

Step 1: First, we must have the normalized tables put into csv files. We can do this by adding this to the end of the ipynb file. For example, df_provider is one of the normalized tables I have so this will be a new csv file.

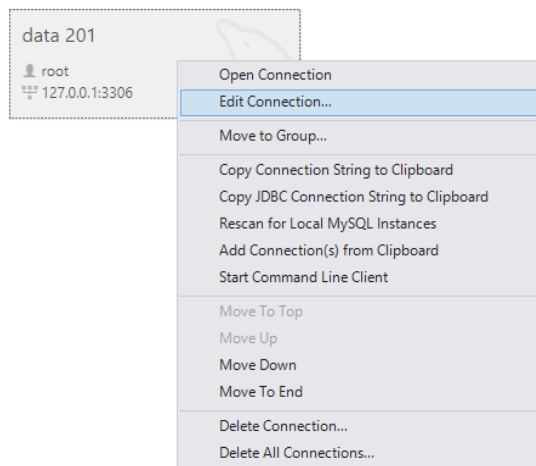
```
# Export each DataFrame to CSV
df_provider.to_csv('provider.csv', index=False)
df_address.to_csv('provider_address.csv', index=False)
df_classification.to_csv('provider_classification.csv', index=False)
df_service_totals.to_csv('service_totals.csv', index=False)
df_drug_services.to_csv('drug_services.csv', index=False)
df_medical_services.to_csv('medical_services.csv', index=False)
df_bene_demo.to_csv('beneficiary_demographics.csv', index=False)
df_bene_cc.to_csv('beneficiary_conditions.csv', index=False)

print("All normalized tables exported successfully!")
```

This will take a bit to run but afterwards these csv files will be in the directory your ipynb file is.

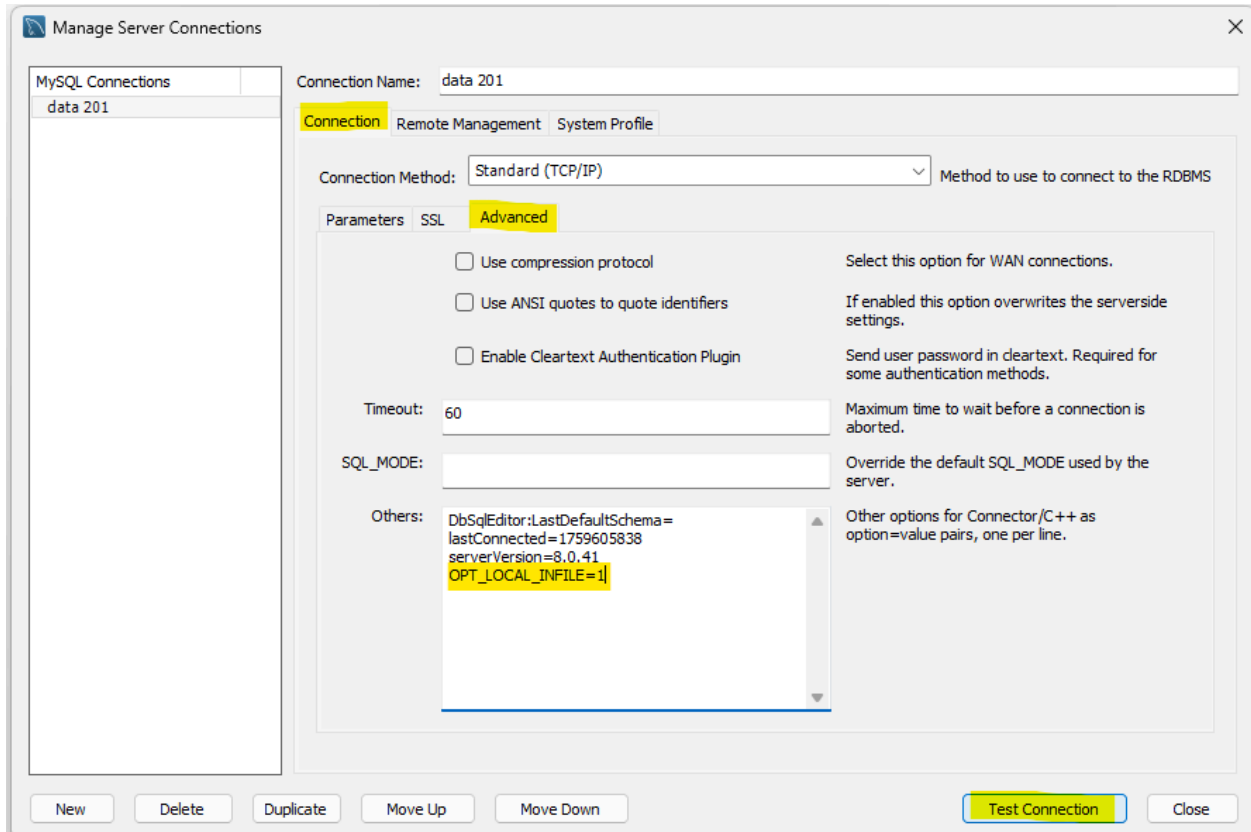
Step 2: Open MySQL and edit the connection you are working in to allow the ability to load in csv files. Right click on the connection you are using, and click on the edit connection option.

