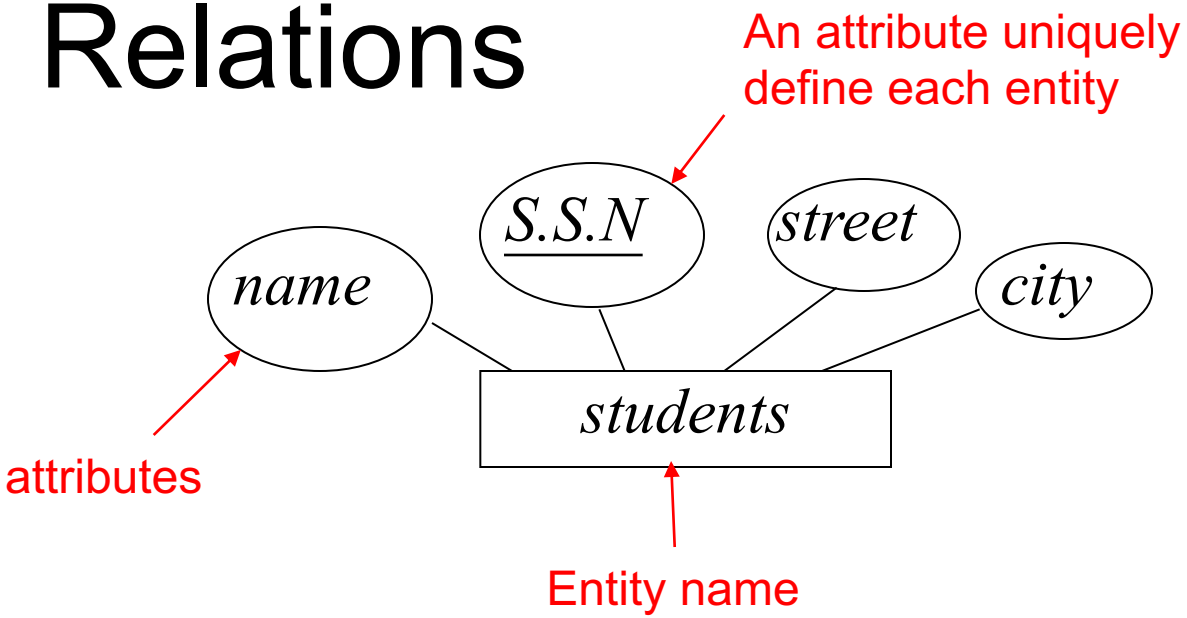
CSCI4333 Database Design & Implement

Lecture Seven – Relational Model 2

Instructor: Dr. Yifeng Gao

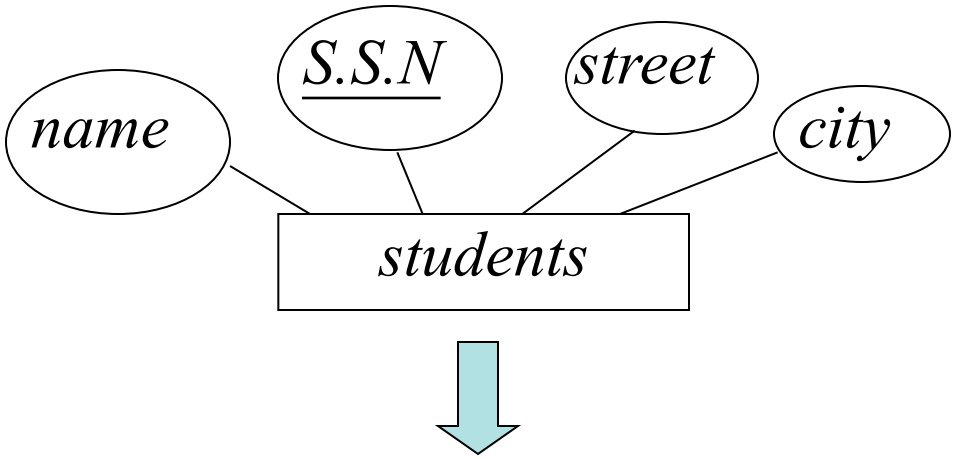
Relations


An attribute uniquely define each entity
S.S.N
street
name
city
students
attributes
Entity name

Relations

A **relation** is a more concrete construction, of something we have seen before, the ER diagram.

A relation is (just!) a table!

