

INTRODUCTION TO SQL

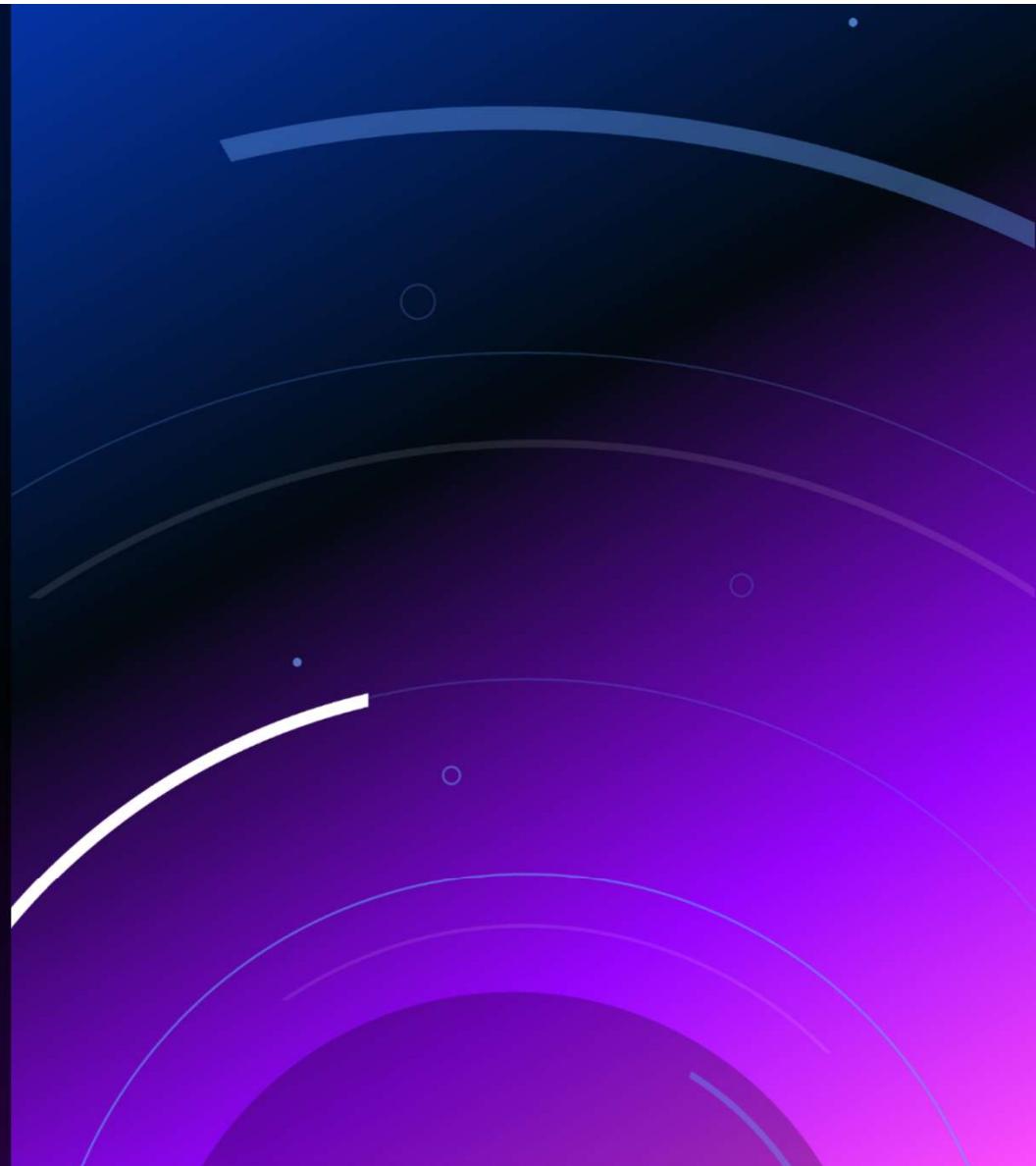
QUANTITATIVE COMPUTING LAB

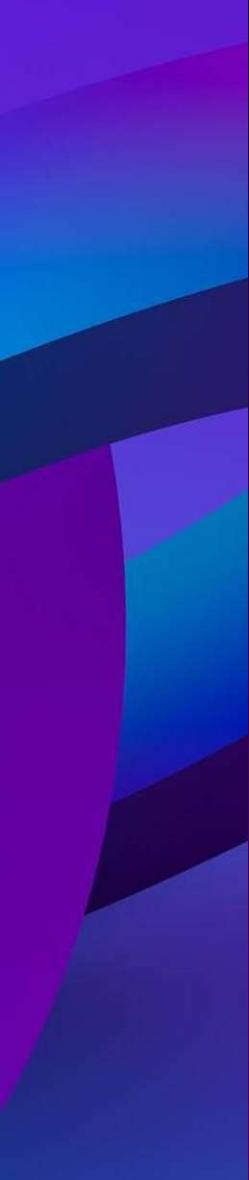
DANIEL PEREZ
GRADUATE FELLOW



After this session you'll be able to:

- Use DBeaver to manage SQL DBMS
- Create a database
- Define table elements
- Import tables and data into your project





AGENDA

- ✓ Databases Overview
- ✓ Database Management Systems
 - ✓ Hands-On
- ✓ Data models
- ✓ Relational Databases
- ✓ SQL
 - ✓ Data transaction Language
 - ✓ Create and import tables
 - ✓ Hands-On

DATABASES

A database is a collection of related data with an implicit meaning.



IMPLICIT PROPERTIES

1. Represents some aspect of the real world
2. Has a logically coherent collection of data with some inherent meaning
3. It has specific purpose.
Preconceived application with a user target.



