TEXAS
Health and Human Services

Texas Department of State Health Services

# SHARP Expansion for Jurisdictional COVID Data

## *Snowflake Reference Guide*

Version 1.0
As of 2/15/2022

# Contents


1. **Introduction to Snowflake & User Guide**
2. **Getting Started: First-time Login to your Snowflake Reader Account**
3. **The Basics: Snowflake Layout/Navigation**
4. **Accessing Data: Queries and Examples**
   - Configure your database and schema
   - Download data
5. **Additional Resources**


TEXAS Health and Human Services

Texas Department of State Health Services

# Introduction to Snowflake & User Guide

# How to leverage Snowflake and this User Guide

1

## A CLOUD-BASED DATA WAREHOUSE

**Snowflake** is a cloud-based data warehouse used to store Texas public health data including **COVID case data and COVID labs data.**

2

## HOW YOU'LL ACCESS COVID DATA

You have been granted a **Snowflake reader account*** which will enable you to view and query COVID case and COVID labs data for your region or jurisdiction.

3

## USING YOUR SNOWFLAKE READER ACCOUNT TO PULL COVID DATA

Snowflake uses the **SQL programming language to run queries**. You can run a query for all COVID data or narrow down the query (e.g., positive COVID cases during April 2021).

4

## THIS USER GUIDE

This document shows you **how to log into** your Snowflake reader account for the first time, **set up your workspace**, and **run queries**. Check out the **'Additional Resources' section** for sample queries and a glossary of terms.

*Additional Information on Snowflake Reader Accounts

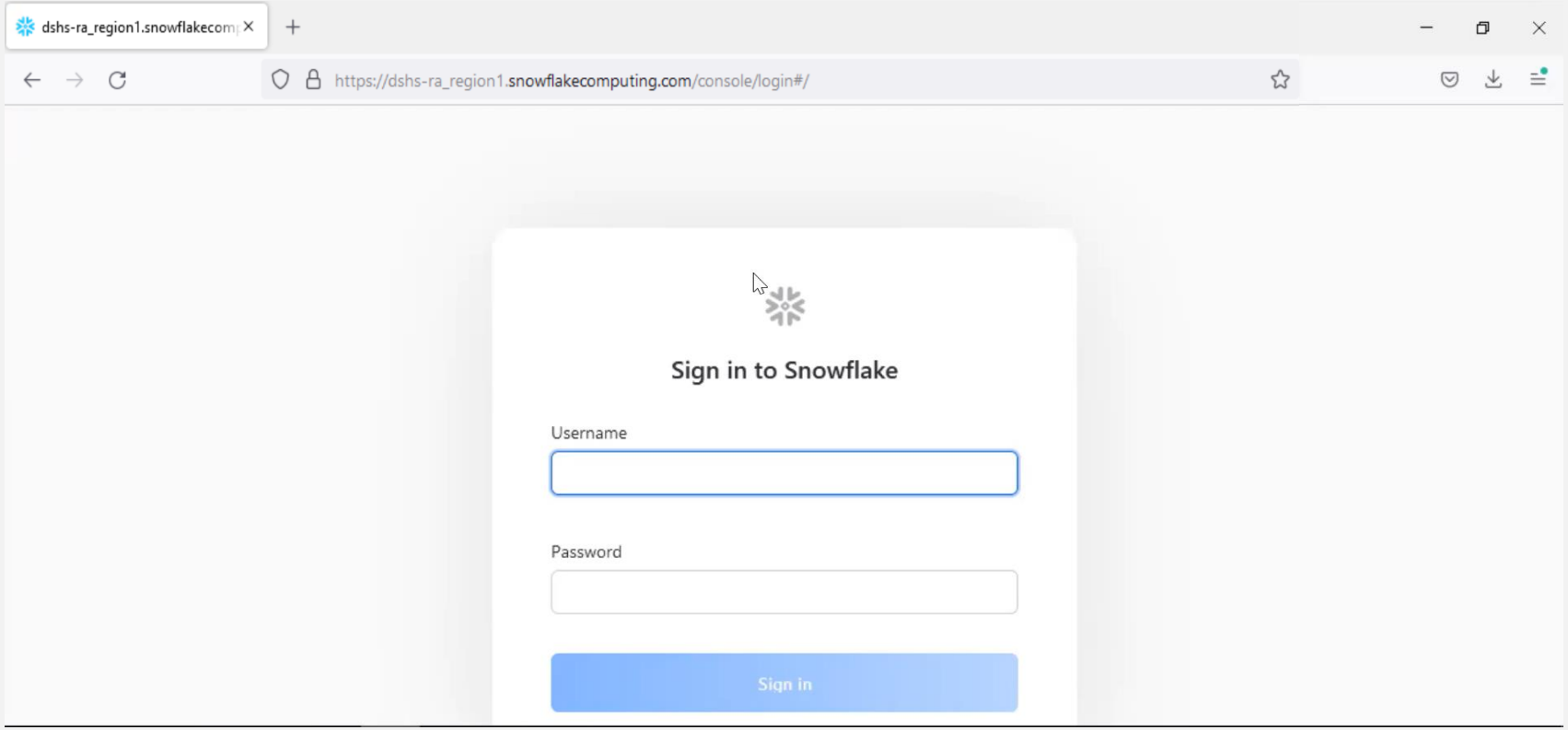
# Getting Started: How to login for the first time to your Snowflake Reader Account

# Step 1A: Navigate to Snowflake

**Pro tip!**

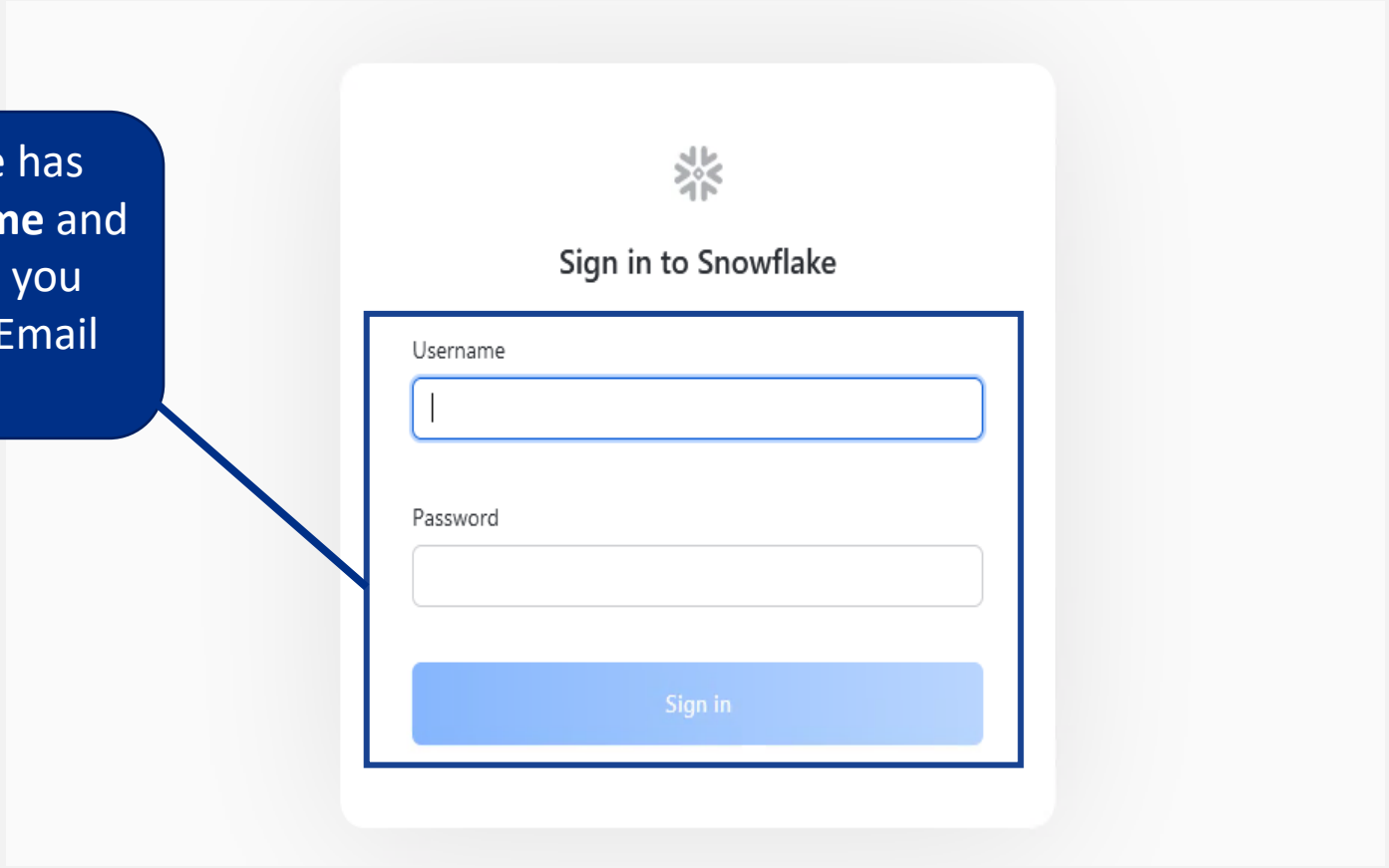
Use the latest version of Microsoft Edge for the best experience.

Navigate to the **Snowflake URL listed in the welcome email from HHS staff** by either clicking on the link or copying and pasting the URL into an internet browser. (Click here for a list of support internet browsers)



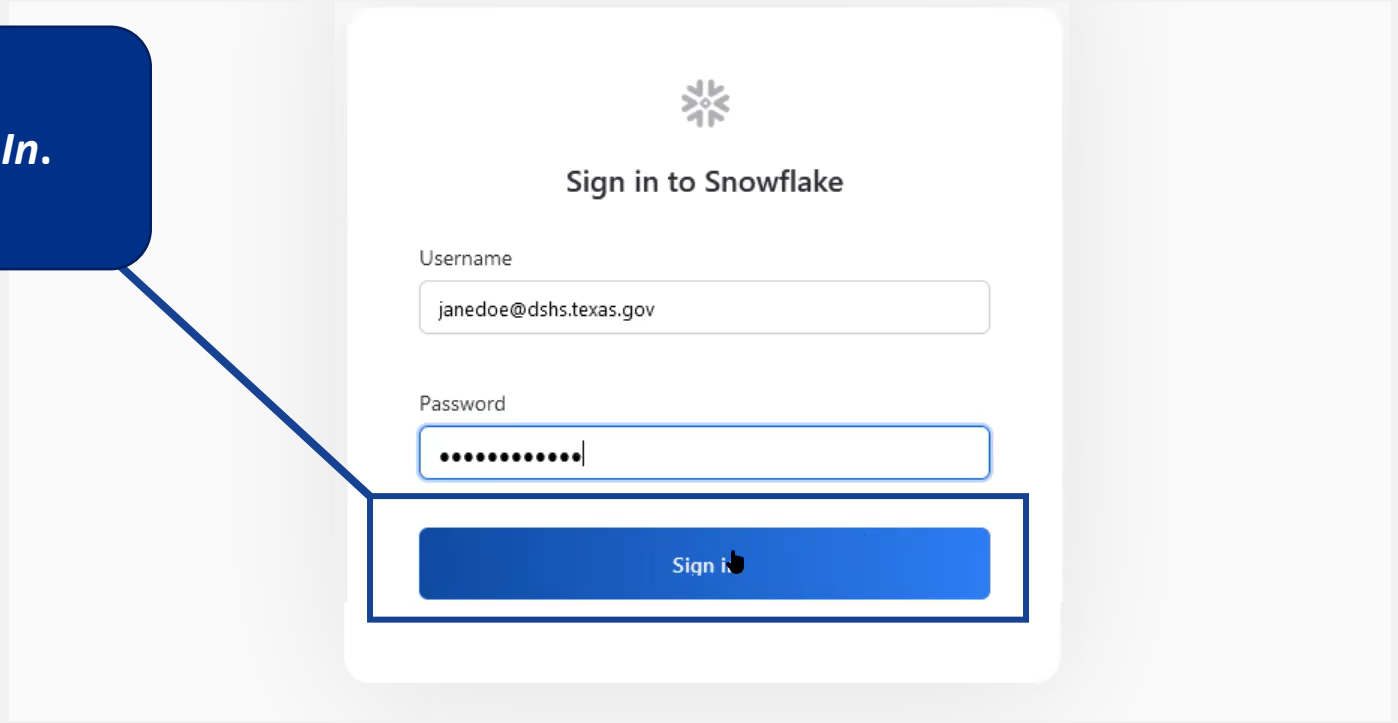
# Step 1B: Sign Into Snowflake

Once the Snowflake page has loaded, enter your **Username** and the **temporary password** you received in the Welcome Email from HHS staff.



# Step 1C: Sign Into Snowflake

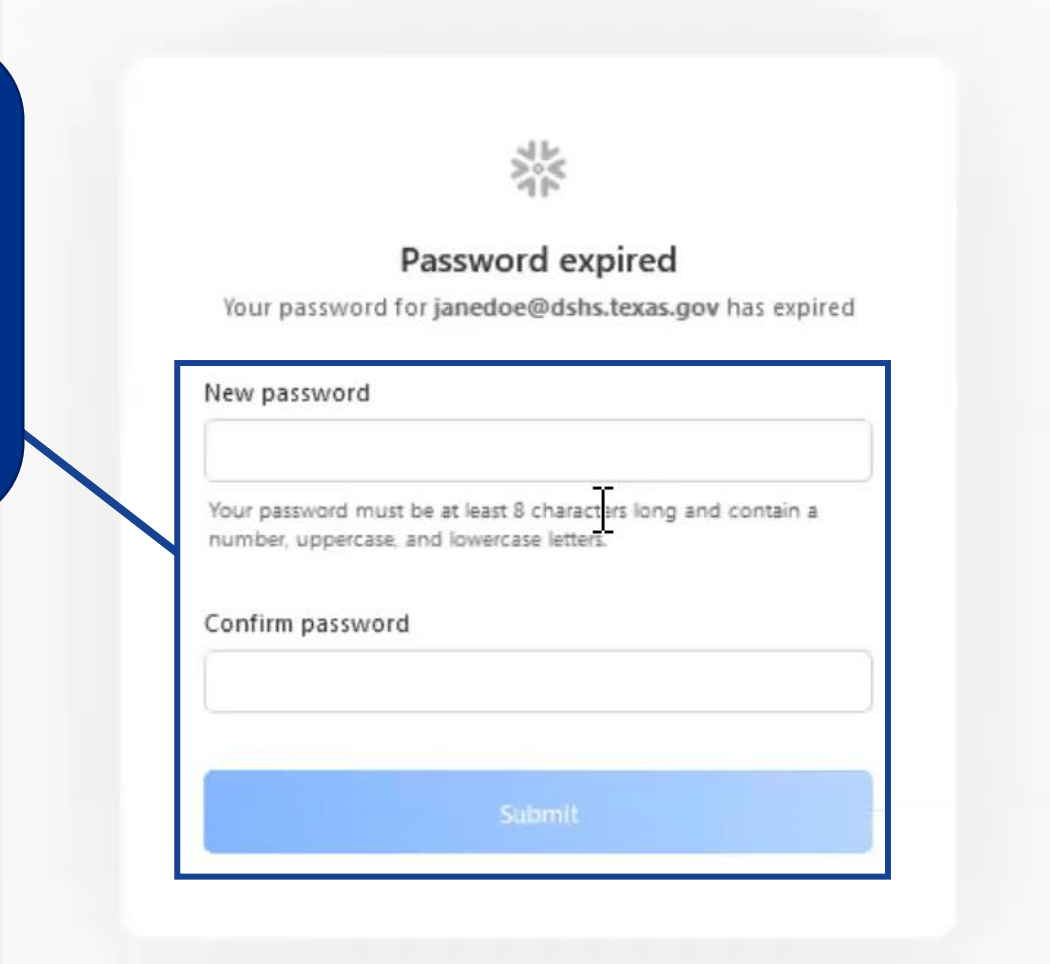
Click ***Sign In***.



