

SQL: Structured Query Language

SQL Terminology

- SQL helps _____ efficiently
 - query data
 - update data
 - manipulate data
- Python works with SQL to _____.
 - connect databases together
 - execute SQL queries
- SQL is _____

a language used for relational databases
- SQL is used for _____ querying data

Data Terminology

- data is _____

a collection of facts (words, numbers) and/or pictures
- data is a critical asset for _____ business
- data needs to be _____ because it is important.
 - secured
 - stored
 - access quickly

Database Terminology

- database is _____
 - a repository of data
 - a program that stores data
- database provides functionality for _____ data.
 - adding
 - modifying
 - querying
- relational database is _____

data stored in tabular form — columns and rows
- columns contain _____ properties of an item
- An example of a property of an item could be _____
 - last name
 - first name
- table is _____

a collection of related things
- T/F Relations between tables can exist.

True
- DBMS (Database Management System) is a _____ software used to manage databases
- Other words that can be used interchangeably for database are _____
 - database server
 - database system
 - data server
 - DBMS

RDBMS (Relational DBMS) Terminology

- RDBMS is _____

a set of software tools that controls the data
- Data in RDBMS can be controlled by actions like _____
 - accessibility
 - organization
 - storage
- Examples of RDBMS _____
 - MySQL
 - Oracle Database
 - IBM Db2
- Basic SQL Commands are _____
 - Create a table
 - Insert
 - Select
 - Update
 - Delete

COUNT

Reference: Database and SQL for Data Science with Python, IBM Coursera

- Count is a _____ built-in database function
- Count retrieves the number of _____ rows
- The syntax for querying the total rows of a given table is _____


```
select COUNT(*) from tablename;
```
- The syntax for querying the total rows of a given column equaling x is _____


```
select COUNT(column_name) from tablename
where column_name = 'x';
```
- The syntax for retrieving unique values in a column is _____


```
select DISTINCT column_name from tablename;
```
- The output of this syntax:


```
Select DISTINCT COUNTRY from MEDALS
where MEDALTYPE = 'GOLD';
```

 is _____.

a list of unique countries that received gold medals.

Limit

- Limit restricts the number of _____ retrieved from the database

rows
- The Syntax:


```
select * from tablename LIMIT 10;
```

 will result in _____.

the first 10 rows in a table
- The Syntax:


```
select * from MEDALS
where YEAR = 2018 LIMIT 5;
```

 will result in _____.

five rows or less of the year 2018

Insert Statement

- The INSERT statement is used to add _____ to a table.

new rows
- T/F the INSERT statement is DML statement.

True
- The syntax of the INSERT statement is _____


```
INSERT INTO tablename
(column_a, column_b, ..., column_n)
VALUES (value_r1, value_r2, ..., value_rn);
      
```

$$\vdots$$

$$(value_rn, value_rn2, ..., value_rnn);$$

WHERE Clause

- The WHERE clause can restrict the _____ result set
- T/F The WHERE clause always requires a predicate.

True
- T/F The predicate always evaluates to True, False, or Unknown.

True
- The syntax for using WHERE clause is _____


```
select column_a, column_b, ..., column_n from tablename
where predicate;
```
- An example of using the WHERE clause is _____


```
select book_id, title from Book WHERE book_id = 'B1';
```
- The 6 comparison operators are _____
 - =
 - >
 - <
 - >=
 - <=
 - <> (not equal to)
- The predicate is also known as _____ condition

Update Statement

① T/F UPDATE statement is a DML Statement.
True

② UPDATE statement is used to _____ alter or modify the data.

③ The syntax for UPDATE Statement is _____.

```
UPDATE tablename SET column_a = 'new_a', Column_b = 'new_b',
... column_n = 'new_n' where predicate;
```

Optional for specificity

④ The syntax to change first name and last name to Lakshmi Katta where AUTHOR-ID = A2 is _____.

```
UPDATE AUTHOR set LASTNAME = 'KATTA', FIRSTNAME = 'LAKSHMI'
where AUTHOR-ID = 'A2';
```

DELETE Statement

① T/F DELETE statement is DML Statement.
True

② The syntax for DELETE statement by a column_name is _____.

```
DELETE FROM tablename
Where column_name in ('property_a', 'property_b', ..,
'property_n');
```

③ The result of the code

```
DELETE FROM AUTHOR where AUTHOR-ID in ('A2', 'A3');
is _____.
```

Deletes a row where AUTHOR-ID is either A2 or A3.

④ T/F If you do not specify a where clause, all the rows will be deleted.
True

⑤ T/F You do not use DELETE to delete columns.
True

⑥ The general syntax for deleting columns is _____

```
ALTER TABLE tablename DROP COLUMN column_name;
```

⑦ T/F MySQL, PostgreSQL have the same syntax for dropping multiple columns but SQL Server and Oracle database each have their own syntaxes.
True.

