

Essentials Lab 5

Exercise 1

- 1) Click **[Databases]** in the **Navigation Ribbon**.
- 2) Click **[Create]** to create a new database. If you don't see the **[Create]** option, you may need to "drill up" by clicking the word **[Databases]** in the **Breadcrumb Trail**.
- 3) Enter the name **USDA_NUTRIENT_STDREF** into the **[Name]** field of the **[Create Database]** dialog.
- 4) Enter a comment in the comment field (OPTIONAL).
- 5) Click **[Finish]**.

The screenshot shows the Snowflake Databases page. At the top, there is a navigation ribbon with icons for Databases (highlighted with a purple circle and number 1), Shares, Warehouses, Worksheets, and History. Below the ribbon, the title 'Databases' is displayed, followed by the sub-instruction 'Manage your databases from this page.' A purple circle with number 2 highlights the 'Create...' button. Below the buttons, there is a table listing existing databases:

Database	Origin	Creation Time	Owner
SNOWFLAKE	SNOWFLAKE.ACC...	5/31/19 7:44:01 AM	
SNOWFLAKE_SAMPLE_DATA	SFC_SAMPLES.SA...	5/31/19 7:45:16 AM	ACCOUNTADMIN

The screenshot shows the 'Create Database' dialog box. It has two input fields: 'Name*' containing 'USDA_NUTRIENT_STDREF' (highlighted with a purple circle and number 3) and 'Comment' containing 'My nutrition database' (highlighted with a purple circle and number 4). At the bottom right are 'Cancel' and 'Finish' buttons, with 'Finish' highlighted with a purple circle and number 5.

Essentials Lab 5

Exercise 2

- 1) Drill into your new database by clicking on the blue hyperlinked name in the list of databases.
- 2) Use the **Breadcrumb Trail** to confirm that you have drilled into your new database.
- 3) Use the blue underline to confirm that you are on the **[Tables]** tab.
- 4) Click **[Create]** to begin creating a new table.

The screenshot shows the Snowflake web interface with the 'Databases' tab selected. The title bar says 'snowflake'. Below it are navigation icons for 'Databases', 'Shares', 'Warehouses', 'Worksheets', and 'History'. The main content area is titled 'Databases' with the sub-instruction 'Manage your databases from this page.' Below this are buttons for 'Create...', 'Clone...', 'Drop...', and 'Transfer Ownership'. A table lists three databases: 'USDA_NUTRIENT_STDREF' (selected, highlighted with a purple circle containing '1'), 'SNOWFLAKE_SAMPLE_DATA', and 'SNOWFLAKE'. The columns are 'Database', 'Origin', 'Creation Time', 'Owner', and 'Comments' (partially visible).

Database	Origin	Creation Time	Owner	Comments
USDA_NUTRIENT_STDREF		6:22:04 PM	ACCOUNTADMIN	My first database
SNOWFLAKE_SAMPLE_DATA	SFC_SAMPLES.SA...	5/31/19 7:45:16 AM	ACCOUNTADMIN	TPC-H Sample Data
SNOWFLAKE	SNOWFLAKE.AC...	5/31/19 7:44:01 AM		

The screenshot shows the 'Tables' tab selected in the Snowflake interface. The breadcrumb trail shows 'Databases > USDA_NUTRIENT_STDREF' (highlighted with a purple circle containing '2'). Below are tabs for 'Tables' (selected), 'Views', 'Schemas', 'Stages', 'File Formats', and 'Sequences'. A row of buttons includes '+ Create...', '+ Create Like...', 'Clone...', 'Load Data...', 'Drop...', and 'Transfer...'. A table lists tables: 'Name', 'Schema', 'Creation Time', and 'Owner'. The first table in the list is highlighted with a purple circle containing '4'.

Name	Schema	Creation Time	Owner



Essentials Lab 5

Exercise 3

- 1) Name your table **FD_GROUP_INGEST**.
- 2) Leave the default value of **PUBLIC** as the schema.
- 3) Enter a comment (OPTIONAL).
- 4) Click **[Add]** to add a column to the table.
- 5) Name the first column **FDGRP_CD** and set the **[Type]** to **VARCHAR(6)**.
- 6) Click **[Add]** to add another column to the table.
- 7) Name the second column **FDGRP_DESC** and set the **[Type]** to **VARCHAR(62)**.
- 8) Click **[Finish]**.

