MySQL - RDBMS

Trainer: Mr. Nilesh Ghule

Normalization

- Concept of table design: Table, Structure, Data Types, Width, Constraints, Relations.
- Goals:
  - Efficient table structure. → data processing
  - Avoid data redundancy i.e. unnecessary duplication of data (to save disk space).
  - Reduce problems of insert, update & delete.
- Done from input perspective.
- Based on user requirements.
- Part of software design phase.
- View entire appln on per transaction basis & then normalize each transaction separately.
- Transaction Examples:
  - Banking, Rail Reservation, Online Shopping.

Normalization

- For given transaction make list of all the fields.
- Strive for atomicity.
- Get general description of all field properties.
- For all practical purposes we can have a single table with all the columns. Give meaningful names to the table. & columns
- Assign datatypes and widths to all columns on the basis of general desc of fields properties.
- Remove computed columns.
- Assign primary key to the table.
- At this stage data is in un-normalized form.
- UNF is starting point of normalization.

Normalization

- 1. Remove repeating group into a new table.
- 2. Key elements will be PK of new table.
- 3. (Optional) Add PK of original table to new table to give us Composite PK.
  - Repeat steps 1-3 infinitely -- to remove all repeating groups into new tables.
  - This is **1-NF**. No repeating groups present here. One to Many relationship between two tables.

Normalization

- 4. Only table with composite PK to be examined.
- 5. Those columns that are not dependent on the entire composite PK, they are to be removed into a new table.
- 6. The key elements on which the non-key elements were originally dependent, it is to be added to the new table, and it will be the PK of new table.
  - Repeat steps 4-6 infinitely -- to separate all non-key elements from all tables with composite primary key.
  - This is **2-NF**. Many-to-Many relationship.

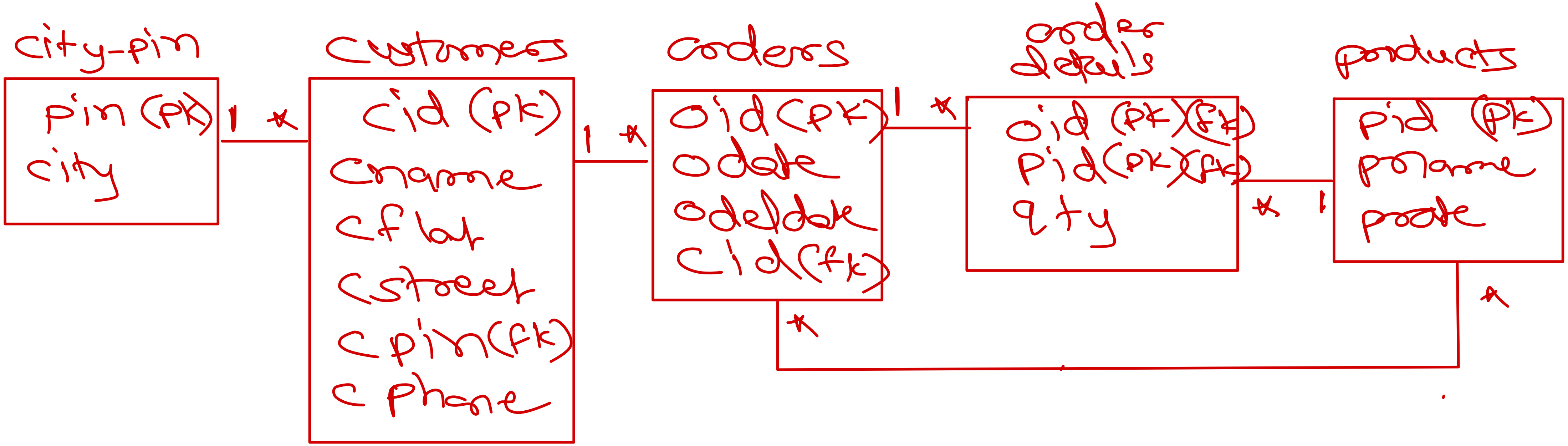
Normalization

- 7. Only non-key elements are examined for inter-dependencies.
- 8. Inter-dependent cols that are not directly related to PK, they are to be removed into a new table.
- 9. (a) Key ele will be PK of new table.
- 9. (b) The PK of new table is to be retained in original table for relationship purposes.
  - Repeat steps 7-9 infinitely to examine all non-key eles from all tables and separate them into new table if not dependent on PK.
  - This is **3-NF**.

Normalization

- To ensure data consistency (no wrong data entered by end user).
- Separate table to be created of well-known data. So that min data will be entered by the end user.
- This is BCNF or 4-NF.

ER - diagram


city-pin
pin (PK)
city
1
*
customers
cid (pk)
cname
cflat
cstreet
cpin(fk)
c phone
1
*
orders
oid (PK)
odate
odeldate
cid(fk)
1
*
order details
oid (PK)(fk)
pid(PK)(fk)
qty
*
1
products
pid (PK)
pname
prate
*
*

De-normalization

- Normalization will yield a structure that is non-redundant.
- Having too many inter-related tables will lead to complex and inefficient queries.
- To ensure better performance of analytical queries, few rules of normalization can be compromised.
- This process is de-normalization.

Commonly done in data warehousing.

DWH is rearranging data so that analytical queries will execute faster.

Common Schemas in DWH

1. star schema
2. snowflake
3. fact constellation

SQL Keys

🔑 An SQL key is either a single column (or attribute) or a group of columns that can uniquely identify rows (or tuples) in a table.

🔑 Super key is a single key or a group of multiple keys that can uniquely identify tuples in a table.

🔑 Candidate key is a single key or a group of multiple keys that uniquely identify rows in a table.

🔑 Primary key is the Candidate key selected by the database administrator (designer) to uniquely identify tuples in a table.

🔑 Alternate keys are those candidate keys which are not the Primary key.

🔑 Foreign key is an attribute which is a Primary key in its parent table, but is included as an attribute in another host table.

PK → empno

CK → empno

SK	SK		SK		SK	SK	SK	SK
empno	ename	addr	job	dept	aadhar	passport	email	phone
1	A	〰	analyst	10	1234	11	a@x.com	〰
2	B	〰	analyst	10	2345	22	b@x.com	〰
3	C	〰	mgr	20	3456	33	c@x.com	〰
4	D	〰	mgr	20	4567	44	d@x.com	〰
5	E	〰	clerk	30	5678	55	e@x.com	〰

key
↓
super key
↓
(unique) Candidate key
↓
Primary key

Thank you!

Nilesh Ghule <nilesh@sunbeaminfo.com>