

CITS1402 Relational Database Management Systems

Week 2—Relational Model and Relational Algebra

CITS1402

Week1

Contents

Introduction

Databases

Relational Model

How to Program

Show the terminal

```
nasi0029@C02FW0D3ML85 DreamHome % sqlite3 DreamHomeRental.db
SQLite version 3.39.5 2022-10-14 20:58:05
Enter ".help" for usage hints.
sqlite> █
```

```
sqlite> .tables
Branch                PrivateOwner          Registration          StaffPropCnt
Client                PropertyForRent        Staff                  Viewing
sqlite> █
```

```
sqlite> SELECT *
...> FROM Branch;
B005|22 Deer Rd|London|SW1 4EH
B007|16 Argyll|Aberdeen|AB2 3SU
B003|163 Main St|Glasgow|G11 9QX
B100|32 Manse Rd|Bristol|BS99 1NZ
B002|56 Clover Dr|London|NW10 6EU
sqlite> █
```

```
sqlite> .headers on ←
sqlite> SELECT *
...> FROM Branch;
branchNo|street|city|postcode
B005|22 Deer Rd|London|SW1 4EH
B007|16 Argyll|Aberdeen|AB2 3SU
B003|163 Main St|Glasgow|G11 9QX
B100|32 Manse Rd|Bristol|BS99 1NZ
B002|56 Clover Dr|London|NW10 6EU
... █
```

Branch

branchNo	street	city	postCode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

Are these rows/records unique?

Yes. But why?

Does the order matter?

Staff

Can staffNo be NULL?

Why are the column names unique?

staffNo	fName	lName	position	sex	DOB	salary	branchNo
SL21	John	White	Manager	M	1-Oct-45	30000	B005
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003
SG14	David	Ford	Supervisor	M	24-Mar-58	18000	B003
SA9	Mary	Howe	Assistant	F	19-Feb-70	9000	B007
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003
SL41	Julie	Lee	Assistant	F	13-Jun-65	9000	B005

Does the order mater?

What values can “salary” have?

```
select fname, lname, city
from Staff Natural Join Branch
where position = 'Manager'
```

How do we “know” this query
will return all records?

Agenda

- ✓ **Connection between mathematical relations and relations in the relational model.**
- ✓ **How tables are used to represent data.**
- ✓ **Properties of database relations.**
- ✓ **Terminology of relational model.**
- ✓ **How to identify CK, PK, and FKs.**
- ✓ **Meaning of entity integrity and referential integrity.**

Agenda

- ✓ **Connection between mathematical relations and relations in the relational model.**
- ✓ How tables are used to represent data.
- ✓ Properties of database relations.
- ✓ Terminology of relational model.
- ✓ How to identify CK, PK, and FKs.
- ✓ Meaning of entity integrity and referential integrity.

Mathematical Definition of Relation

Consider two **sets**, D_1 & D_2 , where

$$D_1 = \{2, 4\} \text{ and } D_2 = \{1, 3, 5\}.$$

Cartesian Product: $D_1 \times D_2$

set of all ordered pairs,

first element is member of D_1 and

second element is member of D_2 .

$$D_1 \times D_2 = \{(2, 1), (2, 3), (2, 5), (4, 1), (4, 3), (4, 5)\}$$

Mathematical Definition of Relation

- ◆ Any **subset** of Cartesian product is a **relation, R**;

$$D_1 \times D_2 = \{(2, 1), (2, 3), (2, 5), (4, 1), (4, 3), (4, 5)\}$$

$$R = \{(2, 1), (4, 1)\}$$

May specify which pairs are in **relation** using some condition for selection; e.g.
second element is 1:

$$R = \{(x, y) \mid x \in D_1, y \in D_2, \text{ and } y = 1\}$$

first element is always twice the second:

$$S = \{(x, y) \mid x \in D_1, y \in D_2, \text{ and } x = 2y\}$$

$$S = \{(2, 1)\}$$

Mathematical Definition of Relation

Consider three sets D_1, D_2, D_3 with Cartesian Product $D_1 \times D_2 \times D_3$; e.g.

$$D_1 = \{1, 3\} \qquad D_2 = \{2, 4\} \qquad D_3 = \{5, 6\}$$

$$D_1 \times D_2 \times D_3 = \{(1,2,5), (1,2,6), (1,4,5), (1,4,6), \\ (3,2,5), (3,2,6), (3,4,5), (3,4,6)\}$$

Any **subset** of these ordered triples is a **relation**.

