Postgres and pgAdmin

DATA ANALYTICS BOOTCAMP

Install PostgreSQL and pgAdmin

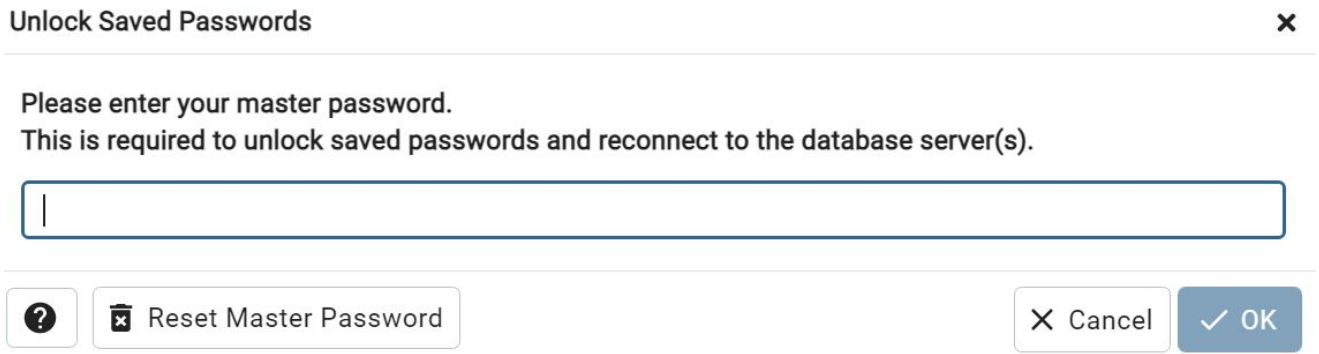
- Download and install the latest version of PostgreSQL for your OS: https://www.enterprisedb.com/downloads/postgres-postgresql-downloads
  - For Mac and Windows 64-bit, this will be Version 15.1
  - For Windows 32-bit, this will be Version 10.23
  - Linux users will need to use the following link: https://www.postgresql.org/download/. You will also need to install pgAdmin. Installation instructions can be found here: https://www.pgadmin.org/download/.
- During installation you can accept all the defaults **except** you can uncheck **Stack Builder** on the **Select Components** menu
- When asked for a database password, use **postgres**
- Confirm the port number is **5432**

Very important to use this password!

*Note: Since this is not secure data, we can set our passwords to **postgres**. This is not something you would do with secure data in a production environment*

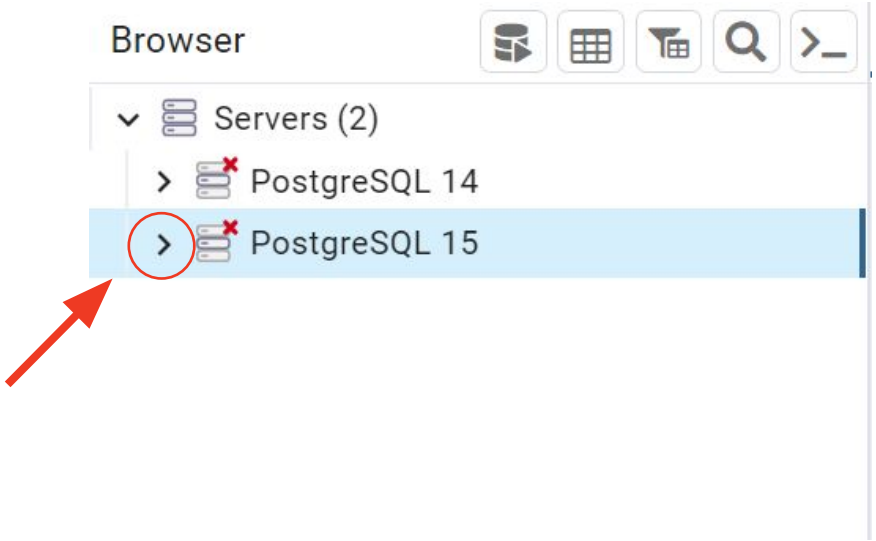
Launch pgAdmin

- Either search for **pgAdmin** in the search bar or navigate to it in your file browser
- Once launched, it will prompt you for a password, which you should have set to **postgres** during the install.



