CompSci 516 Database Systems
SQL Installation

Spring 2022

Slides prepared by Yuchao Tao

- Install postgres + Load data + practice SQL queries

  Link to the MovieLens data:
  https://www2.cs.duke.edu/courses/spring22/compsci516/DataForClass/

Install Postgres

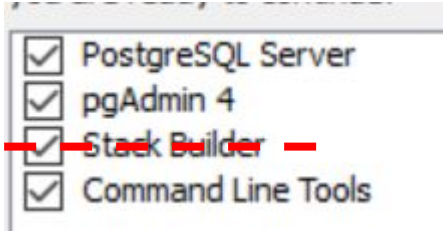
- If you are using Mac or Windows, and would like to use the GUI by pgadmin, follow slide #4-14
- If you are using Linux, or if you want to use Virtual Machine (VM) on Mac or Windows, follow slide #15-21

Postgres GUI Installation and Data Import

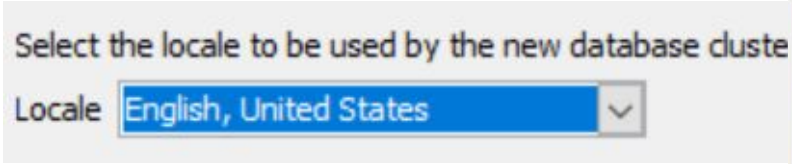
Step 1: Follow the link to download PostgreSQL. Choose ~~10.10~~ **the most updated version** for you platform.

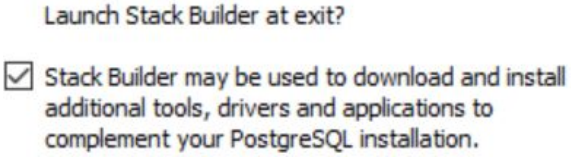
https://www.enterprisedb.com/downloads/postgres-postgresql-downloads

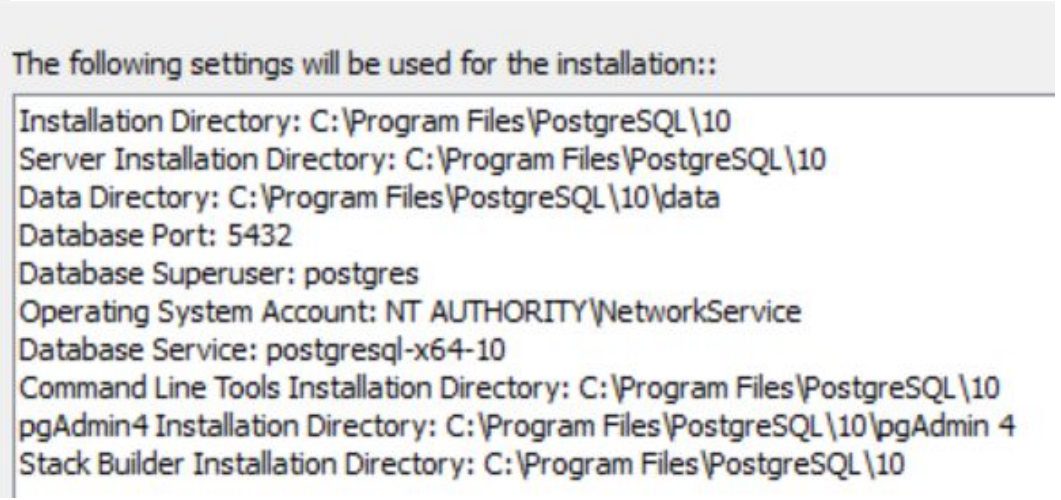
PostgreSQL Version	Windows x86-64	Mac OS X	Linux x86-32	Linux x86-64	Windows x86-32
11.5	Download	Download	N/A	N/A	N/A
10.10	Download	Download	Download	Download	Download
9.6.15	Download	Download	Download	Download	Download
9.5.19	Download	Download	Download	Download	Download
9.4.24	Download	Download	Download	Download	Download
9.3.25 (Not Supported)	Download	Download	Download	Download	Download