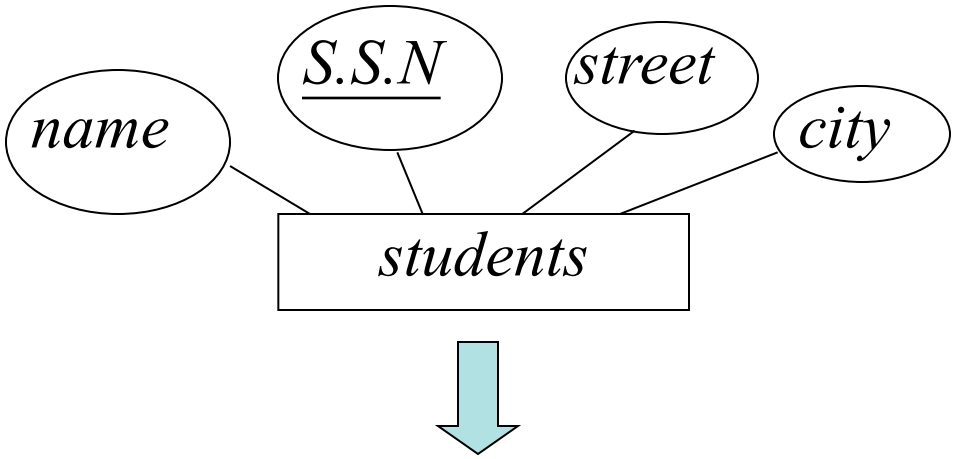


name	*S.S.N*	*street*	*city*
Lisa	1272	10th	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

Students

Relations

Differences between **entities** must be expressed in terms of attributes.

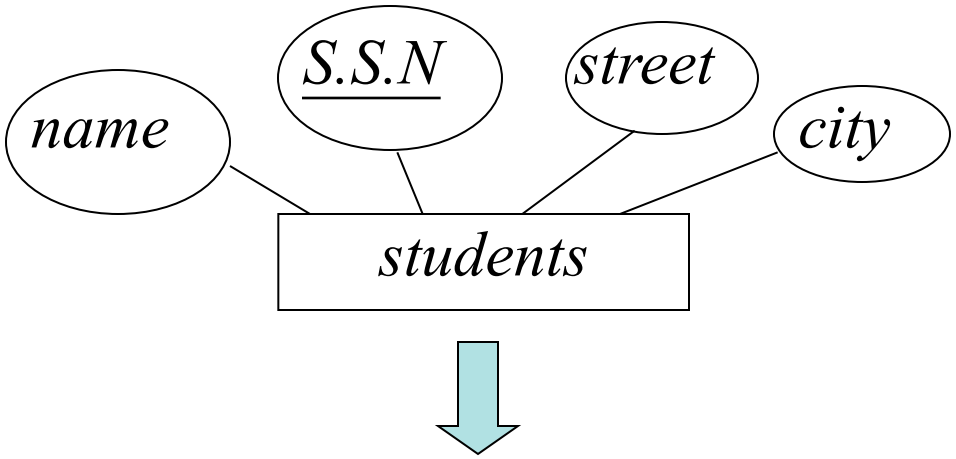


entities

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

Keys

Differences between **entities** must be expressed in terms of attributes.



Key is a set of attribute that allow us to identify uniquely an entity in the entity set.

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

Keys

SSN

(SSN, Name)

(SSN, Name, street)

(SSN, Name, street, city)

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

We have multiple keys in a table. But not all keys are very useful!
Types of key:
1. **superkey:** a set of one or more attributes which, taken collectively, allow us to identify uniquely an entity in the entity set

2. **candidate key:** A superkey for which no subset is a superkey.

Keys

SSN:
candidate key
(SSN, Name)
superkey
(SSN, Name, street)
superkey
(SSN, Name, street, city)
superkey

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

We have multiple keys in a table. But not all keys are very useful!
Types of key:

superkey: a set of one or more attributes which, taken collectively, allow us to identify uniquely an entity in the entity set
candidate key: A superkey for which no subset is a superkey.

Keys

Look at the following table. Answer following questions:

Make	Model	Owner	State	License #	VIN #
Ford	Focus	Mike	CA	SD123	34724
BMW	Z4	Joe	CA	JOE	55725
Ford	Escort	Sue	AZ	TD4352	75822
Honda	Civic	Bert	CA	456GHf	77924

Keys

Look at the following table. Answer following questions:

(State, License#, VIN#) Key? Superkey? Candidate Key?

(Make, Model, Owner) Key? Superkey? Candidate Key?

(State, License#) Key? Superkey? Candidate Key?

(VIN#) Key? Superkey? Candidate Key?