# Step 1D: Change your password

Create a new password for your Snowflake account.
**The new password must be 8 characters long and contain a number, uppercase and lowercase letters**.

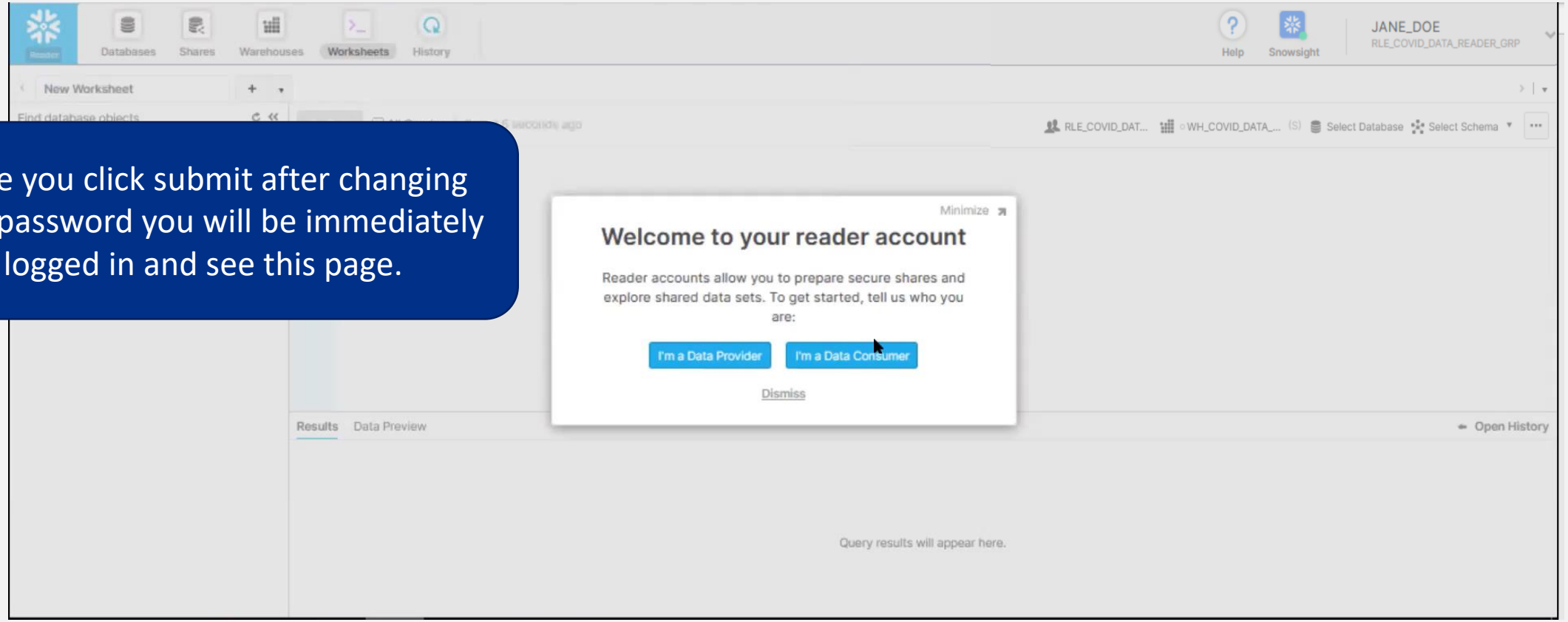
Click ***Submit*** **to** update your password.




**Note!** Creating a new password is a one-time step when you log in for the first time.

# Step 1E: First Time Log In

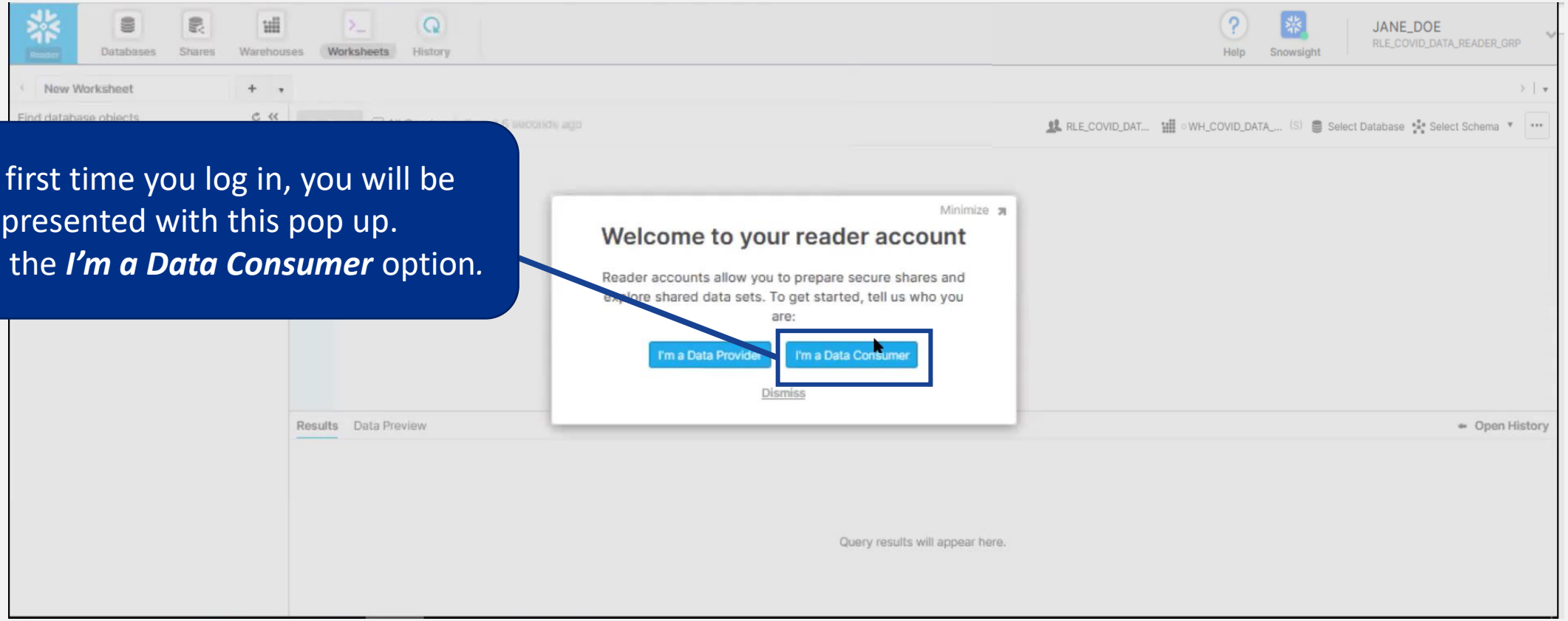
Once you click submit after changing your password you will be immediately logged in and see this page.



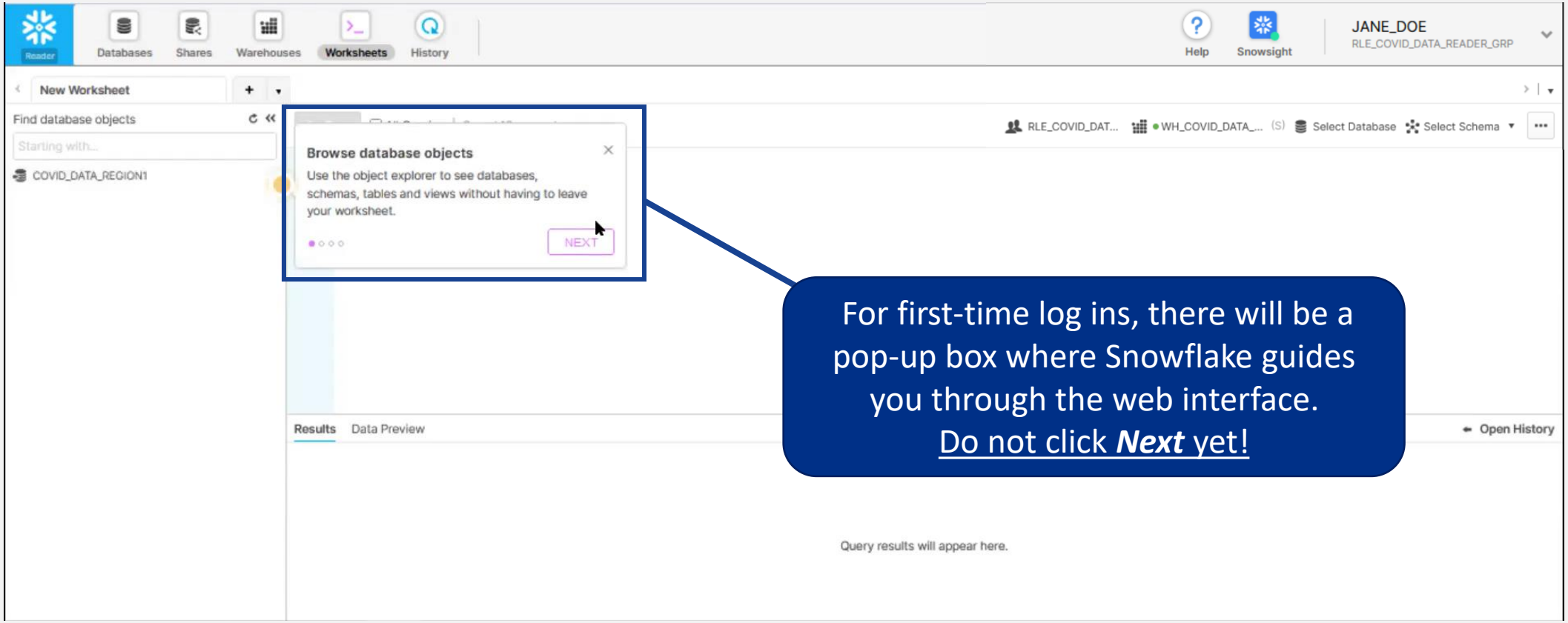
# The Basics: Snowflake Layout & Navigation

# Step 2A: First Time Log In Account Prompt

The first time you log in, you will be presented with this pop up.
Select the ***I'm a Data Consumer*** option.

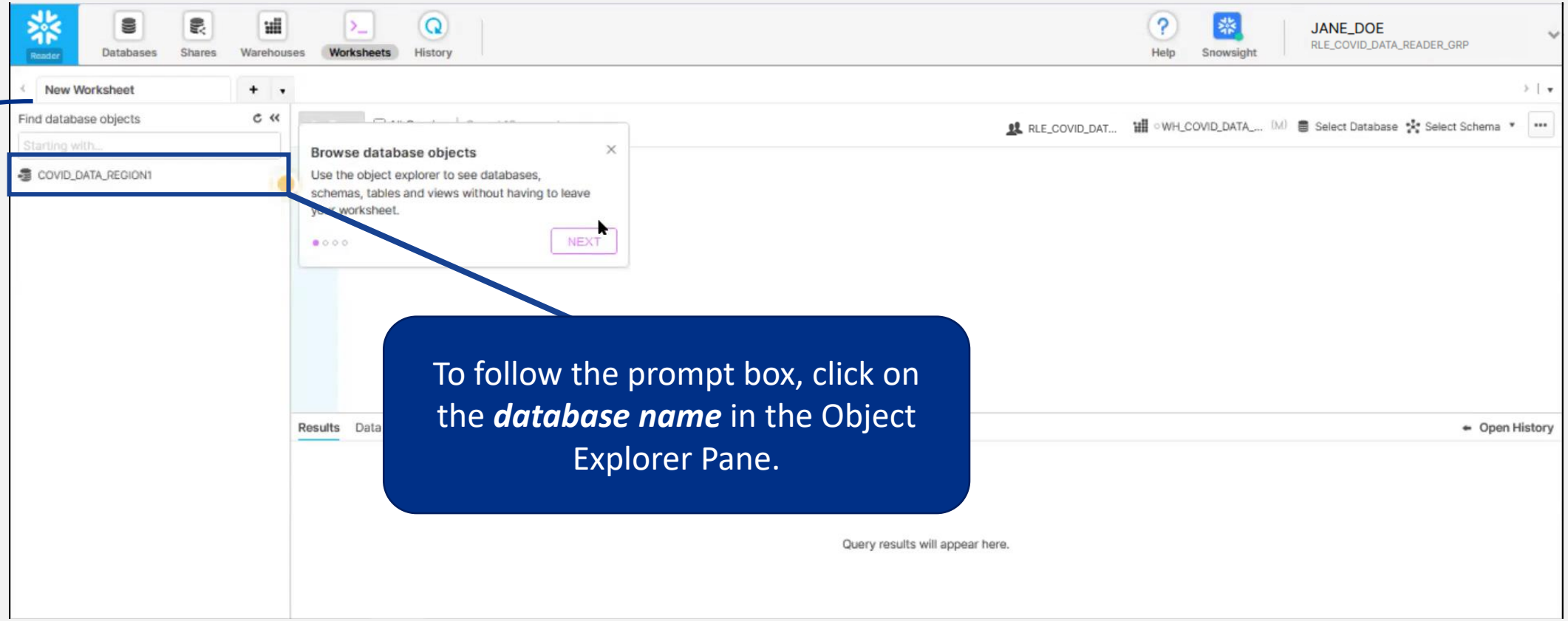


# Step 2B: First Time Log In Navigation Tour



# Step 2C: Snowflake Reader Account Navigation

A Snowflake **database** is a collection of data stored in Snowflake.

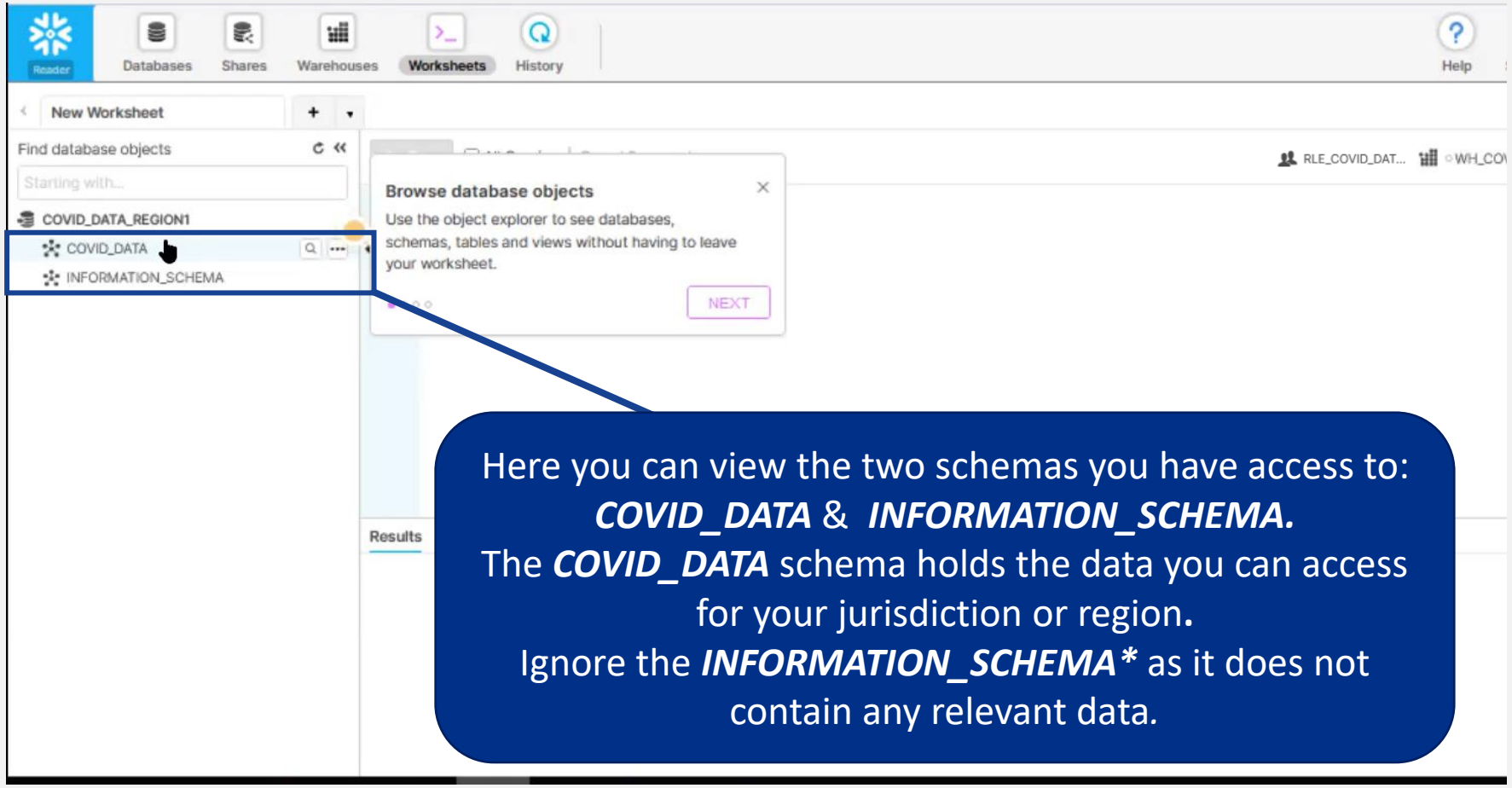



This is the Object Explorer Pane

To follow the prompt box, click on the ***database name*** in the Object Explorer Pane.

