Out[18]:

pitching_count
49430

Written Questions

W1

Question

- *Define the concept of immutable column and key.*
- *Why do some sources recommend that a primary key should be immutable?*
- *How would to implement immutability for a primary key in a table?*

Answer

***Immutable column**: a column in a database table that cannot be updated or deleted once the value is assigned.*

***Immutable key**: value of a primary key or unique key cannot be modified or deleted.*

***Primary key**: is unqiue representation of each rows. When tables are connected, referential integrity and data consisency need to be enforced.*

***Implement**: by setting column as IMMUTABLE PRIMARY KEY.*

```
CREATE TABLE student (
    student_id varchar(128) IMMUTABLE PRIMARY KEY
    )
```

W2

Question

Views are a powerful concept in relational database management systems. List and briefly explain 3 benefits of/reasons for creating a view.

Answer

***1**: Due to **securtity concerns**, view can be personalized collection of virtual relations.*

***2**: It's possible to **support a large number** of views on the top of actual relations, because it's not precomputed and stored. Take place of with clause.*

***3**: Database system stored view as definition. Therefore, content of view is recomputed when we query and **not out of date**.*

W3

Question

Briefly explain the concepts of procedural language and declarative language. SQL is primarily a declarative language. SQL added procedure language capabilities in functions, procedures and triggers. What is a reason for this addition?

Answer

***Procedural DML**: specify what need and how to get those data.*

***Declarative DML**:specify what need without specifying how to get those data.*

***capabilities in functions, procedures, and triggers**: They are more advanced features of SQL and taken for more complexed operations. They provide access to general-purpose programming language, like write reusable code, perform transformation, enforce data consistency, etc.*

W4

Question

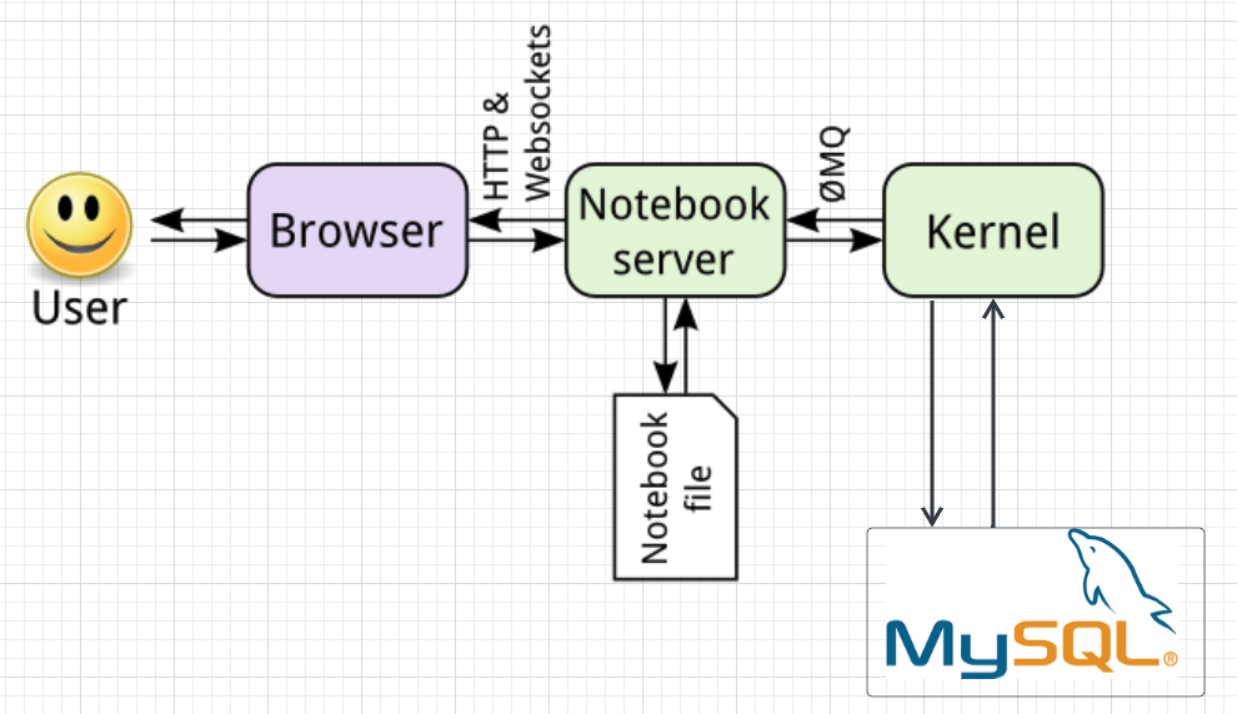
The following diagram is a simple representation of the architecture of a Jupyter notebook using MySQL. Is this a two-tier architecture or a three-tier architecture? Explain your answer briefly.

