

Module overview: SQL

# Getting set up for SQL

# Overview

In this module, we'll be using:



MySQL Workbench

In the next lesson, **Database concepts**, we'll go through the **steps of downloading, installing, and launching** MySQL Workbench and MySQL Community Server on **Windows** and **Mac**.

The various **CSVs, SQL files, and queries** that we'll need to create and transform our database are provided with this document. However, they can also be **downloaded from the walk-throughs** at the point we will need them.

Throughout the module, we will need to use the provided CSVs, SQL files, and queries to create or recreate our database. We will look at **how we use** these files and queries.

# Overview

In this module, we'll also be using:



ANACONDA.



jupyter

In the **SQL in production** lesson, we'll introduce Jupyter Notebooks and go through the **steps of downloading, installing, and launching** Anaconda and Jupyter Notebooks on **Windows** and **Mac**.

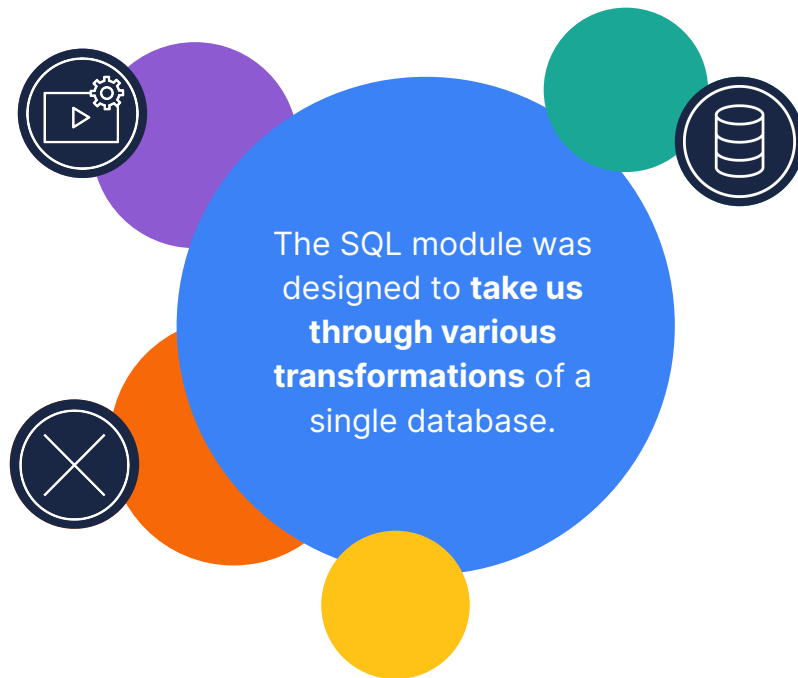
Included in the lessons beyond this point are **[Walk-through]s** in which queries are run in **MySQL Workbench**, as well as the accompanying **[Notebook]s** in which we can execute the same queries in **Jupyter Notebooks**.

Throughout the module, we will be able to interact with and manipulate our database **the same way** using either **MySQL Workbench** or **Jupyter Notebooks**.

# Completing the practical learning activities

The **[Walk-through]** learning activities will teach us **how to query** and **transform our database**.

In order for our **queries to work**, we'll need to **follow along** and **complete all** of the walk-through activities in either **MySQL Workbench or Jupyter Notebooks**.



We'll need to **use the same database** throughout this module and **only recreate** the dataset **when prompted**.

If we do not follow these instructions, many of our queries **may not seem to work**.

# The effect of database transformations

*For example, we will learn how to remove all data from a table in SQL.*

Once we've **removed the data** from the table, any query thereafter will **return an empty table**. In these cases, we'll be prompted with the necessary steps to recreate the database. It's crucial to **follow these instructions closely** so that future queries work.

The screenshot shows the MySQL Workbench interface. The 'Query 1' tab is active, displaying the following SQL code:

```
3 TRUNCATE
4   united_nations.Access_to_Basic_Services;
5
6 SELECT *
7 FROM
8   united_nations.Access_to_Basic_Services
9 LIMIT 100;
```

The 'Result Grid' is empty, showing the following columns: Region, Sub\_region, Country\_name, Time\_period, Pct\_managed\_drinking\_water\_servi..., Pct\_managed\_sanitation\_servi..., and Est\_populatio... The table name 'Access\_to\_Basic\_Services' is visible in the 'Table: Access\_to\_Basic\_Services' section. The 'Action Output' section shows the execution of the query, with a response of '0 row(s) returned' and a duration of '0.0027 sec / 0.000000'.

# The effect of database transformations

For example, we want to divide a larger dataset into smaller tables and link them using relationships.

The **walk-throughs** often build on one another, so if we don't complete them in order, we may end up in a situation like this:

We created the

**Geographic\_location** table in a previous walk-through, and therefore, if we **skipped** it, we would get an **error**.