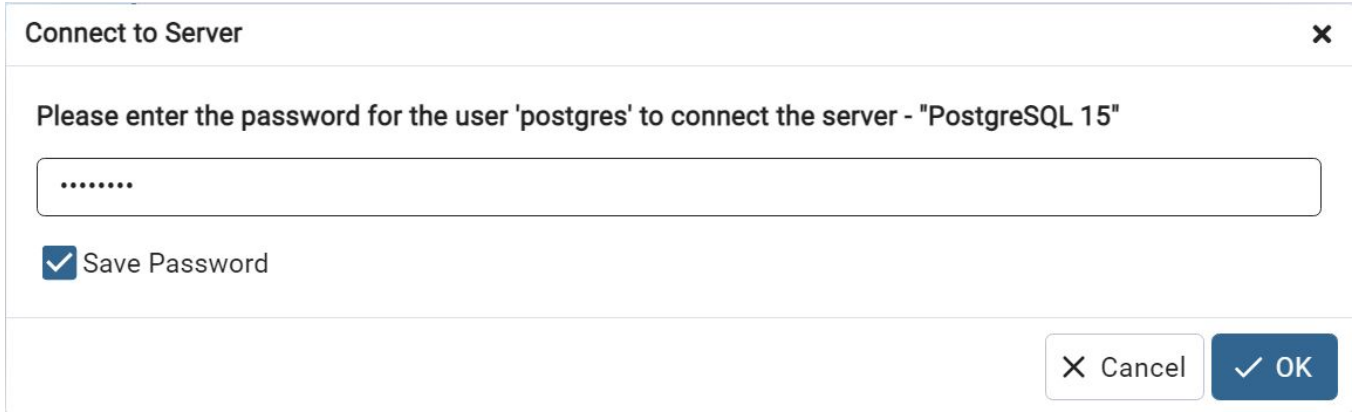
Access Server

- Click the arrow next to **PostgreSQL 15** in the left hand menu



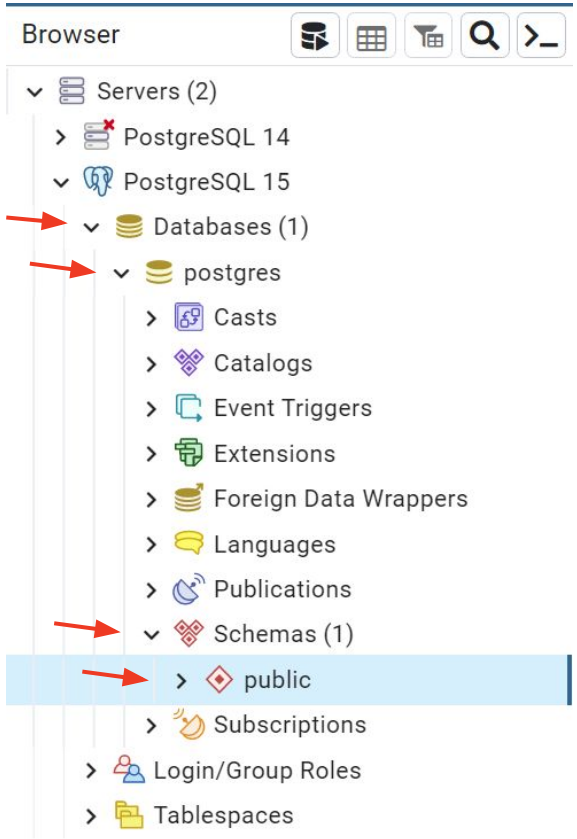
Access Server

- You will be prompted to enter a password for the server, which is also **postgres**
- Check the **Save Password** box so you won't be asked for a password at this stage again
- Click **OK**



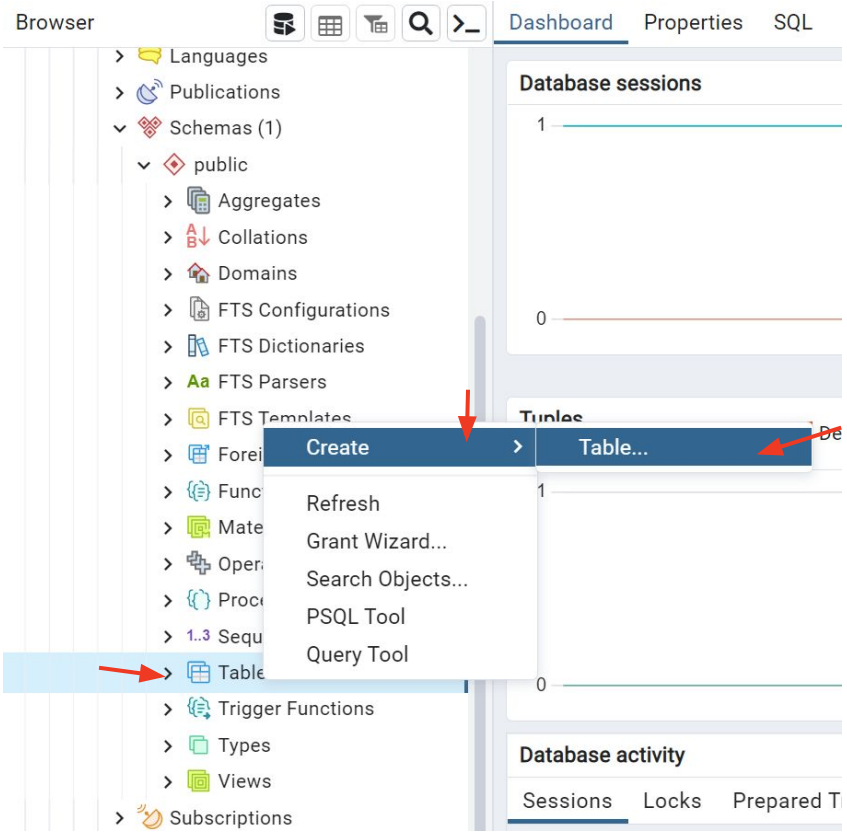
Create a Table

- Click the arrow next to **Databases**
- Click the arrow next to **postgres**
- Click the arrow next to **Schemas**
- Click the arrow next to **public**



Create a Table

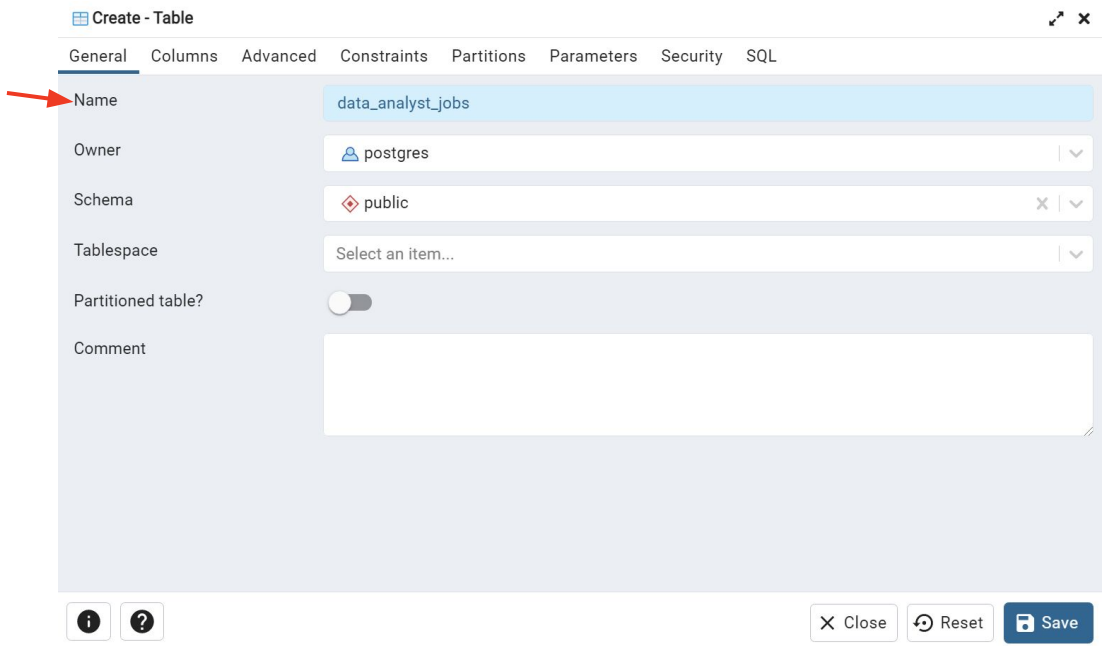
- Right-click on **Tables**
- Hover over **Create**
- Select **Table**



