

Assuming familiarity with tables, rows, columns



Pizzas



Pizzas

Customers



Pizzas

Customers

Orders



Data modelling Relational data model

Pizzas

(name, price)

Customers

(name, address)

Orders

(customer, date, pizza, quantity, price)

Data modelling Relational data model **Primary key**

Pizzas

p_id	name	price
------	------	-------

Customers

(name, address)

Orders

(customer, date, pizza, quantity, price)

Data modelling Relational data model Primary key Natural key Surrogate key

Pizzas

Customers

Orders

o_id customer date pizza quantity price

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy

Pizzas

Customers

Orders

o_id | customer | date | pizza | quantity | price

Data modelling Relational data model Primary key Natural key Surrogate key **Domain Redundancy**

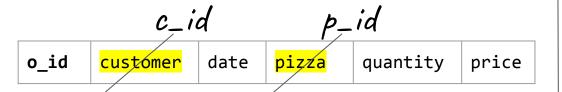
Pizzas

p_id	name	price
------	------	-------

Customers



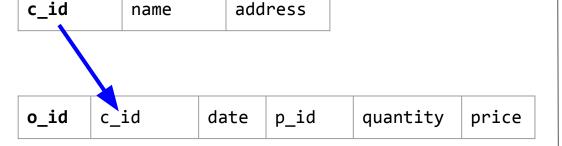
Orders



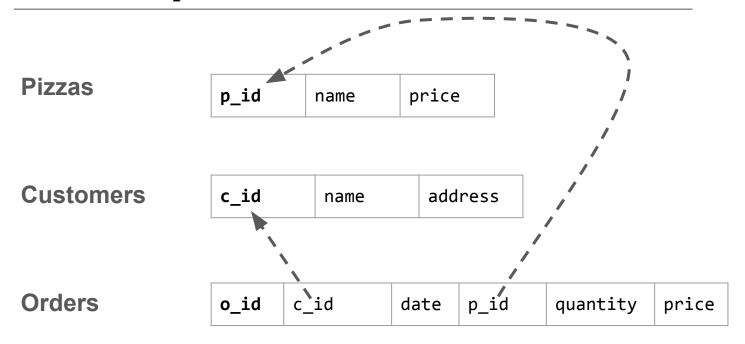
Pizzas

Customers

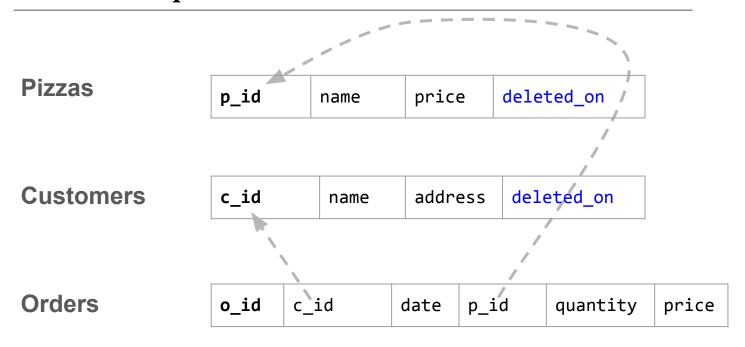
Orders



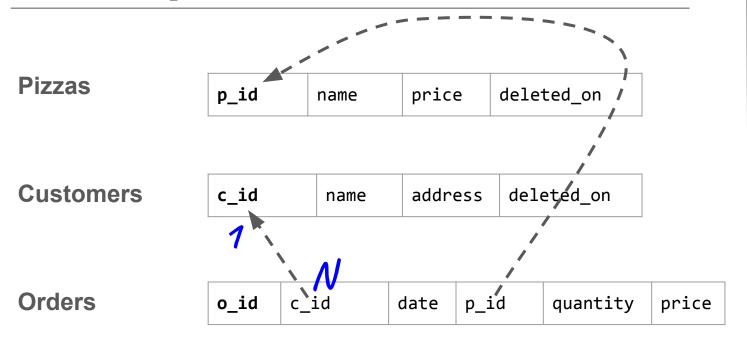
Data modelling
Relational data model
Primary key
Natural key
Surrogate key
Domain
Redundancy
Foreign key
Referential Integrity
Quality of data over time



Data modelling
Relational data model
Primary key
Natural key
Surrogate key
Domain
Redundancy
Foreign key
Referential Integrity
Quality of data over time
On delete/update: Cascade



