Not just reading and filtering data anymore
  - Starting to answer complex questions