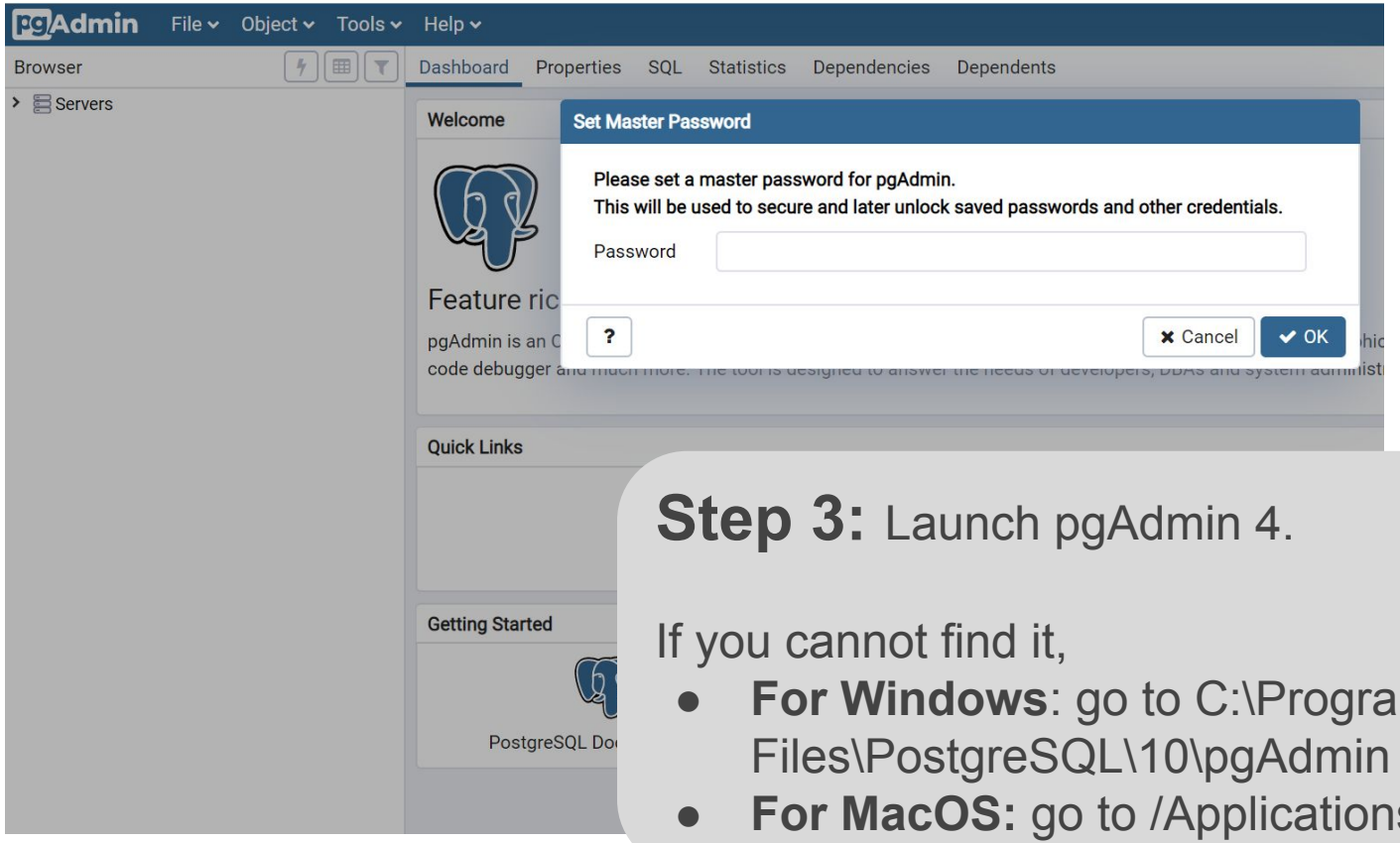


Step 2: Install components. You need a server, a client pgAdmin4, and CLI tools. You can opt out Stack Builder. Keep your postgres password.