Make	Model	Owner	State	License #	VIN #
Ford	Focus	Mike	CA	SD123	34724
BMW	Z4	Joe	CA	JOE	55725
Ford	Escort	Sue	AZ	TD4352	75822
Honda	Civic	Bert	CA	456GHf	77924

Keys

Look at the following table. Answer following questions:

(State, License#, VIN#): superkey
(Make, Model, Owner): Not a key
(State, License#): candidate key
(VIN#): candidate key

Make	Model	Owner	State	License #	VIN #
Ford	Focus	Mike	CA	SD123	34724
BMW	Z4	Joe	CA	JOE	55725
Ford	Escort	Sue	AZ	TD4352	75822
Honda	Civic	Bert	CA	456GHf	77924

Keys

Look at the following table. Answer following questions:

(State, License#, VIN#): superkey
(Make, Model, Owner): Not a key
(State, License#): candidate key
(VIN#): candidate key

A table can have multiple candidate keys!

Make	Model	Owner	State	License #	VIN #
Ford	Focus	Mike	CA	SD123	34724
BMW	Z4	Joe	CA	JOE	55725
Ford	Escort	Sue	AZ	TD4352	75822
Honda	Civic	Bert	CA	456GHf	77924

Keys

Look at the following table. Answer following questions:

(State, License#, VIN#): superkey
(Make, Model, Owner): Not a key
(State, License#): candidate key
(VIN#): candidate key

A table can have multiple candidate keys! But each table we only choose one.

Make	Model	Owner	State	License #	VIN #
Ford	Focus	Mike	CA	SD123	34724
BMW	Z4	Joe	CA	JOE	55725
Ford	Escort	Sue	AZ	TD4352	75822
Honda	Civic	Bert	CA	456GHf	77924

Keys

Different Types of Key:

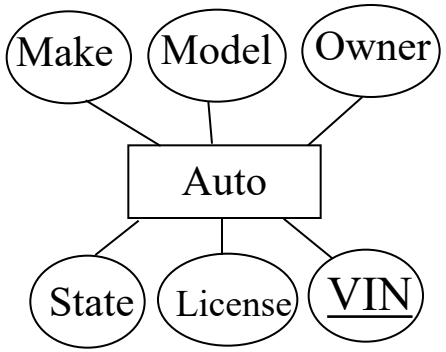
superkey: a set of one or more attributes which, taken collectively, allow us to identify uniquely an entity in the entity set

candidate key: A superkey for which no subset is a superkey.

primary key is a candidate key (there may be more than one) chosen by the DB designer to identify entities in an entity set.

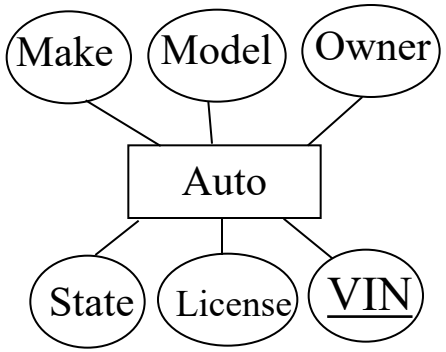
Keys

- The **primary key** is denoted in an ER diagram by underlining.



Keys

- The **primary key** is denoted in an ER diagram by underlining.



Note that a good choice of primary key is very important!

For example, it is usually much faster to search a database by the primary key, than by any other key (we will see why later).

Go back to relation

A **relation** is a more concrete construction, of something we have seen before, the ER diagram.

A relation is (just!) a table!

name	*S.S.N*	*street*	*city*
Lisa	1272	10^{th}	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9^{th}	Mission
Sue	5555	Coria	Brownsville

Go back to relation

A **relation** is a more concrete construction, of something we have seen before, the ER diagram.

A relation is (just!) a table!

name	*S.S.N*	*street*	*city*
Lisa	1272	10^{th}	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9^{th}	Mission
Sue	5555	Coria	Brownsville

A **relation** consists of a **relational schema** and a **relational instance**.

Go back to relation

A **relation** is a more concrete construction, of something we have seen before, the ER diagram.

A relation is (just!) a table!

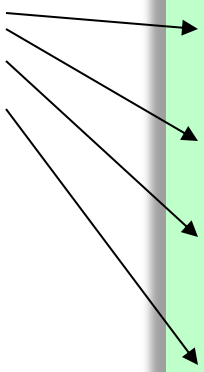
name	*S.S.N*	*street*	*city*
Lisa	1272	10th	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

A **relation** consists of a **relational schema** and a **relational instance**.