Data modelling
Relational data model
Primary key
Natural key
Surrogate key
Domain
Redundancy
Foreign key
Referential Integrity
Quality of data over time
On delete/update: Cascade
Soft delete

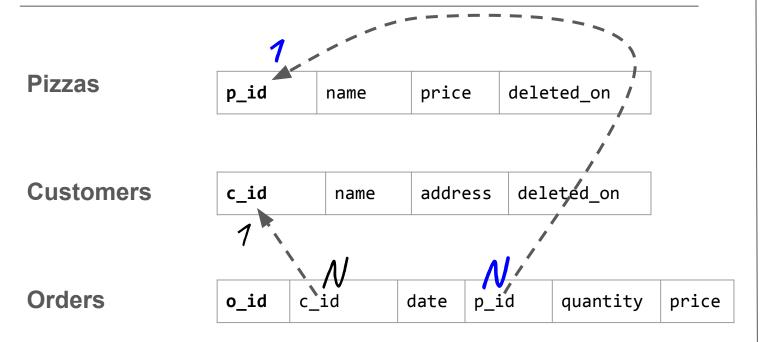


Data modelling
Relational data model
Primary key
Natural key
Surrogate key
Domain
Redundancy
Foreign key
Referential Integrity
Quality of data over time
On delete/update: Cascade
Soft delete
Cardinality

1:N

N:N

1:1



Data modelling

Cardinality 1:1

1:N

N:N

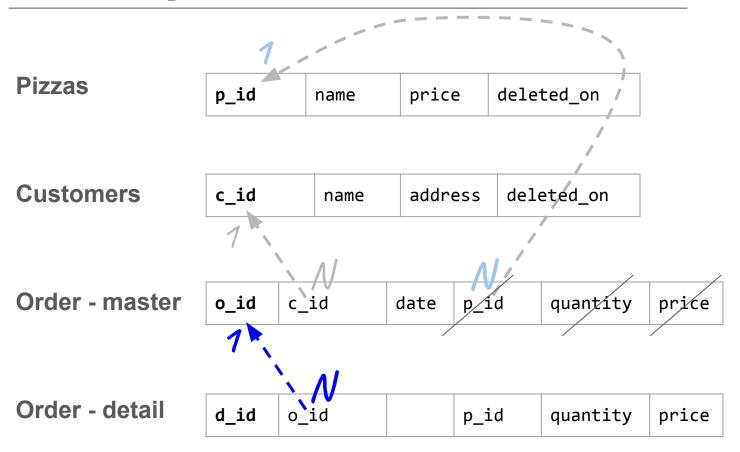
Relational data model Primary key Natural key Surrogate key Domain Redundancy Foreign key Referential Integrity Quality of data over time On delete/update : Cascade Soft delete

Pizzas deleted_on p_id price name **Customers** deleted on c id address name Order - master o_id c id date p<u>/</u>id quant⁄ity pri⁄ce Order - detail d id p_id quantity price

Data modelling

Relational data model
Primary key
Natural key
Surrogate key
Domain
Redundancy
Foreign key
Referential Integrity
Quality of data over time
On delete/update: Cascade
Soft delete
Cardinality

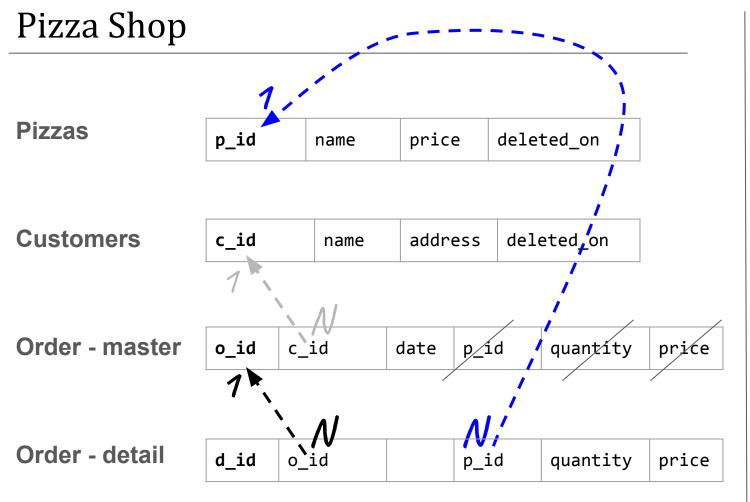
1:1 1:N N:N
Many to many
Master-detail tables