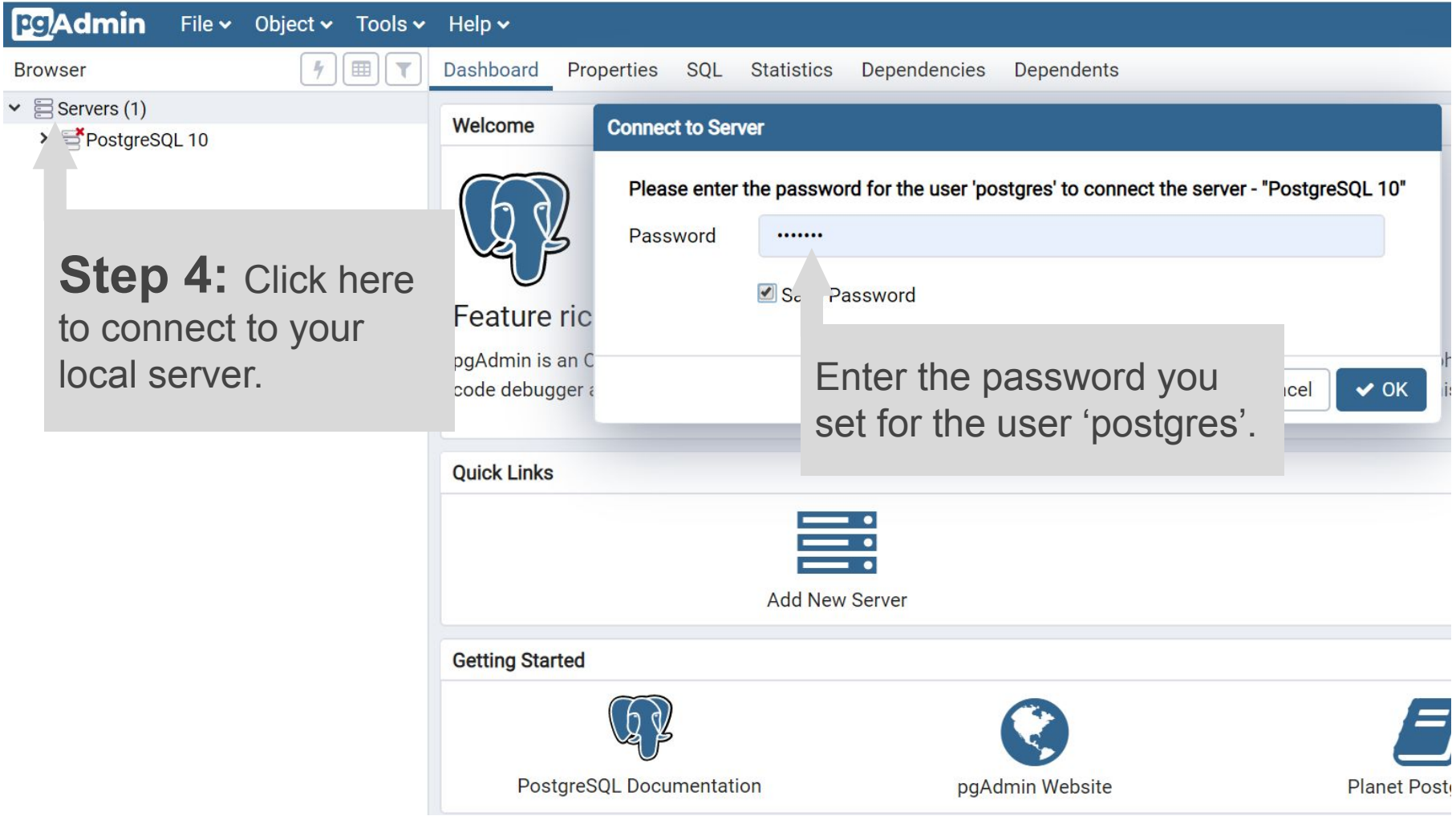


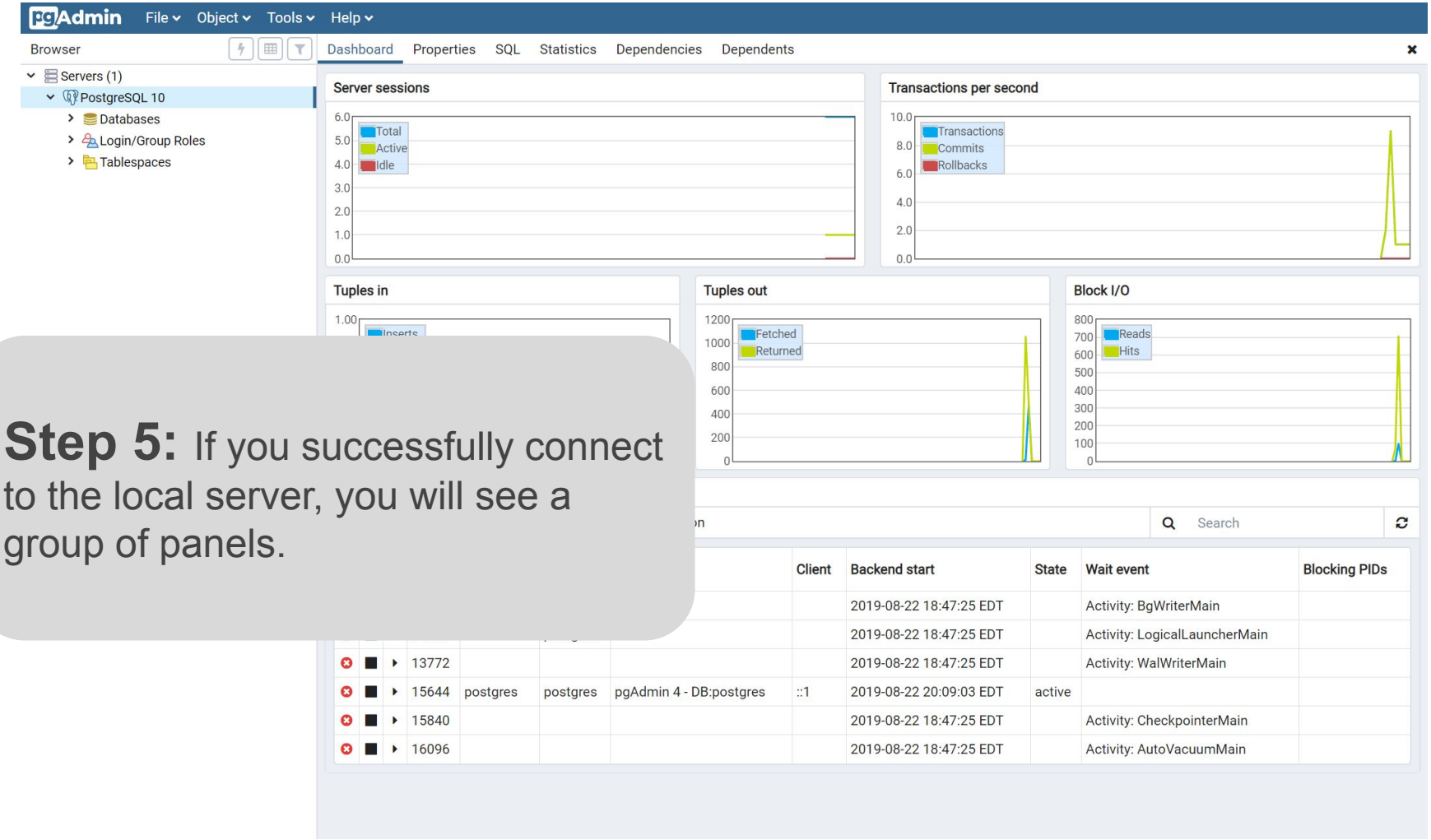


Step 3: Launch pgAdmin 4.