A **relation schema** is essentially a list of column names with their data types. In this case…

students(*name* : string, *S.S.N* : string, *street* : string, *city* : string)

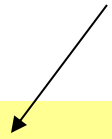
• A **relation instance** is made up of zero or more **tuples** (rows, records)



name	*S.S.N*	*street*	*city*
Lisa	1272	Main	Fairfax
Bart	5592	Apple	Manassas
Lisa	7552	Ox	Fairfax
Sue	5555	Lee	Vienna

students(*name* : string, *S.S.N* : string, *street* : string, *city* : string)

A schema specifies a relation's name.



students(*name* : string, *S.S.N* : string, *street* : string, *city* : string)

A schema specifies a relation's name.

students(*name* : string, *S.S.N* : string, *street* : string, *city* : string)

A schema also specifies the name of each **field**, and its domain. **Fields** are often referred to as columns, attributes, dimensions

A schema specifies a relation's name.

students(*name* : string, *S.S.N* : string, *street* : string, *city* : string)

A schema also specifies the name of each **field**, and its domain. **Fields** are often referred to as columns, attributes, dimensions

primary key: just like ER diagram. Remember that the primary key may be a combination of two or more fields.

A minor, but important point about relations, they are unordered.

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

=

name	*S.S.N*	*city*	*street*
Lisa	1272	Mcallen	10th
Bart	5592	Edinburg	Sugar
Lisa	7552	Mission	9th
Sue	5555	Brownsville	Coria

A minor, but important point about relations, they are unordered.

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

=

name	*S.S.N*	*city*	*street*
Lisa	1272	Mcallen	10th
Bart	5592	Edinburg	Sugar
Lisa	7552	Mission	9th
Sue	5555	Brownsville	Coria

This is not a problem, since we refer to fields by name.

A minor, but important point about relations, they are unordered.

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

=

name	*S.S.N*	*city*	*street*
Lisa	1272	Mcallen	10th
Bart	5592	Edinburg	Sugar
Lisa	7552	Mission	9th
Sue	5555	Brownsville	Coria

This is not a problem, since we refer to fields by name.

However sometimes, we refer to the fields by their column number, in which case the ordering becomes important. I will point this out when we get there.

A minor, but important point about relations, they are unordered.

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	Mcallen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

=

name	*S.S.N*	*city*	*street*
Lisa	1272	Mcallen	10th
Bart	5592	Edinburg	Sugar
Lisa	7552	Mission	9th
Sue	5555	Brownsville	Coria

This is not a problem, since we refer to fields by name.

However sometimes, we refer to the fields by their column number, in which case the ordering becomes important. I will point this out when we get there.

Also, the tuples are unordered too!

Note that every tuple in our instance is unique. This is not a coincidence. The definition of relation demands it.

name	*S.S.N*	*street*	*city*
Lisa	1272	Main	Fairfax
Bart	5592	Apple	Manassas
Lisa	7552	Ox	Fairfax
Sue	5592	Lee	Vienna

The number of tuples = **cardinality** of the relation

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	McAllen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

The number of tuples = **cardinality** of the relation

Of course, we don't count the row of the attribute name!

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	McAllen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

The number of tuples = **cardinality** of the relation

Of course, we don't count the row of the attribute name!

In this table, **cardinality** is 4

name	*S.S.N*	*street*	*city*
Lisa	1272	10th	McAllen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

The number of fields is called the **degree** (or **arity**, or dimensionality of the relation).

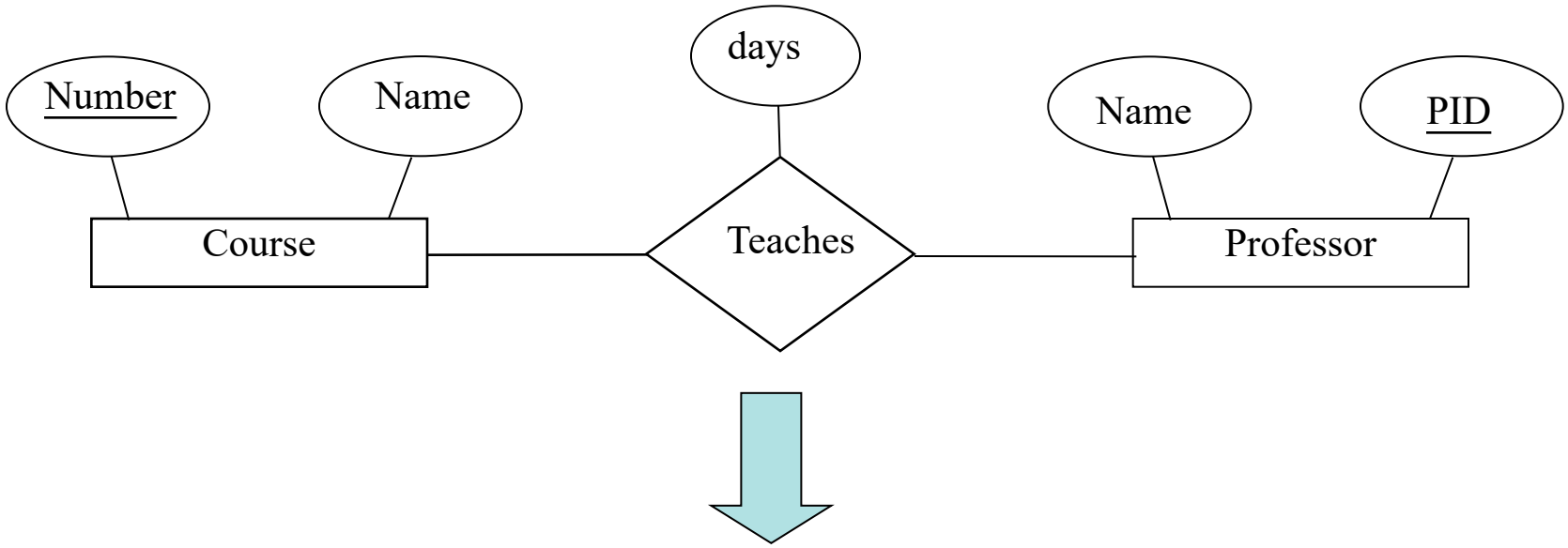
name	*S.S.N*	*street*	*city*
Lisa	1272	10th	McAllen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