 *#this img can run, but cannot show after export html working, thus, I use import Image instead.*

In [19]:
```
#above <img src="./jupyter-notebook.png"> can run, but cannot show after export html w
from IPython.display import Image

Image("jupyter-notebook.png")
```

Out[19]:



Answer

***Three-tier achitecture**: web browser as front-end communicate with application server(notebook server). Notebook server communicate with mysql to access the data. Web browser does not contain any direct database calls.*

W5

Question

- *Consider a US Social Security Number. An example is "012-34-6789".*
- *The data type is character string.*
- *The relational model requires that columns (attributes) are from a domain.*
- *Use the Social Security Number example to explain the difference between a type and a domain.*

Answer

***Type** of SSN define it's a string with fixed length of nine.*

***Domain** acts as contraint on values that it can take to enforce referential integrity and authorization. SSN is unique with fixed 3-2-4 number pattern. People with alaries must own SSN. Not all users can read, insert, update, delete information of SSN in the tables. Domain can differentiate among the users.*

W6

Question

Briefly explain the differences between:

- *Database stored procedure*
- *Database function*
- *Database trigger*

Answer

Procedure

- *Can change data.*
- *Sometime return value by use in and out argument.*
- *Call to be executed repeatedly. Not automatic.*
- *Allow same name as long as arguments are different.*

Function

- *Cannot change data.*
- *Always return a table or a number.*
- *Called with prefixed name with inputs in a statement. Not automatic.*
- *Allow same name as long as arguments are different.*

Trigger

- *Can change data.*
- *Not return value.*
- *Carried out automatically on certain events such as insertion, deletion, or update in a specified relation.*
- *In different schemas can have same name.*

W7

Question

Briefly explain:

- *Natural join*
- *Equi-join*
- *Theta join*
- *Self-join*

Answer

***Natural join**: match tuples with same value for all matched columns, and only keep one copy of replicated columns.*

***Equi-join**: return tuples from tables when values in the matched columns are equal.*

***Theta join**: use comparison operator other than equality to match tuples from tables.*

***Self-join**: join with self. Able to put one table in two distinct tables like employee and manager. Processing hierarchy. Generating pairs of rows based on condition.*

W8

Question

Briefly explain the difference between a unique (key) constraint and a primary key constraint?

Answer

***Primary key**: uniquely identify rows, not null.*

***Unqiue key**: a column or a set of columns are unique, not need to be primary identifier, allow null.*

W9

Question

Give two reasons for using an associative entity to implement a relationship instead of using a foreign key.

Answer

***1**: when foreign key is not sufficient to establish relationship between two entities, associative can used as additional attributes stored and associated with the relationship.*

***2**: to implement many to many relationship between two or more entities.*

W10

Question

Briefly explain the concepts of:

- *Conceptual model*
- *Logical model*
- *Physical model*

For data modeling.

Answer

***Conceptual model**: establish high-level, static business structures and concepts.*

***Logical model**: define entity types, data attributes and relationships between entities.*

***Physical model**: internal schema database design with tables, columns, keys, and indexes.*

W11

Question

Briefly explain the concepts of:

- *Data manipulation language*
- *Data definition language*

Given an example statement in SQL for DML and for DDL.

Answer

***Data manipulation language**: manipulate data in a database. EXA, insert, update, delete, retrieve data.*

```
insert into s23_w4111_hw2_bq2150.name_basics_all (nconst,
primaryName, birthYear) values ('nm0203893', 'Ton Ha', '1960')
```

***Data definition language**: manage database structure. EXA, create, modify, delete database schema(table, index).*

```
DROP TABLE name_basics_all
```

W12

Question

Codd's 4th rule is:

Rule 4 - Dynamic online catalog based on the relational model:

The data base description is represented at the logical level in the same way as ordinary data, so that authorized users can apply the same relational language to its interrogation as they apply query to the regular data.

Explain what this means, and use SQL to provide examples.

Answer

Database schema, stored in the same relational database system, should be accessible to authorized users in the same way as the regular data. Example below.

```
SELECT * FROM INFORMATION_SCHEMA.COLUMNS
where TABLE_SCHEMA='s23_w4111_hw2_bq2150' And
TABLE_NAME='name_basics_all';
```

W13

Question

The formal definition of a theta join is

$$r \bowtie_\theta s = \sigma_\theta(r \times s) \tag{1}$$

Briefly explain the definition and give an example.

Why is the fact that the relational algebra is closed is important to this definition? /

Answer

EXPLAIN: Result will include all columns from r and s, and rows that satisfy theta condition.

EXA:In cartesian-product of r × s, attribute names can only come from instructor and teaches.

Thus, it satisfies closed that result of operators is a defined set and produced from the element within the same set.

W14

Question

Consider two different statements in the relational algebra or SQL.

Despite being different statements, the statements may be equivalent. Briefly explain what this means.

Answer

We use different syntax, but result in the same output.

For example, in HW2P2, "Please write an equivalent query that does not use anti-join".

W15

Question

Consider the following relation definitions.

$$Customers(ID, last\_name, first\_name) \tag{2}$$

$$Accounts(ID, balance, customer\_ID) \tag{3}$$

What is problem with using natural join on the two tables?

Answer

Accounts.ID are view as same column with Customers.ID to pair up in natural join, which is actually not correct.

Accounts.customers_ID will be the correct column to join. We can clarify it by add ON statement

Entity Relationship Modeling

ER-1

Question

This question tests your ability to "bottom up" model or "reverse engineering" a SQL schema to produce an explanatory ER-diagram.

Use Lucidchart to draw a Crow's Foot notation diagram representing the following SQL.

You can use the simple table names, e.g. `students` *instead of* `s23_w4111_midterm.students`.

```
drop schema if exists s23_midterm;

create schema s23_midterm;

use s23_midterm;

drop table if exists departments;
create table if not exists departments
(
    dept_code varchar(4)   not null
        primary key,
    dept_name varchar(128) not null
);

drop table if exists instructors;
create table if not exists instructors
(
    UNI        varchar(12)  not null
```