Create Table

Table Name * **FD_GROUP_INGEST**

Schema Name **PUBLIC**

Comment

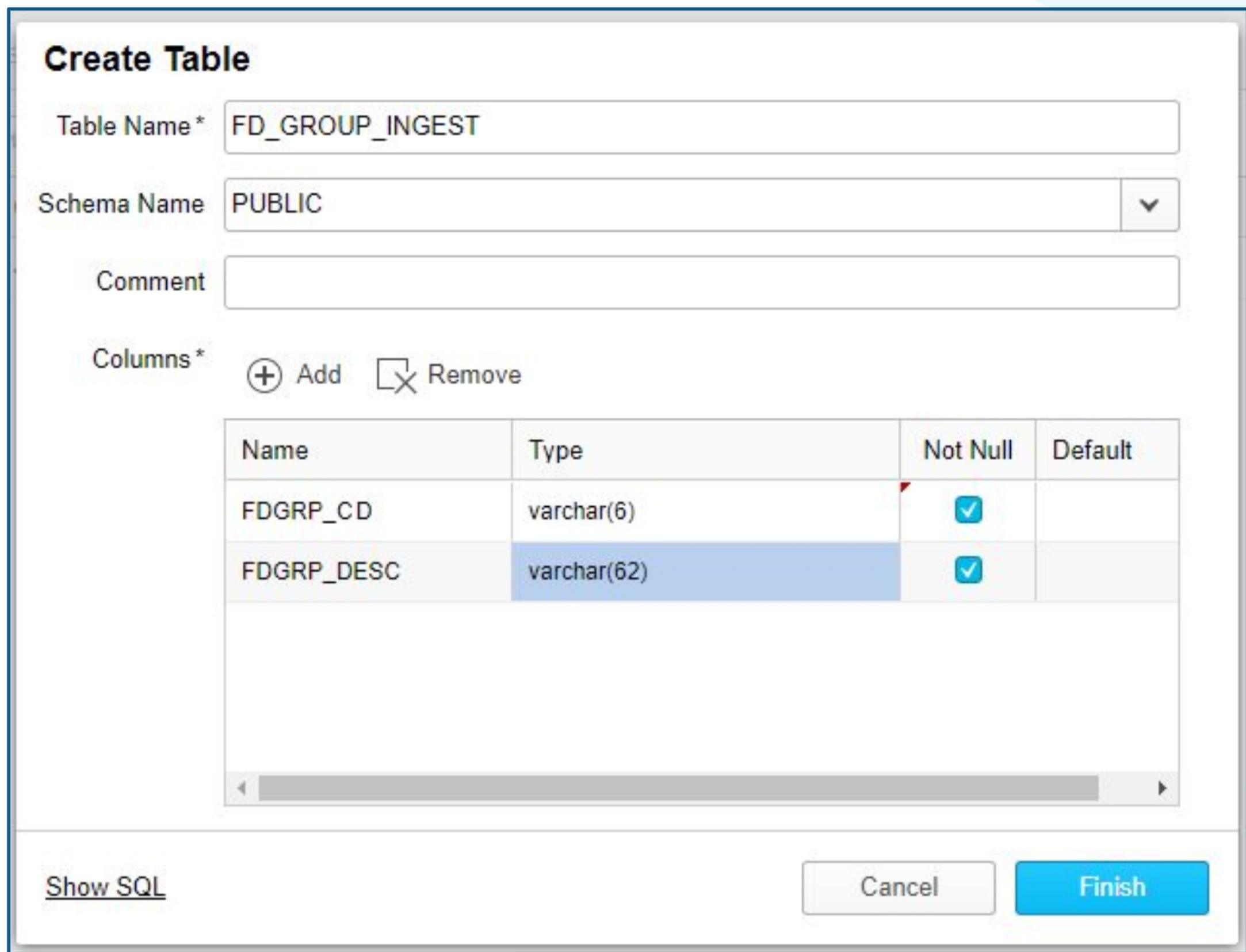
Columns *

Add **Remove**

Name	Type	Not Null	Default
FDGRP_CD	varchar(6)	<input checked="" type="checkbox"/>	
FDGRP_DESC	varchar(62)	<input checked="" type="checkbox"/>	

[Show SQL](#)

Cancel **Finish**



Essentials Lab 5

Exercise 4

- 1) Confirm via the breadcrumb that you are still in your USDA database.
- 2) Click the **File Formats** tab.
- 3) Click **[Create]** to begin creating a new file format.

NOTE:

Your file format describes the type of file you want to load. In the case of the USDA files we need to let Snowflake know that you will be loading a file with carets as column delimiters.

The screenshot shows the Snowflake web interface. At the top, there's a navigation bar with icons for Databases, Shares, Warehouses, Worksheets, and History. Below the navigation bar, the breadcrumb path reads "Databases > USDA_NUTRIENT_STDREF". A purple circle with the number 1 is placed over the word "USDA_NUTRIENT_STDREF". The main content area has tabs for Tables, Views, Schemas, Stages, File Formats (which is highlighted with a blue underline and a purple circle with the number 2), and Sequences. Below the tabs is a toolbar with buttons for Create..., Clone..., Edit..., Drop..., and Transfer Ownership. A purple circle with the number 3 is placed over the "Create..." button. The main table has columns for File Format, Schema, Type, Creation Time, and Owner. There is one row visible in the table.

File Format	Schema	Type	Creation Time	Owner



Essentials Lab 5

Exercise 5

- 1) Name your file format **USDA_FILE_FORMAT**.
- 2) Leave the default schema of **PUBLIC**.
- 3) Choose **CSV** as the format type, even though the file does not have comma separated values.
- 4) Leave the default compression method of **Auto**.
- 5) For **[Column separator]** - first choose the value **Other**, then input the caret symbol using **(Shift+6)** on your keyboard.
- 6) Click **[Finish]**.