DATABASE ADVANTAGES

- BETTER DATA INTEGRATION Improves data handling and reduces redundancy
- STORAGE IS MORE SECURE Provides better privacy and security policies
- FASTER DATA ACCESS Produce quick answers to data queries

DATABASE MANAGEMENT SYSTEM (DBMS)

Computerized system that enables users to create, protect, maintain and share a database

Defining

Data Types

Data structures

Data constraints

Constructing

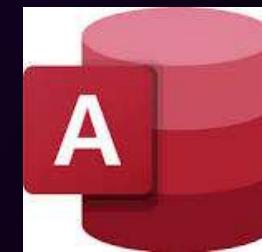


Manipulating

Functions

Querying to retrieve data

Updating to reflect changes



DATABASE MANAGEMENT SYSTEMS

DBMS



Postgres Database



ORACLE

HANDS-ON

DOWNLOAD THE COURSE MATERIALS

GIT Repository

The screenshot shows a GitHub repository page for 'CMC-QCL / Introduction-to-SQL'. The 'Code' tab is selected. A purple box highlights the 'Code' tab in the top navigation bar. Another purple box highlights the 'Download ZIP' button in the 'Clone' section of the right sidebar.

CMC-QCL / Introduction-to-SQL Public

Code Issues Pull requests Actions Projects Security Insights

main · 1 Branch 0 Tags

Go to file Code

Clone

HTTPS GitHub CLI

<https://github.com/CMC-QCL/Introduction-to-SQL>

Clone using the web URL

Open with GitHub Desktop

Download ZIP

File	Description	Last Commit
jedpe Removed MySQL-specific content		
DBeaver Installation.pdf	Added installation	
LICENSE	Initial commit	
README.md	Create README.md	
SQL Workshop.pdf	Removed MySQL-specific	
state_computer_data_insert.sql	Removed MySQL-specific	
state_crime.csv	Added dataset for the workshop	2 years ago
state_people_insert.sql	Added dataset for the workshop	2 years ago
state_workforce.csv	Added dataset for the workshop	2 years ago



<https://dbeaver.io/download/>



DBeaver

DBeaver 24.0.5 - <QCLWorkshop> QCLWorkshop

File Edit Navigate Search SQL Editor Database Window Help

Database Navigator Projects

state_computer_data

```
SELECT * FROM state_computer_data scd
-- Create a table
CREATE TABLE state_computer_data (
    State TEXT,
    Persons_per_household REAL,
    Households_with_computer REAL,
    Households_with_internet REAL
);

-- Insert Values
INSERT INTO state_computer.data
VALUES("Alabama", 2.55, 85.5, 76.4)
```

state_computer_data 1

State	Persons_per_household	Households_with_computer	Households_with_internet
1 Alabama	2.55	85.5	76.4
2 Alaska	2.8	94.1	85.5
3 Arizona	2.68	91.7	84.1
4 Arkansas	2.52	86.2	73
5 California	2.95	93	86.7
6 Colorado	2.56	93.9	87.6
7 Connecticut	2.53	90.8	85.5
8 Delaware	2.57	91.6	85
9 District of Columbia	2.3	91.8	82.6
10 Florida	2.65	91.5	83
11 Georgia	2.7	90.2	81.3
12 Hawaii	3	91.2	84.8
13 Idaho	2.68	91.8	82.7
14 Illinois	2.57	89.9	82.7
15 Indiana	2.52	88.7	80.1
16 Iowa	2.4	89	80.8
17 Kansas	2.51	90	81.8
18 Kentucky	2.49	86.4	78.4
19 Louisiana	2.61	85.6	75.5
20 Maine	2.32	89.7	82.1
21 Maryland	2.67	92.4	86.4
22 Massachusetts	2.52	91.4	86.4
23 Michigan	2.47	89.6	81.5
24 Minnesota	2.49	91.6	84.8
25 Mississippi	2.62	83.8	71.5
26 Missouri	2.46	89	80.2

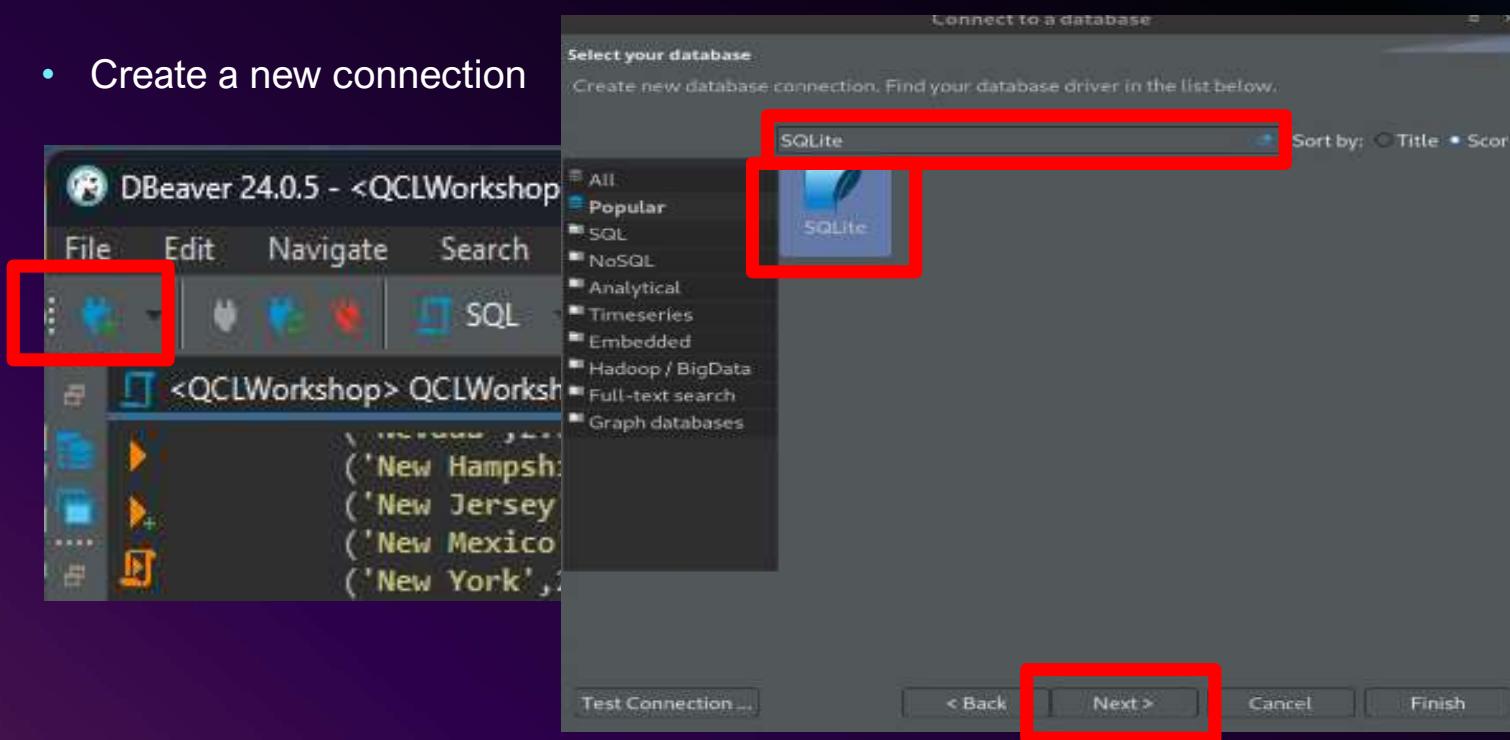
PST en Writable Smart Insert 3:1:12 Set 0|0 ...