Data modelling

Relational data model
Primary key
Natural key
Surrogate key
Domain
Redundancy
Foreign key
Referential Integrity
Quality of data over time
On delete/update: Cascade
Soft delete
Cardinality

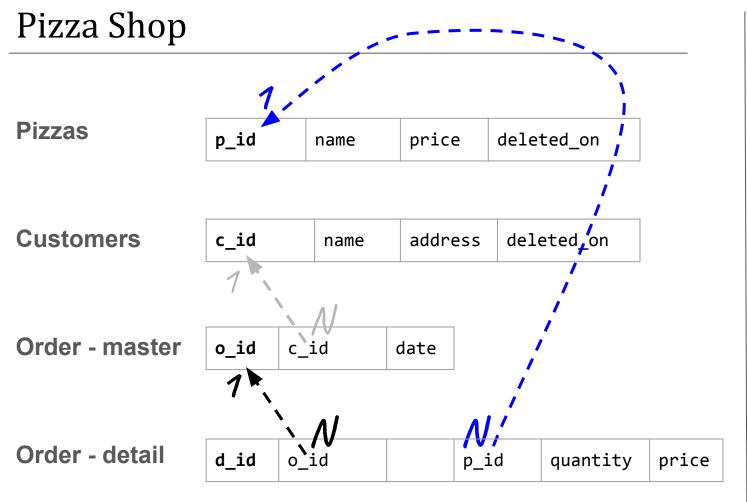
1:1 1:N N:N
Many to many
Master-detail tables



Data modelling

Relational data model
Primary key
Natural key
Surrogate key
Domain
Redundancy
Foreign key
Referential Integrity
Quality of data over time
On delete/update: Cascade
Soft delete
Cardinality

1:1 1:N N:N Many to many Master-detail tables



Data modelling

Relational data model
Primary key
Natural key
Surrogate key
Domain
Redundancy
Atomicity
Foreign key
Referential Integrity
Quality of data over time
On delete/update: Cascade
Soft delete

Cardinality
1:1 1:N N:N

Many to many
Master-detail tables

Pizzas

Customers

Order - master

o_id c_id date

Order - detail

Data modelling
Relational data model
Primary key
Natural key
Surrogate key
Domain
Redundancy
Atomicity
Foreign key
Referential Integrity
Quality of data over time
On delete/update: Cascade
Soft delete
Cardinality

1:1 1:N N:N
Many to many

Master-detail tables

```
CREATE TABLE pizza (
  p_id SERIAL PRIMARY KEY,
  name VARCHAR (100),
  price INTEGER
);
```

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update : Cascade Soft delete Cardinality 1:1 1:N $N \cdot N$ Many to many

DDL CREATE, SERIAL, VARCHAR REFERENCES,

Master-detail tables

http://sqlfiddle.com/

#!9/8ab956

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update : Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many Master-detail tables

DDL CREATE, SERIAL, VARCHAR REFERENCES,

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update : Cascade Soft delete Cardinality 1:1 1:N $N \cdot N$ Many to many

Master-detail tables

DDL, DML CREATE, SERIAL, VARCHAR REFERENCES, INSERT

http://sqlfiddle.com/

#!9/062125

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update : Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many Master-detail tables

DDL, DML CREATE, SERIAL, VARCHAR REFERENCES, INSERT MULTIPLE INSERT

http://sqlfiddle.com/#!9/062125

```
INSERT INTO order master
                                          CREATE TABLE order_master (
    (c id, order date)
                                            c id INTEGER REFERENCES
                                            customer(c id),
VALUES
    (1, '2019-01-02');
                  customer
                  c id
                                                    address
                             name
                             Iron man
                                                    59th Street, Broadway
                             Dr Strange
                                                    1774 Ripakar Street
```

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1:N 1.1 $N \cdot N$ Many to many Master-detail tables

DDL, **DML** CREATE, SERIAL, VARCHAR REFERENCES,

MULTIPLE INSERT Referential Integrity

INSFRT

http://sqlfiddle.com/#!9/062125

customer

c_id	name	address
1	Iron man	59th Street, E

order_master

o_id	c_id	order_date
1	1	2019-01-02T00:00:00Z

DELETE FROM customer WHERE c_id = 1;