Create File Format

Name * **USDA_FILE_FORMAT**

Schema Name **PUBLIC**

Format Type **CSV**

Compression Method **Auto**

Column separator **^**

Row separator **New Line**

Header lines to skip **0**

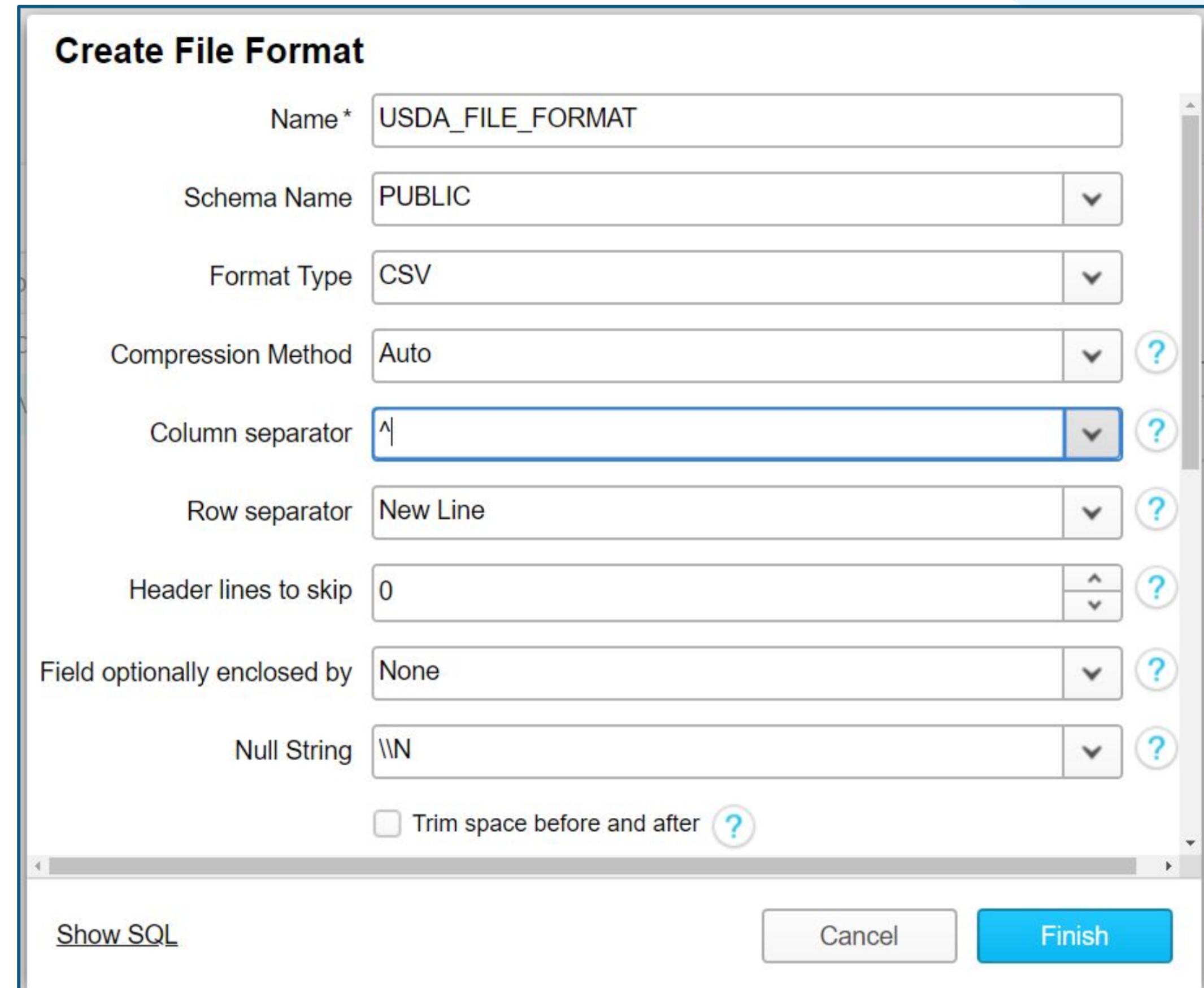
Field optionally enclosed by **None**

Null String **\N**

Trim space before and after **?**

[Show SQL](#)