UNIVERSAL DATABASE TOOL

DBeaver Community is a free cross-platform **database tool** for developers, database administrators, analysts, and everyone working with data. It supports all popular SQL databases like MySQL, MariaDB, PostgreSQL, SQLite, Apache Family, and more.

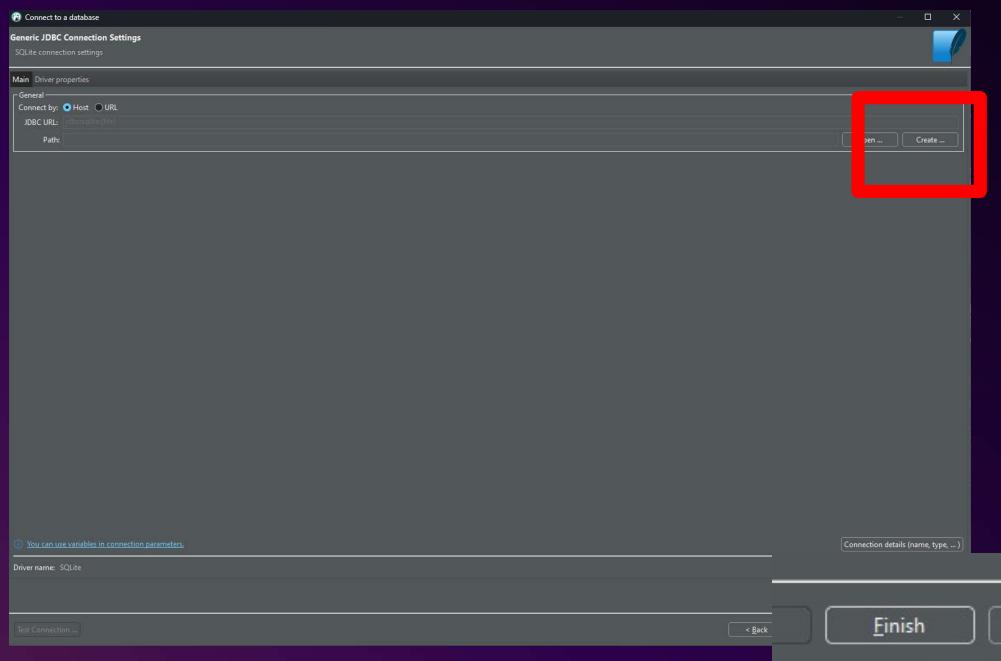
CREATE A CONNECTION

- Create a new connection



CREATE A DATABASE

- Create a new database



The screenshot shows the 'Choose database file' dialog box and the 'Database Navigator' window. The 'Choose database file' dialog shows 'My Documents' as the current folder. The 'Database Navigator' window shows a tree view of the 'QCLWorkshop.db' database, which contains 'Tables', 'Views', 'Indexes', 'Sequences', 'Table Triggers', and 'Data Types'. Under 'Data Types', four items are listed: '123 INTEGER', '123 REAL', '123 NUMERIC', and '123 TEXT'. The number '2' is located in the bottom right corner of the 'Database Navigator' window.

Name	Date modified	Type	Size
state_computer_data_insert	5/19/2024 11:31 PM	DBeaver	2 KB
state_crime	5/19/2024 11:31 PM	Microsoft Excel C...	7 KB
state_people_insert	5/19/2024 11:31 PM	DBeaver	2 KB

DATA MODELS

1. High level or conceptual data models
2. Low-level or physical data models

Relational data model

	CustomerID	CustomerName	Status
Row => Tuples/Records	1	James	Active
	2	Jackie	Active
	3	Robert	Inactive

Column =>
Attribute/field

STORAGE CLASSES AND DATA TYPES

- NULL
- INTEGER
- REAL
- TEXT
- BLOB

Integer	Character	Boolean / Flag	Date and time	Float / Decimal
StudentID	StudentName	StatusActive	DOB	GPA
1	James	1 - True	2002-03-12	2.2
2	Jackie	1 - True	2005-01-19	3.6
3	Robert	0 - False	2000-06-06	3.8