NASHVILLE SOFTWARE SCHOOL

Create a Table

- In the **General** tab, name the table **data_analyst_jobs**

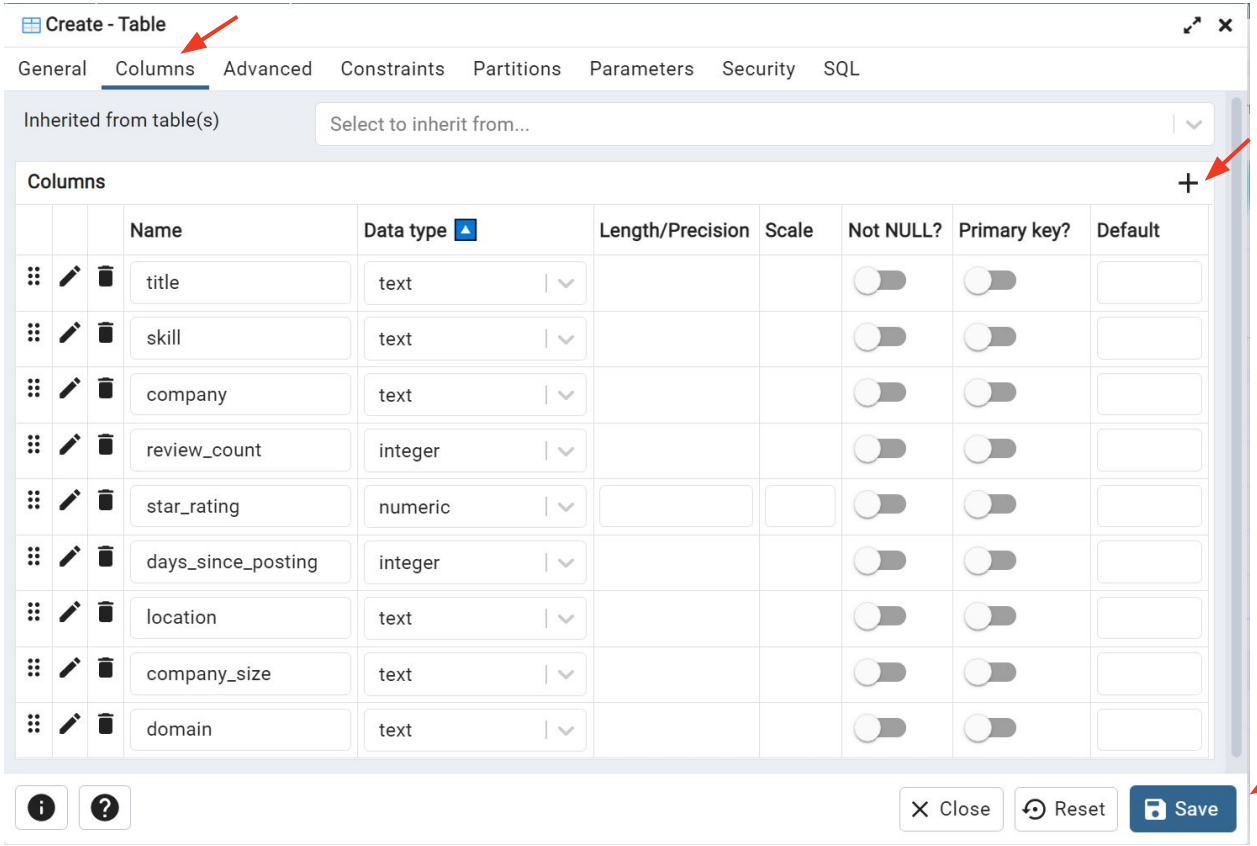


NASHVILLE SOFTWARE SCHOOL

Create a Table

- Under the **Columns** tab define your columns with the **names** and **data types** shown on the next slide
  - The names need to match what you see **EXACTLY**; be careful not to put in any extra spaces at the beginning or end
  - **Click the + symbol** in the upper right hand corner of the window to create a new column
  - The only fields you should have to edit are **Name** and **Data type**
  - These column names and types correspond to the columns in the csv we will be using
- Click **Save** when you are done

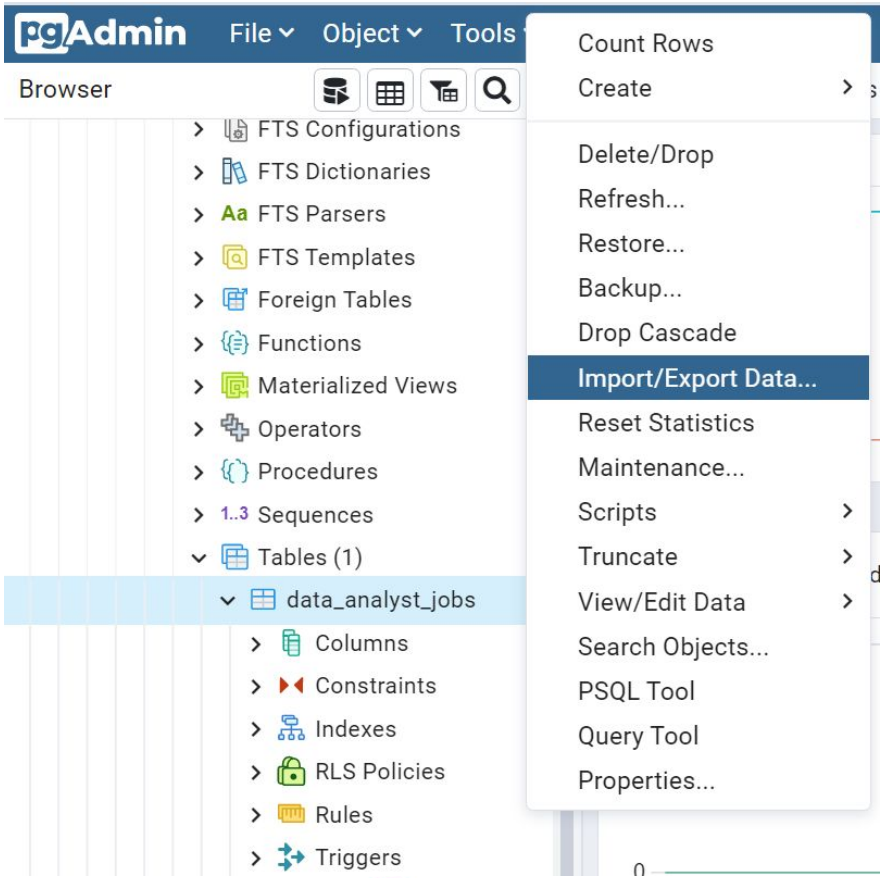
Create a Table



Create a Table

After clicking Save...