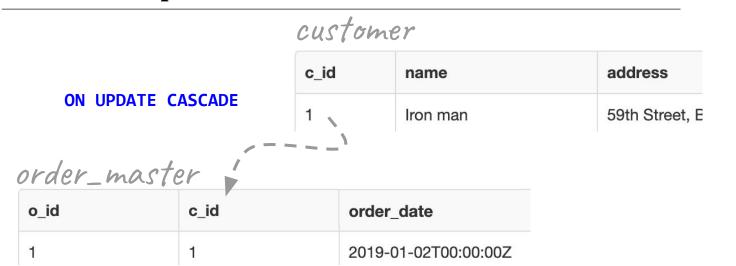
Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update : Cascade Soft delete Cardinality 1:1 1:N $N \cdot N$ Many to many Master-detail tables

DDL, DML CREATE, SERIAL, VARCHAR

REFERENCES, INSERT

MULTIPLE INSERT
Referential Integrity
DELETE FROM

http://sqlfiddle.com/#!9/062125



UPDATE customer SET c_id = 17 WHERE c_id = 1;

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many Master-detail tables

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE

http://sqlfiddle.com/

#!9/9577dd

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many

Master-detail tables

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

http://sqlfiddle.com/#!9/9577dd

```
SELECT * FROM pizza;
```

```
SELECT COUNT(*) FROM customer; WHERE
```

SELECT * FROM customer ORDER BY name;

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update : Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Master-detail tables

SELECT, COUNT, ORDER BY

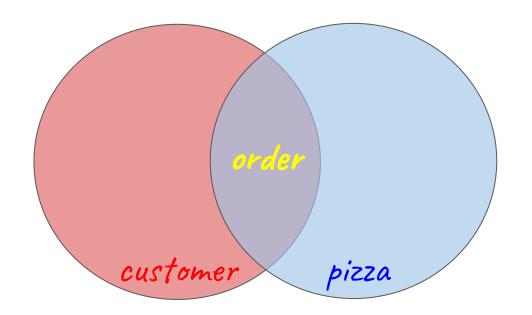
```
SELECT order date, COUNT(*)
  FROM order_master
 GROUP BY order date;
HAVING COUNT(*) > 1
HAVING order date = '2019-01-02'
```

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many Master-detail tables

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

SELECT, COUNT, ORDER BY GROUP BY, HAVING

http://sqlfiddle.com/#!9/9577dd



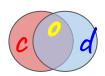
Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update : Cascade Soft delete Cardinality 1:1 N:N 1:N Many to many

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Master-detail tables

SELECT, COUNT, ORDER BY GROUP BY, HAVING INNER JOIN

http://sqlfiddle.com/#!9/9577dd



Find customers who have made some orders

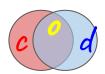
Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update : Cascade Soft delete Cardinality 1:1 1:N $N \cdot N$ Many to many

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Master-detail tables

SELECT, COUNT, ORDER BY GROUP BY, HAVING INNER JOIN

http://sqlfiddle.com/#!9/9577dd



Find customers who have made some orders

SELECT c_id, name FROM customer WHERE c_id IN

(SELECT c_id FROM order_master)

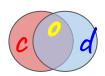
Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1:N 1.1 $N \cdot N$ Many to many

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Master-detail tables

SELECT, COUNT, ORDER BY GROUP BY, HAVING INNER JOIN Sub query

http://sqlfiddle.com/#!9/9577dd



Find customers who have made some orders

SELECT customer.c_id, customer.name

FROM customer

INNER JOIN order_master

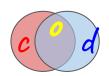
ON customer.c_id = order_master.c_id

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1:N 1.1 $N \cdot N$ Many to many Master-detail tables