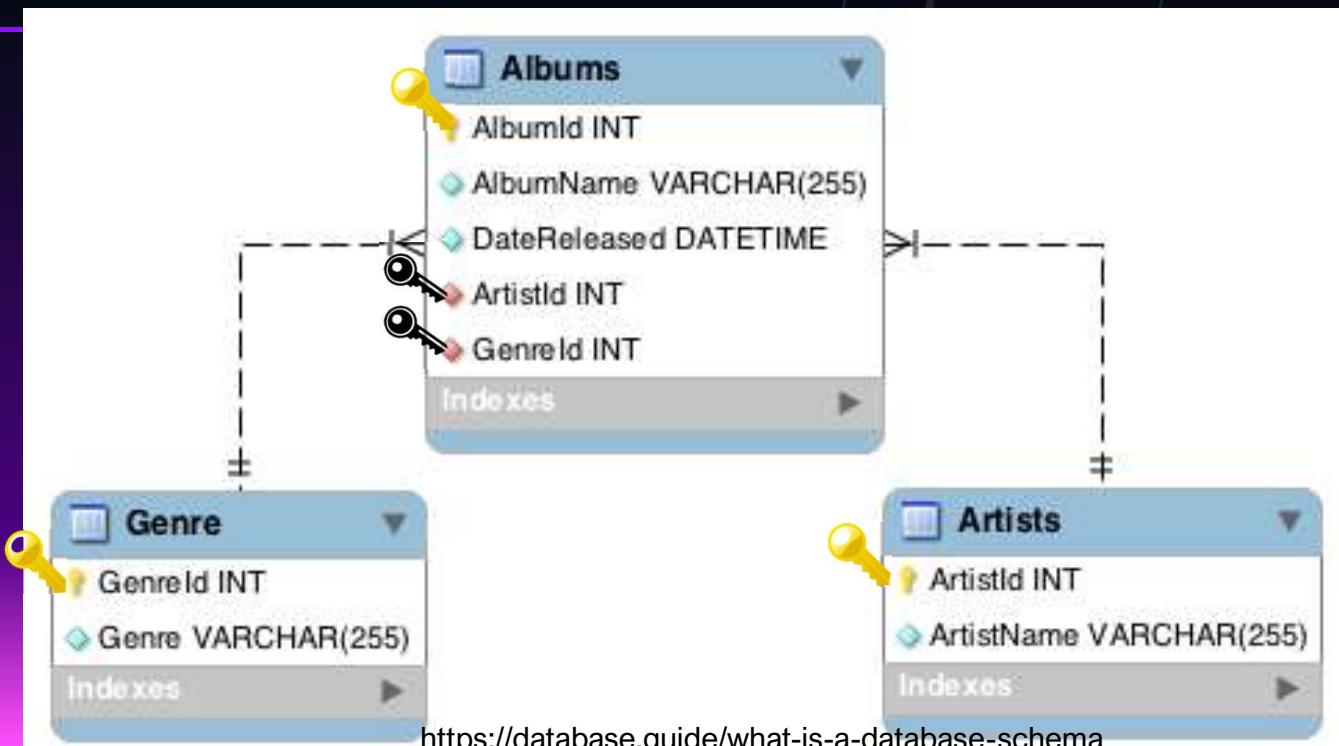
RELATIONAL DATABASES

StudentID	StudentName	Status
1	James	Active
2	Jackie	Active
3	Robert	Inactive

StudentID	ClassID	Semester
1	1005	Summer24
2	1006	Summer23
3	1009	Fall21

ClassID	Title	ClassNum
1005	Intro to Art History	500
1006	Intro to SQL	501
1009	Intro to Databases	300

DATABASE SCHEMA



TODAY'S DATA

State Crime CSV

- Information on the crime rates and totals for states across the United States for a wide range of years
- Reports go from 1960 to 2019 (only used 2010, 2014 and 2019)
- https://github.com/CMC-QCL/Introduction-to-SQL/blob/main/state_crime.csv
<<https://corgis-edu.github.io>>

State Demographics CSV and SQL Files

- state_computer_data.sql, state_workforce.csv, state_people.sql
- Summarized information obtained about state demographics in the US between 2015 - 2019 via the United States Census Bureau
- Just the summarized data as of 2019
- https://corgis-edu.github.io/corgis/csv/state_demographics/

SQL

STRUCTURED QUERY LANGUAGE

QUERY

An application program accesses the database by sending queries for the data to the DBMS