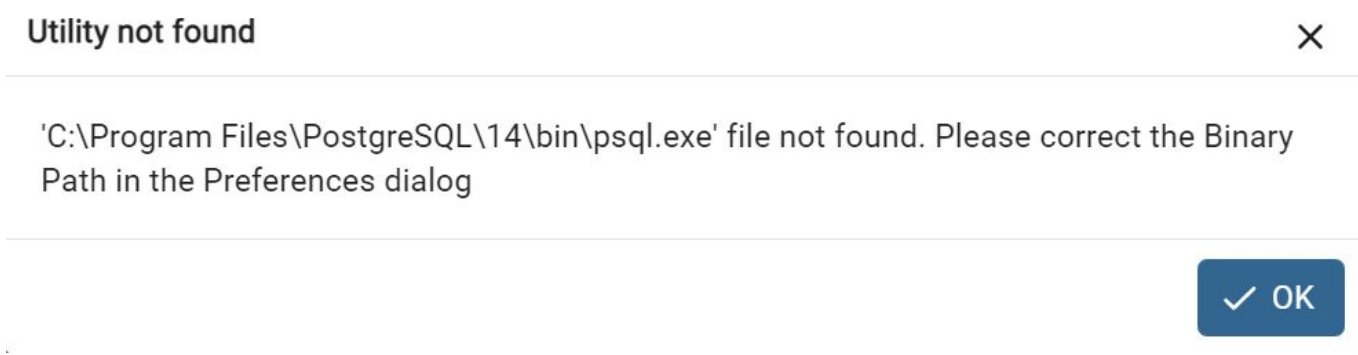
- Right-click on the **data_analyst_jobs** table
- Select **Import/Export Data...**



NASHVILLE SOFTWARE SCHOOL

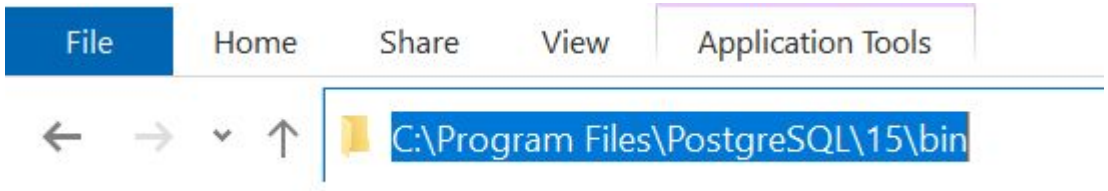
Utility not found

- If you get this error message (or anything similar that mentions the Binary Path), you must take extra steps to set the binary path.



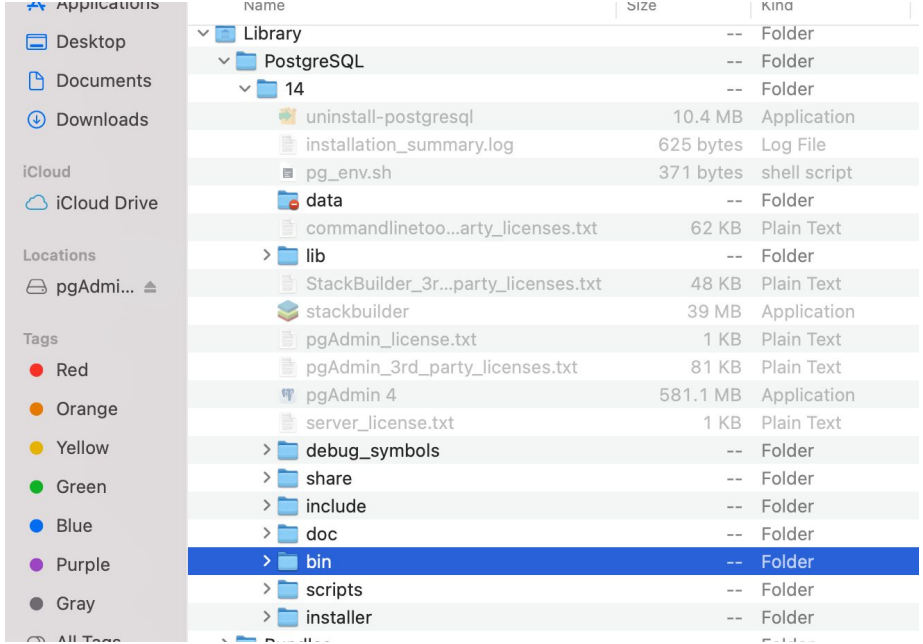
Utility not found

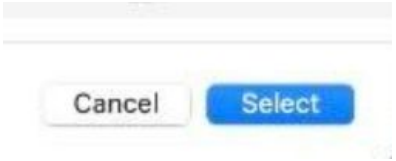
- On Windows, navigate to the **Program Files** folder
- Find the **PostgreSQL** folder and inside. There should be a **15** folder with a **bin** folder inside of that (14 for Macs)
- Click in the address bar and copy the **entire path** once you are inside the **bin** folder
- It should look something like this:



Utility not found

- On Mac, navigate to the **Macintosh HD** folder
- From there go to **Library → PostgreSQL → 15** and select **bin** (14 for Macs)
- Click **Select** in the bottom right corner