MySQL Connections ⓘ ⓘ

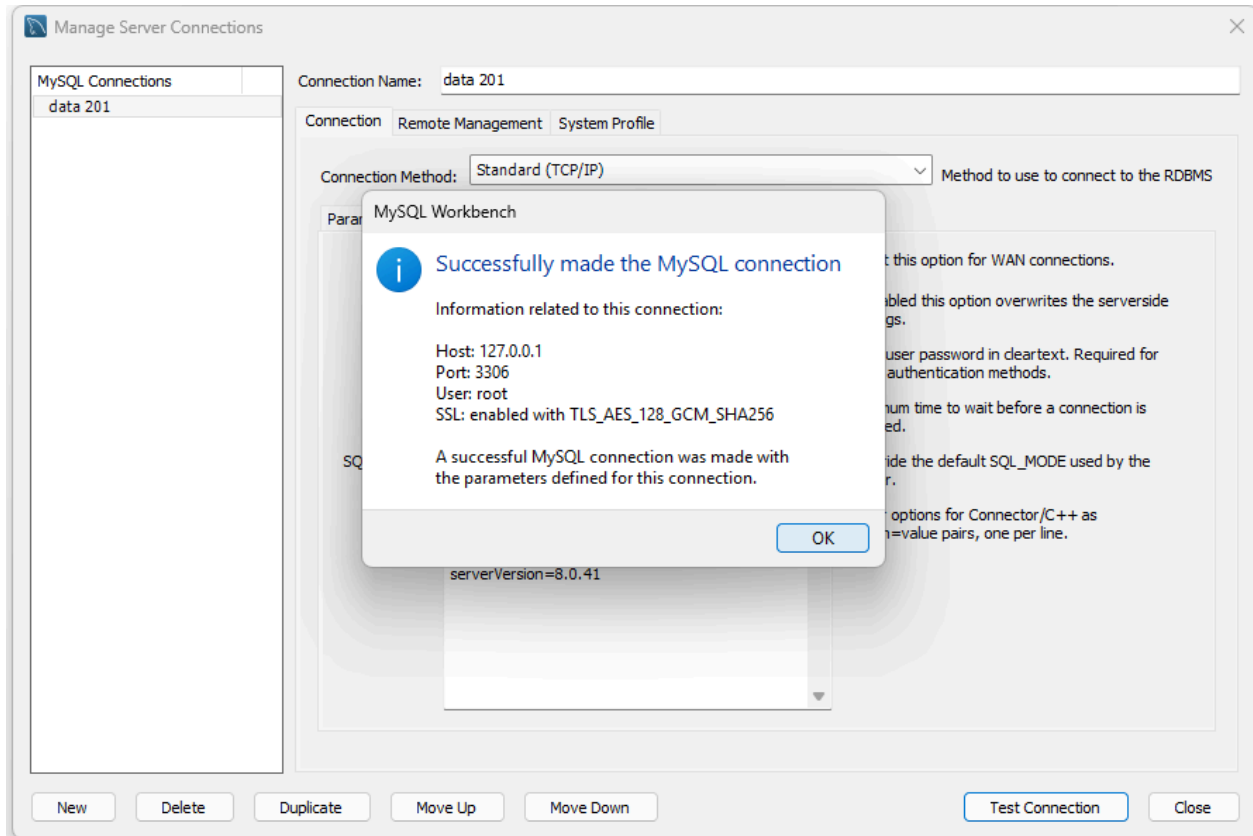


Step 3: Click on the Advanced tab and add the following to the Others section. This allows us to directly import csv file data to our database