The number of fields is called the **degree** (or **arity**, or dimensionality of the relation).

Below we have a table of degree 4.

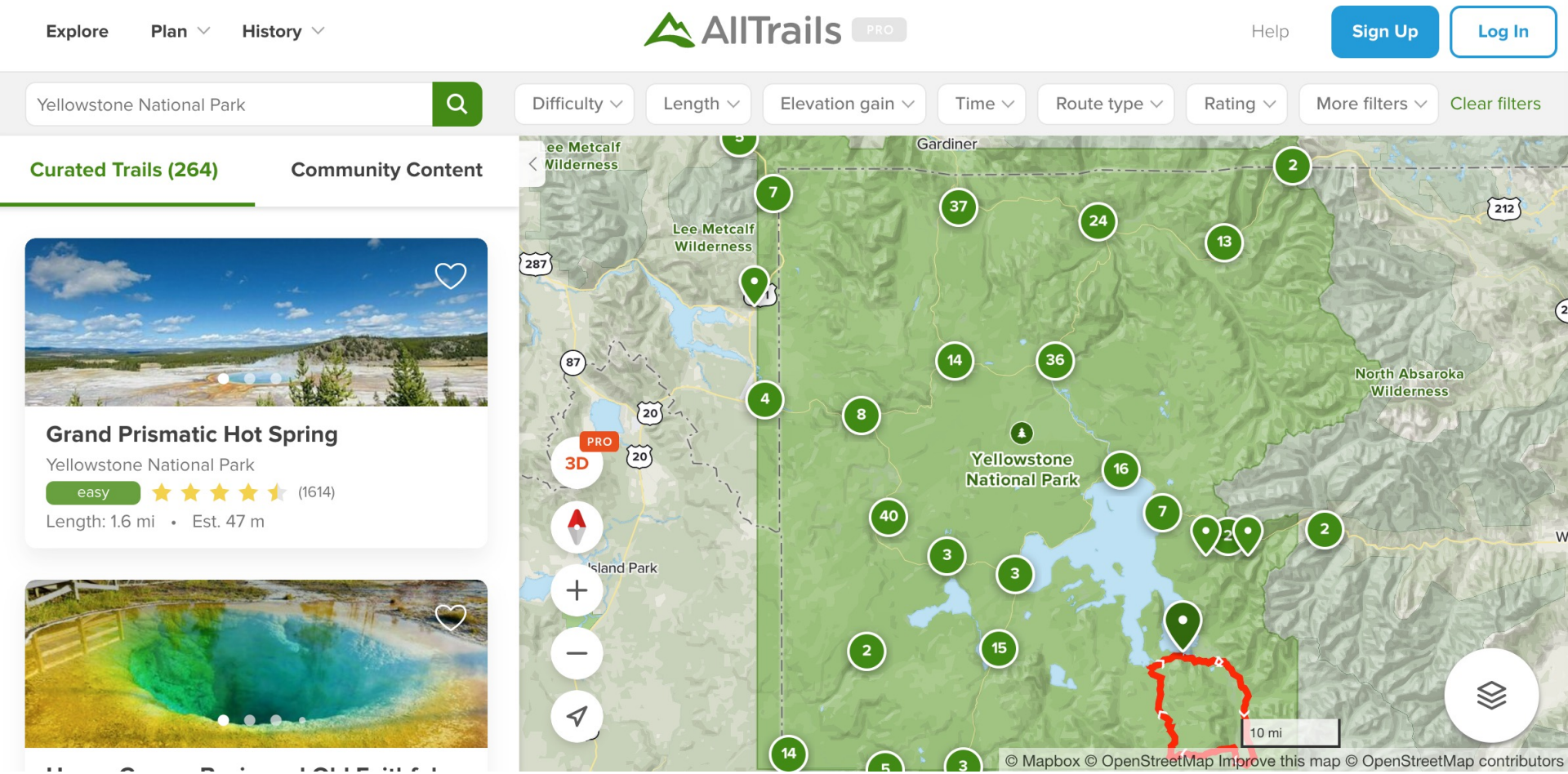
name	*S.S.N*	*street*	*city*
Lisa	1272	10th	McAllen
Bart	5592	Sugar	Edinburg
Lisa	7552	9th	Mission
Sue	5555	Coria	Brownsville