# Step 2D: Snowflake Reader Account Navigation

A **schema** represents how data is stored in a database.

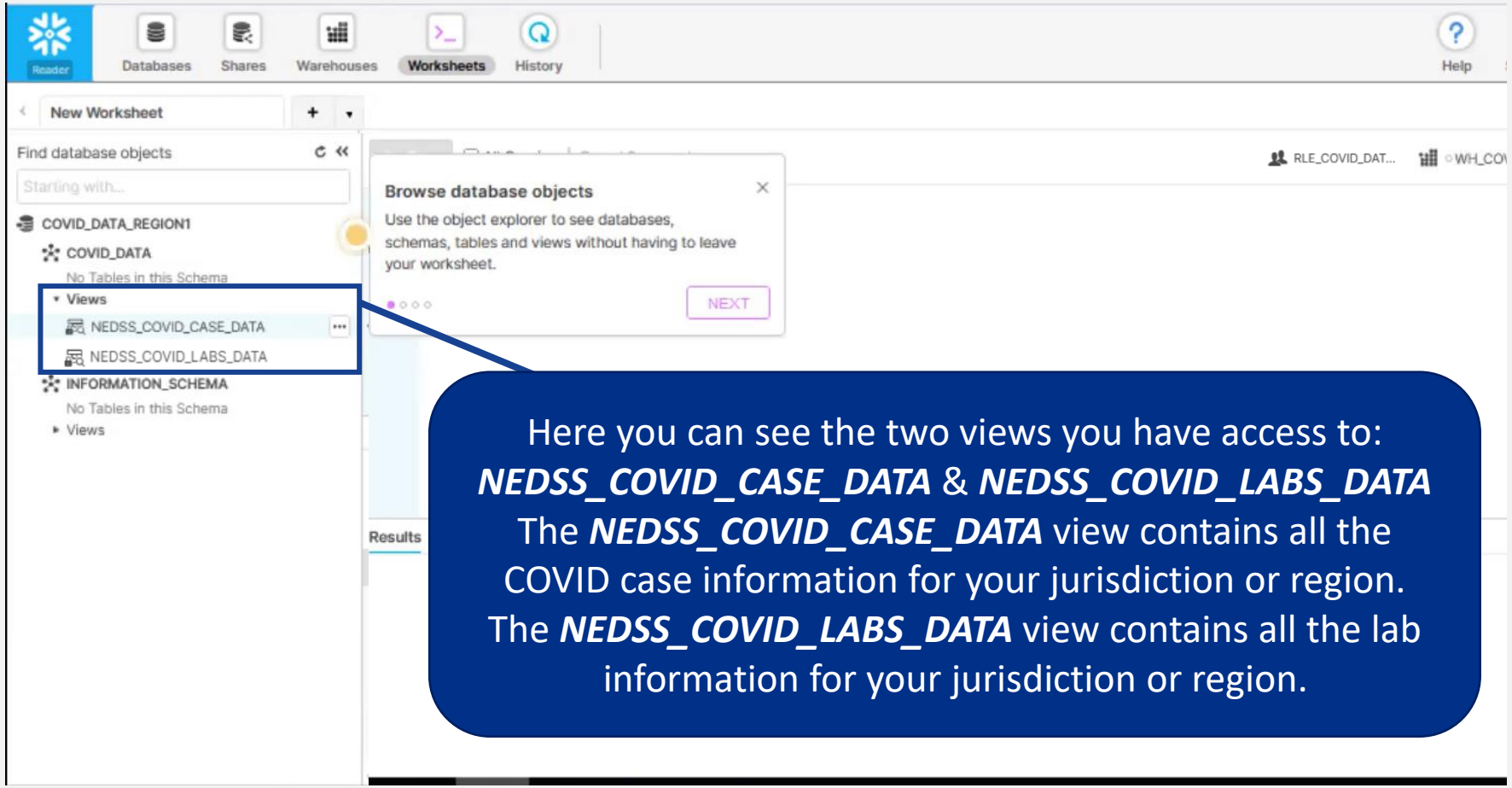



Here you can view the two schemas you have access to: ***COVID_DATA*** & ***INFORMATION_SCHEMA.***
The ***COVID_DATA*** schema holds the data you can access for your jurisdiction or region**.**
Ignore the ***INFORMATION_SCHEMA**** as it does not contain any relevant data.

*Additional detail on INFORMATION SCHEMA

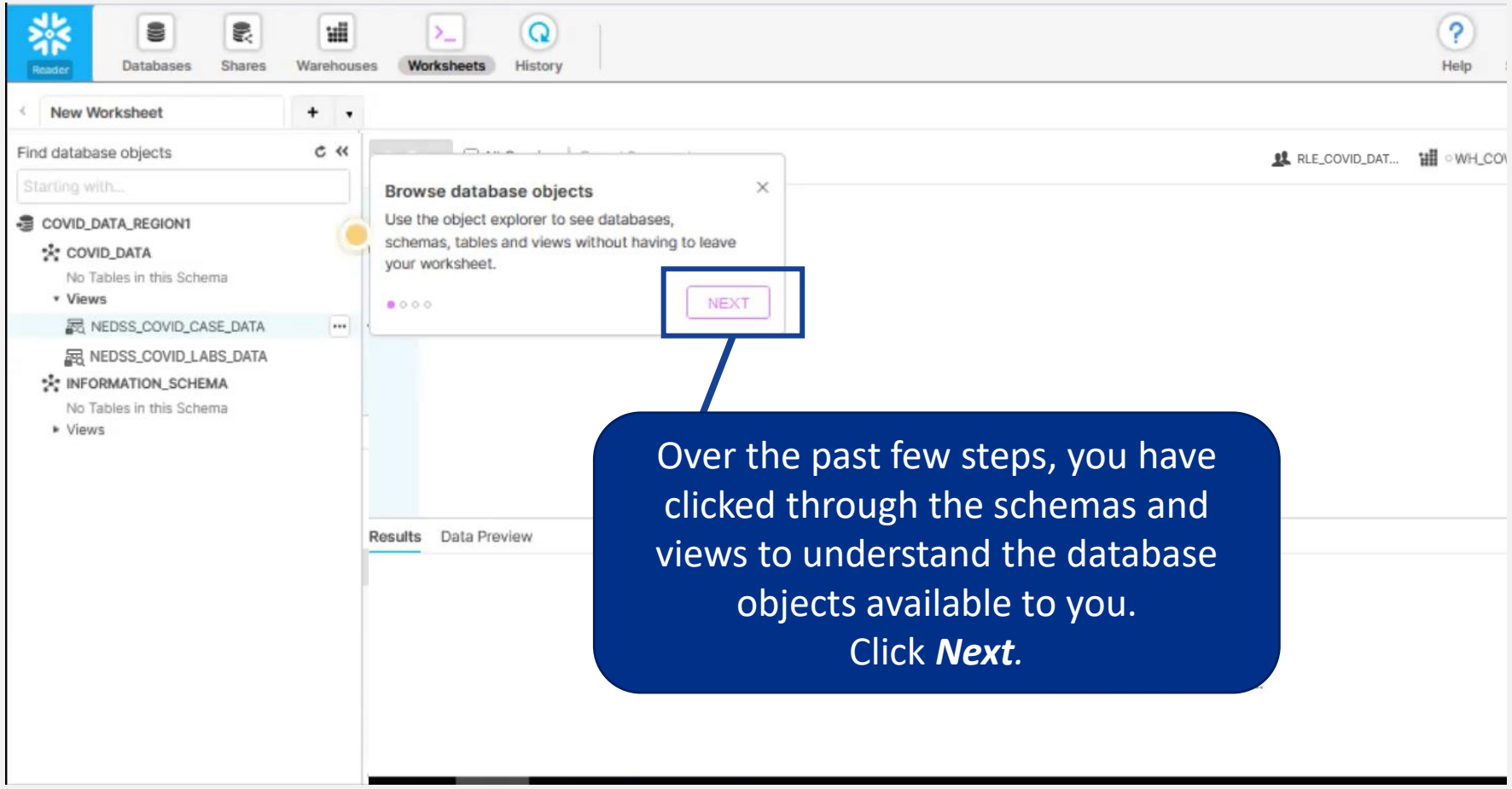
# Step 2E: Snowflake Reader Account Navigation

A **view** is a virtual table that contains data from one or more Snowflake tables.




Here you can see the two views you have access to: ***NEDSS_COVID_CASE_DATA*** & ***NEDSS_COVID_LABS_DATA***
The ***NEDSS_COVID_CASE_DATA*** view contains all the COVID case information for your jurisdiction or region.
The ***NEDSS_COVID_LABS_DATA*** view contains all the lab information for your jurisdiction or region.

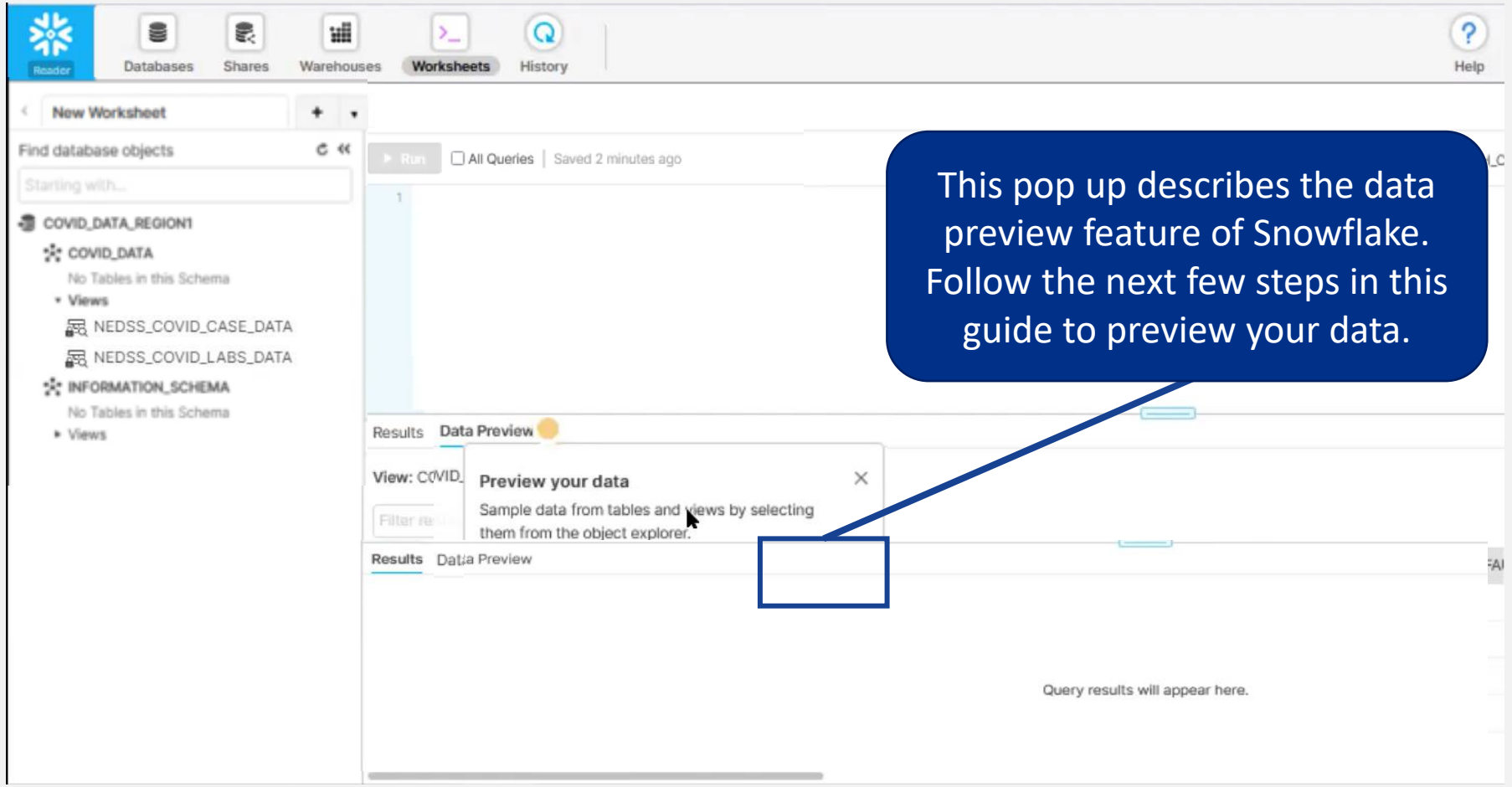
# Step 2F: Snowflake Reader Account Navigation



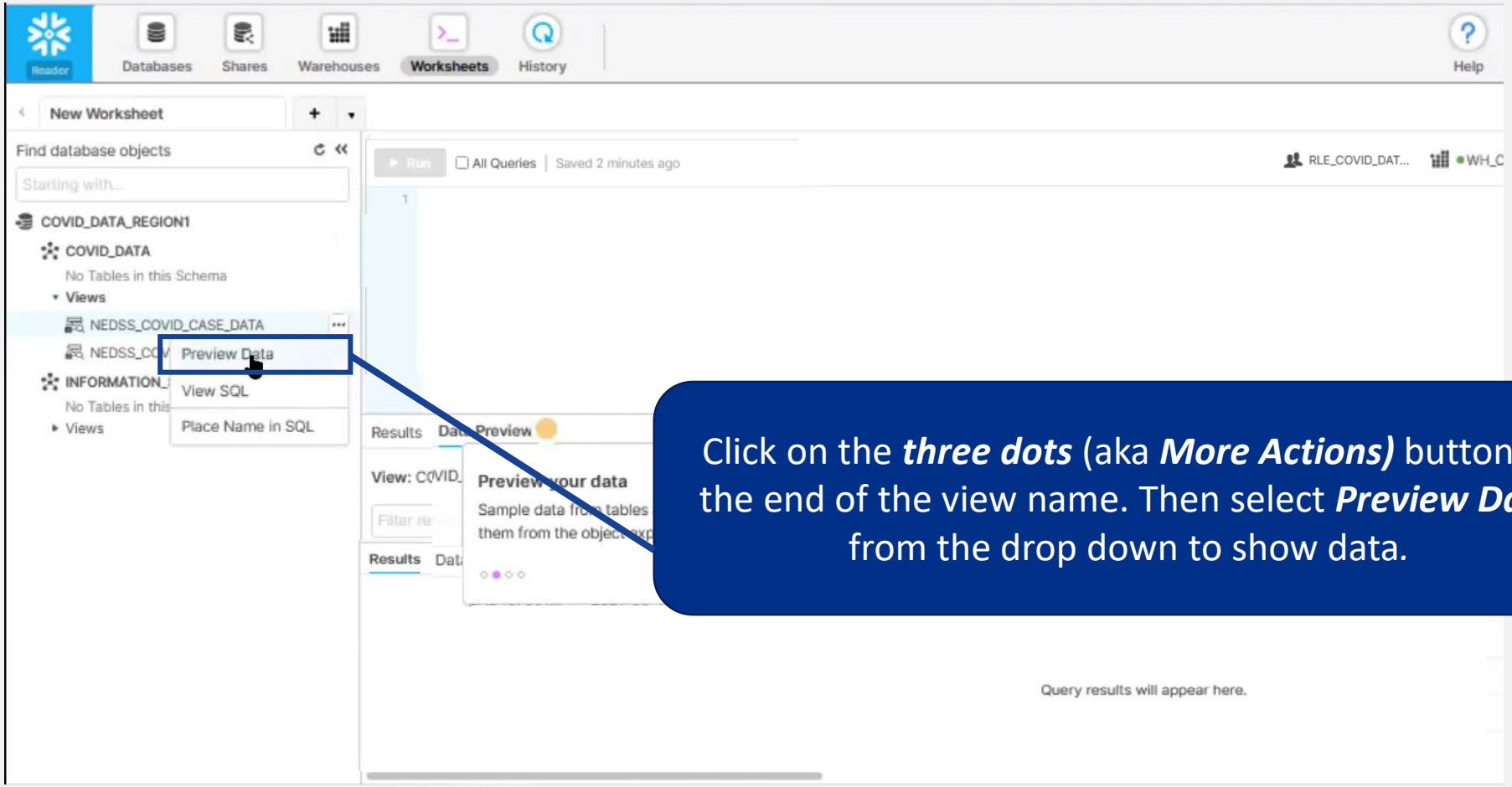

Over the past few steps, you have clicked through the schemas and views to understand the database objects available to you.
Click ***Next***.