Mathematical Definition of Relation

Cartesian product of n sets (D_1, D_2, \dots, D_n) is:

$$D_1 \times D_2 \times \dots \times D_n = \{(d_1, d_2, \dots, d_n) \mid d_1 \in D_1, d_2 \in D_2, \dots, d_n \in D_n\}$$

usually written as:

$$\prod_{i=1}^n$$

D_i

Any set of n -tuples from this Cartesian Product is a relation on the n sets.

Agenda

- ✓ **Connection between mathematical relations and relations in the relational model.**
- ✓ **How tables are used to represent data.**
- ✓ Properties of database relations.
- ✓ Terminology of relational model.
- ✓ How to identify CK, PK, and FKs.
- ✓ Meaning of entity integrity and referential integrity.

So...how does this apply to a database

Branch

branchNo	street	city	postCode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

Staff

staffNo	fName	lName	position	sex	DOB	salary	branchNo
SL21	John	White	Manager	M	1-Oct-45	30000	B005
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003
SG14	David	Ford	Supervisor	M	24-Mar-58	18000	B003
SA9	Mary	Howe	Assistant	F	19-Feb-70	9000	B007
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003
SL41	Julie	Lee	Assistant	F	13-Jun-65	9000	B005

So...how does this apply to a database?

Let A_1, A_2, \dots, A_n , be **attributes** (columns) with **domains** (possible values) D_1, D_2, \dots, D_n

The set $\{A_1:D_1, A_2:D_2, \dots, A_n:D_n\}$ is a **relation schema**

Relation R is the mapping from attribute to domain
 $(A_1:d_1, A_2:d_2, \dots, A_n:d_n)$

We can think of a relation in the relational model as any subset of the **Cartesian Product** of the **domains** of the **attributes**

How does this apply to a database?

The set $\{A_1:D_1, A_2:D_2, \dots, A_n:D_n\}$ is a **relation schema**

{branchNo: BranchNumbers, street: StreetNumbers,
city: CityNames, postcode: Postcodes}

Attribute	Domain Name	Meaning	Domain Definition
branchNo	BranchNumbers	The set of all possible branch numbers	character: size 4, range B001–B999
street	StreetNames	The set of all street names in Britain	character: size 25
city	CityNames	The set of all city names in Britain	character: size 15
postcode	Postcodes	The set of all postcodes in Britain	character: size 8
sex	Sex	The sex of a person	character: size 1, value M or F
DOB	DatesOfBirth	Possible values of staff birth dates	date, range from 1-Jan-20, format dd-mmm-yy
salary	Salaries	Possible values of staff salaries	monetary: 7 digits, range 6000.00–40000.00

How does this apply to a database?

Relation R is the mapping from attribute to domain

$(A_1:d_1, A_2:d_2, \dots, A_n:d_n)$

{(branchNo: B005, street: 22 Deer Rd
city: London, postcode: SW1 4EH)}

A **relation instance**

A **table** is a physical representation of a relation

Branch

branchNo	street	city	postCode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

Database Relations

Relation schema

Named relation defined by a set of **attribute** and **domain** name pairs.

Relational database schema

Set of relation schemas,
each with a distinct name.

$R = \{R_1, R_2, \dots, R_m\}$

DreamHomeSchema = {Staff, Branch, ...}

Agenda

- ✓ **Connection between mathematical relations and relations in the relational model.**
- ✓ **How tables are used to represent data.**
- ✓ **Properties of database relations.**
- ✓ Terminology of relational model.
- ✓ How to identify CK, PK, and FKs.
- ✓ Meaning of entity integrity and referential integrity.

Properties of Relations

Relation name is **distinct** from all other relation names in relational schema.

Each cell of relation contains **exactly one** atomic (single) value.

Each attribute has a **distinct** name.

Values of an attribute are all from the same **domain**.

Properties of Relations



Each tuple is distinct; there are no duplicate tuples.



Order of attributes has no significance.



Order of tuples has no significance, theoretically.



Why is this so exciting?

Properties of Relations

Each tuple is distinct; there are no duplicate tuples.

Order of attributes has no significance.

Order of tuples has no significance, theoretically.