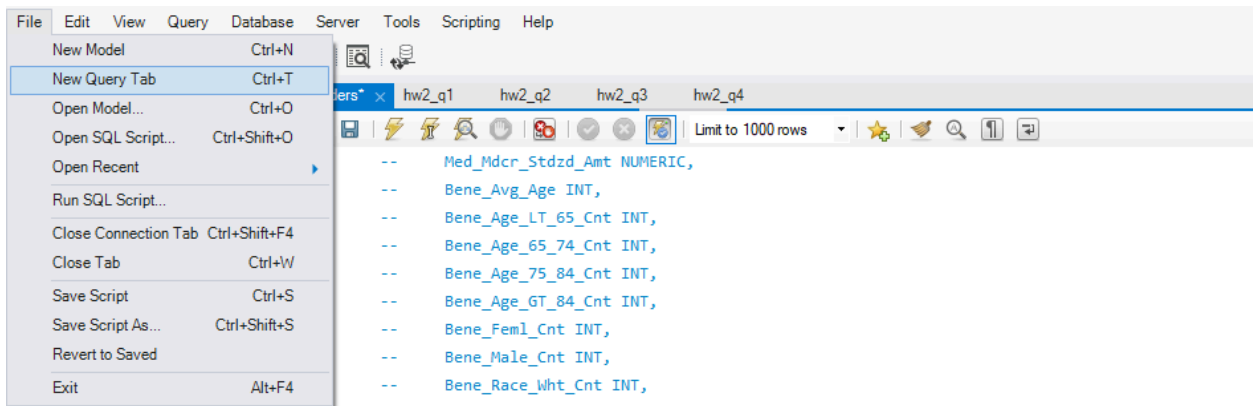
```
OPT_LOCAL_INFILE=1
```



Step 4: After adding that you must click on the Test Connection Button. The following prompt should pop up afterwards and that means it was successful. You can click Okay afterwards and close the window.



Step 5: Open the connection like usual by double clicking on it and open a new Query tab.



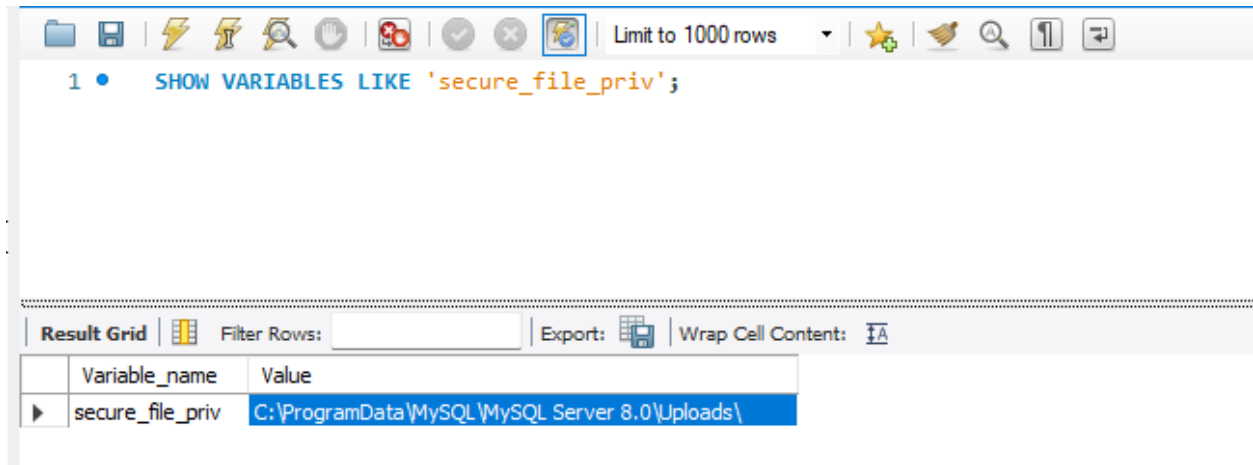
Step 6: On this query tab, you will create all the tables and database you will use (we will be using the same database for all tables from your ipynb file and the other datasets)

For example:

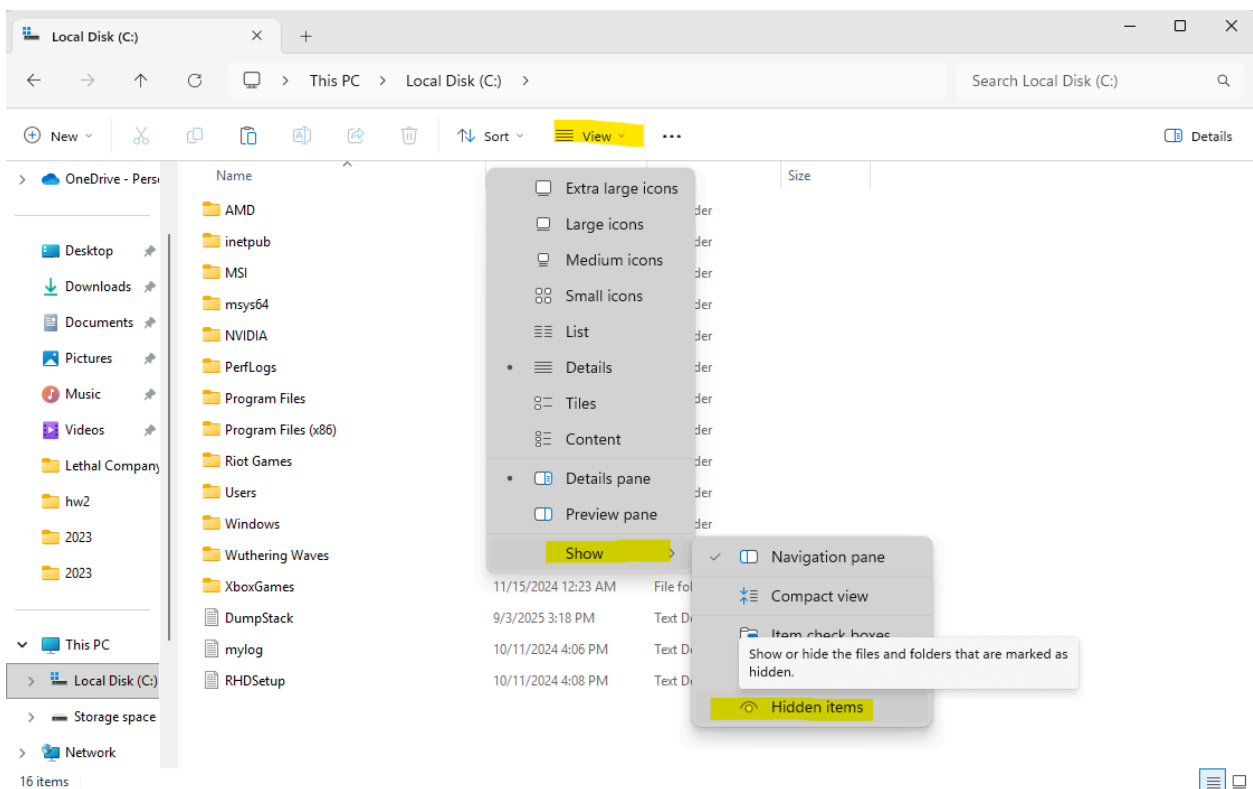
```
1 • DROP DATABASE IF EXISTS cms;
2 • CREATE DATABASE cms;
3 • USE cms;
4
5 • CREATE TABLE provider (
6     Rndrng_NPI INT,
7     Rndrng_Privr_Last_Org_Name TEXT,
8     Rndrng_Privr_First_Name TEXT,
9     Rndrng_Privr_MI TEXT,
10    Rndrng_Privr_Crdntls TEXT,
11    Rndrng_Privr_Ent_Cd TEXT,
12    Year INT,
13    PRIMARY KEY (Rndrng_NPI, Year)
14 );
15
16 • CREATE TABLE provider_address (
17     Rndrng_NPI INT,
18     Rndrng_Privr_St1 TEXT,
19     Rndrng_Privr_St2 TEXT,
20     Rndrng_Privr_City TEXT,
21     Rndrng_Privr_State_Abrvtn TEXT,
22     Rndrng_Privr_State_FIPS TEXT,
23     Rndrng_Privr_Zip5 TEXT,
24     Rndrng_Privr_Cntry TEXT,
25     Year INT,
26     PRIMARY KEY (Rndrng_NPI, Year)
27 );
```

Step 7: Create another new Query tab and Copy and paste this line and run it to get the path we will be pasting the csv files with the normalized tables. You will get the following output.