•**Example**



Professor(*PID* : int, *name* : string)
Course(Number: int, Name: string)
Teaches(*PID* : int, Number: int,days: date)

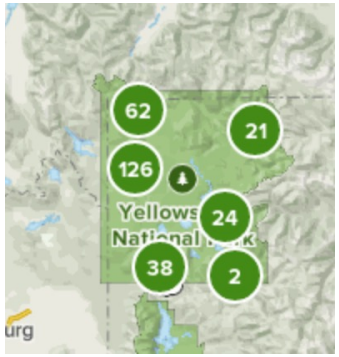
Think about our previous example



Think about our previous example

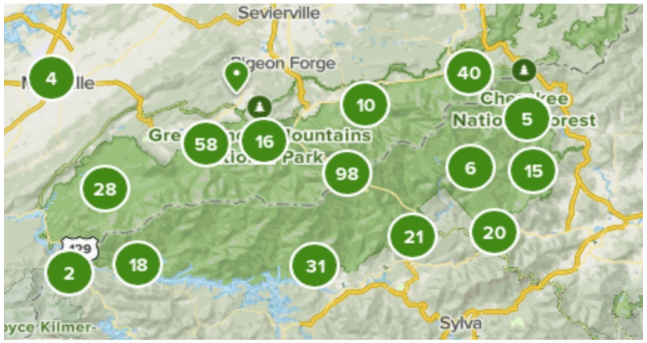
Name	State	Established
Yellow Stone	WY	1872
Great Smoky Mountain	TN	1934
Acadia	ME	1916

Yellow Stone



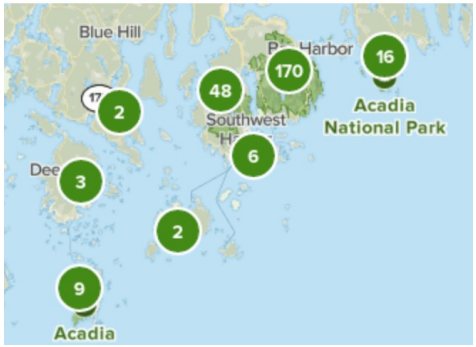
264 trails

Great Smoky Mountain



373 trails

Acadia



224 trails

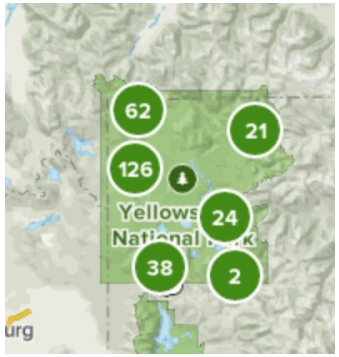
Think about our previous example

Suppose the schema is:
(Name, State, Established, Trail Name)

degree?
cardinality?

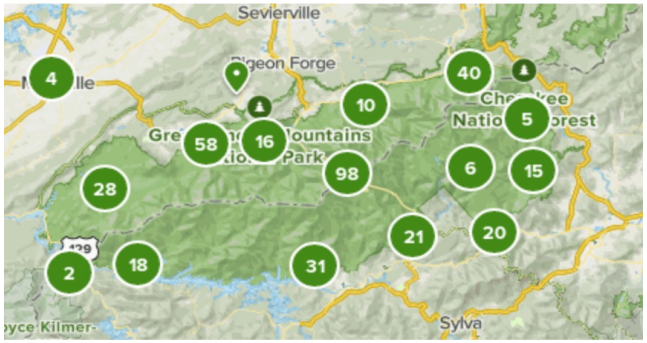
Name	State	Established
Yellow Stone	WY	1872
Great Smoky Mountain	TN	1934
Acadia	ME	1916