*Additional detail on INFORMATION SCHEMA

# Step 2G: Viewing Data

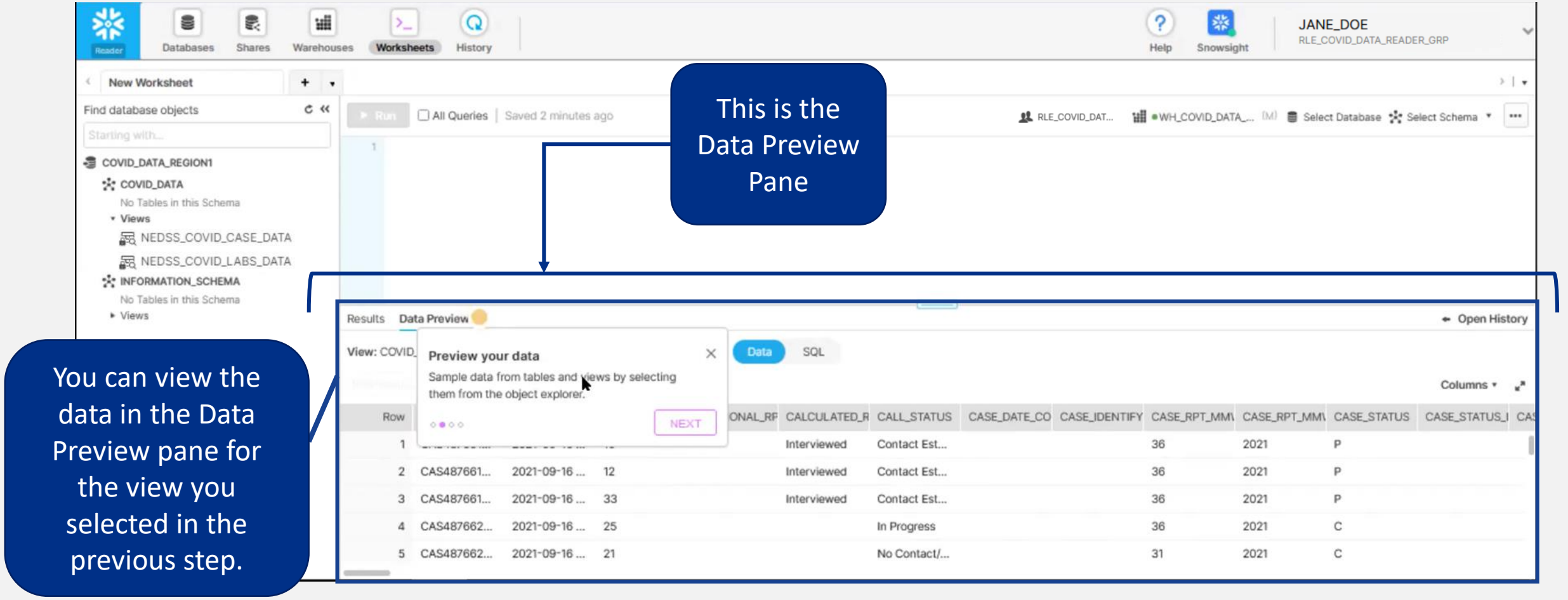


# Step 2H: Viewing Data

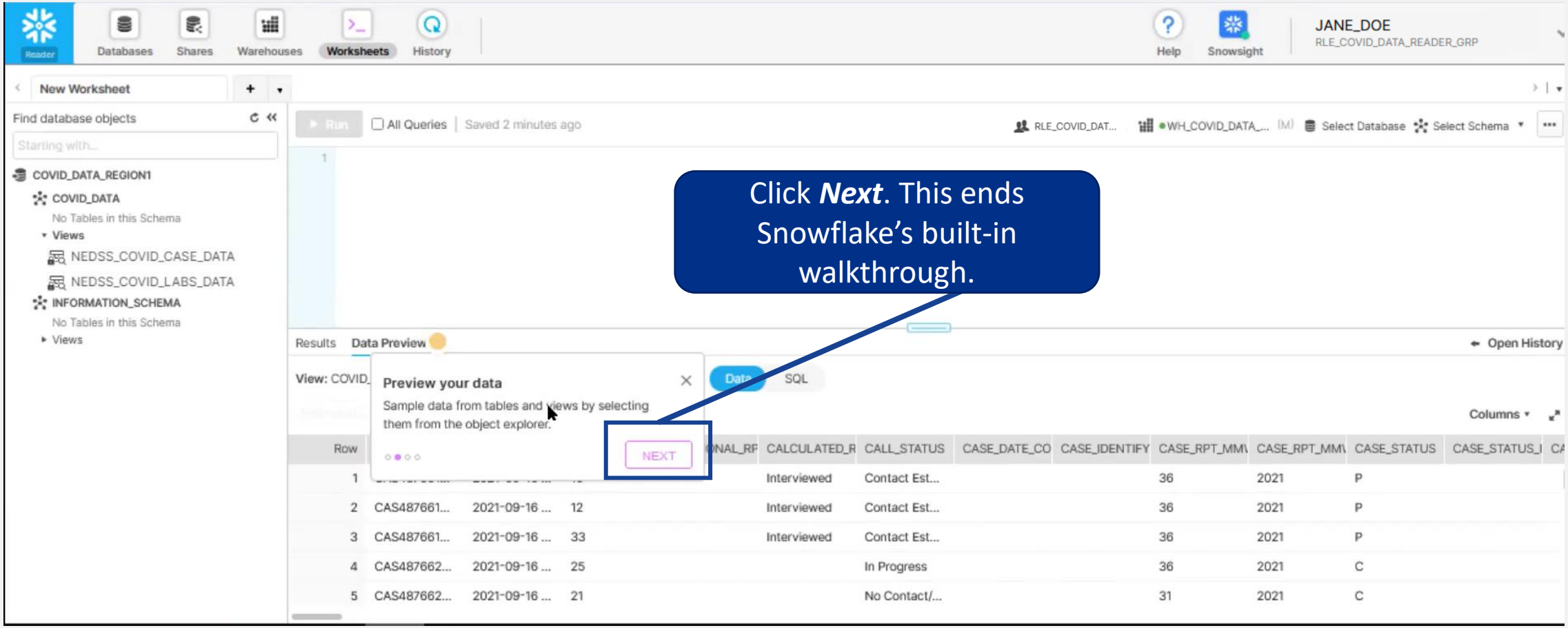



Click on the ***three dots*** (aka ***More Actions)*** button at the end of the view name. Then select ***Preview Data*** from the drop down to show data.

# Step 2I: Viewing Data



# Step 2J: Viewing Data

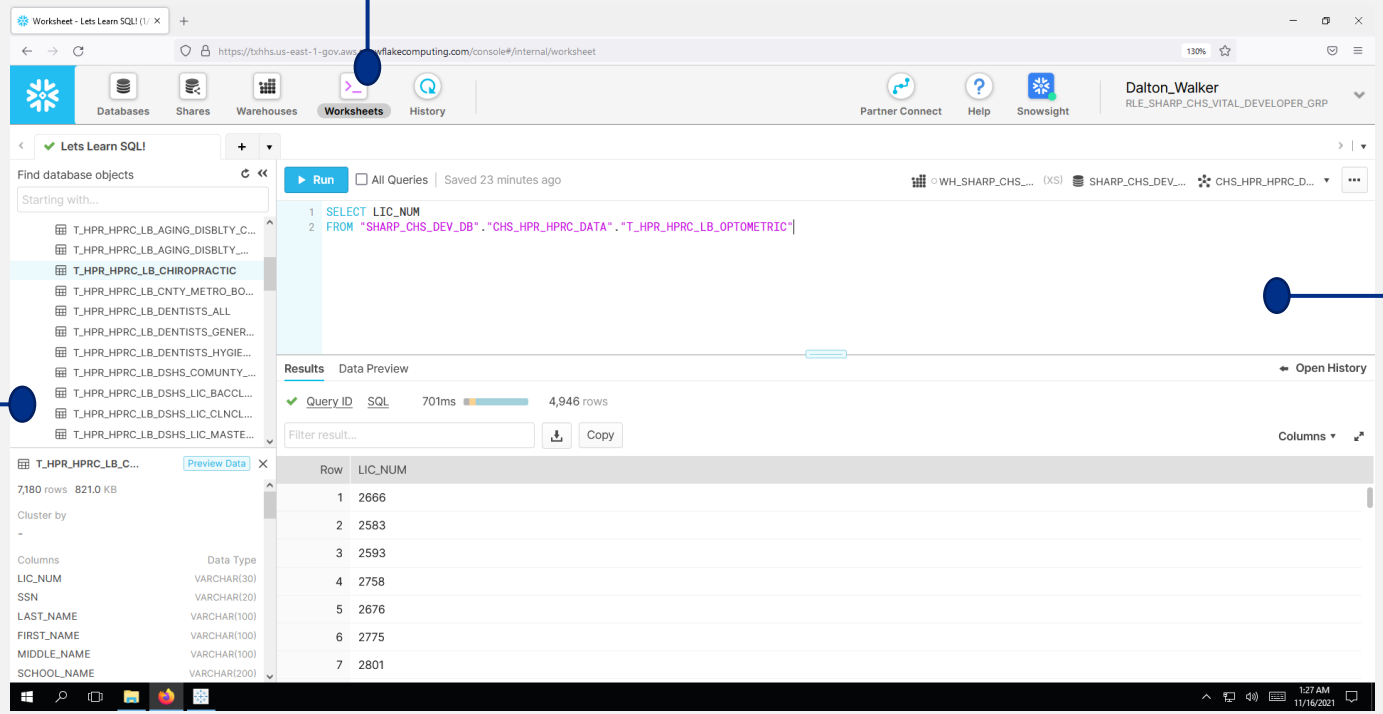


# Overview: Snowflake User Interface

Once you have logged into the Snowflake Reader Account user interface*, using the top ribbon you can navigate through databases, worksheets, history, and additional Snowflake objects. However, conventional users will spend most of their time on the worksheet screen, where you can write queries and analyze data.

## Tables and Views

- This list displays the database, schema, tables, and views on the left-hand side of the screen.
- All tables and views follow a naming convention that enables users to understand the scope of the data from the title.




## Snowflake Ribbon

- The Snowflake Ribbon consists of tabs for Databases, Shares, Warehouses, Worksheets, and History.
- Queries are written in the Worksheets tab.

## SQL Editor

- Write queries in order to answer questions and retrieve your data.
- Queries are written in SQL Syntax.
- Query results will appear in the lower panel. This panel can be used to preview tables or bring up results from the query expressed above.

