SQL = Structured Query Language

* Set of query languages that all share same structure and basic commands

SQL is a language that allows you to communicate to your computer to pull or push information from a database

Basic Commands:

SELECT

UPDATE

DELETE  
INSERT

WHERE

R DB MS = Relational Database Management System

The most commonly used query is the select statement

To select one column from a table, say

Select column from table

To select all columns from a table, say

Select \* from table

To Select multiple columns from table, say

Select col1, col2, … coln

SQL is a a declarative language (as opposed to an imperative language)

With Databricks SQL, you can query your data lake, visualize the results, and easily share insights with your team.

To make a multi-line comment, same as Python syntax

Most databases are indexed so searches can be speed up (eliminating unnecessary checks)

Every table has a primary key which is a field whose column values uniquely identify each record in the table

SELECT Statement

Can be used to get data from single table or multiple table (in the latter case using JOINS)

“Select” specifies which columns you want to select data from, the “From” specifies which table(s) you are selecting the columns from.

* Can be used with DISTINCT clause to return all unique rows in a column
* AS statement creates an Alias (for column for duration of query)

COUNT function

Count (x), where x = \* or x= column\_name

* Returns the number of rows for which the column(s) or expression is non-null

WHERE Clause

Specify the rows you want to select

* Logical operators AND / OR can be used
* BETWEEN clause can be used for dates

ORDER BY Clause

Specifies column according to which the table will be ordered   
(Has to have a type equipped with relational operator <)

* Automatically sorts by ASC (only DESC needs to be specified)

LIMIT Clause

Restrict number of records that are returned

MIN( x\_) / MAX ( x) function

Allows selecting the minimum and maximum values from columns

WHERE LIKE

Allows searching of strings - pattern matching in texts

* “%” Symbol is a placeholder for 0 or more character
* ‘\_’ (underscore) can be used to replace characters

RLIKE

* Regular Expressions (own type of language) - ReGex (Finite state machine)
  + Way of looking for Regular patternings across language

GROUP BY

* Groups records that have the same values (in a column) into summary rows

So it seems like normally you can select COLUMNS from a table, (hence the select statement),

But you can also select VALUES from a table (as that is what functions return)

**Note: There are tons of different functions that will be automatically generated (by only typing in a snippet of SQL code)**

There are 1,000,000 records in bp17701850