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

SELECT, COUNT, ORDER BY GROUP BY, HAVING INNER JOIN Sub query

http://sqlfiddle.com/#!9/9577dd



Find customers who have made some orders

SELECT t1.c_id, t1.name

FROM customer AS t1

INNER JOIN order_master AS t2

ON t1.c_id = t2.c_id

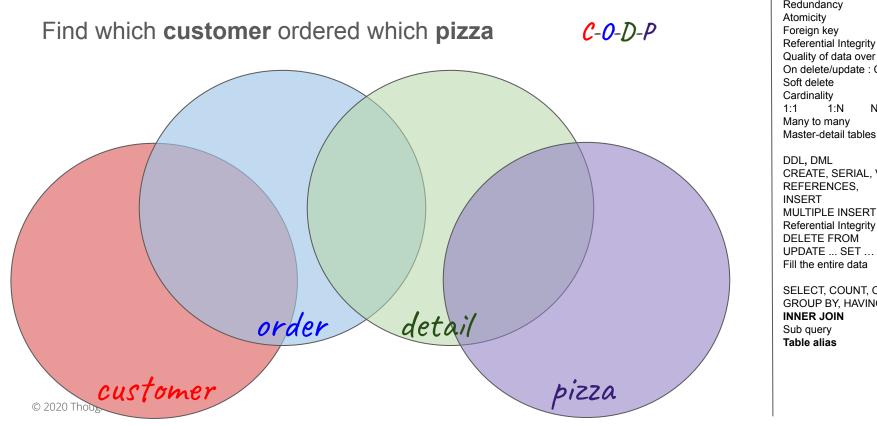
Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Master-detail tables

SELECT, COUNT, ORDER BY GROUP BY, HAVING INNER JOIN Sub query Table alias

http://sqlfiddle.com/#!9/9577dd



Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many

DDL, DML CREATE, SERIAL, VARCHAR REFERENCES. INSFRT MULTIPLE INSERT Referential Integrity DELETE FROM UPDATE ... SET ... WHERE Fill the entire data

SELECT, COUNT, ORDER BY **GROUP BY. HAVING INNER JOIN** Sub query Table alias

Pizza Shop

http://sqlfiddle.com/#!9/9577dd

Find which **customer** ordered which **pizza**

C-0-D-P

SELECT customer.c_id, customer.name, pizza.name

FROM customer

INNER JOIN order_master ON customer.c_id = order_master.c_id

INNER JOIN order_detail ON order_master.o_id = order_detail.o_id

INNER JOIN pizza ON order_detail.p_id = pizza.p_id

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1:N 1.1 $N \cdot N$ Many to many Master-detail tables

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Pizza Shop

http://sqlfiddle.com/#!9/9577dd

Find customers who have **NOT** made *any* orders

SELECT c_id, name FROM customer WHERE c_id **NOT** IN

(SELECT c_id FROM order_master)

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1:N 1.1 $N \cdot N$ Many to many

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Master-detail tables

Find customers who have **NOT** made *any* orders

SELECT t1.c_id, t2.name

FROM customer AS t1

INNER JOIN order master AS t2

ON t1.c_id = t2.c_id;

Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1.1 1·N $N \cdot N$ Many to many Master-detail tables

Data modelling

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Find customers who have **NOT** made *any* orders

SELECT *

FROM customer AS t1

LEFT JOIN order_master AS t2

ON $t1.c_id = t2.c_id$;

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many Master-detail tables

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Find customers who have **NOT** made *any* orders

SELECT *

FROM customer AS t1

LEFT JOIN order_master AS t2

ON t1.c_id = t2.c_id

WHERE t2.o_id IS NULL;

Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1:N 1.1 $N \cdot N$ Many to many Master-detail tables

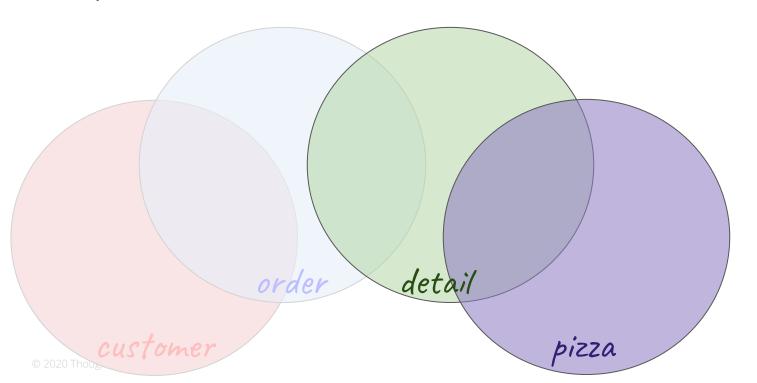
Data modelling

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Pizza Shop

http://sqlfiddle.com/#!9/9577dd

Find pizzas which have **NEVER** been ordered



Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1:1 1:N $N \cdot N$ Many to many