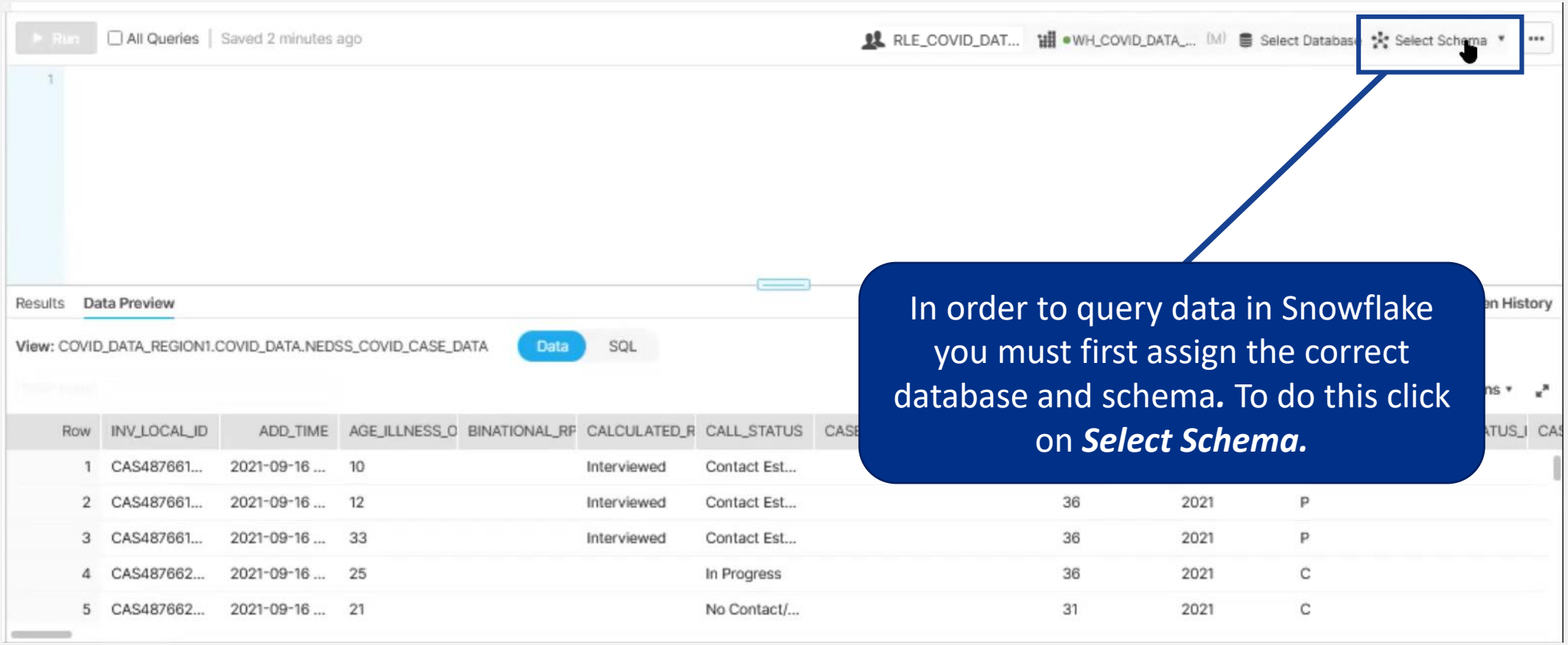
*Complete overview can be found here

# Accessing Data: Queries and Examples

# Step 3A: Configure Database and Schema

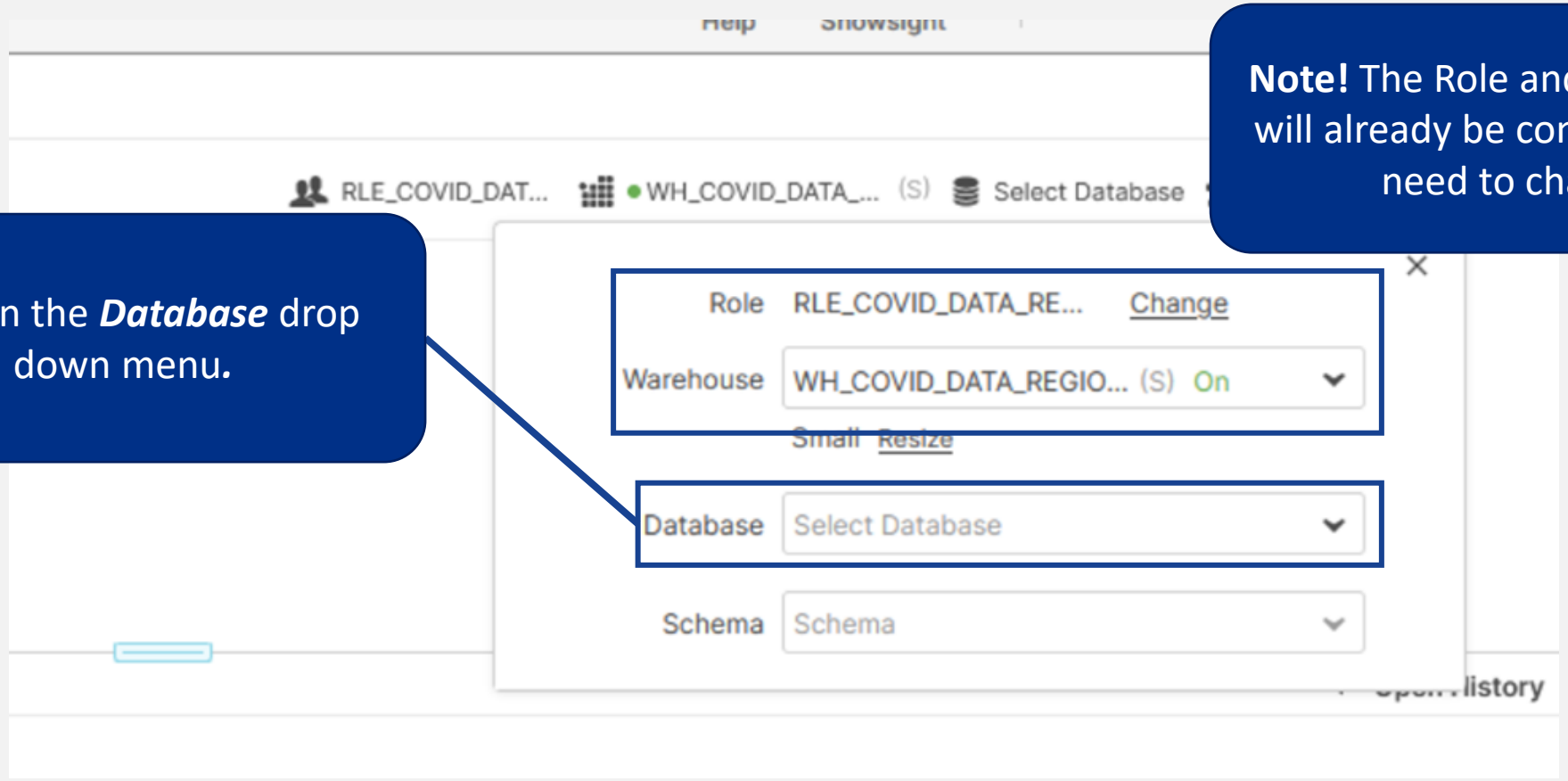
**Pro tip!**

Make sure the browser window is maximized to see the full database and schema ribbon!



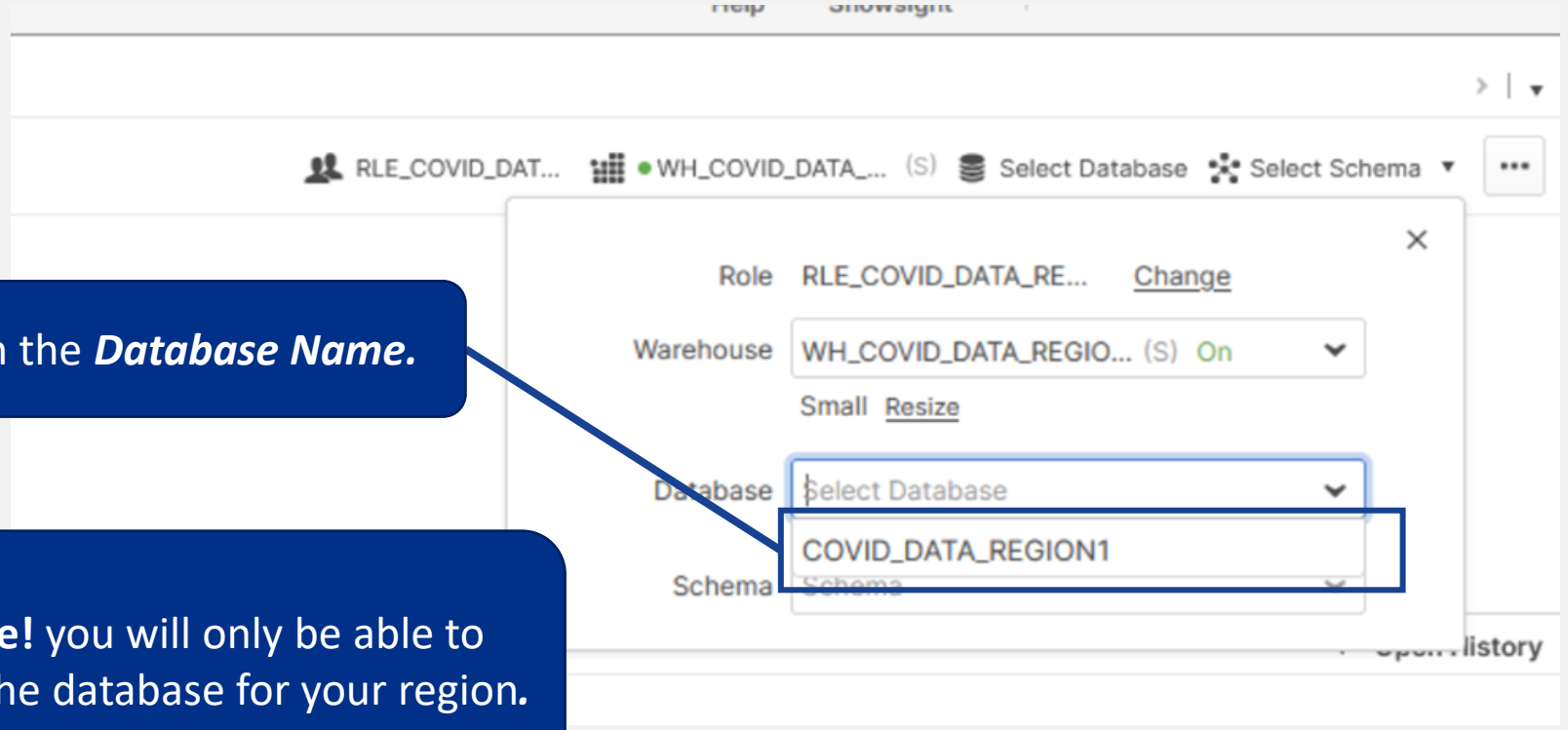
# Step 3B: Configure Database and Schema

Click on the ***Database*** drop down menu.




**Note!** The Role and Warehouse will already be configured – no need to change.

# Step 3C: Configure Database and Schema

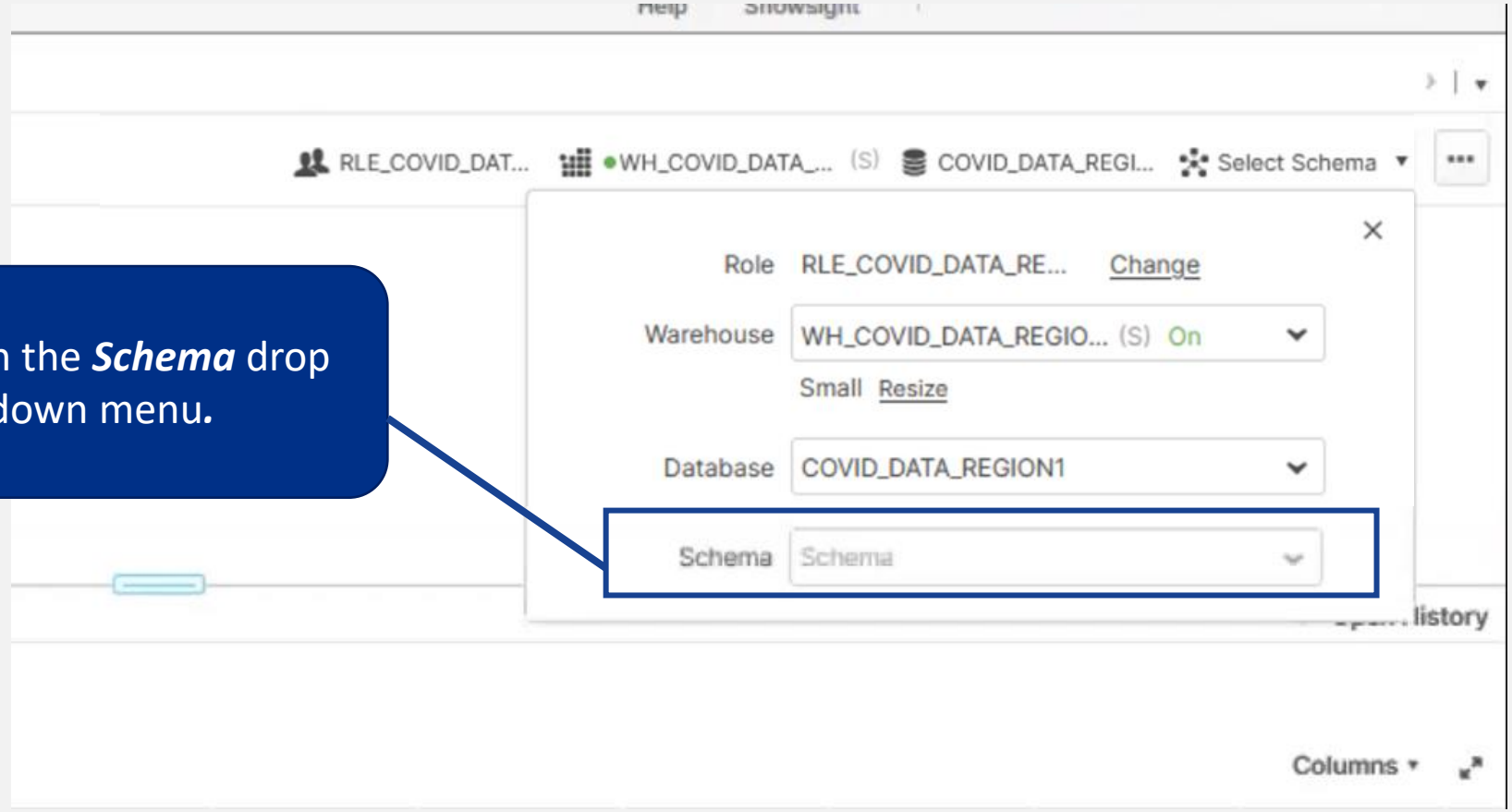



Click on the ***Database Name.***

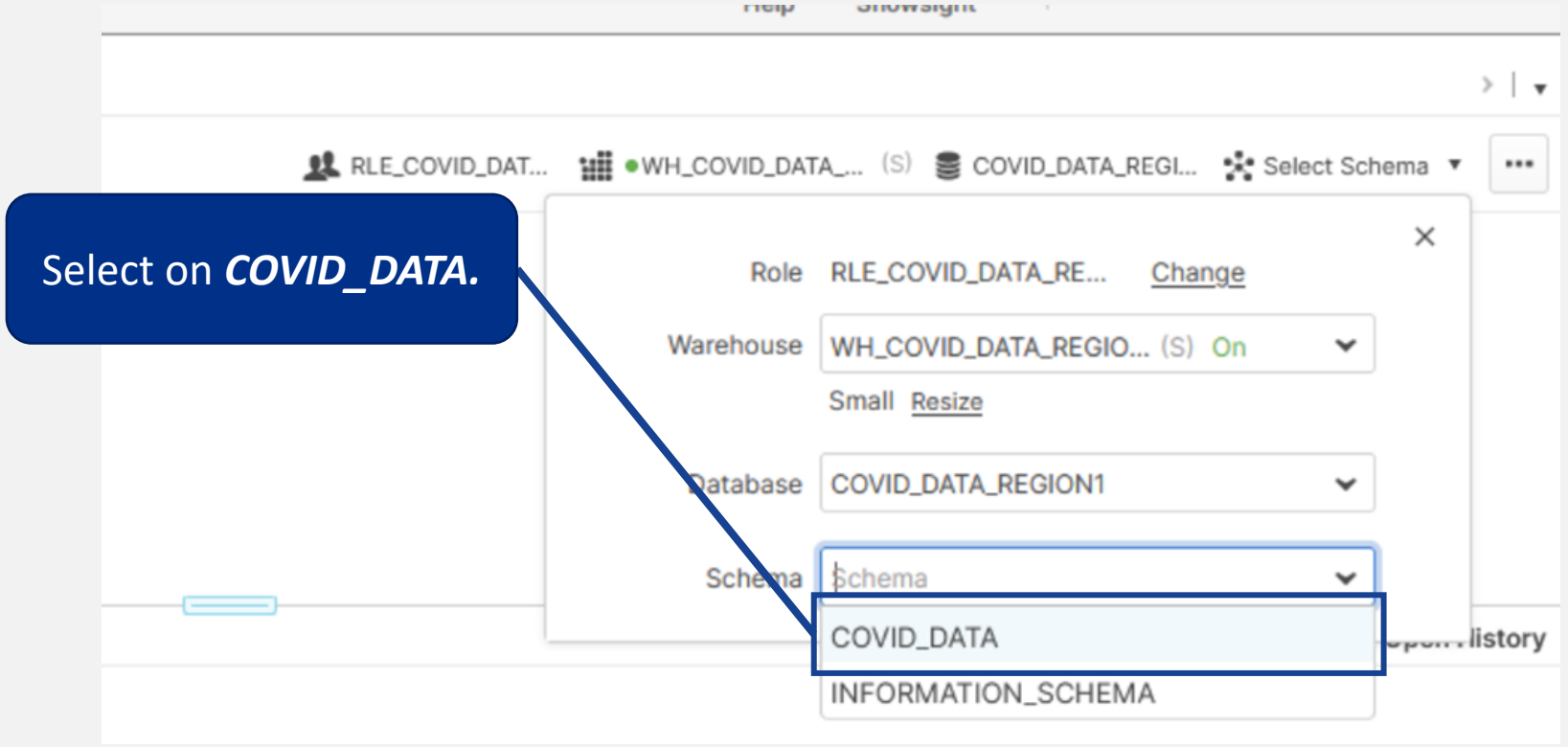
**Note!** you will only be able to view the database for your region.

# Step 3D: Configure Database and Schema

Click on the ***Schema*** drop down menu.