Why is this so exciting?

We can apply set operations to relations

Agenda

- ✓ **Connection between mathematical relations and relations in the relational model.**
- ✓ **How tables are used to represent data.**
- ✓ **Properties of database relations.**
- ✓ **Terminology of relational model.**
- ✓ **How to identify CK, PK, and FKs.**
- ✓ **Meaning of entity integrity and referential integrity.**

Relational Model Terminology

A **relation** is a **table** with **columns** and **rows**.

Only applies to logical structure of the database, not the physical structure.

Attribute is a named **column** of a relation.

Domain is the set of allowable values for one or more **attributes**.

Examples of Attribute Domains

Attribute	Domain Name	Meaning	Domain Definition
branchNo	BranchNumbers	The set of all possible branch numbers	character: size 4, range B001–B999
street	StreetNames	The set of all street names in Britain	character: size 25
city	CityNames	The set of all city names in Britain	character: size 15
postcode	Postcodes	The set of all postcodes in Britain	character: size 8
sex	Sex	The sex of a person	character: size 1, value M or F
DOB	DatesOfBirth	Possible values of staff birth dates	date, range from 1-Jan-20, format dd-mmm-yy
salary	Salaries	Possible values of staff salaries	monetary: 7 digits, range 6000.00–40000.00

Relational Model Terminology

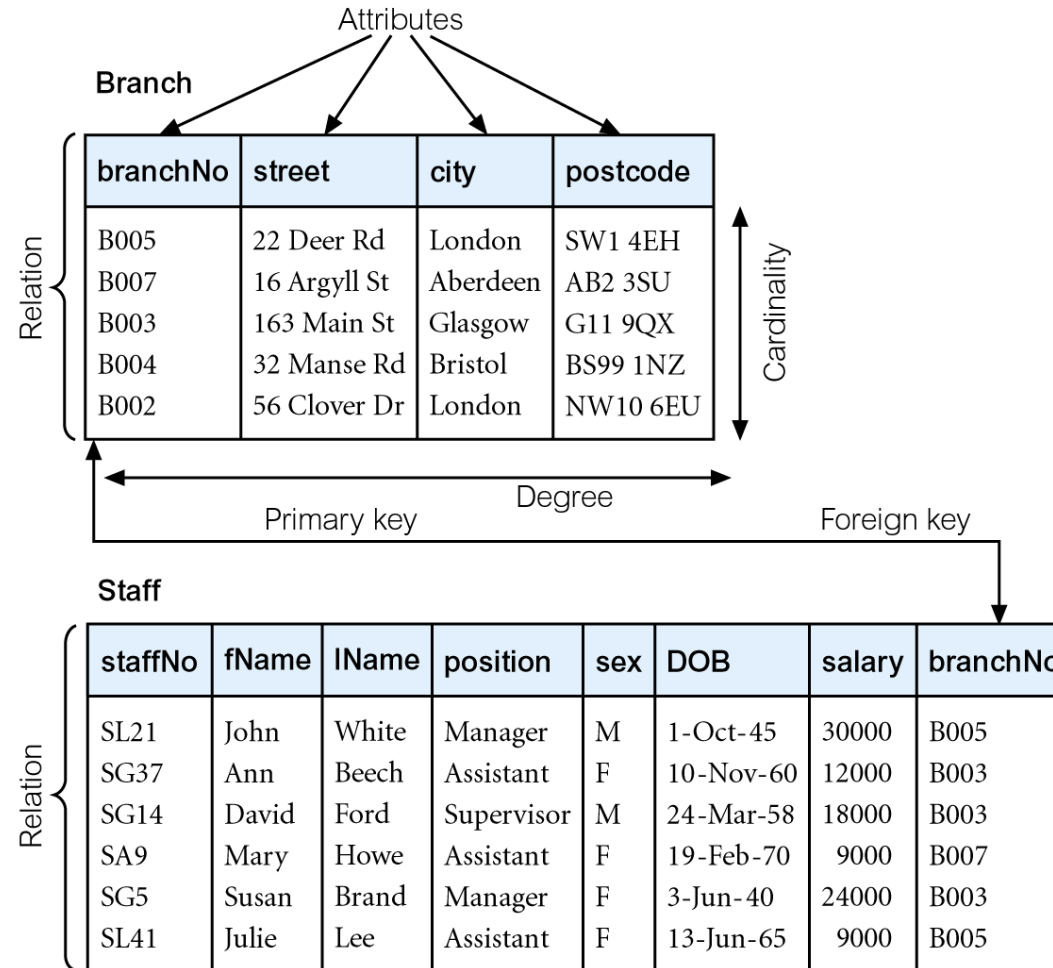
Tuple is a **row** of a relation.