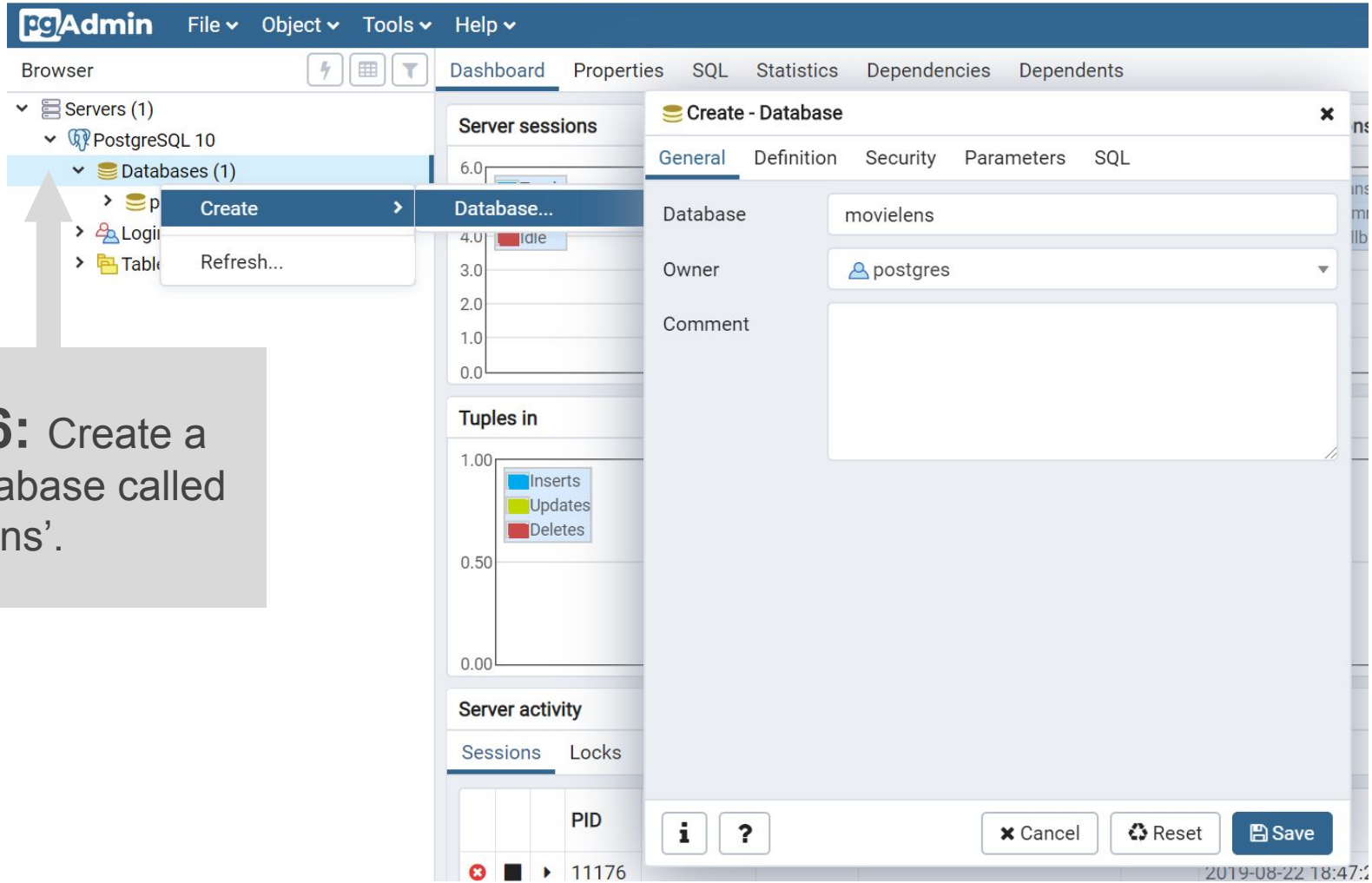
If you cannot find it,

- **For Windows**: go to C:\Program Files\PostgreSQL\10\pgAdmin 4\bin
- **For MacOS:** go to /Applications/PostgreSQL 10