Yellow Stone



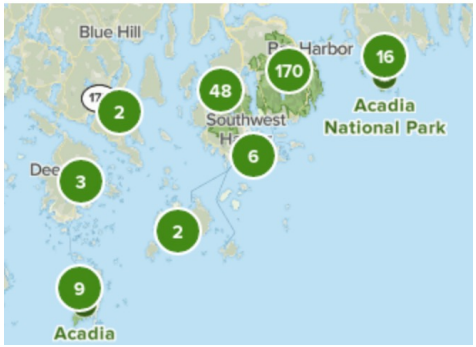
264 trails

Great Smoky Mountain



373 trails

Acadia



224 trails

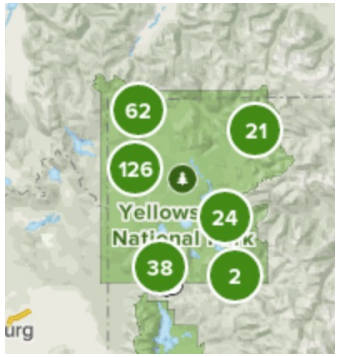
Think about our previous example

Suppose the schema is:
(Name, State, Established, Trail Name)

degree? 4
cardinality? 264+ 373+ 224= 861

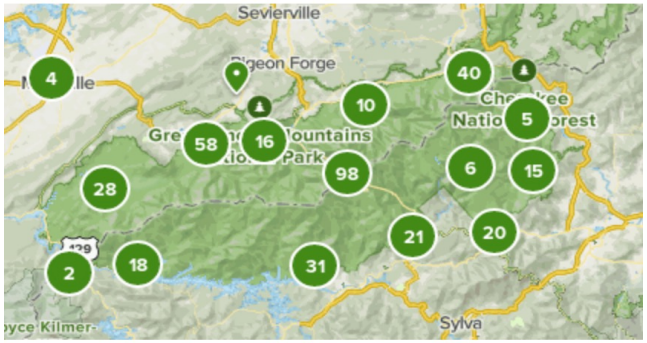
Name	State	Established
Yellow Stone	WY	1872
Great Smoky Mountain	TN	1934
Acadia	ME	1916

Yellow Stone



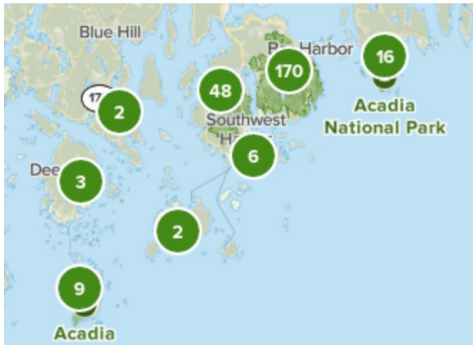
264 trails

Great Smoky Mountain



373 trails

Acadia



224 trails

Think about our previous example

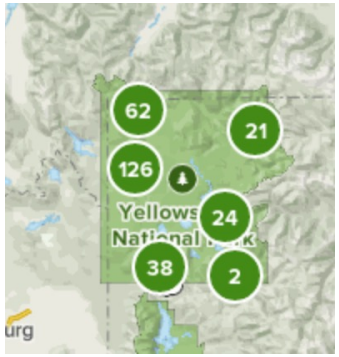
Suppose the schema of Table A is:
(Name, State, Established)

Table B is:
(Name, Trail Name)

Table A: degree? cardinality? Table B: degree? cardinality?

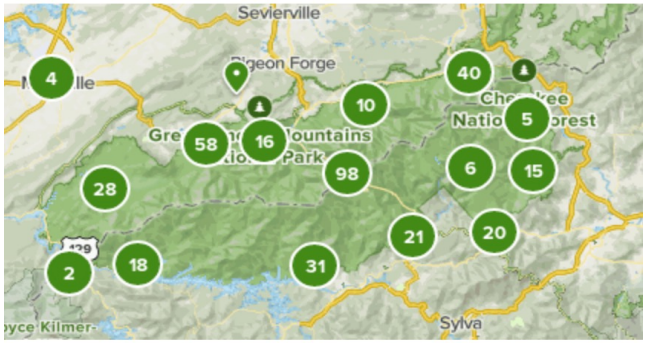
Name	State	Established
Yellow Stone	WY	1872
Great Smoky Mountain	TN	1934
Acadia	ME	1916