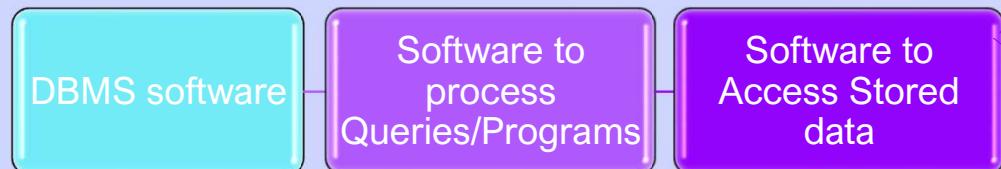


<Conceptual level>

Database System

- Application Programs/Queries

<Internal schema>



<The three-schema architecture>



<External level>



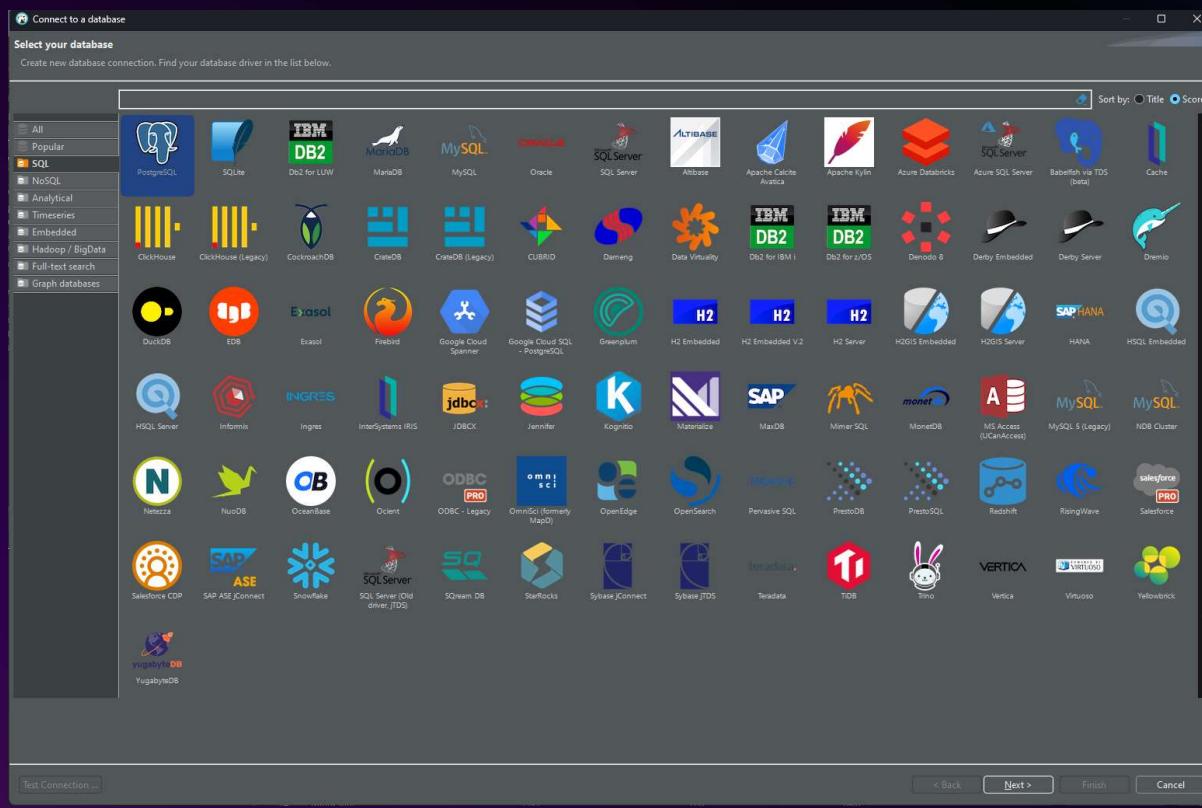
STRUCTURED QUERY LANGUAGE

(SQL) is a programming language for storing and processing information in a relational database.

Why do we use it?

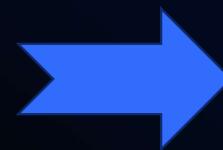
- Integrates well with different programming languages
- Embed SQL queries with Java to build high-performing data processing applications
- Fairly easy to learn as it uses common English keywords in its statements

SQL SYNTAX



DATA TRANSACTION LANGUAGE

- Definition (DDL)
 - CREATE, DROP, ALTER, TRUNCATE
- Manipulation (DML)
 - INSERT, UPDATE, DELETE
- Query (DQL)
 - SELECT, JOIN
- Data Control or transaction control (DCL/TCL)
 - GRANT, REVOKE



Create
Read
Update
Delete

DATA DEFINITION & CONSTRAINTS

DEFINE SQL DATA TYPES

COLUMN<ATTRIBUTE> DATA_TYPE,

>>FOR STATE COMPUTER DATA

PERSONS_PER_HOUSEHOLD INT,
HOUSEHOLDS_WITH_COMPUTER INT,
HOUSEHOLDS_WITH_INTERNET INT

DEFINE KEYS

COLUMN DATA_TYPE PRIMARY KEY

DEFINE CONSTRAINTS FOR DATA TYPES

STATE VARCHAR(50) PRIMARY KEY,
PERSONS_PER_HOUSEHOLD INT CHECK
(PERSONS_PER_HOUSEHOLD > 0 AND
PERSONS_PER_HOUSEHOLD <= 10),
HOUSEHOLDS_WITH_COMPUTER DECIMAL(10,2),
HOUSEHOLDS_WITH_INTERNET DECIMAL(10,2)

FOLLOW LOCALLY

CREATE COMMENTS

```
-- THIS IS A SINGLE LINE COMMENT
```

```
/* [QCL WORKSHOP]
* INTRODUCTION TO SQL
* THIS IS A MULTILINE COMMENT
*/
```