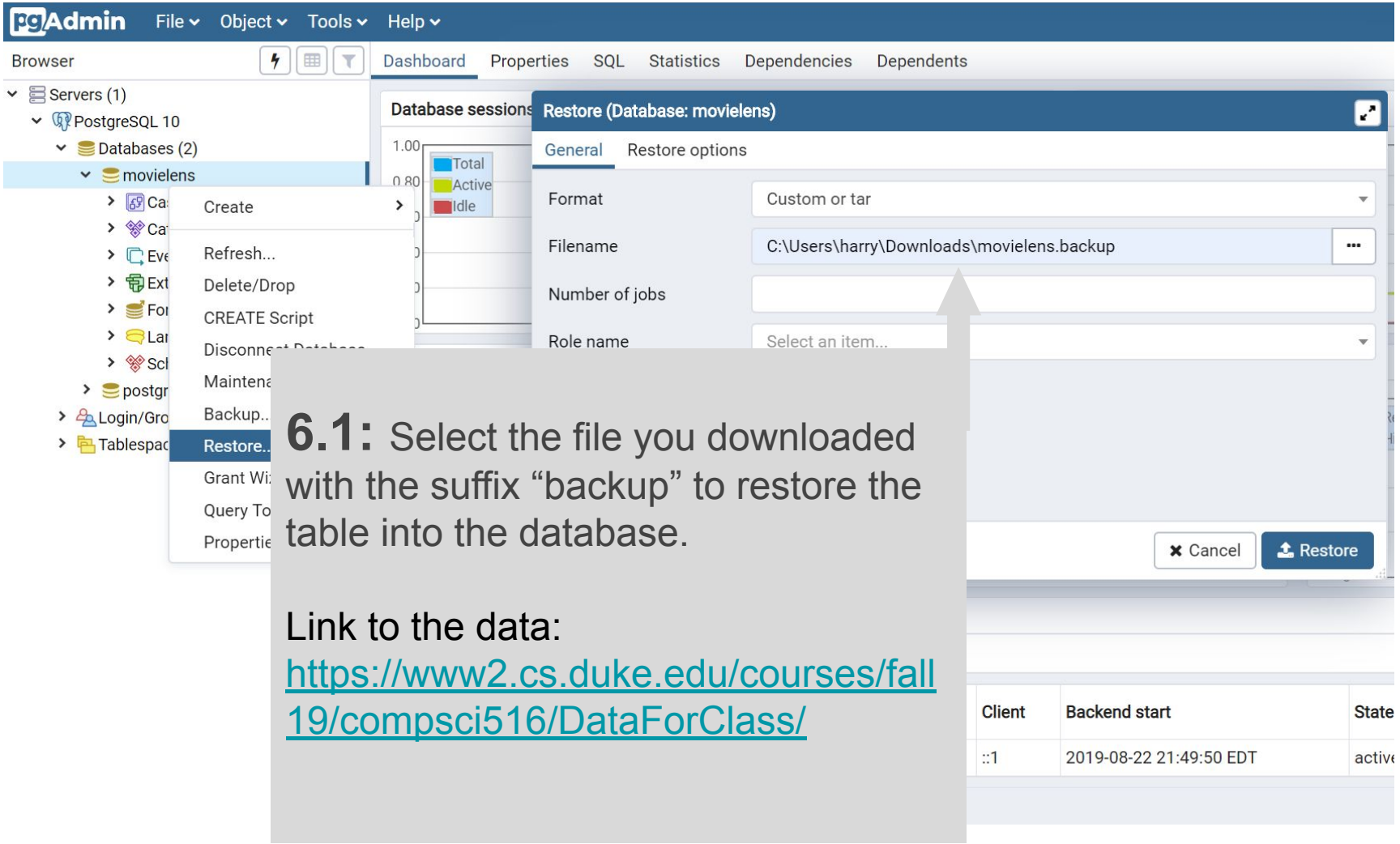

pgAdmin
File
Object
Tools
Help
Browser
Dashboard
Properties
SQL
Statistics
Dependencies
Dependents
Servers (1)
PostgreSQL 10
Step 4: Click here to connect to your local server.
Welcome
Connect to Server
Please enter the password for the user 'postgres' to connect the server - "PostgreSQL 10"
Password
Password
OK
Enter the password you set for the user ‘postgres’.
Quick Links
Add New Server
Getting Started
PostgreSQL Documentation
pgAdmin Website



Step 5: If you successfully connect to the local server, you will see a group of panels.