Degree is the number of **attributes** in a relation.

Cardinality is the number of **tuples** in a relation.

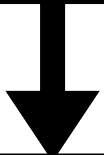
Relational Database is a collection of normalized relations with distinct relation names.

Instances of Branch and Staff Relations



Relational Model Terminology

The structure of a relation, and a specification of the domains is called the intension



Remains fixed
e.g. Branch (branchNo, street, city, postcode)

The tuples/rows are called the extension of a relation

Changes over time
e.g. (B005,22 Deer Rd, London, SW1 4EH)

Alternative Terminology for Relational Model

Formal terms	Alternative 1	Alternative 2
Relation	Table	File
Tuple	Row	Record
Attribute	Column	Field

Agenda

- ✓ Connection between mathematical relations and relations in the relational model.
- ✓ How tables are used to represent data.
- ✓ Properties of database relations.
- ✓ Terminology of relational model.
- ✓ **How to identify CK, PK, and FKs.**
- ✓ Meaning of entity integrity and referential integrity.

Relational Keys

Superkey

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

Candidate Key

- Superkey (K) such that no proper subset is a superkey within the relation.
- In each tuple of R, values of K uniquely identify that tuple (uniqueness).
- No proper subset of K has the uniqueness property (irreducibility).

Relational Keys

Viewing

clientNo	propertyNo	viewDate	comment
CR56	PA14	24-May-08	too small
CR76	PG4	20-Apr-08	too remote
CR56	PG4	26-May-08	.
CR62	PA14	14-May-08	no dining room
CR56	PG36	28-Apr-08	

Relational Keys

Primary Key

- Candidate key selected to identify tuples uniquely within relation.