Relational Model

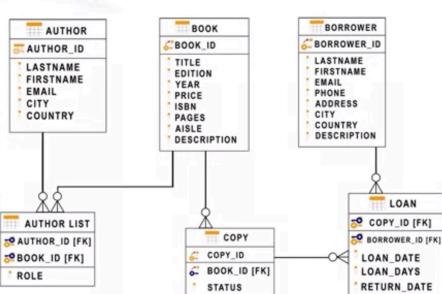
① The relational model is the most used data model because it allows for _____ data independence

② The three types of independence preserved are _____

1. logical independence
2. physical data independence
3. physical storage independence

③ An entity relationship (ER) data model is _____ an alternative to relational data model

④ This is an example of an ER (entity relationship diagram):



⑤ ERD is used to represent _____ entities called table and their relationships.

⑥ T/F The ER model is used as a tool to design relational databases.
True

⑦ In an ER model, an entity is an _____ object

⑧ T/F Entities in an ER model exists independently of any entities in the database.

⑨ The building blocks of an ER diagram are _____.

(2)

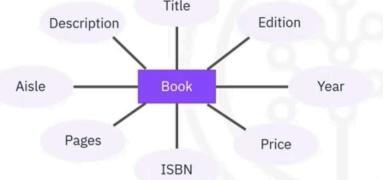
1. entities
2. attributes

ER Diagram

① An Entity can be _____.

Noun : person, place, or thing

(2)



In this diagram, the entity is _____

Book

(2b) In this diagram, the attributes are _____

Title, Edition, etc

(3) Attributes are _____

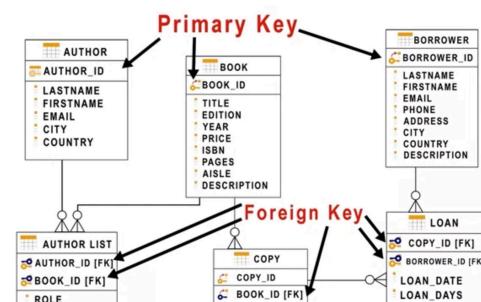
data elements that characterize the entity

Entity-Relationship Model

① Entity becomes a _____ in the database.
table

② Attributes become the _____ in a table.
Columns

Primary keys and foreign keys



① T/F Each table is assigned a primary key.

True.

② The primary key uniquely identifies each

_____ of a table.

tuple or row

③ The primary key prevents _____.

duplication of data

④ Foreign keys are _____ primary keys defined in other tables

⑤ The foreign key helps create a _____ link between tables

DDL vs DML

① You can interchange row with _____ tuple

② You can interchange column with _____ attribute

③ DDL stands for _____ data definition language

④ DML stands for _____ data manipulation language

⑤ DDL statements are used to _____ objects (such as tables).

1. define
2. change
3. drop

⑥ Common DDL statement types are _____.

1. CREATE
2. ALTER
3. TRUNCATE
4. DROP

Common DDL statements

① CREATE is used for _____

1. creating tables
2. defining its columns

② ALTER is used for _____

1. Adding / dropping columns
2. modifying their datatypes

③ TRUNCATE is used for _____

deleting data in a table but not the table itself

④ DROP is used for _____

deleting tables

DML statements

① DML statements are used to _____.

read and modify data

② Read and modify data with _____.
CRUD operations

③ CRUD stands for _____.

Create, Read, Update & Delete rows

Common DML statements

① Common DML statements are _____.

1. INSERT
2. SELECT
3. UPDATE
4. DELETE

② INSERT is used for _____.

Inserting a row or several rows of data into a table.

③ SELECT is used for selecting data in a _____.

row(s)

④ UPDATE is used for editing _____.

row(s)

⑤ DELETE is used for deleting data in a _____.

CREATE TABLE Statement

① The syntax for CREATE TABLE is _____.

```
CREATE TABLE tablename (
    column_1 datatype optional parameters,
    column_2 datatype,
    ...
    column_n datatype
);
```

② A sample code:

```
CREATE TABLE provinces(
    id CHAR(2) PRIMARY KEY NOT NULL,
    name VARCHAR(24) NOT NULL,
    population BIGINT
);
```

③ The CHAR(2) tells us the datatype is character string for id and has a _____ fixed length of 2.

④ The varchar is a character string of a _____ variable length.

⑤ VARCHAR(24) means the variable character length can be _____ up to 24.

ALTER TABLE Statement

① The syntax for ALTER TABLE is _____.