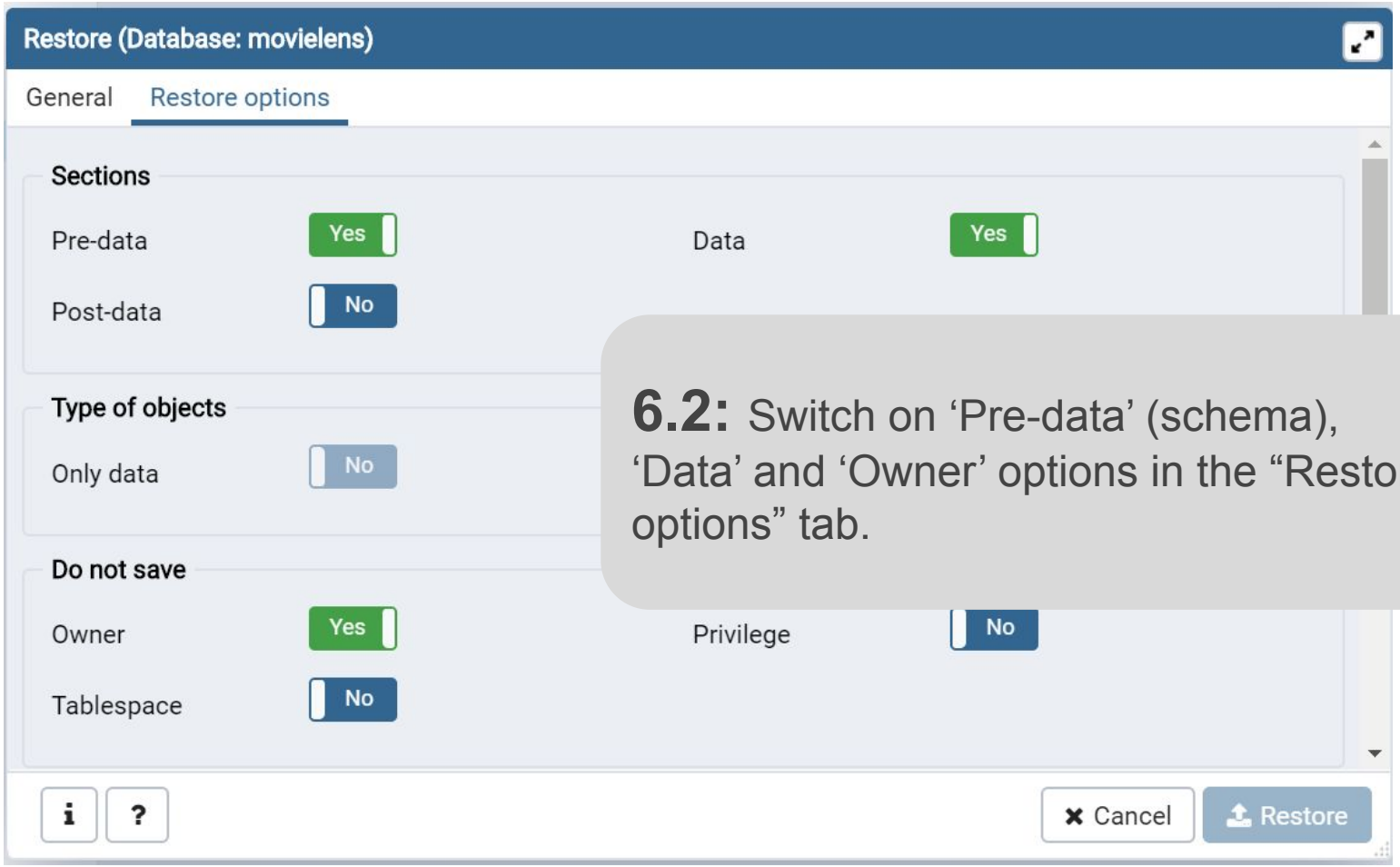


Step 6: Create a new database called ‘movielens’.