Cancel **Finish**



Essentials Lab 5

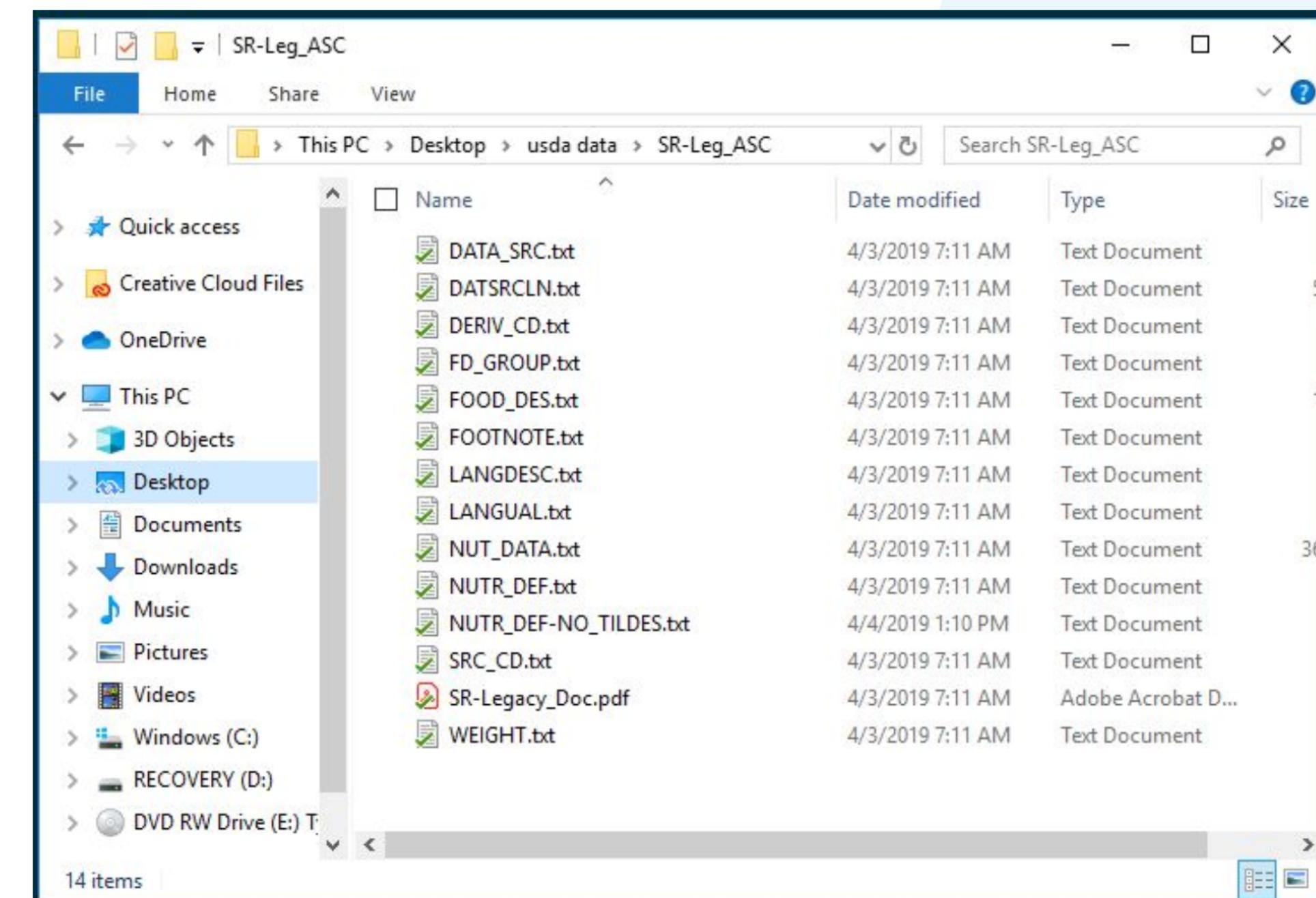
Exercise 6

- 1) Download USDA data files and the USDA specification documents.

https://www.ars.usda.gov/ARSUserFiles/80400525/Data/SR-Legacy/SR-Leg_ASC.zip

- 2) Unzip/extract the files to a local drive.

- 3) Make a note of the files' location so that you can load from that location in the next exercise.



Essentials Lab 5

Exercise 7

- 1) Use the breadcrumb trail to confirm you are in your USDA database.
- 2) Click the **[Tables]** tab.
- 3) Click the background of the row representing your new ***FD_GROUP_INGEST*** table.
- 4) With the ***FD_GROUP_INGEST*** row highlighted, click the **[Load Data]** option.

The Load wizard will open.

Databases > **USDA_NUTRIENT_STDREF** 1

Tables 2 Views Schemas Stages File Formats Sequences

+ Create... + Create Like... Clone... Load Data... Drop... Transfer Ownership

Table Name	Schema	Create Time	Owner	Rows	Size
FD_GROUP_INGEST 3	PUBLIC	7:15:19 PM	ACCOUNTADMIN		