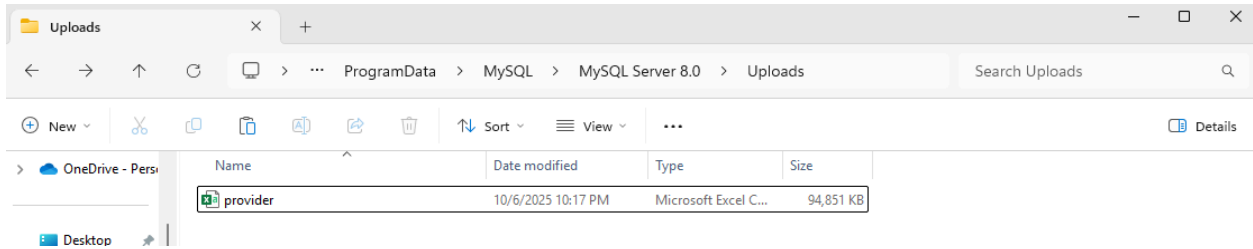
```
SHOW VARIABLES LIKE 'secure_file_priv';
```



Step 8: Go to file explorer and make sure you can view hidden files by clicking on View→Show→Hidden items. This makes it so you can access the Program Data file if it is not there before.



Step 9: Copy and paste all csv files with normalized tables on the file location path we got before. Program Data→MySQL→MySQL Server 8.0 →Uploads



Step 9: Comment that line out and copy and paste the following code. This will import data from csv file to the table we created.

```
USE cms; -- this is the name of the database use previous
SET GLOBAL local_infile = 1;
LOAD DATA LOCAL INFILE 'C:\\ProgramData\\MySQL\\MySQL Server
8.0\\Uploads\\provider.csv' -- this is the path for the file
INTO TABLE provider -- make sure correct csv file data is loaded to
the correct table
FIELDS TERMINATED BY ','
ENCLOSED BY '"'
LINES TERMINATED BY '\n'
IGNORE 1 ROWS; -- ignores first row which is the column names
```

After running it you should get that it successfully has imported data like the following screenshot. Change the table name and csv name you are importing to repeat process for all tables you have. You can double check that the rows all got imported correctly in the yellow highlighted section.

```

1  -- SHOW VARIABLES LIKE 'secure_file_priv'
2
3  USE cms; -- this is the name of the database use previous
4  SET GLOBAL local_infile = 1;
5  LOAD DATA LOCAL INFILE 'C:\\ProgramData\\MySQL\\MySQL Server 8.0\\Uploads\\provider.csv' -- this is the path for the file
6  INTO TABLE provider -- make sure correct csv file data is loaded to the correct table
7  FIELDS TERMINATED BY ','
8  ENCLOSED BY '"'
9  LINES TERMINATED BY '\n'
10 IGNORE 1 ROWS; -- ignores first row which is the column names
11