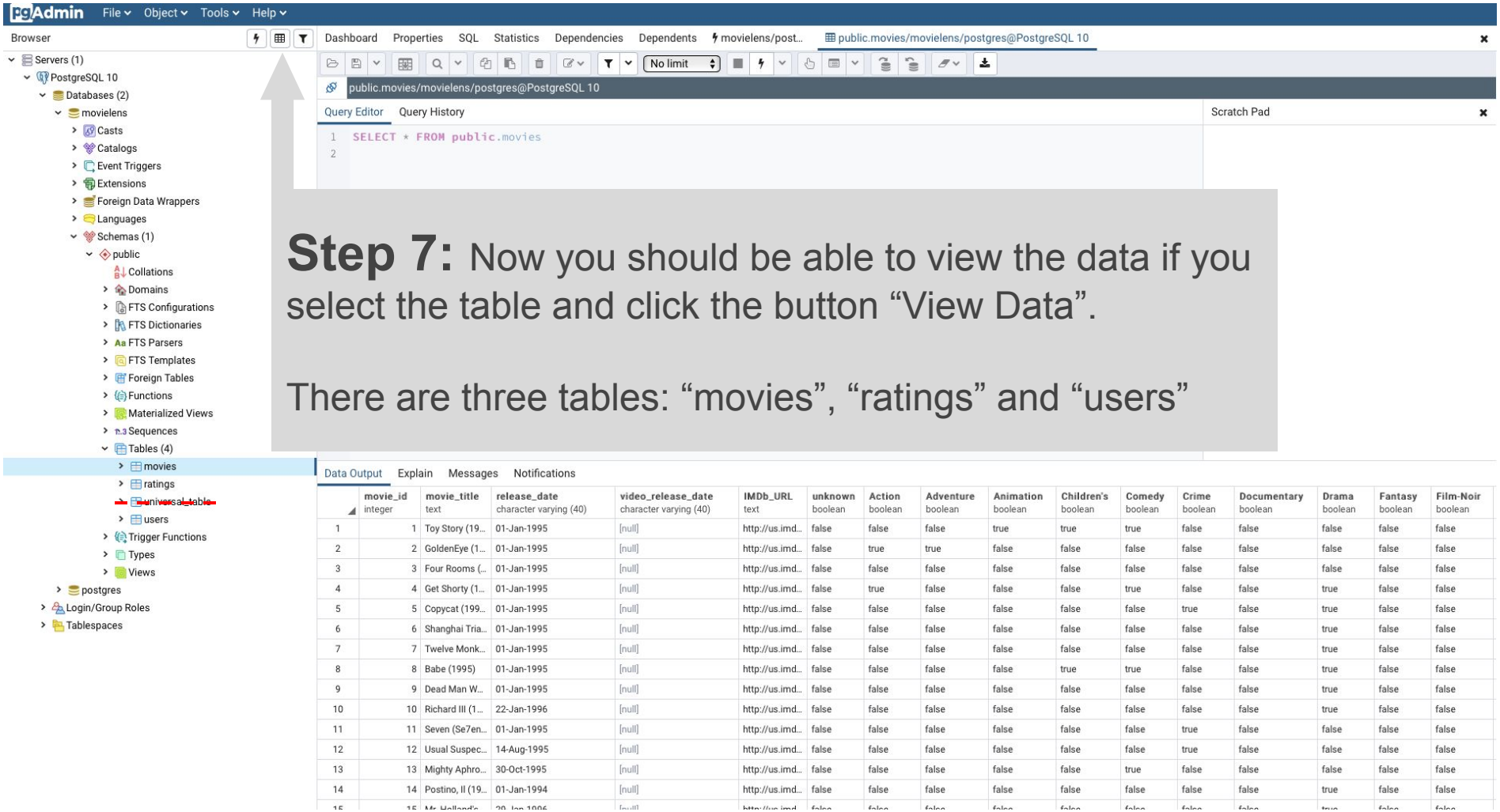


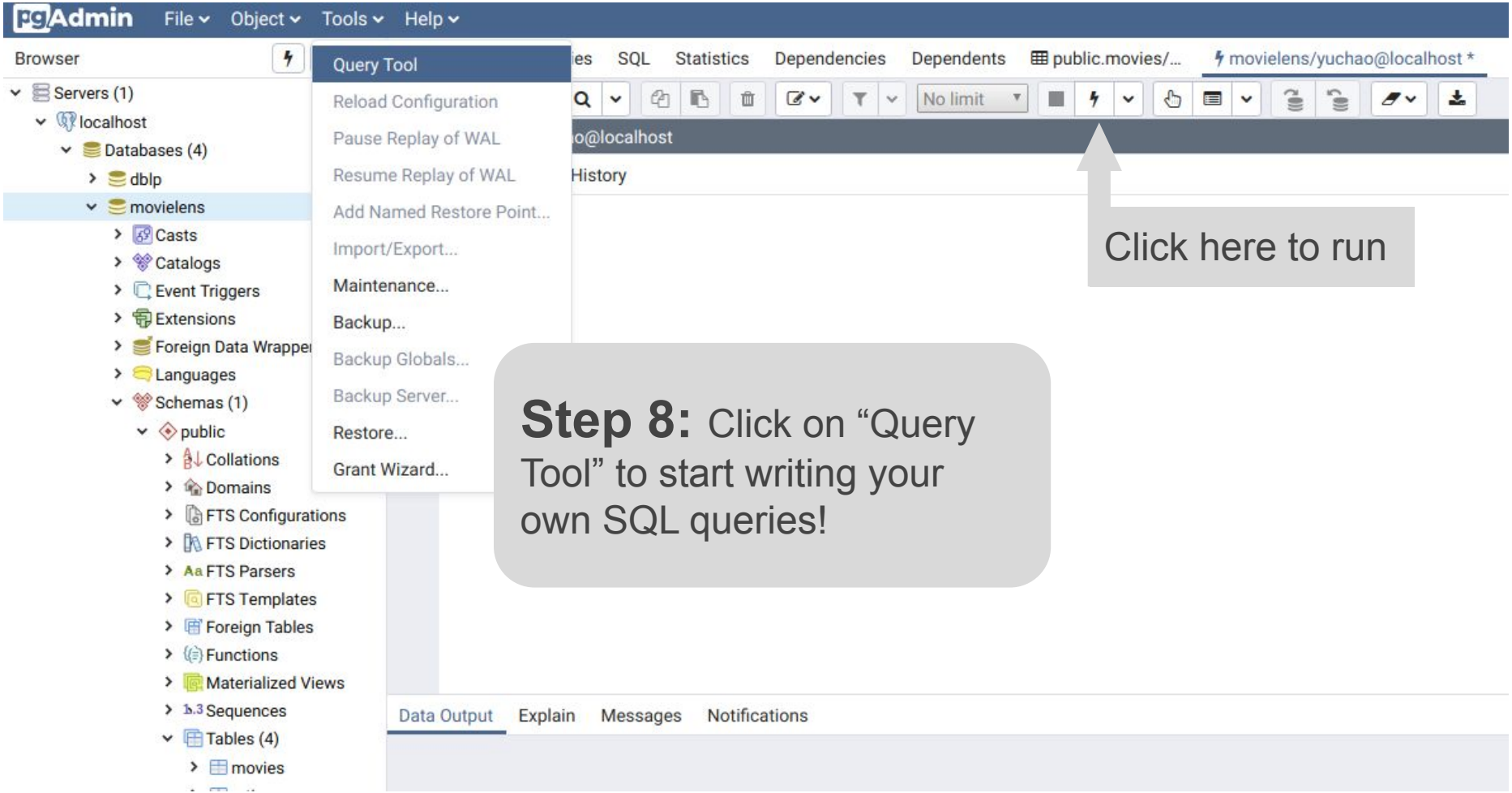
6.1: Select the file you downloaded with the suffix “backup” to restore the table into the database.

Link to the data: https://www2.cs.duke.edu/courses/fall19/compsci516/DataForClass/



6.2: Switch on ‘Pre-data’ (schema), ‘Data’ and ‘Owner’ options in the “Restore options” tab.