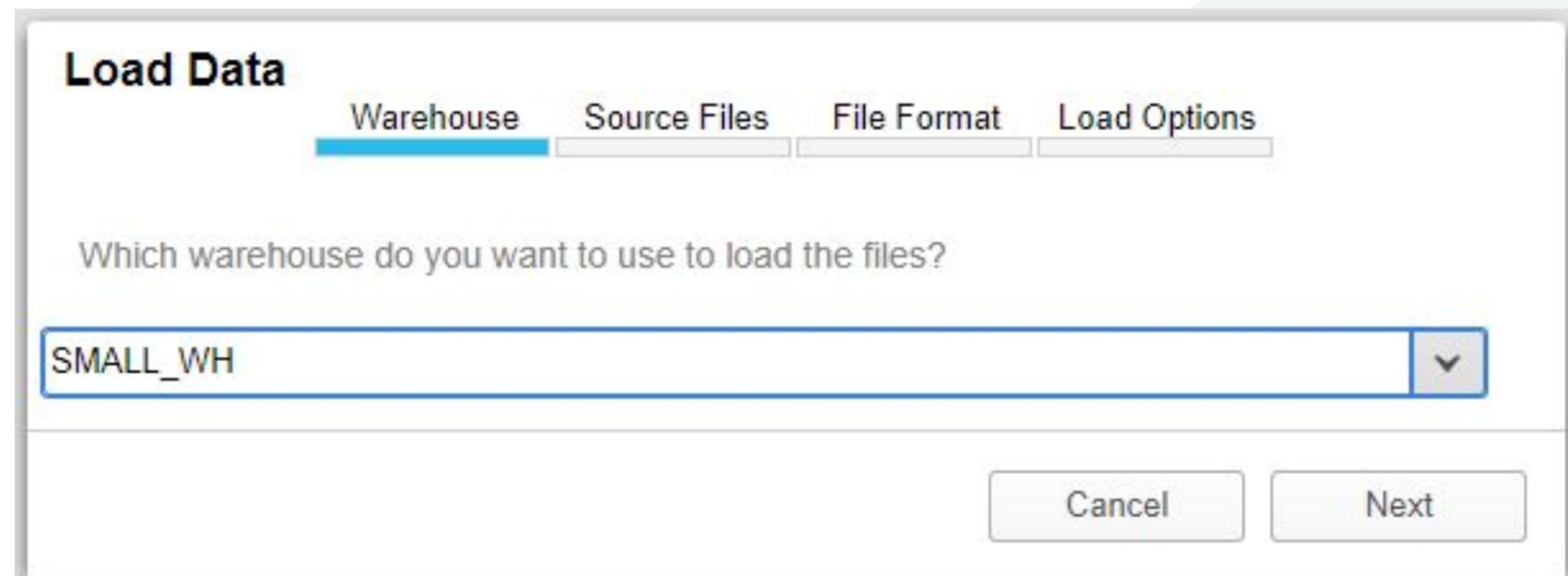
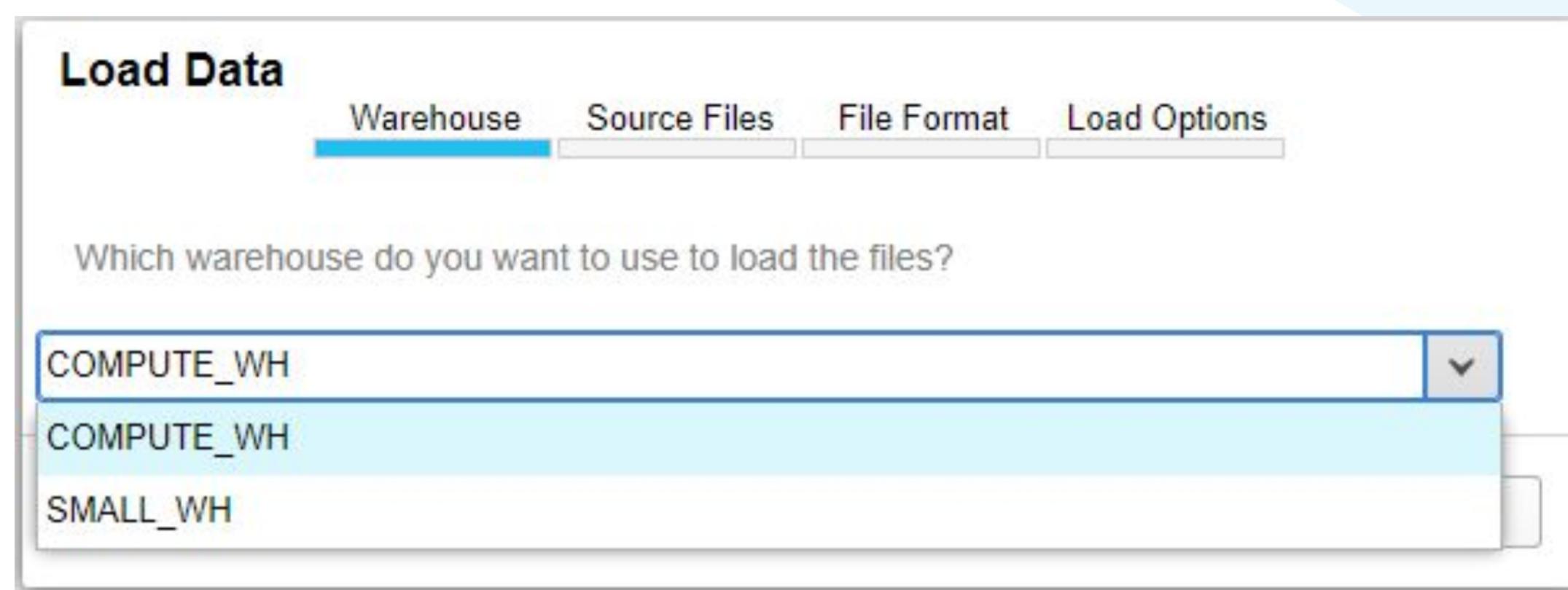
Essentials Lab 5

Exercise 8

- 1) Choose the **Warehouse** you would like to have run the load.

Remember that the warehouse is just the compute power applied to a process. It has no effect on data storage location.

- 2) Click **[Next]**.



Essentials Lab 5

Exercise 9

- 1) Choose the first of two options:
[Load files from your computer]
- 2) Click **[Select Files...]** and navigate to the **FD_GROUP.txt** file that you downloaded in Exercise 6 of this lab.
- 3) Click **[Next].**
- 4) Select **USDA_FILE_FORMAT** (if this is your only file format, it will be selected for you).
- 5) Click **[Next].**