```

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Contact Help Stopgap

#	Time	Action	Message	Duration / Batch
1	22:10:50	SHOW VARIABLES LIKE 'secure_file_priv'	1 row(s) returned	0.016 sec / 0.000 sec
2	22:21:36	DROP DATABASE IF EXISTS cms	3 row(s) affected	0.047 sec
3	22:21:36	CREATE DATABASE cms	1 row(s) affected	0.000 sec
4	22:21:36	USE cms	0 row(s) affected	0.000 sec
5	22:21:36	CREATE TABLE provider_addresses (Riding_NPI INT, Riding_Privd_Last_Org_Name TEXT, Riding_Privd_Fmt_Name TEXT, Riding_Privd_Mt TEXT, Riding_Privd_CityState TEXT, Riding_Privd_Stk_Cat TEXT, ...)	0 row(s) affected	0.032 sec
6	22:21:36	CREATE TABLE provider_classification (Riding_NPI INT, Riding_Privd_32 TEXT, Riding_Privd_City TEXT, Riding_Privd_State_Abbrev TEXT, Riding_Privd_State_ZIPS TEXT, ...)	0 row(s) affected	0.015 sec
7	22:21:36	CREATE TABLE provider_addresses (Riding_NPI INT, Riding_Privd_Last_Org_Name TEXT, Riding_Privd_Fmt_Name TEXT, Riding_Privd_Mt TEXT, Riding_Privd_CityState TEXT, Riding_Privd_Stk_Cat TEXT, ...)	0 row(s) affected	0.031 sec
8	22:21:36	CREATE TABLE service_totals (Riding_NPI INT, Tot_HCPCS_Cat INT, Tot_Service INT, Tot_Service FLOAT, Tot_Sentinel FLOAT, Tot_Major_Accred_Amt FLOAT, Tot_Major_Pmt_Amt FLOAT, ...)	0 row(s) affected	0.016 sec
9	22:21:36	CREATE TABLE org_services (Riding_NPI INT, Org_Serv_Int TEXT, Org_Tot_HCPCS_Cat FLOAT, Org_Tot_Service FLOAT, Org_Tot_Sentinel FLOAT, Org_Sentinel FLOAT, Org_Major_A...	0 row(s) affected	0.016 sec
10	22:21:36	CREATE TABLE service_addresses (Riding_NPI INT, Mail_Serv_Int TEXT, Mail_Tot_HCPCS_Cat FLOAT, Mail_Tot_Service FLOAT, Mail_Tot_Sentinel FLOAT, Mail_Major_A...	0 row(s) affected	0.016 sec
11	22:21:36	CREATE TABLE service_addresses (Riding_NPI INT, Bene_Age_Int INT, Bene_Age_1T_85_Cat FLOAT, Bene_Age_85_74_Cat FLOAT, Bene_Age_74_64_Cat FLOAT, Bene_Age_64_54_Cat FLO...	0 row(s) affected	0.016 sec
12	22:21:36	CREATE TABLE bene_clinical_conditions (Riding_NPI INT, Bene_CC_BH_AJCHD_OncCD_V1_Flt FLOAT, Bene_CC_BH_AJCHD_OncCD_V1_Flt FLOAT, Bene_CC_BH_Tobacco_V1_Flt FLOAT, Bene_C...	0 row(s) affected	0.031 sec
13	22:21:46	USE cms	0 row(s) affected	0.000 sec
14	22:21:46	SET GLOBAL local_infile = 1	0 row(s) affected	0.000 sec
15	22:21:46	LOAD DATA LOCAL INFILE 'C:\\ProgramData\\MySQL\\MySQL Server 8.0\\Uploads\\provider.csv' -- this is the path for the file INTO TABLE provider -- make sure correct csv file data is loaded to the correct table FIELDS T...	Error Code: 2. File 'C:\\ProgramData\\MySQL\\MySQL Server 8.0\\Uploads\\provider.csv' not found (OS error 2 - No such file or directory)	0.000 sec
16	22:22:45	USE cms	0 row(s) affected	0.000 sec
17	22:22:45	SET GLOBAL local_infile = 1	0 row(s) affected	0.000 sec
18	22:32:25	LOAD DATA LOCAL INFILE 'C:\\ProgramData\\MySQL\\MySQL Server 8.0\\Uploads\\provider.csv' -- this is the path for the file INTO TABLE provider -- make sure correct csv file data is loaded to the correct table FIELDS...	2415213 row(s) affected Records: 2415213 Deleted: 0 Skipped: 0 Warnings: 0	18.516 sec

Note: If there is a yellow warning sign instead of the green check mark, it is also okay as long as you can see the time in seconds on the very left. You can double check if its loaded in by going to step 10.

101	12:21:53	USE cms	0 row(s) affected	0.000 sec
102	12:21:53	SET GLOBAL local_infile = 1	0 row(s) affected	0.000 sec
103	12:21:53	LOAD DATA LOCAL INFILE 'C:\\ProgramData\\MySQL\\MySQL Server 8.0\\Upd...	27.032 sec	

Step 10: You can check everything is done correctly by doing the following.

```
SELECT * FROM provider; -- check the first 1000 rows
SELECT COUNT(*) FROM provider; -- check count of rows
```

12 • `SELECT * FROM provider;`

Result Grid

Filter Rows:

Edit:

Export/Import:

Wrap Cell Content:

Fetch rows:

	Rndrng_NPI	Rndrng_Privr_Last_Org_Name	Rndrng_Privr_First_Name	Rndrng_Privr_MI	Rndrng_Privr_Crdntls	Rndrng_Privr_Ent_Cd	Year
▶	1003000126	ENKESHAFI	ARDALAN		M.D.	I	2019
	1003000126	Enkeshafi	Ardalan		M.D.	I	2023
	1003000134	CIBULL	THOMAS	L	M.D.	I	2019
	1003000134	Cibull	Thomas	L	M.D.	I	2023
	1003000142	KHALIL	RASHID		M.D.	I	2019

12 • `SELECT * FROM provider;`

13

14 • `SELECT COUNT(*) FROM provider;`

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

	COUNT(*)
▶	2415213