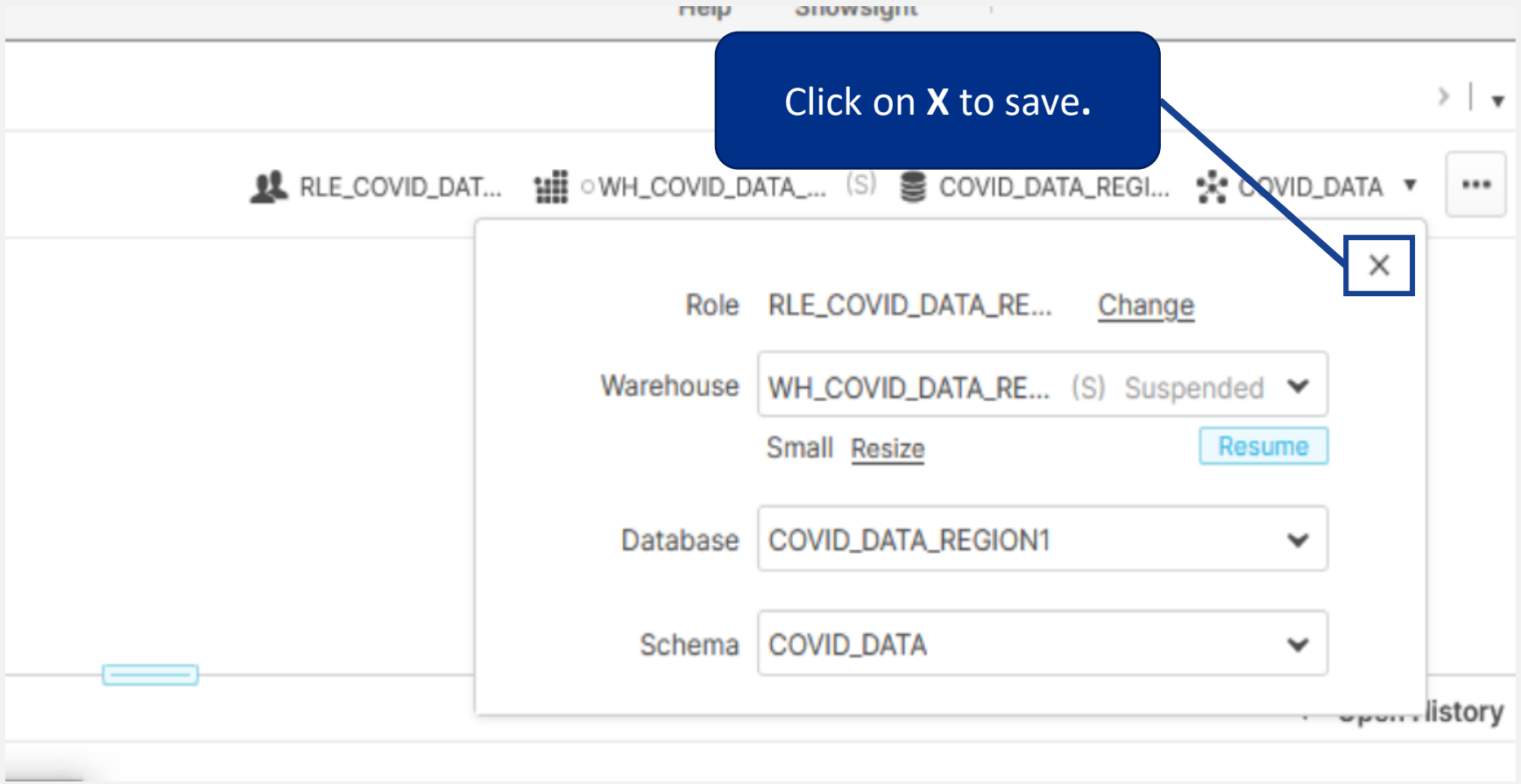


# Step 3E: Configure Database and Schema

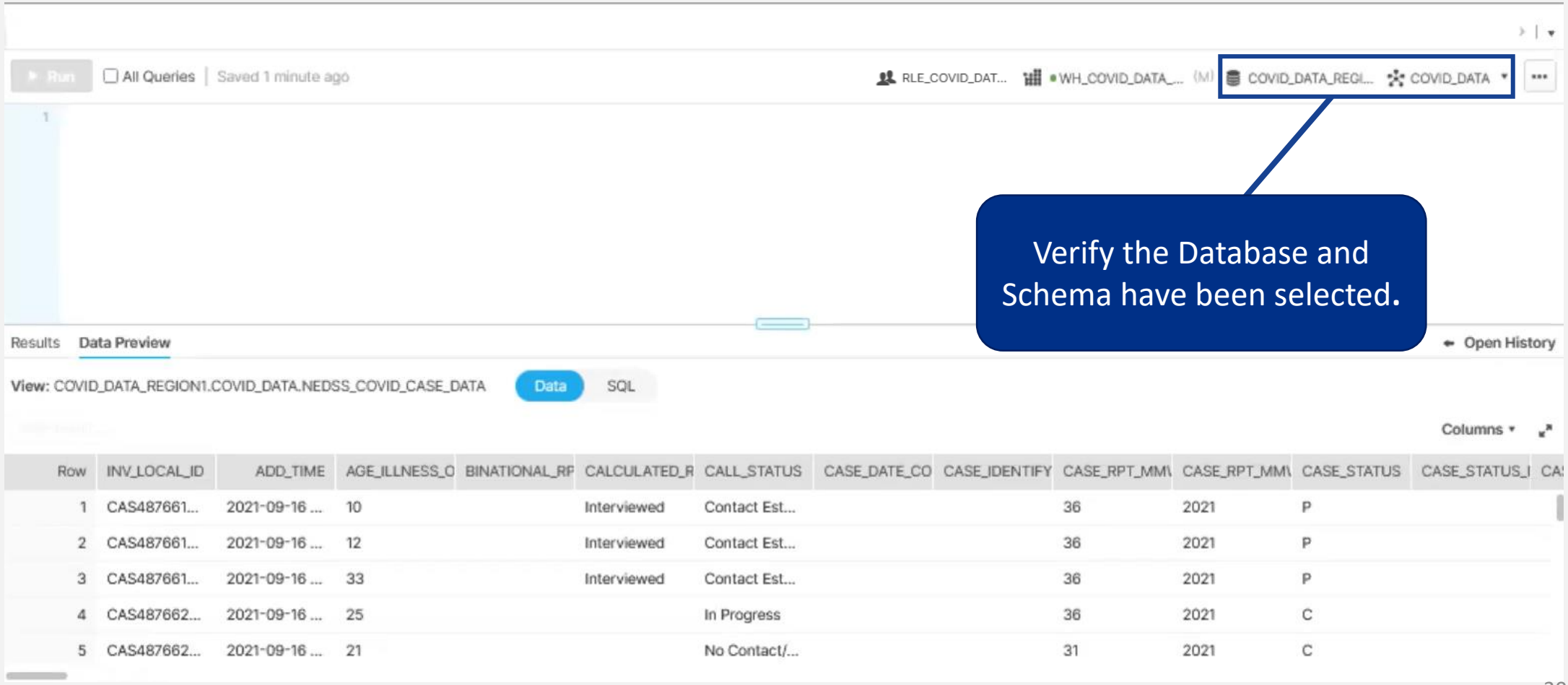



Select on ***COVID_DATA.***

# Step 3F: Configure Database and Schema



# Step 3G: Configure Database and Schema

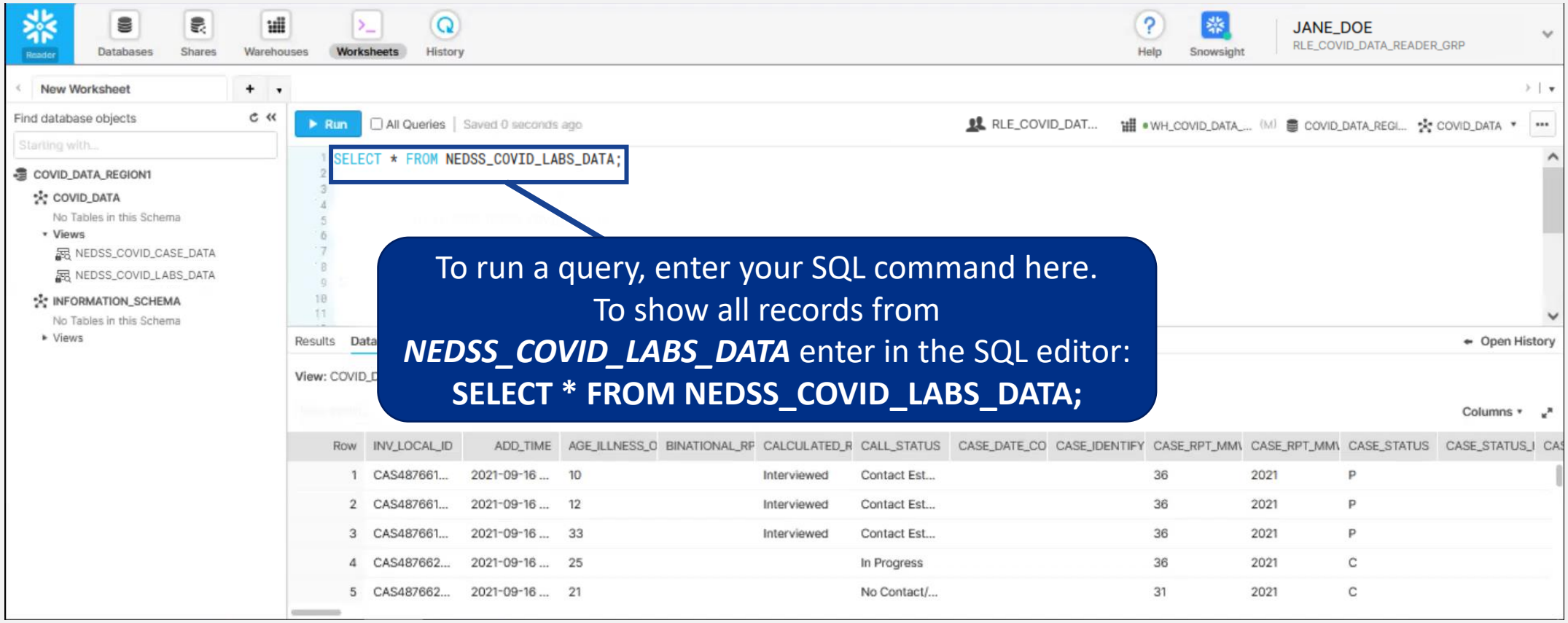



| Row | INV_LOCAL_ID | ADD_TIME | AGE_ILLNESS_O | BINATIONAL_RP | CALCULATED_R | CALL_STATUS | CASE_DATE_CO | CASE_IDENTIFY | CASE_RPT_MM\ | CASE_RPT_MM\ | CASE_STATUS | CASE_STATUS_I | CA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | CAS487661... | 2021-09-16 ... | 10 | | Interviewed | Contact Est... | | | 36 | 2021 | P | | |
| 2 | CAS487661... | 2021-09-16 ... | 12 | | Interviewed | Contact Est... | | | 36 | 2021 | P | | |
| 3 | CAS487661... | 2021-09-16 ... | 33 | | Interviewed | Contact Est... | | | 36 | 2021 | P | | |
| 4 | CAS487662... | 2021-09-16 ... | 25 | | | In Progress | | | 36 | 2021 | C | | |
| 5 | CAS487662... | 2021-09-16 ... | 21 | | | No Contact/... | | | 31 | 2021 | C | | |

# Step 3H: View all data with a Snowflake Query

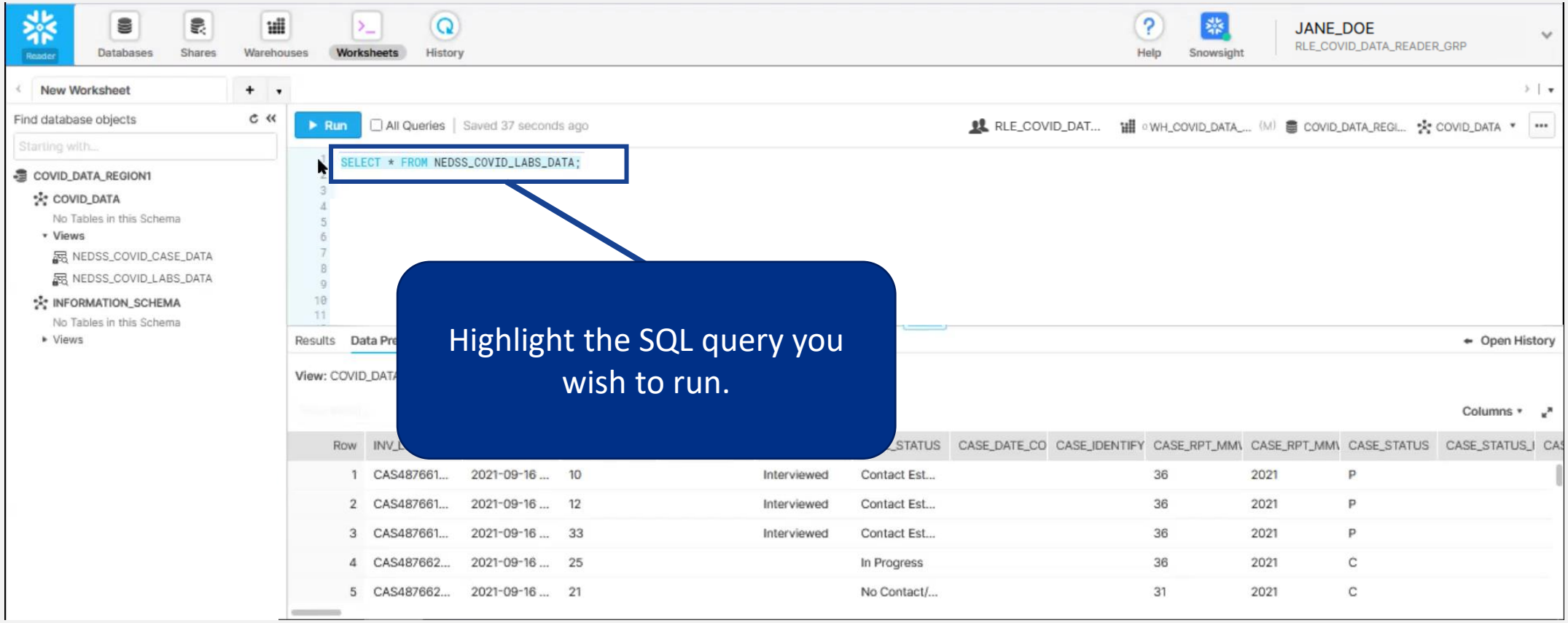
**Pro tip!**

Check out the **Additional Resources** section of this guide for SQL basics and more sample queries.

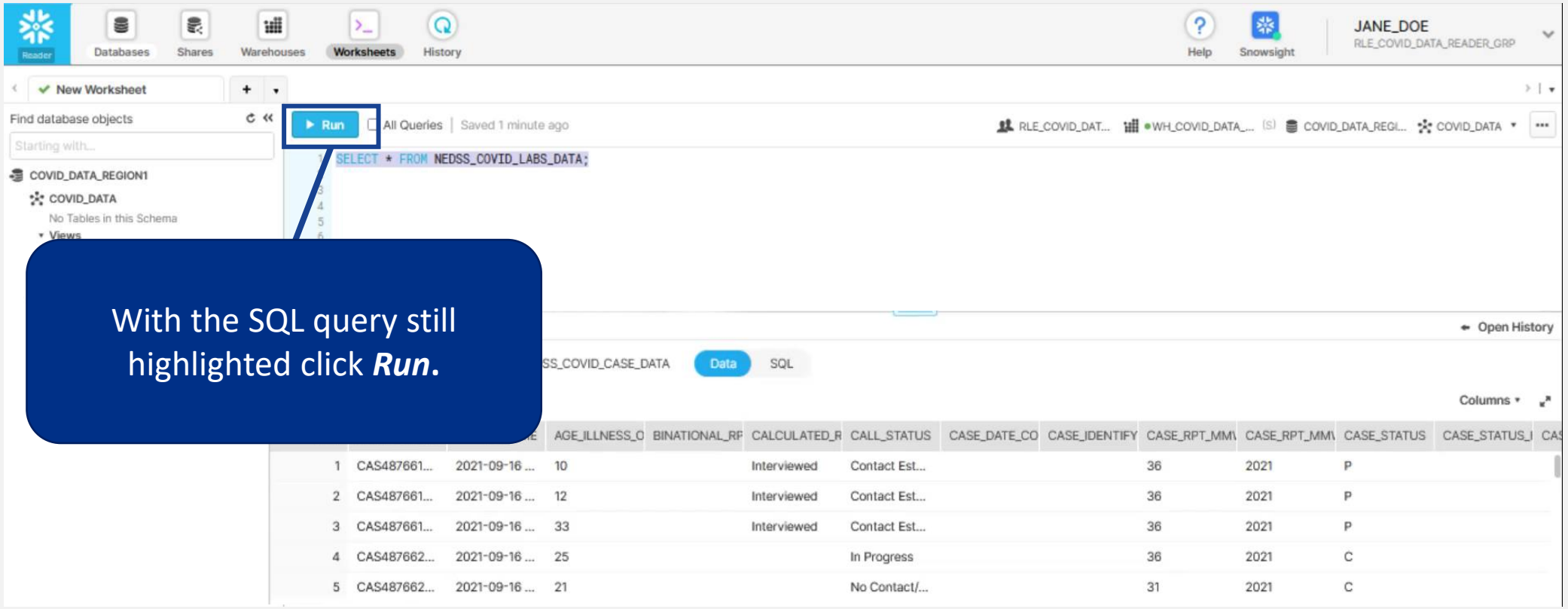



To run a query, enter your SQL command here.
To show all records from
***NEDSS_COVID_LABS_DATA*** enter in the SQL editor:
**SELECT * FROM NEDSS_COVID_LABS_DATA;**