Load Data

Warehouse **Source Files** File Format Load Options

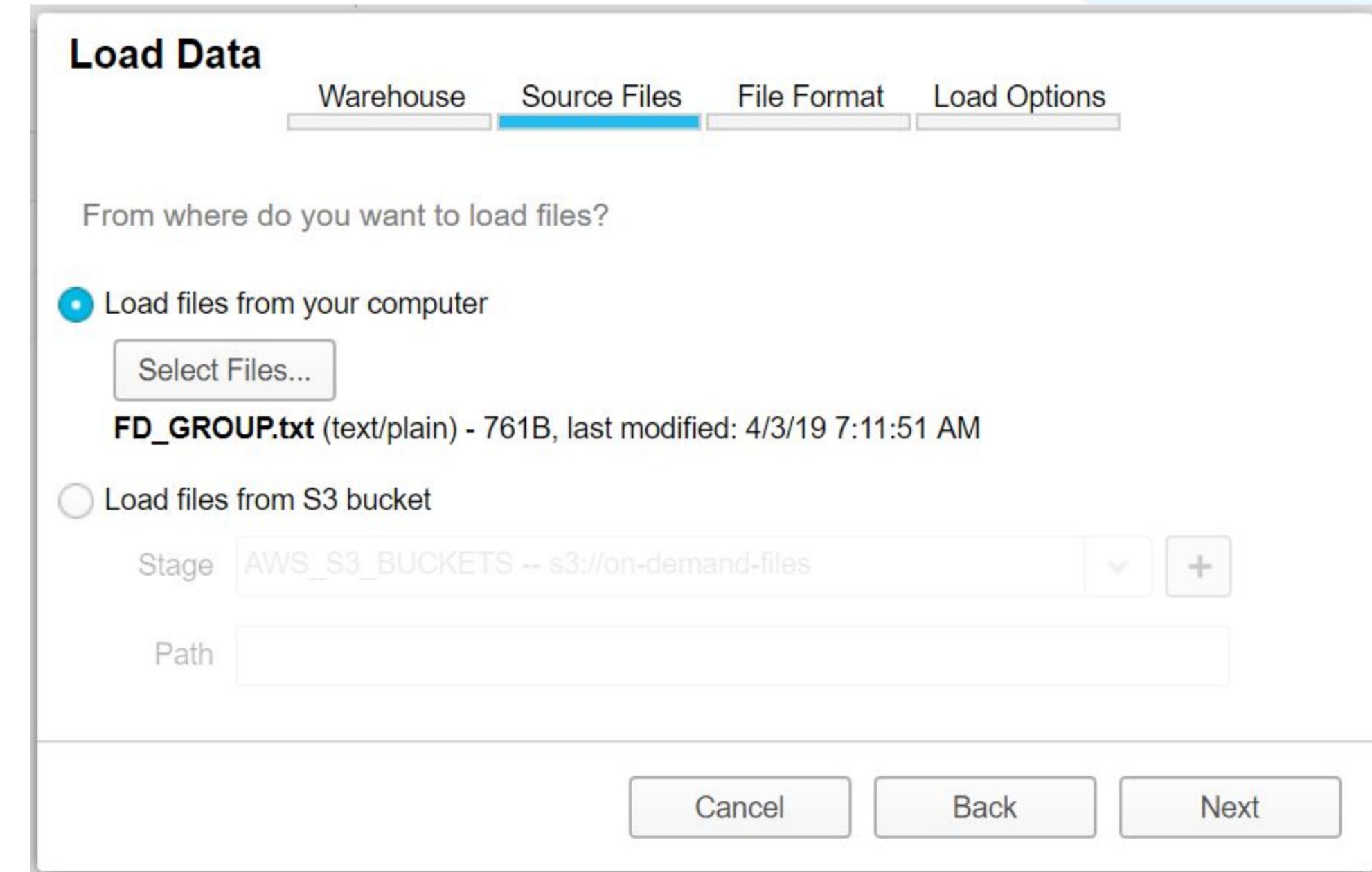
From where do you want to load files?

Load files from your computer
Select Files...
FD_GROUP.txt (text/plain) - 761B, last modified: 4/3/19 7:11:51 AM