The screenshot shows a database management interface with a 'SCHEMAS' panel on the left and a 'Query 1' editor on the right. The 'Access\_to\_Basic\_Services' table is highlighted in the schemas panel. The SQL query in the editor is as follows:

```

1 CREATE TABLE united_nations.Basic_services (
2   Country_name VARCHAR(37),
3   Time_period INTEGER,
4   Pct_managed_drinking_water_services NUMERIC(5,2),
5   Pct_manage_sanitation_services NUMERIC(5,2),
6   PRIMARY KEY (Country_name, Time_period),
7   FOREIGN KEY (Country_name) REFERENCES Geographic_Location (Country_name)
8 );

```

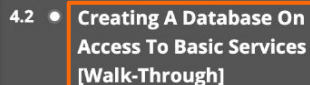
The 'Action Output' panel at the bottom shows the execution results of the query. The first 15 steps completed successfully, but the final step (16) failed with an error.

	Time	Action	Response
✓ 1	13:10:23	TRUNCATE TABLE united_nations.Acc...	0 row(s) affected
✓ 2	13:10:52	TRUNCATE TABLE united_nations.Acc...	0 row(s) affected
✓ 3	13:10:52	SELECT * FROM united_nations.Acces...	0 row(s) returned
✓ 4	13:14:15	USE united_nations	0 row(s) affected
✓ 5	13:14:42	SHOW SESSION VARIABLES LIKE 'lowe...	OK
✓ 6	13:14:42	SHOW DATABASES	OK
✓ 7	13:14:45	SHOW SESSION VARIABLES LIKE 'lowe...	OK
✓ 8	13:14:45	SHOW COLUMNS FROM 'united_natio...	OK
✓ 9	13:14:48	PREPARE stmt FROM 'INSERT INTO 'u...	OK
✓ 10	13:14:49	DEALLOCATE PREPARE stmt	OK
✓ 11	13:15:19	USE united_nations	0 row(s) affected
✓ 12	13:15:19	SELECT * FROM united_nations.Acces...	100 row(s) returned
✓ 13	13:18:38	USE united_nations	0 row(s) affected
✓ 14	13:18:38	TRUNCATE united_nations.Access_to...	0 row(s) affected
✓ 15	13:18:38	SELECT * FROM united_nations.Acces...	0 row(s) returned
✗ 16	17:16:44	CREATE TABLE united_nations.Basic_s...	Error Code: 1824. Failed to open the referenced table 'Geographic_Locati...

The error message in the final row indicates that the table 'Geographic\_Location' referenced in the foreign key does not exist, which is the cause of the failure.

# When and how to download a new dataset

As we transform our database, we will need a **CSV** or new **SQL file** to “reset” our database to a usable format. We can find these files by clicking on the popup in the walk-through video.



4.2 • Creating A Database On Access To Basic Services [Walk-Through]

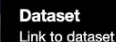
4.3 ○ Data Query Language (DQL) [Slides]

4.4 ○ Useful SQL Tools [Slides]

4.5 ○ SELECT And SELECT WHERE [Walk-Through]

4.6 ○ Comparison Operators

4.8 ○ Using Logical And Comparison Operators [Walk-Through]