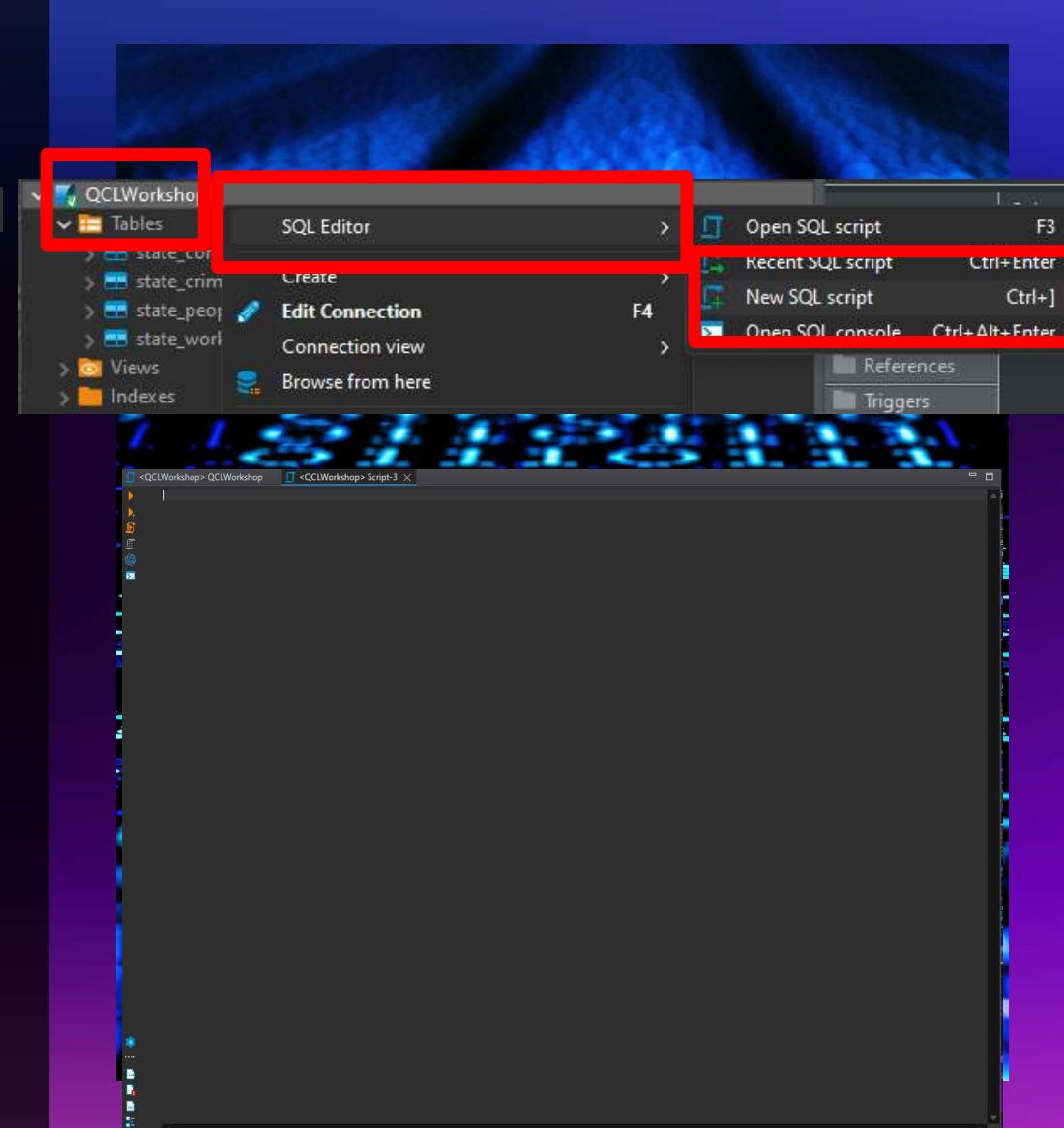
CREATE TABLE

```
CREATE TABLE
TEST_TABLE(DATA_DEFINITION)
```

```
CREATE TABLE TEST_TABLE(
TEST_COLUMN TEXT PRIMARY KEY
)
```

DROP TABLE

```
DROP TABLE TEST_TABLE
```

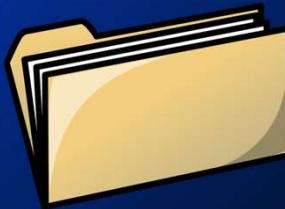


Create the table state_computer_data

Name	Value
Updated Rows	50
Query	<pre>CREATE TABLE state_computer_data (State TEXT PRIMARY KEY, Persons_per_household DECIMAL(10,2), Households_with_computer DECIMAL(10,2), Households_with_internet DECIMAL(10,2))</pre>
Start time	Mon Jun 10 18:26:34 PDT 2024
Finish time	Mon Jun 10 18:26:34 PDT 2024

INSERT VALUES

```
-- INSERT VALUES
INSERT INTO STATE_COMPUTER_DATA
VALUES("ALABAMA", 2.55, 85.5,
76.4)
```



INSERT MULTIPLE VALUES

```
-- STATE COMPUTER DATA INSERT
INSERT INTO STATE_COMPUTER_DATA
(STATE, PERSONS_PER_HOUSEHOLD, HOUSEHOLDS_WITH_COMPUTER, HOUSEHOLDS_WITH_INTERNET)
VALUES ('ALASKA', 2.8, 94.1, 85.5),
('ARIZONA', 2.68, 91.7, 84.1),
('ARKANSAS', 2.52, 86.2, 73.0),
('CALIFORNIA', 2.95, 93.0, 86.7),
('COLORADO', 2.56, 93.9, 87.6),
('CONNECTICUT', 2.53, 90.8, 85.5),
('DELAWARE', 2.57, 91.6, 85.0),
('DISTRICT OF COLUMBIA', 2.3, 91.8, 82.6),
('FLORIDA', 2.65, 91.5, 83.0),
('GEORGIA', 2.7, 90.2, 81.3),
('HAWAII', 3.0, 91.2, 84.8),
('IDAHO', 2.68, 91.8, 82.7),
('ILLINOIS', 2.57, 89.9, 82.7),
('INDIANA', 2.52, 88.7, 80.1),
('IOWA', 2.4, 89.0, 80.8),
('KANSAS', 2.51, 90.0, 81.8),
('KENTUCKY', 2.49, 86.4, 78.4),
('LOUISIANA', 2.61, 85.6, 75.5),
('MAINE', 2.32, 89.7, 82.1),
('MARYLAND', 2.67, 92.4, 86.4),
('MASSACHUSETTS', 2.52, 91.4, 86.4),
('MICHIGAN', 2.47, 89.6, 81.5),
('MINNESOTA', 2.49, 91.6, 84.8),
('MISSISSIPPI', 2.62, 83.8, 71.5),
('MISSOURI', 2.46, 89.0, 80.2),
('MONTANA', 2.39, 88.9, 80.7),
('NEBRASKA', 2.45, 90.0, 83.4),
('NEVADA', 2.67, 92.5, 83.2),
('NEW HAMPSHIRE', 2.46, 93.0, 87.7),
('NEW JERSEY', 2.69, 91.4, 85.8),
('NEW MEXICO', 2.63, 85.9, 74.6),
('NEW YORK', 2.59, 89.6, 82.8),
('NORTH CAROLINA', 2.52, 89.1, 80.7),
('NORTH DAKOTA', 2.3, 89.8, 80.7),
('OHIO', 2.43, 89.1, 82.0),
('OKLAHOMA', 2.58, 88.6, 78.6),
('OREGON', 2.51, 93.0, 85.9),
('PENNSYLVANIA', 2.45, 88.0, 81.5),
('RHODE ISLAND', 2.47, 89.1, 84.0),
('SOUTH CAROLINA', 2.54, 88.3, 78.2),
('SOUTH DAKOTA', 2.43, 88.5, 80.7),
('TENNESSEE', 2.52, 87.1, 78.4),
('TEXAS', 2.85, 91.0, 81.9),
('UTAH', 3.12, 95.3, 87.5),
('VERMONT', 2.3, 89.9, 81.5),
('VIRGINIA', 2.61, 91.1, 83.9),
('WASHINGTON', 2.55, 93.8, 88.3),
('WEST VIRGINIA', 2.42, 84.2, 76.0),
('WISCONSIN', 2.39, 89.4, 82.5),
('WYOMING', 2.46, 91.8, 83.4);
```