Load files from S3 bucket
Stage AWS_S3_BUCKETS -- s3://on-demand-files

Path

Cancel Back Next

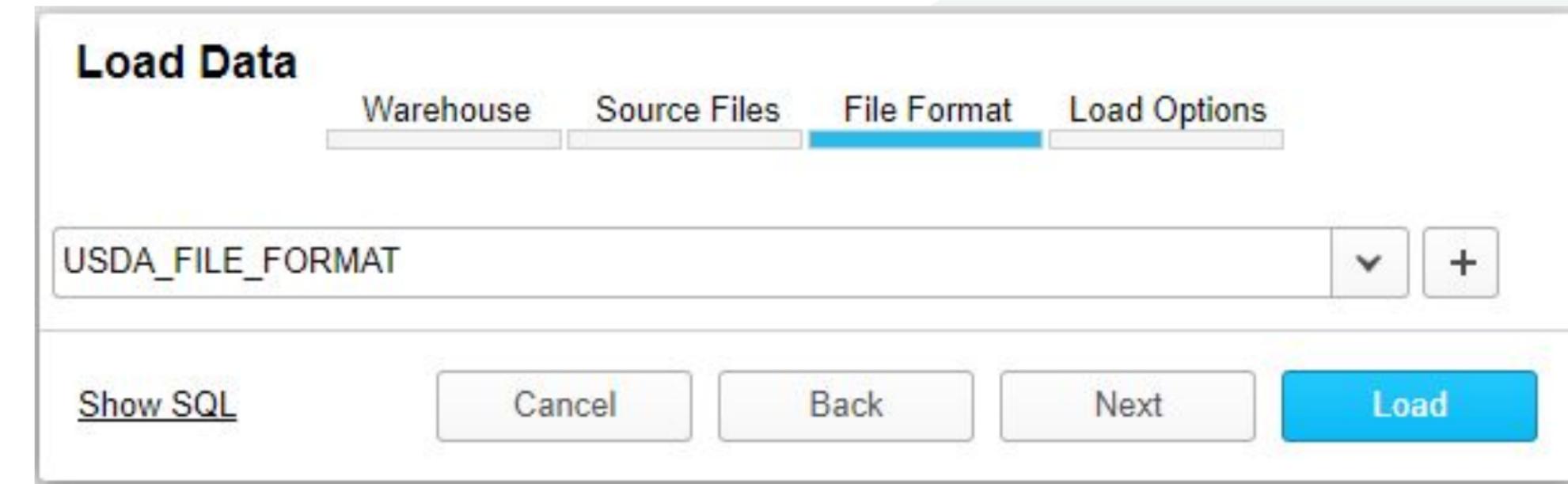


Load Data

Warehouse Source Files **File Format** Load Options

USDA_FILE_FORMAT

Show SQL Cancel Back Next **Load**



Essentials Lab 5

Exercise 10

- 1) Leave the default option:
[Stop loading, rollback and return the error]
- 2) Click **[Load]**.
- 3) Wait for confirmation.

NOTE:

If your file does not load and you don't understand the issue, you may want to try some of the other load options.

For example, the **"Continue loading..."** option would allow you to see which rows successfully loaded and allow you deduce possible issues in the rows that did not.

Load Data

Warehouse Source Files File Format **Load Options**

What should the load do if it encounters an error while parsing a file?

Do not load any data in the file
 Stop loading, rollback and return the error
 Do not load any data in the file if the error count exceeds:
Threshold 0

Continue loading valid data from the file

[Show SQL](#) Cancel Back **Load**

Load Results

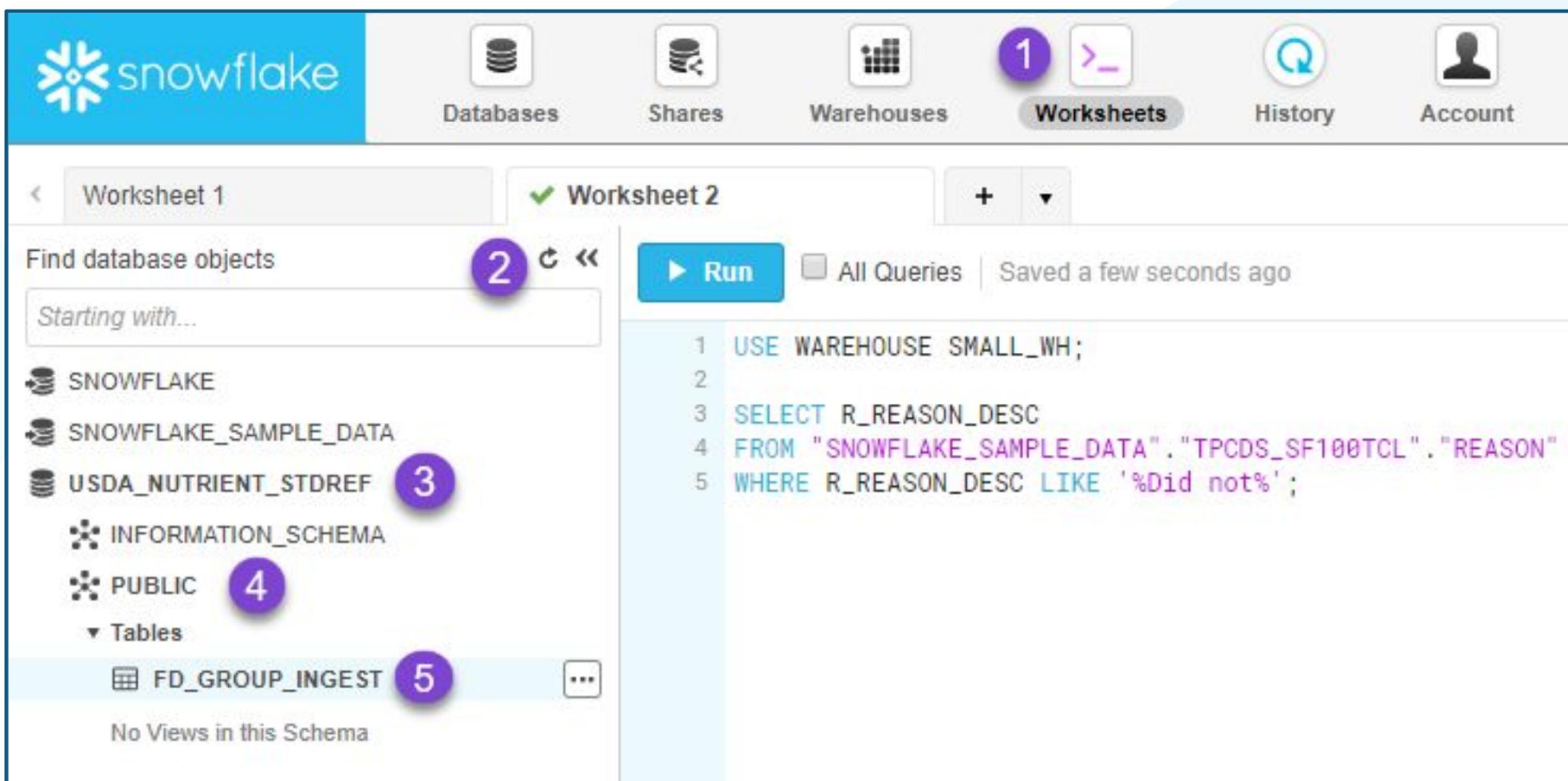
Loaded	File	Rows Parsed	Rows Loaded
✓	FD_GROUP.txt	25	25