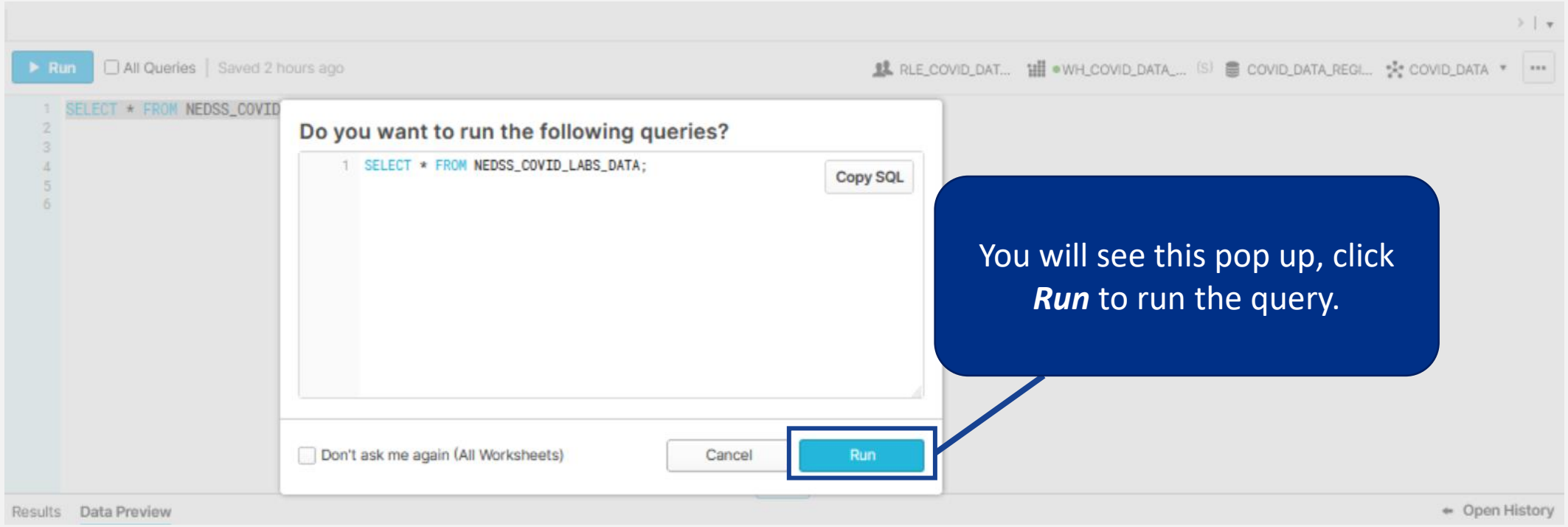
# Step 3I: View all data with a Snowflake Query



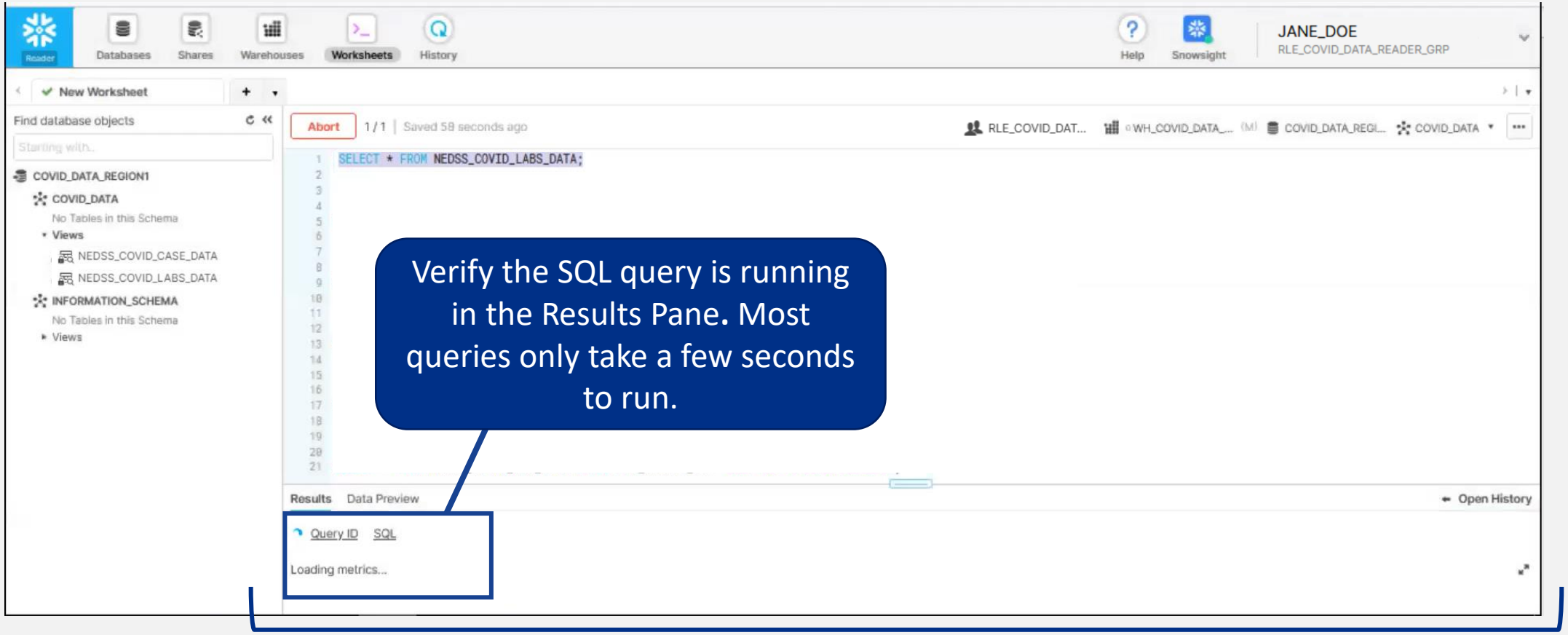
# Step 3J: View all data with a Snowflake Query



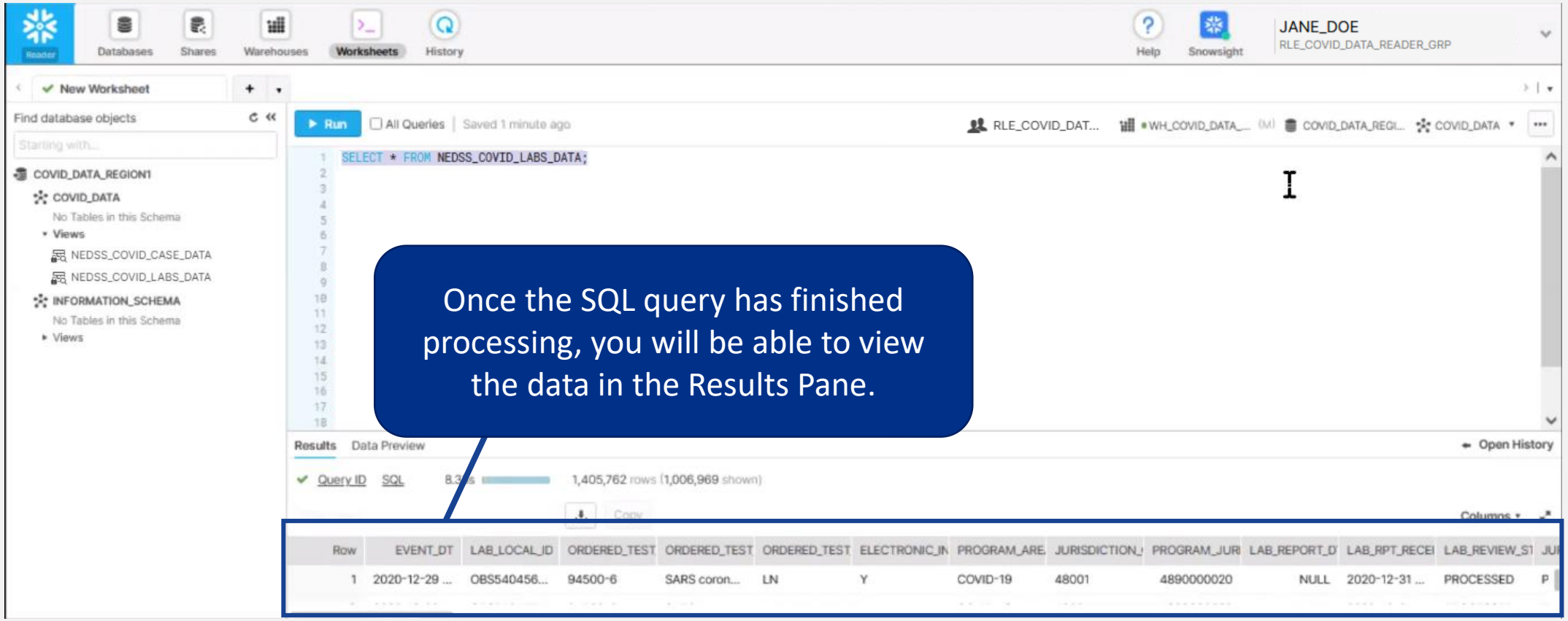
# Step 3K: View all data with a Snowflake Query



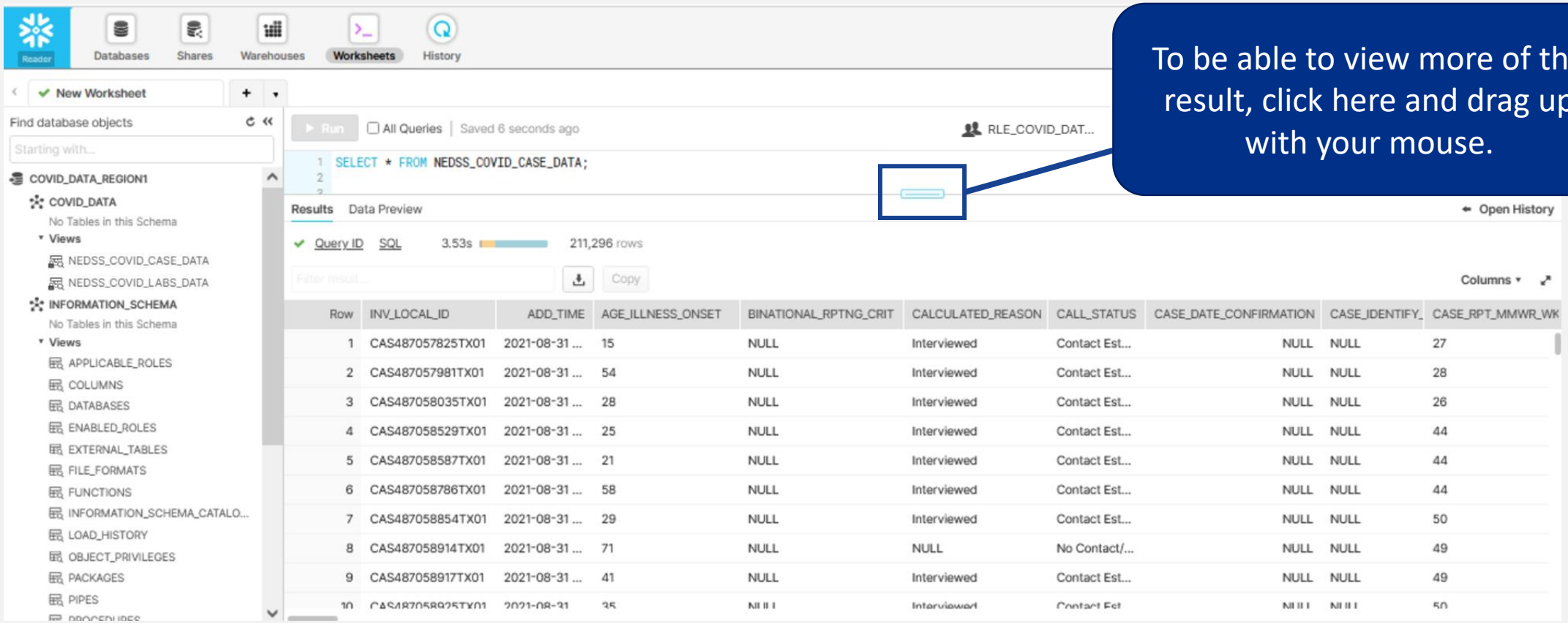
# Step 3L: View all data with a Snowflake Query