*<QCLWorkshop> QCLWorkshop state_computer_data

Properties Data ER Diagram

state_computer_data Enter a SQL expression to filter results (use Ctrl+Space)

Grid Text Record

	State	Persons_per_household	Households_with_computer	Households_with_internet
1	Alabama	2.55	85.5	76.4
2	Alaska	2.8	94.1	85.5
3	Arizona	2.68	91.7	84.1
4	Arkansas	2.52	86.2	73
5	California	2.95	93	86.7
6	Colorado	2.56	93.9	87.6
7	Connecticut	2.53	90.8	85.5
8	Delaware	2.57	91.6	85
9	District of Colur	2.3	91.8	82.6
10	Florida	2.65	91.5	83
11	Georgia	2.7	90.2	81.3
12	Hawaii	3	91.2	84.8
13	Idaho	2.68	91.8	82.7
14	Illinois	2.57	89.9	82.7
15	Indiana	2.52	88.7	80.1
16	Iowa	2.4	89	80.8
17	Kansas	2.51	90	81.8
18	Kentucky	2.49	86.4	78.4
19	Louisiana	2.61	85.6	75.5
20	Maine	2.32	89.7	82.1
21	Maryland	2.67	92.4	86.4
22	Massachusetts	2.52	91.4	86.4
23	Michigan	2.47	89.6	81.5
24	Minnesota	2.49	91.6	84.8
25	Mississippi	2.62	83.8	71.5
26	Missouri	2.46	89	80.2
27	Montana	2.39	88.9	80.7
28	Nebraska	2.45	90	83.4
29	Nevada	2.67	92.5	83.2
30	New Hampshire	2.46	93	87.7
31	New Jersey	2.69	91.4	85.8
32	New Mexico	2.63	85.9	74.6
33	New York	2.59	89.6	82.8
34	North Carolina	2.52	89.1	80.7
35	North Dakota	2.3	89.8	80.7
36	Ohio	2.43	89.1	82
37	Oklahoma	2.58	88.6	78.6
38	Oregon	2.51	93	85.9
39	Pennsylvania	2.45	88	81.5
40	Rhode Island	2.47	89.1	84
41	South Carolina	2.54	88.3	78.2

Refresh Save Cancel Export data 200 51 51 row(s) fetched - 0.001s, on 2024-06-09 at 23:17:08

HANDS-ON

CREATE TABLE AND INSERT DATA

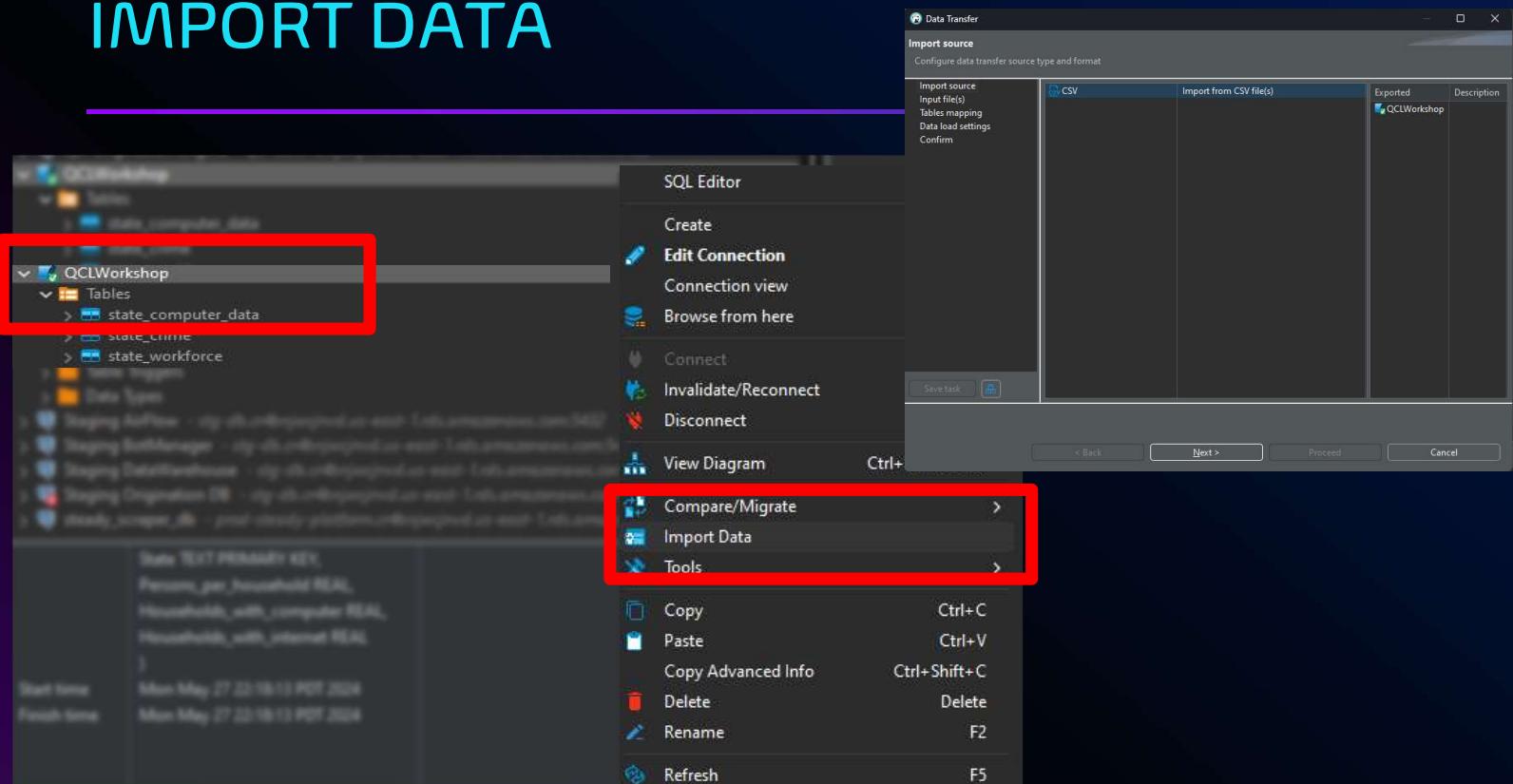
Create the table defined in the schema:

state_people	
ABC	state
123	employment_firms_total
123	age_percent_under_18_years
123	age_percent_65_and_older

Insert data into the new table with the file
`state_people_insert.sql`

- Right click on database >Import Data
- Import source
- Select Input file
- Data load settings
- Confirmation

IMPORT DATA



 Data Transfer

Import source
Configure data transfer source type and format

Import source
Input file(s)
Tables mapping
Data load settings
Confirm

CSV Import from CSV file(s) Exported Description

QCLWorkshop

Save task 

< Back Next > Proceed Cancel

*<QCLWorkshop> QCLWorkshop state_computer_data state_people state_workforce

Properties Data ER Diagram

state_workforce | Enter a SQL expression to filter results (use Ctrl+Space)

Grid	State	Population_Percent_Change	Population_2014	Population_2010	Education_High_School_or_Higher	Education
1	Alabama	2.6	5,024,279	4,779,736	86.2	
2	Alaska	3	733,391	710,231	92.8	
3	Arizona	13.9	7,151,502	6,392,017	87.1	
4	Arkansas	3.5	3,011,524	2,915,918	86.6	
5	California	6.1	39,538,223	37,253,956	83.3	
6	Colorado	14.5	5,773,714	5,029,196	91.7	
7	Connecticut	-10.2	3,605,944	3,574,097	90.6	
8	Delaware	8.4	989,948	897,934	90	
9	District of Columbia	17.3	689,545	601,723	90.9	
10	Florida	14.2	21,538,187	18,801,310	88.2	
11	Georgia	9.6	10,711,908	9,687,653	87.1	
12	Hawaii	4.1	1,455,271	1,360,301	92	
13	Idaho	14	1,839,106	1,567,582	90.8	
14	Illinois	-11.2	12,812,508	12,830,632	89.2	
15	Indiana	3.8	6,785,528	6,483,802	88.8	
16	Iowa	3.6	3,190,369	3,046,355	92.1	
17	Kansas	2.1	2,937,880	2,853,118	91	
18	Kentucky	3	4,505,836	4,339,367	86.3	
19	Louisiana	2.5	4,657,757	4,533,372	85.2	
20	Maine	1.2	1,362,359	1,328,361	92.6	
21	Maryland	4.7	6,177,224	5,773,552	90.2	
22	Massachusetts	5.3	7,029,917	6,547,629	90.8	
23	Michigan	1	10,077,331	9,883,640	90.8	
24	Minnesota	6.3	5,706,494	5,303,925	93.1	
25	Mississippi	0.3	2,961,279	2,967,297	84.5	
26	Missouri	2.5	6,154,913	5,988,927	89.9	
27	Montana	8	1,084,225	989,415	93.6	
28	Nebraska	5.9	1,961,504	1,826,341	91.4	
29	Nevada	14.1	3,104,614	2,700,551	86.7	
30	New Hampshire	3.3	1,377,529	1,316,470	93.1	
31	New Jersey	1	9,288,994	8,791,894	89.8	
32	New Mexico	1.8	2,117,522	2,059,179	85.6	
33	New York	0.4	20,201,249	19,378,102	86.8	
34	North Carolina	10	10,439,388	9,535,483	87.8	
35	North Dakota	13.3	779,094	672,591	92.6	
36	Ohio	1.3	11,799,448	11,536,504	90.4	
37	Oklahoma	5.5	3,959,353	3,751,351	88	
38	Oregon	10.1	4,237,256	3,831,074	90.7	
39	Pennsylvania	0.8	13,002,700	12,702,379	90.5	
40	Rhode Island	0.6	1,097,379	1,052,567	88.8	

HANDS-ON

IMPORT FILES

Import file named state_workforce

Import file named state_crime

What are the table's attributes?

- ❖ Data Types
- ❖ Constraints
- ❖ Keys

state_computer_data	
ABC	State
123	Persons_per_household
123	Households_with_computer
123	Households_with_internet

state_people	
ABC	State
123	employment_firms_total
123	age_percent_under_18_years
123	age_percent_65_and_older

state_crime	
ABC	State
123	Crime_Year
123	Population
123	Rates_Property_Theft
123	Rates_Violent_Robbery
123	Totals_Property_Theft
123	Totals_Violent_Robbery

state_workforce	
ABC	State
123	Population_Percent_Change
123	Population_2014
123	Population_2010
123	Education_High_School_or_Higher
123	Education_Bachelors_Degree_or_Higher
123	Sales_Retail_Sales
123	Mean_Travel_Time_to_Work
123	Income_Median_Household_Income
123	Income_Per_Capita_Income
123	Income_Persons_Below_Poverty_Level