Essentials Lab 5

Exercise 11

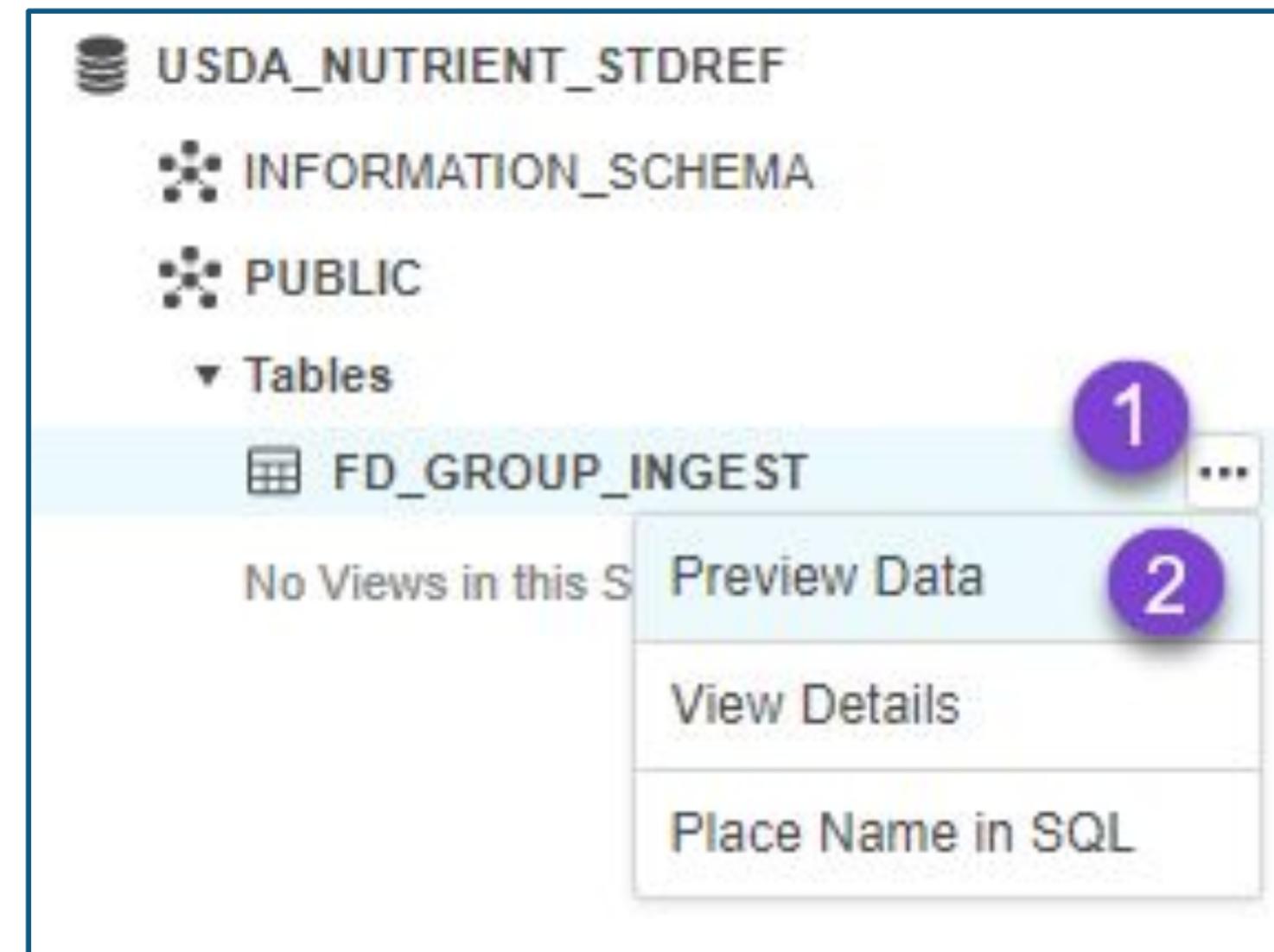
- 1) Click **[Worksheets]** in the **Navigation Ribbon**.
- 2) Use the refresh button to update objects in the **Navigation Tree**.
- 3) Click the **USDA_NUTRIENT_STDREF** database name to unfold its contents.
- 4) Click the **PUBLIC** schema.
- 5) Click **[Tables]**.
- 6) Click **FD_GROUP_INGEST**.



Essentials Lab 5

Exercise 12

- 1) Click [...] on the **FD_GROUP_INGEST** table label.
- 2) Select **[Preview Data]** to view the data you loaded.



OPTIONAL: Consider running a SQL query to look at your newly loaded data.

```
SELECT FDGRP_DESC  
FROM    "USDA_NUTRIENT_STDREF"."PUBLIC"."FD_GROUP_INGEST"  
WHERE FDGRP_DESC LIKE '%Food%' ;
```