# Step 3M: View all data with a Snowflake Query

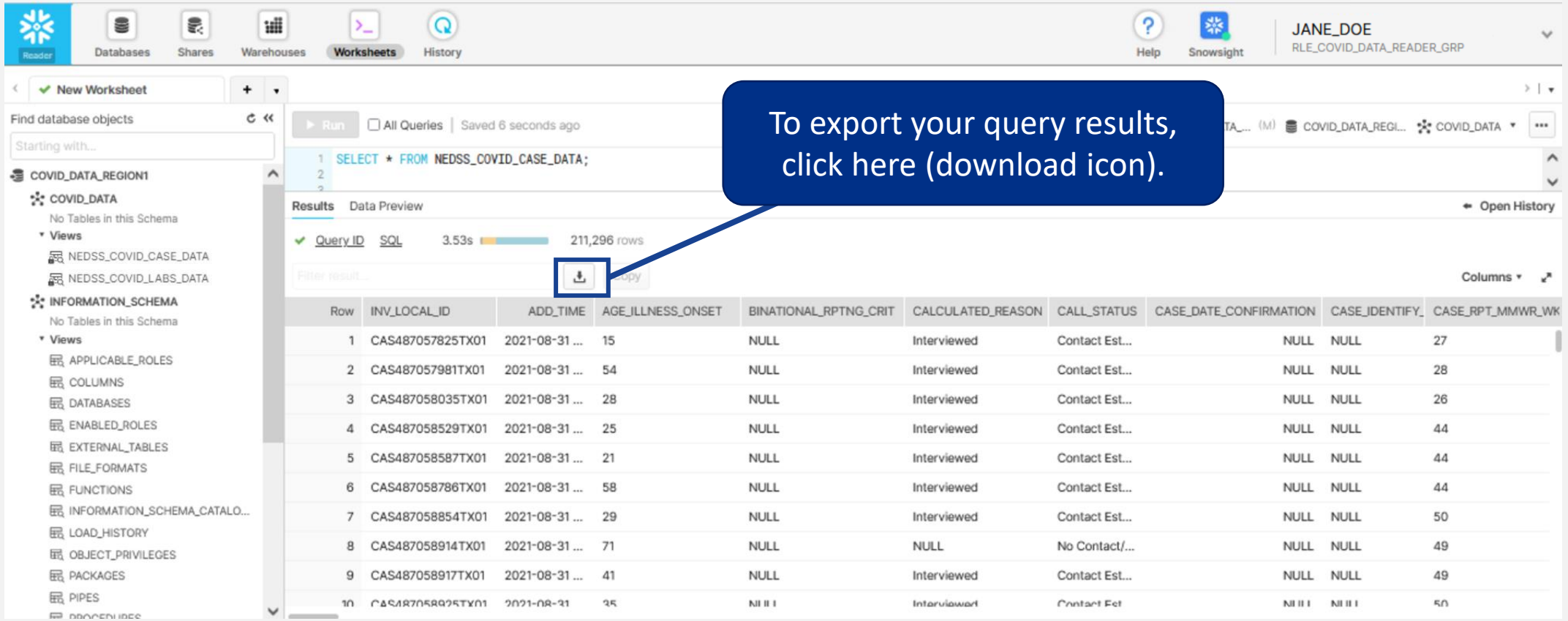


# Step 3N: View all data with a Snowflake Query



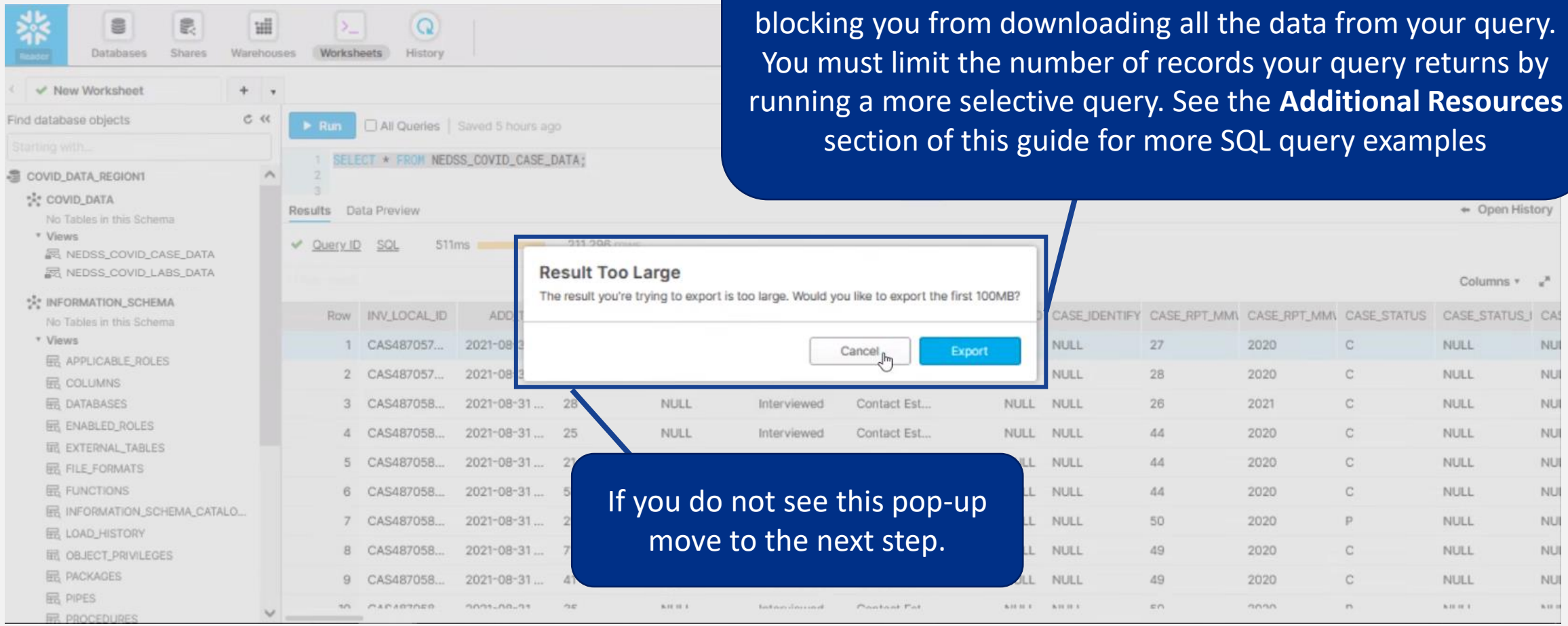

| Row | INV_LOCAL_ID | ADD_TIME | AGE_ILLNESS_ONSET | BINATIONAL_RPTNG_CRIT | CALCULATED_REASON | CALL_STATUS | CASE_DATE_CONFIRMATION | CASE_IDENTIFY_ | CASE_RPT_MMWR_WK |
|---|---|---|---|---|---|---|---|---|---|
| 1 | CAS487057825TX01 | 2021-08-31 ... | 15 | NULL | Interviewed | Contact Est... | NULL | NULL | 27 |
| 2 | CAS487057981TX01 | 2021-08-31 ... | 54 | NULL | Interviewed | Contact Est... | NULL | NULL | 28 |
| 3 | CAS487058035TX01 | 2021-08-31 ... | 28 | NULL | Interviewed | Contact Est... | NULL | NULL | 26 |
| 4 | CAS487058529TX01 | 2021-08-31 ... | 25 | NULL | Interviewed | Contact Est... | NULL | NULL | 44 |
| 5 | CAS487058587TX01 | 2021-08-31 ... | 21 | NULL | Interviewed | Contact Est... | NULL | NULL | 44 |
| 6 | CAS487058786TX01 | 2021-08-31 ... | 58 | NULL | Interviewed | Contact Est... | NULL | NULL | 44 |
| 7 | CAS487058854TX01 | 2021-08-31 ... | 29 | NULL | Interviewed | Contact Est... | NULL | NULL | 50 |
| 8 | CAS487058914TX01 | 2021-08-31 ... | 71 | NULL | NULL | No Contact/... | NULL | NULL | 49 |
| 9 | CAS487058917TX01 | 2021-08-31 ... | 41 | NULL | Interviewed | Contact Est... | NULL | NULL | 49 |
| 10 | CAS487058925TX01 | 2021-08-31 | 35 | NULL | Interviewed | Contact Est | NULL | NULL | 50 |

# Step 3O: Export Data from a Snowflake Query



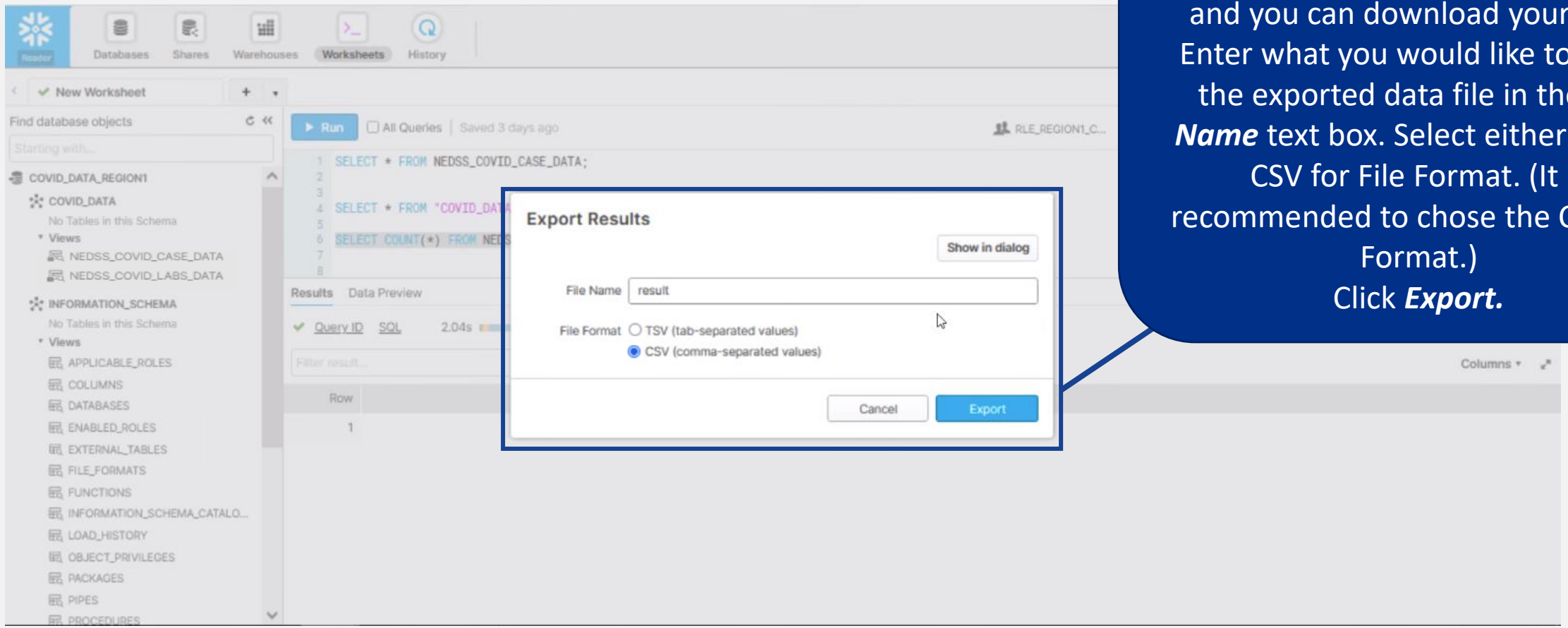
# Step 3P: Export Data from a Snowflake Query

If you see this pop-up, it means Snowflake's export limits are blocking you from downloading all the data from your query. You must limit the number of records your query returns by running a more selective query. See the **Additional Resources** section of this guide for more SQL query examples




If you do not see this pop-up move to the next step.

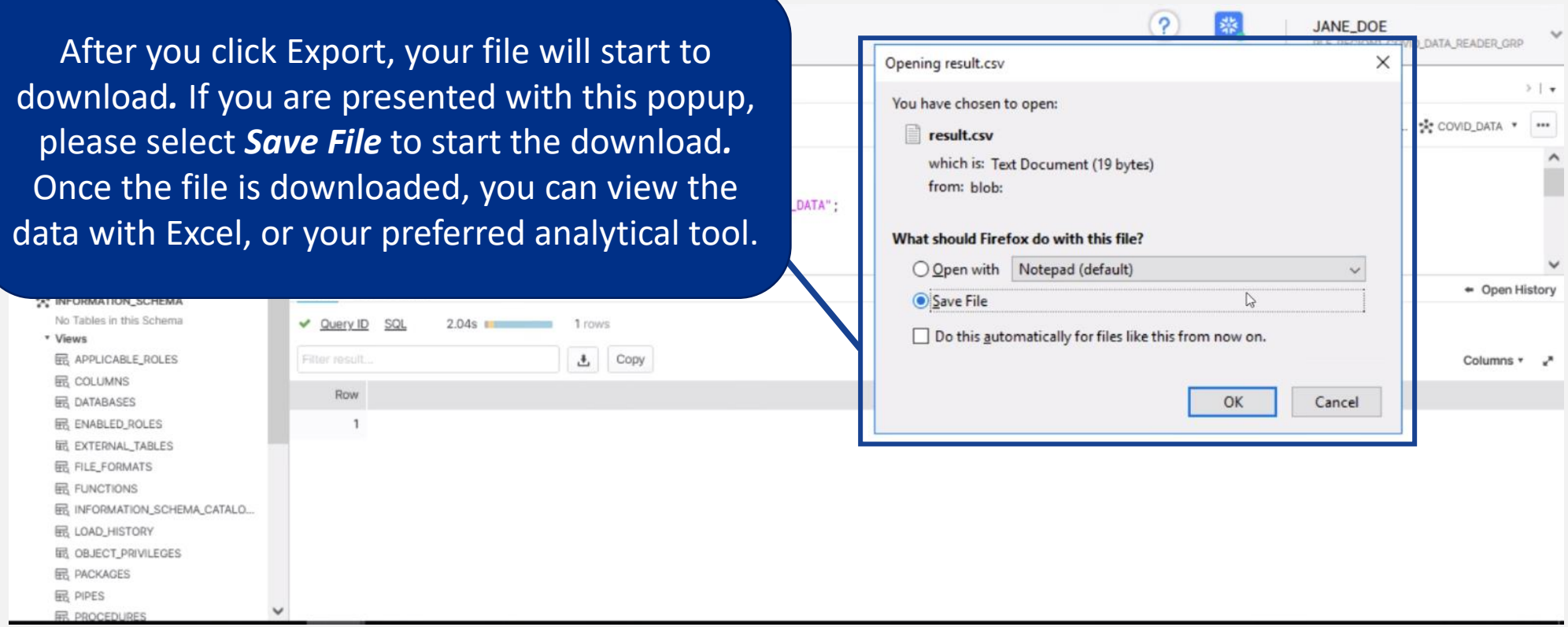
# Step 3Q: Export Data from a Snowflake Query




If you see this pop-up, your results are within the Snowflake export limits and you can download your data. Enter what you would like to name the exported data file in the ***File Name*** text box. Select either TSV or CSV for File Format. (It is recommended to chose the CSV File Format.)
Click ***Export.***

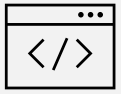
# Step 3R: Export Data from a Snowflake Query

After you click Export, your file will start to download. If you are presented with this popup, please select ***Save File*** to start the download. Once the file is downloaded, you can view the data with Excel, or your preferred analytical tool.



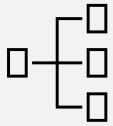
# Additional Resources

# Glossary of Terms

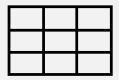


**Structured Query Language (SQL) -** programming language used to run queries in Snowflake.

**Snowflake Database** - is a collection of data stored in Snowflake.



**Snowflake Schema -** represents how data is stored in a database. Users have access to the *COVID_DATA* schema which has all the data that a user can access with their Snowflake Reader account.



**Snowflake Table -** where data is stored in Snowflake. For data security users will not have access to any Snowflake tables at this time.

**Snowflake View -** is a virtual table that contains data from one or more Snowflake tables. Users have access to the NEDSS_COVID_CASE_DATA view which contains all the COVID case information for their jurisdiction or region and the NEDSS_COVID_LABS_DATA view which contains all the lab information for their jurisdiction or region.

# Additional SQL Queries to Try in Snowflake

**Pro tip!**

Anything written after two dashes ( ) in SQL will appear in green and will not execute when the code is run! This allows you to add comments to your SQL code in Snowflake.

**--Queries for NEDSS COVID CASE DATA**

```
SELECT * FROM NEDSS_COVID_CASE_DATA; -- Returns all data from NEDSS_COVID_CASE_DATA

SELECT COUNT(*) FROM NEDSS_COVID_CASE_DATA; -- Returns total number of records from NEDSS_COVID_CASE_DATA

SELECT DISTINCT JURISDICTION_NM FROM NEDSS_COVID_CASE_DATA; -- Returns all the jurisdiction/region names available in NEDSS_COVID_CASE_DATA without any duplicate names appearing in the list.

SELECT * FROM NEDSS_COVID_CASE_DATA WHERE CASE_STATUS='C'; -- Returns all records with confirmed cases status

SELECT * FROM NEDSS_COVID_CASE_DATA WHERE CASE_RPT_MMWR_YR='2021'; -- Returns all records where the case was reported in the year 2021

SELECT * FROM NEDSS_COVID_CASE_DATA WHERE INV_START_DT>='2021-01-01'; -- Returns all records where the case investigation start date was on or past 2021-01-01
```

# Additional SQL Queries to Try in Snowflake

**---Queries for NEDSS COVID LABS DATA**

```
SELECT * FROM NEDSS_COVID_LABS_DATA; -- Returns all data from NEDSS_COVID_LABS_DATA

SELECT COUNT(*) FROM NEDSS_COVID_LABS_DATA; -- Returns total number of records from NEDSS_COVID_LABS_DATA

SELECT DISTINCT JURISDICTION_NM FROM NEDSS_COVID_LABS_DATA;  -- Returns all the jurisdiction/region names available in NEDSS_COVID_LABS_DATA without any duplicate names appearing in the list.

SELECT * FROM NEDSS_COVID_LABS_DATA WHERE RESULT_CATEGORY='Positive'; -- Returns all records where the category of the test result was “Positive”

SELECT * FROM NEDSS_COVID_LABS_DATA WHERE RESULT_TEST_TYPE='PCR'; -- Returns all records where the type of test was “PCR”

SELECT * FROM NEDSS_COVID_LABS_DATA WHERE LAB_UPDATE_DT>='2021-03-11 22:27:52.220'; -- Returns all records where the last updated date of the lab was on or past “2021-03-11 22:27:52.220”

SELECT DISTINCT PERFORM_FACILITY_NAME FROM NEDSS_COVID_LABS_DATA; -- Returns all the testing facility names available in NEDSS_COVID_LABS_DATA without any duplicate names appearing in the list.
```

# Helpful Links

- [Snowflake Getting Started]
- [Additional SQL Syntax]