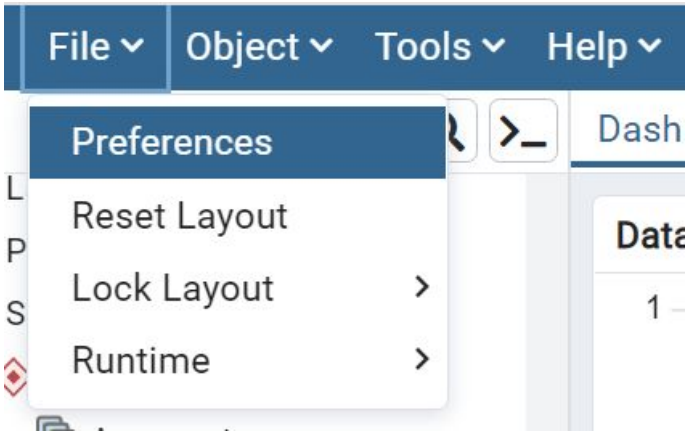




NASHVILLE SOFTWARE SCHOOL

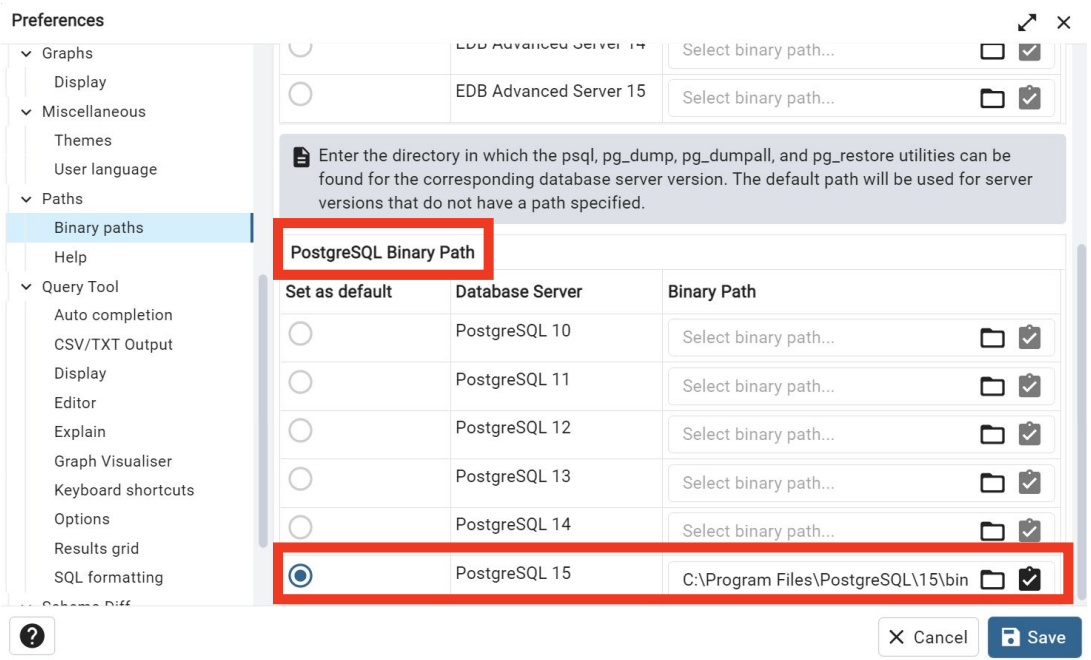
Utility not found

- In pgAdmin, Go to **File** → **Preferences**



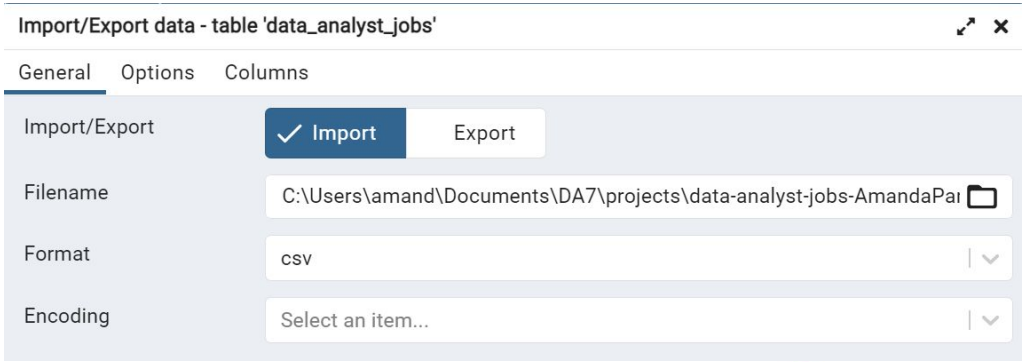
Utility not found

- In the left hand menu, scroll down to **Paths** and click on **Binary Paths**
- Scroll down to the second set of servers called **PostgreSQL Binary Path**
- Paste the path into the **Binary Path** space next to **PostgreSQL 15 (14 for Macs)**
- Check the button next to it under **Set as default**
- Click **Save**



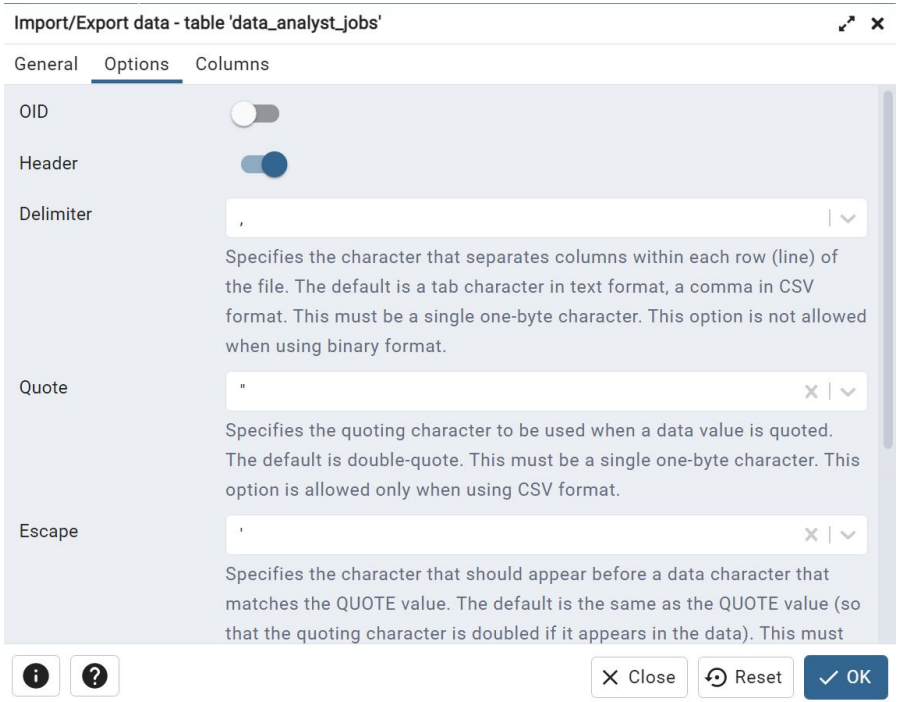
Create a Table

- In the **General** tab, be sure the **Import/Export** option is toggled to **Import**
  Selecting Export will clear out the contents of the CSV and you will have to clone down the data again
- Click on the **folder** icon and navigate to the file **indeed_data_analysts.csv** in the data folder of the assignment repo
- Double click on the file



Create a Table

- In the **Options** tab, change the **Header** button to **blue** to indicate our file has headers for the columns
- Be sure the **Delimiter** is set to “,”
- Click **OK**



NASHVILLE SOFTWARE SCHOOL

Success!