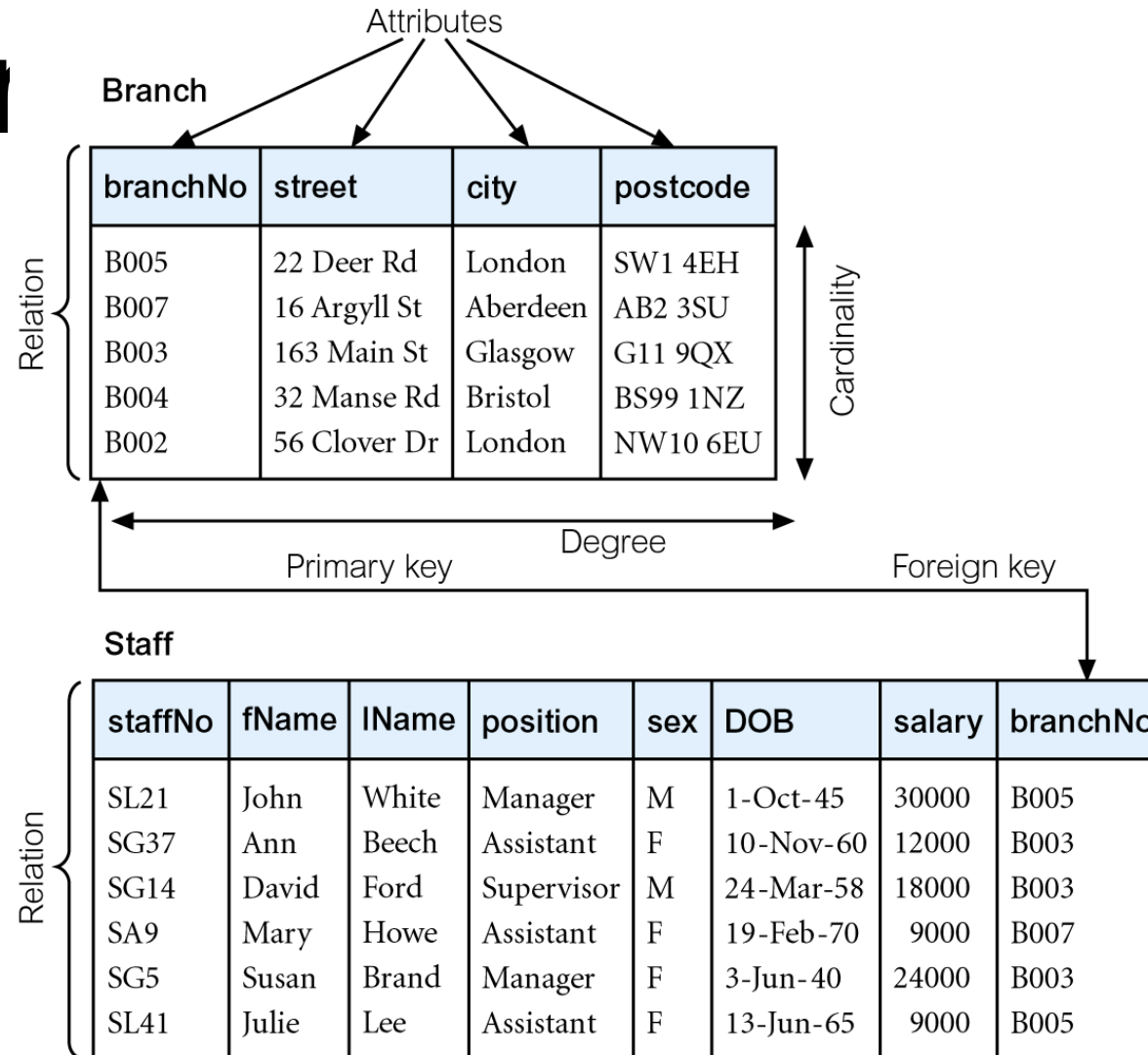
Alternate Keys

- Candidate keys that are not selected to be primary key.

Foreign Key

- Attribute, or set of attributes, within one relation that matches candidate key of some (possibly same) relation.

Instances of Branch and Staff Relation



Agenda

- ✓ Connection between mathematical relations and relations in the relational model.
- ✓ How tables are used to represent data.
- ✓ Properties of database relations.
- ✓ Terminology of relational model.
- ✓ How to identify CK, PK, and FKs.
- ✓ **Meaning of entity integrity and referential integrity.**

Integrity Constraints - NULL

Represents value for an attribute that is currently unknown or not applicable for tuple.

Deals with incomplete or exceptional data.

Represents the absence of a value

is not the same as zero or spaces, which are values.

Relational Keys

Viewing

clientNo	propertyNo	viewDate	comment
CR56	PA14	24-May-08	too small
CR76	PG4	20-Apr-08	too remote
CR56	PG4	26-May-08	.
CR62	PA14	14-May-08	no dining room
CR56	PG36	28-Apr-08	

Integrity Constraints

Entity Integrity

- In a base relation, no attribute of a primary key can be null.