Dataset  
Link to dataset

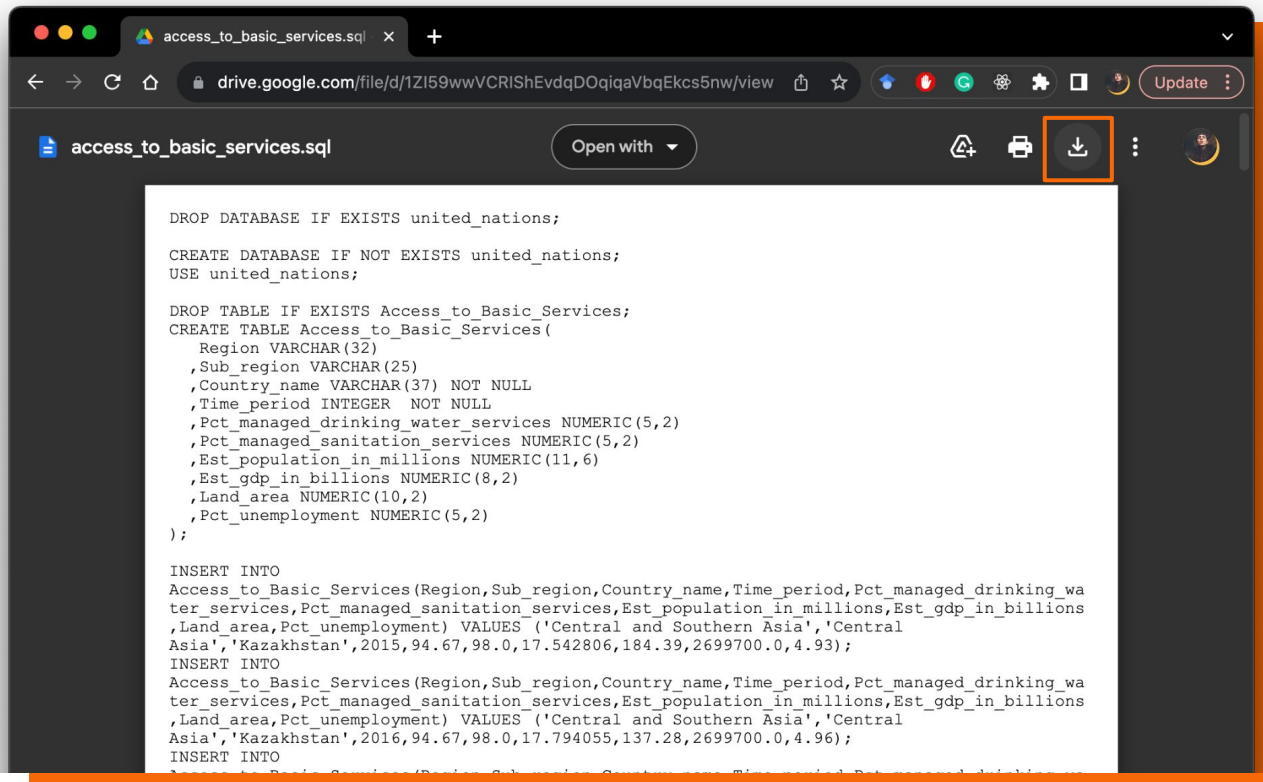


Only **some walk-through videos** will include a dataset or file download since we will need to **persist some of the changes** to the database for the learning activities that follow.

# Download CSVs and SQL files

Clicking on the Dataset popup on a walk-through will **open the file in a new tab**.

We can click on the **download button**, which will **save the file to our local machine**, very often to a Downloads folder.





## Using notebook files and dependencies

*We will receive database files that we need to query in a notebook.*

The notebook environment will **look for the database file in the current folder\***. If the file is not in the same folder as the notebook, the system won't be able to locate it and we will get a "no such table" or similar error message. We need to **make sure we have all the notebook dependencies saved in the same folder as the notebook.**

Basic SQL queries in a notebook [Exercise]

Last Checkpoint: 10 minutes ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted

Python 3 (ipykernel)

Code

First, let's load our sample database:

```

In [1]: 1 # load and activate the SQL extension to allows us to execute SQL in a Jupyter notebook
        2 %load_ext sql

In [2]: 1 # load the Chinook database stored in your local machine.
        2 #Make sure the file is saved in the same folder as this notebook.
        3 %sql sqlite:///chinook.db

In [3]: 1 %%sql
        2
        3 SELECT FirstName
        4 FROM customers
        5 LIMIT 10;
  
```

```

* sqlite:///chinook.db
(sqlite3.OperationalError) no such table: customers
[SQL: SELECT FirstName
FROM customers
LIMIT 10;]
(Background on this error at: https://sqlalche.me/e/20/e3q8)
  
```

Here is a [view](#) of all of our tables in the database:

# Using notebook files and dependencies

*It is possible to specify the correct file path in our code if the file is not saved in the same folder.*

Should the **database be saved in a different folder** we will need to use the absolute or relative file path to locate the folder.

- We use three slashes **sqlite:///** in order to use a relative path.
- We use four slashes **sqlite://** in order to use an absolute path.



```
jupyter Basic SQL queries in a notebook [Exercise] Last Checkpoint: Last Friday at 8:19 AM (unsaved changes) Python 3 (ipykernel)
```

File Edit View Insert Cell Kernel Widgets Help

First, let's load our sample database:

```
In [1]: 1 # load and activate the SQL extension to allows us to execute SQL in a Jupyter notebook
        2 %load_ext sql
```

Absolute file path

```
In [ ]: 1 # load the Chinook database stored in your local machine.
        2 # Use the absolute file path if the file is not stored in the same folder as the notebook
        3 %sql sqlite:///Users/anton/Downloads/Module_3_SQL/Exercises/chinook.db
```

Relative file path

```
In [2]: 1 # Use the relative file path if the file is not stored in the same folder as the notebook
        2 %sql sqlite:///Module_3_SQL/Exercises/chinook.db
```

```
In [3]: 1 %%sql
        2
        3 SELECT FirstName
        4 FROM customers
        5 LIMIT 10;
```

\* sqlite:///Databases/chinook.db  
Done.

```
Out[3]:
```

FirstName
Luis
Leonie
François
Bjorn