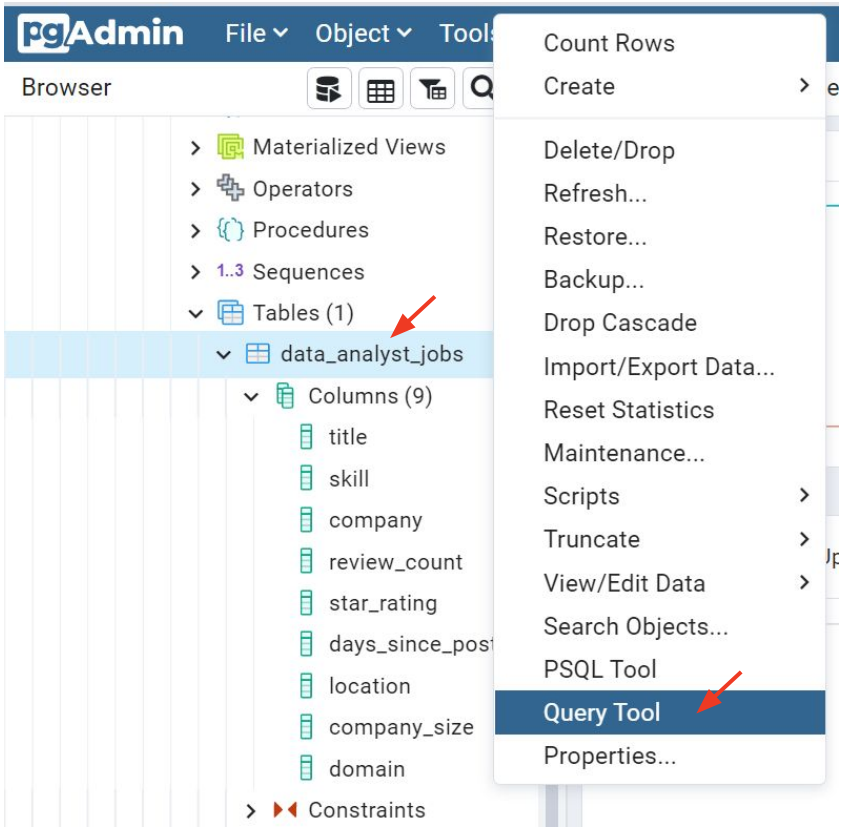
- Your data will import and show the following success message

pgAdmin Key Features & Tips

NASHVILLE SOFTWARE SCHOOL

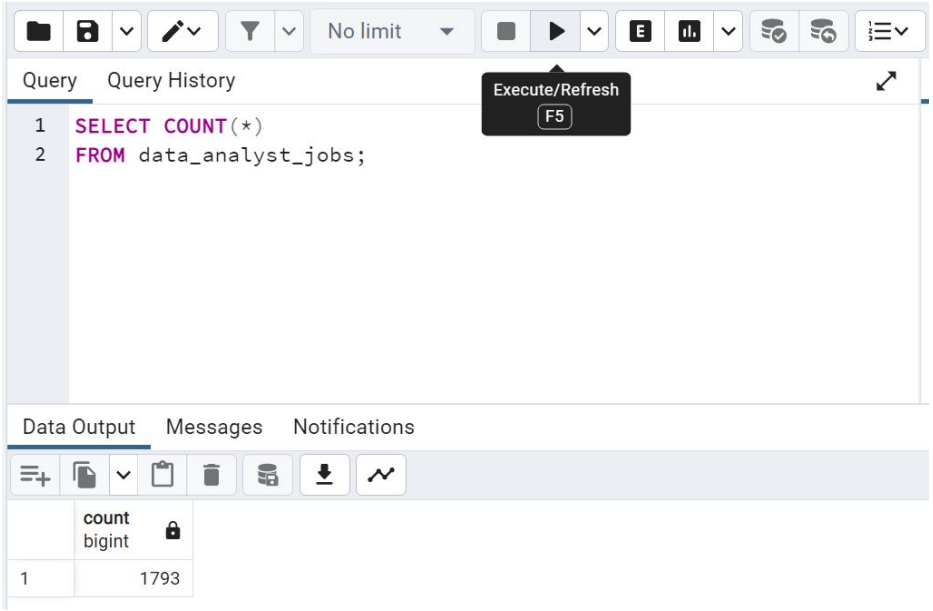
Starting a Script

- To start a new script where you will write your queries, **right-click** on **data_analyst_jobs** and select **Query Tool**



Starting a Script

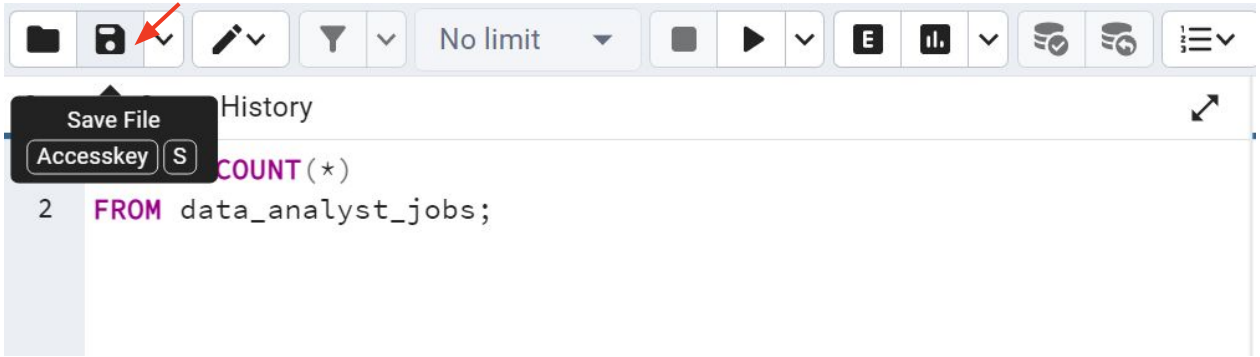
- To test that you have imported successfully, run the query shown to make sure you have the right number of rows
- You're now ready to write your own queries



NASHVILLE SOFTWARE SCHOOL

Saving Your Script

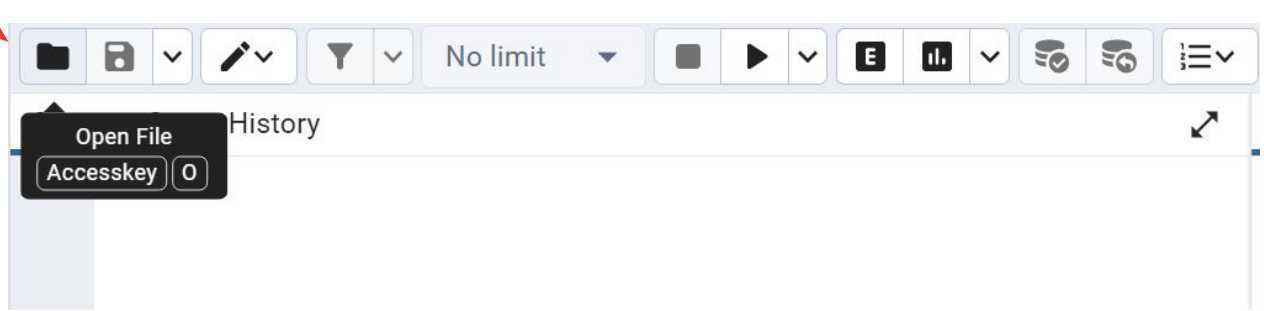
- To save, first click the floppy disk icon at the top left of the query tool