Referential Integrity

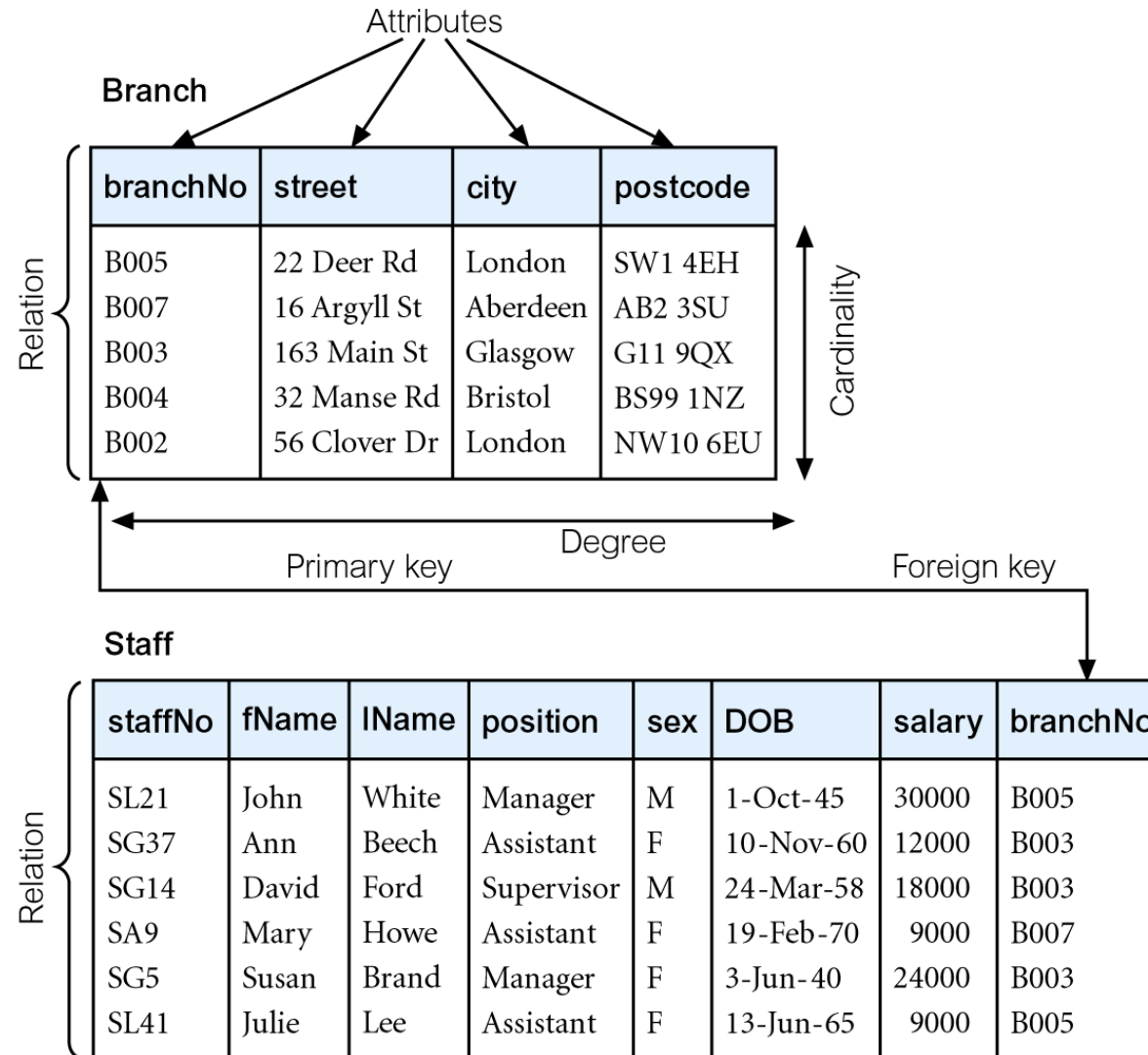
- If foreign key exists in a relation, either foreign key value must match a candidate key value of some tuple in its home relation or foreign key value must be wholly null.

Relational Keys

Viewing

clientNo	propertyNo	viewDate	comment
CR56	PA14	24-May-08	too small
CR76	PG4	20-Apr-08	too remote
CR56	PG4	26-May-08	.
CR62	PA14	14-May-08	no dining room
CR56	PG36	28-Apr-08	

Relational Keys



Integrity Constraints

General Constraints

Additional rules specified by users or database administrators that define or constrain some aspect of the enterprise.

Client cannot view the property on the same day!

Branch

branchNo	street	city	postCode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

Are these rows/records unique?

Yes. But why?

Does the order matter?

Staff

Can staffNo be NULL?

Why are the column names unique?

staffNo	fName	lName	position	sex	DOB	salary	branchNo		salary	branchNo
SL21	John	White	Manager	M	1-Oct-45	30000	B005	;	30000	B005
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003	60	12000	B003
SG14	David	Ford	Supervisor	M	24-Mar-58	18000	B003	58	18000	B003
SA9	Mary	Howe	Assistant	F	19-Feb-70	9000	B007	70	9000	B007
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003	;	24000	B003
SL41	Julie	Lee	Assistant	F	13-Jun-65	9000	B005	;	9000	B005

Does the order mater?

What values can “salary” have?

```
select fname, lname, city
from Staff Natural Join Branch
where position = 'Manager'
```

How do we “know” this query will return all records?

Agenda

- ✓ **Connection between mathematical relations and relations in the relational model.**
- ✓ **How tables are used to represent data.**
- ✓ **Properties of database relations.**
- ✓ **Terminology of relational model.**
- ✓ **How to identify CK, PK, and FKs.**
- ✓ **Meaning of entity integrity and referential integrity.**