Yellow Stone



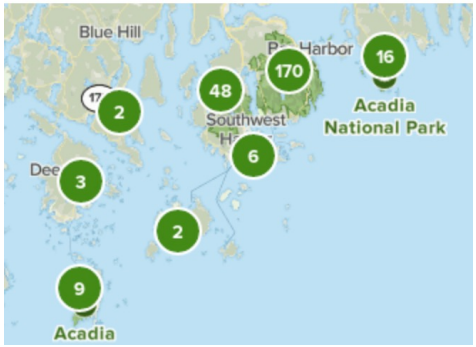
264 trails

Great Smoky Mountain



373 trails

Acadia



224 trails

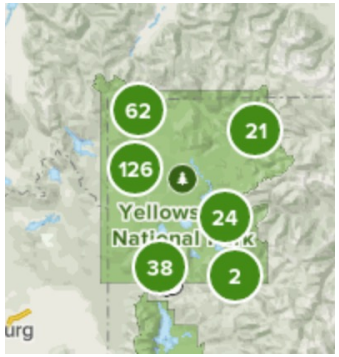
Think about our previous example

Frist design (a big table): degree x cardinality = 4 x 861 = 3444

Two table design:
Table A: degree x cardinality = 3x3
Table B: degree x cardinality = 2x861
Total: 1731

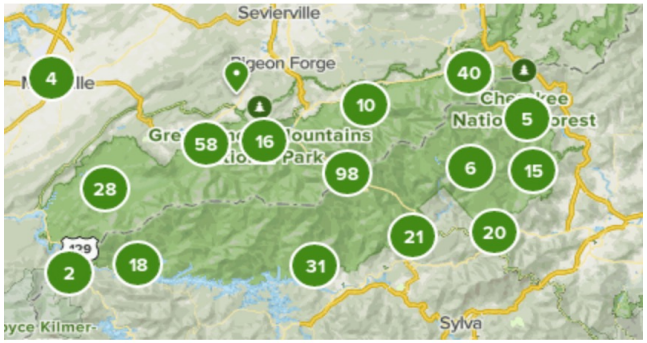
Name	State	Established
Yellow Stone	WY	1872
Great Smoky Mountain	TN	1934
Acadia	ME	1916

Yellow Stone



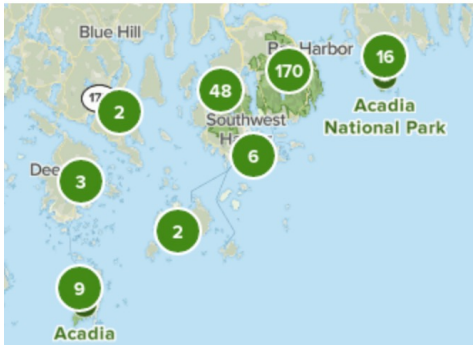
264 trails

Great Smoky Mountain



373 trails

Acadia



224 trails

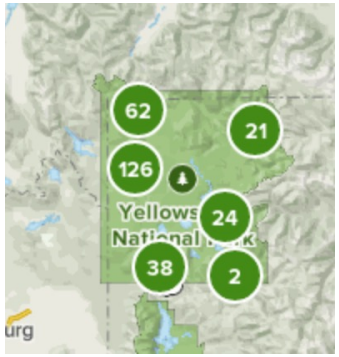
Think about our previous example

Frist design (a big table): degree x cardinality = 4 x 861 = 3444

Two table design:
Table A: degree x cardinality = 3x3
Table B: degree x cardinality = 2x861
Total: 1731
Reduce almost 50% space!

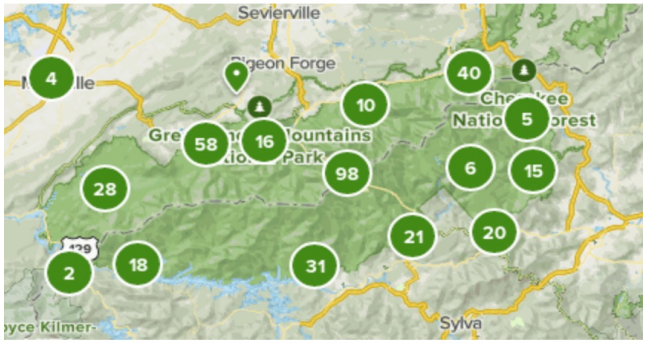
Name	State	Established
Yellow Stone	WY	1872
Great Smoky Mountain	TN	1934
Acadia	ME	1916

Yellow Stone



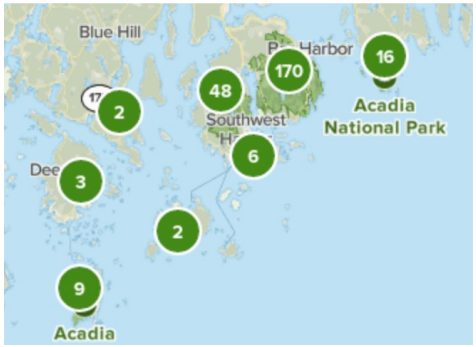
264 trails

Great Smoky Mountain



373 trails

Acadia



224 trails