- If not already created, create a folder called **scripts** within your SQL folder
- Navigate into the **scripts** folder, name the file something meaningful, ensure the file type is ***.sql** and click **Save**

Opening Your Script

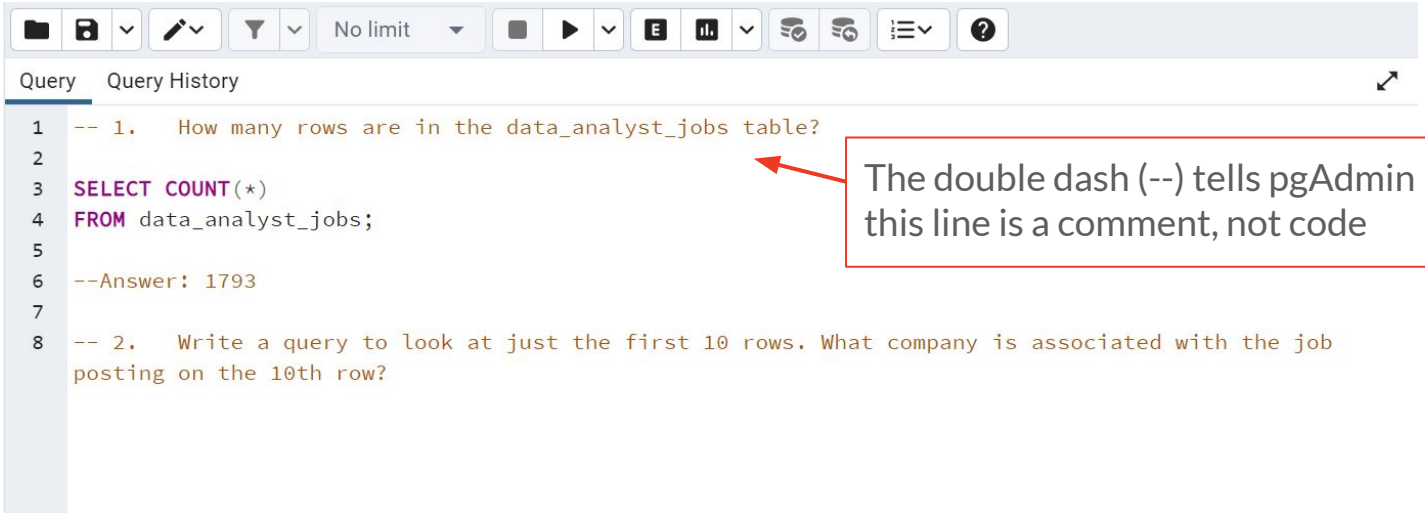
- When you return to work on the script you saved, click on the folder icon



- Navigate back to the **scripts** folder and select your saved script

Formatting Questions and Answers

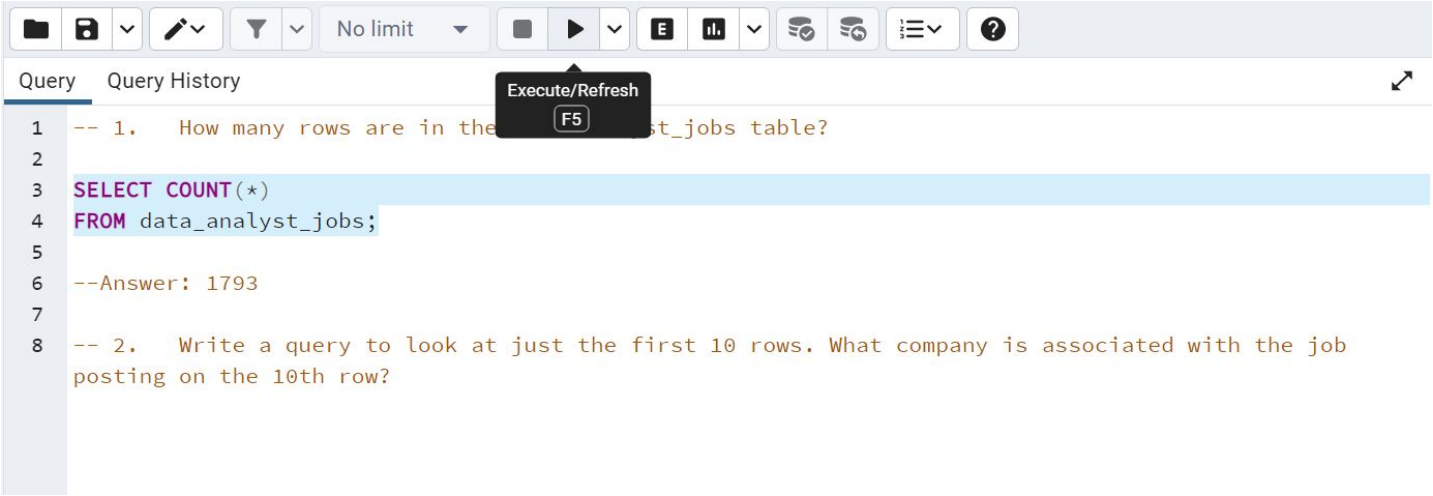
In order for your answers to be clearly labeled and easy to read through, put the question first as a comment, then the answer query, followed by the answer as a comment



NASHVILLE SOFTWARE SCHOOL

Formatting Questions and Answers

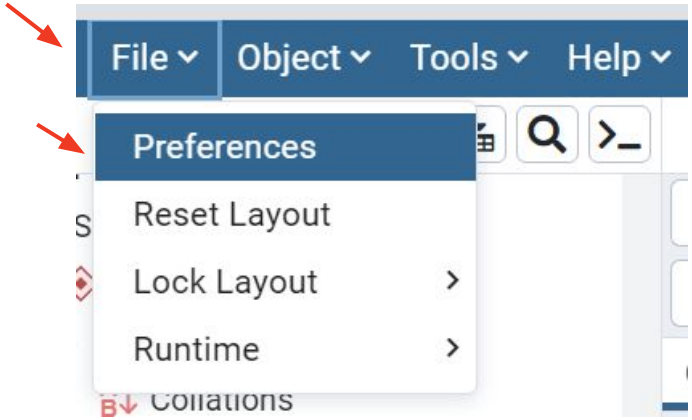
When you have multiple queries in a script, it is good practice to highlight the query you are working on before running it to ensure queries before it won't be run first.