Data modelling

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Master-detail tables

NoSQL

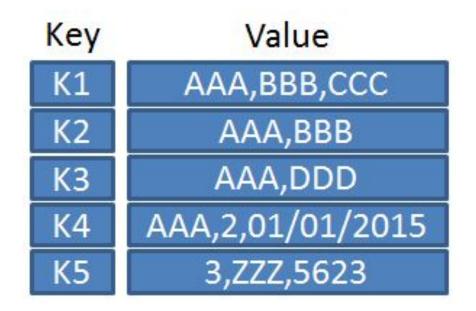
- Non SQL or non relational (sometimes referred to as Not only SQL)
- Different mechanism than the traditional tabular relations
- Highly scalable horizontal scaling
- Used when data is unstructured or the structure might change
- Various types:
 - Key-value
 - Document
 - Time series
 - Graph
 - Wide Column

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1:N 1.1 $N \cdot N$

Many to many Master-detail tables

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Key Value Store



Eg: Redis

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Master-detail tables

Document Store

```
"address": {
         "building": "1007",
         "coord": [ -73.856077, 40.848447 ].
 4
         "street": "Morris Park Ave".
         "zipcode": "10462"
      },
      "borough": "Bronx",
 8
 9
      "cuisine": "Bakerv".
      "grades":
10
         { "date": { "$date": 1393804800000 }, "grade": "A", "score": 2 },
11
12
         { "date": { "$date": 1378857600000 }, "grade": "A", "score": 6 },
         { "date": { "$date": 1358985600000 }, "grade": "A", "score": 10 },
         { "date": { "$date": 1322006400000 }, "grade": "A", "score": 9 },
14
15
         { "date": { "$date": 1299715200000 }, "grade": "B", "score": 14 }
16
17
      "name": "Morris Park Bake Shop",
18
      "restaurant id": "30075445"
19
   }
```

Eg: MongoDB

Data modelling Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many Master-detail tables

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

SELECT, COUNT, ORDER BY GROUP BY, HAVING INNER JOIN Sub query Table alias LEFT JOIN NoSQL

Some more examples

Key - value: DynamoDB

Document: CosmosDB

Timeseries: Influx DB

Graph: Neo4j

Wide column: Cassandra

Relational data model Primary key Natural key Surrogate key Domain Redundancy Atomicity Foreign key Referential Integrity Quality of data over time On delete/update: Cascade Soft delete Cardinality 1.1 1:N $N \cdot N$ Many to many Master-detail tables

Data modelling

DDL, DML
CREATE, SERIAL, VARCHAR
REFERENCES,
INSERT
MULTIPLE INSERT
Referential Integrity
DELETE FROM
UPDATE ... SET ... WHERE
Fill the entire data

Acronyms

ACID

- Atomicity
- Consistency
- Isolation
- Durability

ORM

Object Relational Mapping

Installing MySQL on OSX

- Download the package from:
 https://dev.mysql.com/downloads/mysql/
- Run the installer
- Remember the root password
- For detailed instructions, follow:
 https://dev.mysql.com/doc/mysql-installation-excerpt/8.0/en/osx-installation-pkg.html
- Check if mysql server is running: sudo /usr/local/mysql/support-files/mysql.server status
- Download and install MySQL workbench: https://dev.mysql.com/downloads/workbench/

Installing sqlite on OSX

- Download the package from:
 https://sqlitebrowser.org/blog/version-3-12-1-released/
- Run the installer

Exercise makes everything better!

Let's build something like Spotify or Gaana

Don't worry, we'll keep it simple for now



Requirements

We need to represent the following in our database:

- Users
- Playlists
- Tracks
- Artists

Requirements

Users (user_id,user_name, email)

Playlists (playlist_id,playlist_name, description, user_id,

created_on)

Artist (artist_id,artist_name)

Tracks (track_id,track_name, duration, artist_id)

Playlist_Track (playlist_id, track_id)

Let's create the Database first

- Enter the console: /usr/local/mysql/bin/mysql -u root -p
- Create a Database
 CREATE DATABASE <database_name>
- Create a test user
 CREATE USER 'testuser'@'localhost' IDENTIFIED BY
 'some_password'
- Grant privileges to testuser
 GRANT ALL PRIVILEGES ON <database_name>.* TO
 'testuser'@'localhost'

Create tables

Create tables....

© 2020 ThoughtWorks

Let's add some data

Add some data....

Let's answer some questions

- How many users have registered with our app?
- What playlist is the latest one?
- Tracks count per playlist
- Details of the playlist for a given user name
- Most popular artist's name (Most tracks by that artist)
- Longest track (in terms of time)
- Which users have no playlists?
- Details of playlists which are empty

What to submit

- Create queries
- Insert data queries
- Sql statements for questions in prev