ALTER TABLE tablename

```
ADD COLUMN column_1 datatype,
...
ADD COLUMN column_n
```

);

② T/F The ALTER TABLE does not allow you to perform multiple operations (e.g. adding and modifying columns) in a single statement.

True.

③ The actions you can perform with ALTER TABLE _____.

1. Add or remove columns
2. Modify the data types of columns
3. Add or remove keys
4. Add or remove constraints

④ The syntax for modifying the datatype is _____.

ALTER TABLE tablename

MODIFY column-name datatype;

⑤ T/F The CHAR datatype does not include '(', '+', '*' .

True

⑥ To delete a column, the syntax is _____.

ALTER TABLE tablename

DROP COLUMN column-name;

⑦ The DROP COLUMN is known as a _____ clause.

⑧ The code for deleting a table is _____.

DROP TABLE tablename;

⑨ The code for deleting the data and not the table itself is _____.

TRUNCATE TABLE tablename

IMMEDIATE;

⑩ If IMMEDIATE is used to process the statement immediately and the action cannot be undone.

Retrieving Rows of Data: String Patterns

① If you don't know which value to specify in the predicate, (e.g. not knowing the exact value to search for), we can do _____.

String patterns

② Suppose we know the first name of the author begins with 'R'. The code for using string pattern is _____.

```
SELECT * from tablename
WHERE column LIKE 'R%';
```

③ The '%' serves as a _____ placeholder for other characters.

④ T/F You can put '%' before, middle, and after characters.

True

Range

① For numbers, you can use the operators _____.

1. '>=' ... '<='
2. 'BETWEEN' ... 'AND'

② The syntax to use range are _____.

```
1. SELECT column_name(s) FROM tablename
   WHERE column_name BETWEEN X AND Y;
      (X,Y ∈ R);
```

2. SELECT column_name(s) FROM tablename
WHERE column_name \geq X AND column_name \leq Y;
 $(X,Y \in R)$.

③ T/F Most prefer the first approach because it is easier and quicker to write.

True

A Set of Values

① If you cannot group numbers into a range, you can use _____ in operator

② The syntax for in operator is _____.

SELECT column_name(s) FROM tablename
WHERE column_name IN (value_1, value_2, ..., value_n);

③ The in operator takes a list of expressions to _____ compare against

Sorting Results Set: ORDER BY Clause

① The ORDER BY clause is used in a query to sort the result set by _____.

a specified column

② The syntax for ORDER BY is _____.

SELECT column_name(s) FROM tablename

ORDER BY Column-name;

③ By default, the result set is sorted in _____ order.

ascending

④ The syntax for sorting the result set in descending order is _____.

SELECT column_name(s) FROM tablename
ORDER BY column DESC;

Column Sequence Number

① Another way to specify sort column is to indicate the _____ column sequence number

② Suppose the second column in the table was named pages. The code to sort the result set using column sequence number for pages is _____.

SELECT column_name(s) FROM tablename
ORDER BY 2;

Grouping Result Sets: Eliminating Duplicates, DISTINCT Clause

Grouping Result Sets:

Eliminating Duplicates, DISTINCT Clause

① The DISTINCT clause removes duplicates within a _____ column

② The syntax for using the DISTINCT clause is _____.

SELECT DISTINCT (Column-name)
FROM tablename;

③ To count the number of unique values of a column, use _____.

COUNT function & GROUP BY clause

④ The syntax for GROUP BY clause is _____.

```
SELECT column_name, COUNT(column_name)
FROM tablename GROUP BY column_name;
```

⑤ The GROUP BY clause groups a result into _____.

subsets that has matching values for one or more columns

⑥ Example of the code and output using GROUP BY:

```
db2 => select country, COUNT(country)
          from Author GROUP BY country
Country      2
AU            1
BR            1
CA            3
CN            6
IN            6
RO            3
6 record(s) selected.
```

⑦ T/F First the 'country' is grouped. Then they are counted.

True

⑧ The number '2' is displayed as a column name because the column name is _____.

not directly available in the table

⑨ The second column in the result set was calculated by the _____.

Count function

⑩ We can assign a column name to the result set with this code _____.

```
SELECT column_name, COUNT(column_name)
      AS Column-name_1 FROM tablename
      GROUP BY column_name;
```

⑪ The AS is known as the _____.

AS keyword

Condition on GROUP BY Clause: HAVING Clause

⑫ T/F The WHERE clause works with the entire result set, but the HAVING clause only works with GROUP BY clause.

True

⑬ The syntax for using HAVING clause is _____.

```
SELECT column_name, COUNT(column_name)
      AS Column-name_1 FROM tablename
      GROUP BY column_name HAVING predicate;
```