NASHVILLE SOFTWARE SCHOOL

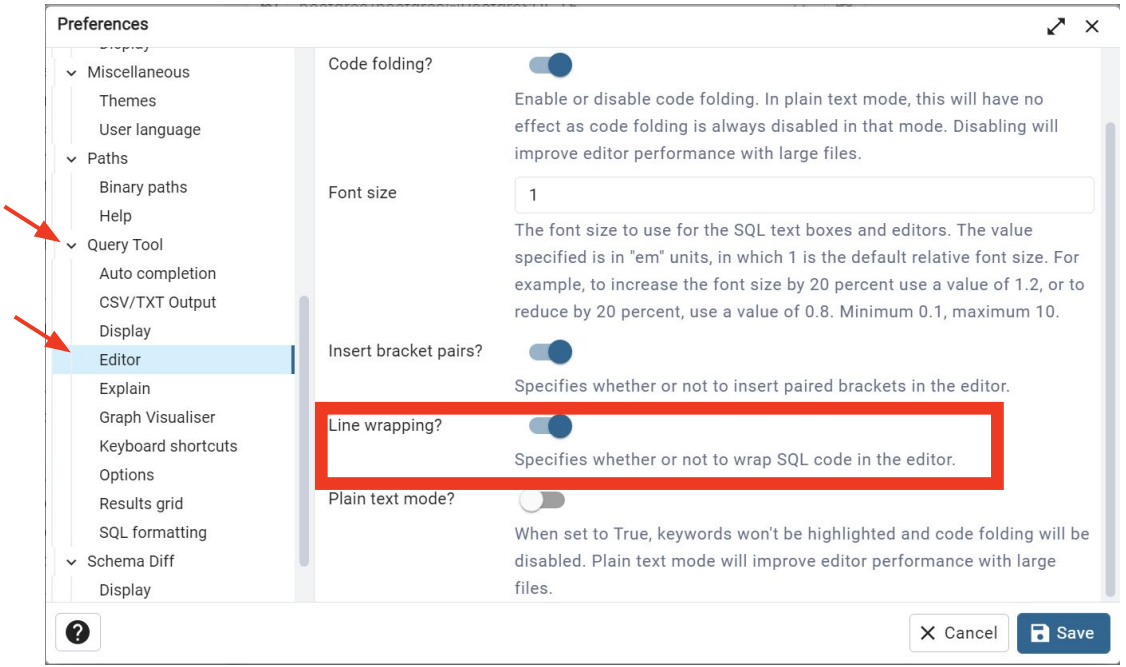
Text Wrapping

- Automatically making your comments and queries wrap text to the next line instead of creating a scroll bar to have to scroll through will make working in pgAdmin a lot easier.
- To turn this setting on, go to **File → Preferences**



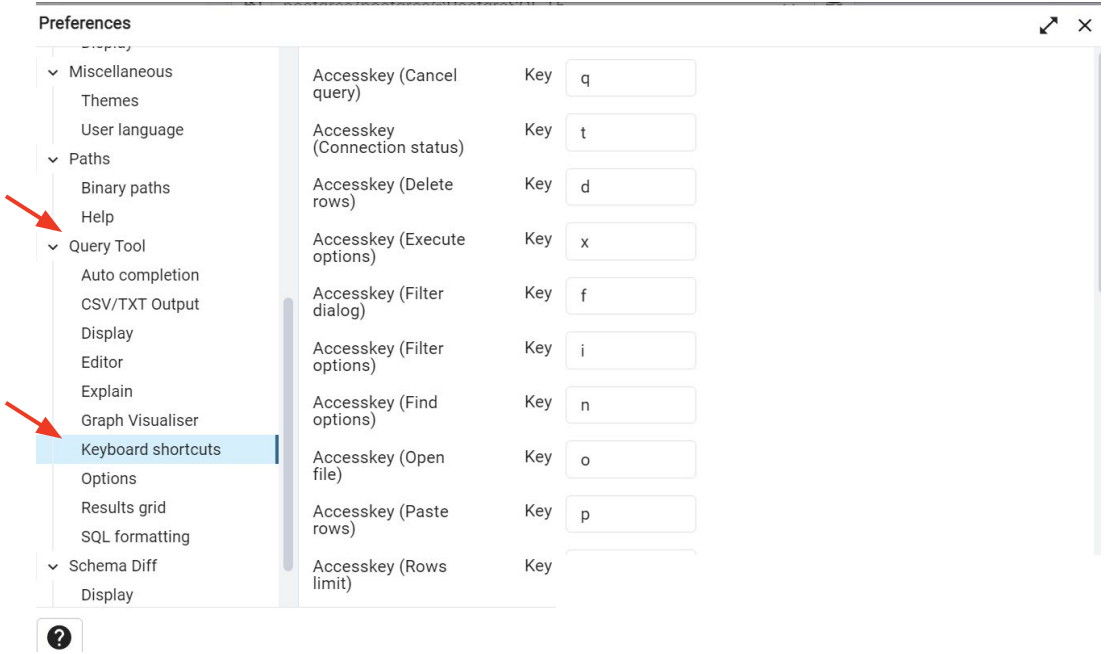
Text Wrapping

- Then select **Editor** under **Query Tool** on the left hand menu and turn on the **Line wrapping?** option.



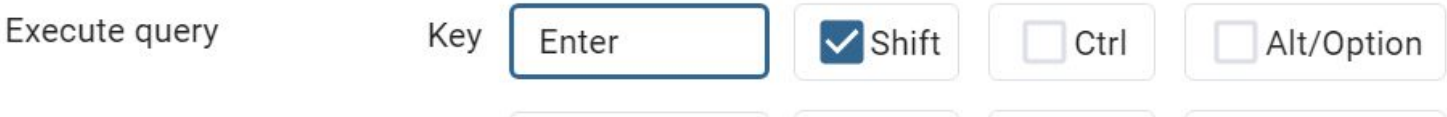
Shortcuts

From the **File** → **Preferences** menu, select **Keyboard shortcuts** on the left hand menu under **Query Tool**



Shortcuts

- From this page, you can see the shortcut keys available and change any to make use easier for you.
- One example that many students choose to change is the **Execute query** shortcut. **F5** is located in an inconvenient spot on the keyboard for frequent use. Changing this shortcut to **Shift + Enter** will make SQL query execution more in line with Jupyter Notebook execution when we get to Python.



Shortcuts

Another helpful shortcut that is not listed on the shortcuts page is **ctrl** + / which allows you to comment out everything in a selection at once.

```
2.  Write a query to look at just the first 10 rows. What company is associated with the job posting on the 10th row?
```

```
-- 2.   Write a query to look at just the first 10 rows. What company is associated with the job posting on the 10th row?
```