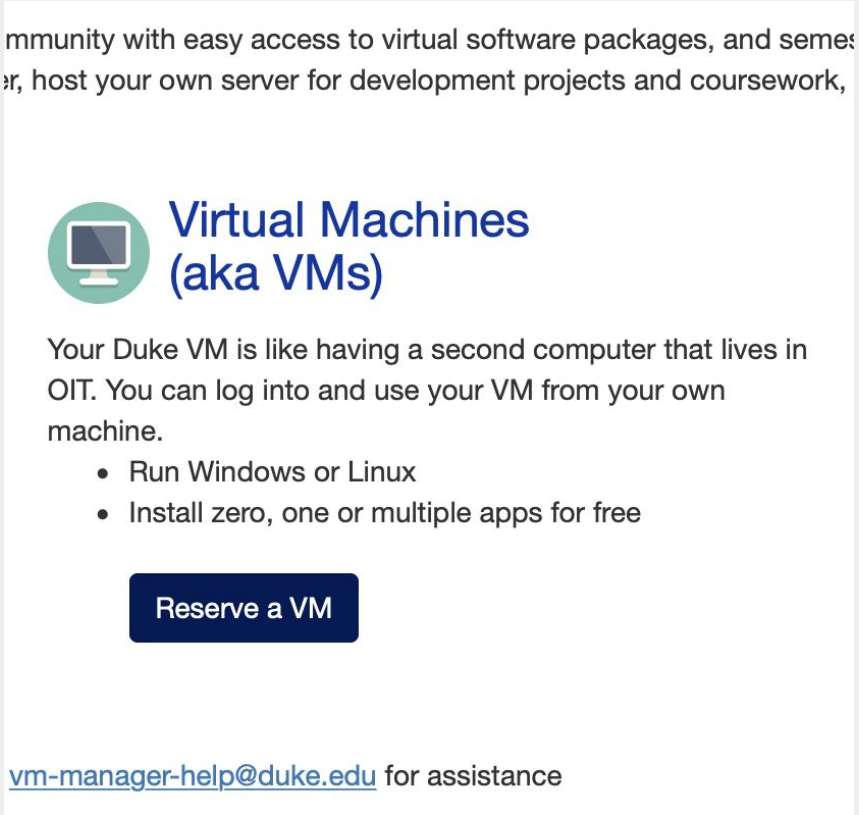

pgAdmin
File
Object
Tools
Help
Browser
Dashboard
Properties
SQL
Statistics
Dependencies
Dependents
movielens/post...
public.movies/movielens/postgres@PostgreSQL 10
No limit
Servers (1)
PostgreSQL 10
Databases (2)
movielens
Casts
Catalogs
Event Triggers
Extensions
Foreign Data Wrappers
Languages
Schemas (1)
public
Collations
Domains
FTS Configurations
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized Views
Sequences
Tables (4)
movies
ratings
universal_table
users
Trigger Functions
Types
Views
postgres
Login/Group Roles
Tablespaces
Query Editor
Query History
Scratch Pad
1 SELECT * FROM public.movies
2
Step 7: Now you should be able to view the data if you select the table and click the button “View Data”.
There are three tables: “movies”, “ratings” and “users”
Data Output
Explain
Messages
Notifications
movie_id integer
movie_title text
release_date character varying (40)
video_release_date character varying (40)
IMDb_URL text
unknown boolean
Action boolean
Adventure boolean
Animation boolean
Children's boolean
Comedy boolean
Crime boolean
Documentary boolean
Drama boolean
Fantasy boolean
Film-Noir boolean
1 1 Toy Story (19... 01-Jan-1995 [null] http://us.imd... false false false true true true false false false false false
2 2 GoldenEye (1... 01-Jan-1995 [null] http://us.imd... false true true false false false false false false false false
3 3 Four Rooms (... 01-Jan-1995 [null] http://us.imd... false false false false false false false false false false false
4 4 Get Shorty (1... 01-Jan-1995 [null] http://us.imd... false true false false false true false false true false false
5 5 Copycat (199... 01-Jan-1995 [null] http://us.imd... false false false false false false true false true false false
6 6 Shanghai Tria... 01-Jan-1995 [null] http://us.imd... false false false false false false false false true false false
7 7 Twelve Monk... 01-Jan-1995 [null] http://us.imd... false false false false false false false false true false false
8 8 Babe (1995) 01-Jan-1995 [null] http://us.imd... false false false false true true false false true false false
9 9 Dead Man W... 01-Jan-1995 [null] http://us.imd... false false false false false false false false true false false
10 10 Richard III (1... 22-Jan-1996 [null] http://us.imd... false false false false false false false false true false false
11 11 Seven (Se7en... 01-Jan-1995 [null] http://us.imd... false false false false false false true false false false false
12 12 Usual Suspec... 14-Aug-1995 [null] http://us.imd... false false false false false false true false false false false
13 13 Mighty Aphro... 30-Oct-1995 [null] http://us.imd... false false false false false true false false false false false
14 14 Postino, Il (19... 01-Jan-1994 [null] http://us.imd... false false false false false false false false true false false


pgAdmin
File
Object
Tools
Help
Browser
Query Tool
Reload Configuration
Pause Replay of WAL
Resume Replay of WAL
Add Named Restore Point...
Import/Export...
Maintenance...
Backup...
Backup Globals...
Backup Server...
Restore...
Grant Wizard...
SQL
Statistics
Dependencies
Dependents
public.movies/...
movielens/yuchao@localhost *
No limit
History
Servers (1)
localhost
Databases (4)
dblp
movielens
Casts
Catalogs
Event Triggers
Extensions
Languages
Schemas (1)
public
Collations
Domains
FTS Configurations
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized Views
Sequences
Tables (4)
movies
Data Output
Explain
Messages
Notifications
Click here to run
Step 8: Click on “Query Tool” to start writing your own SQL queries!

Postgres Installation and data import on VM

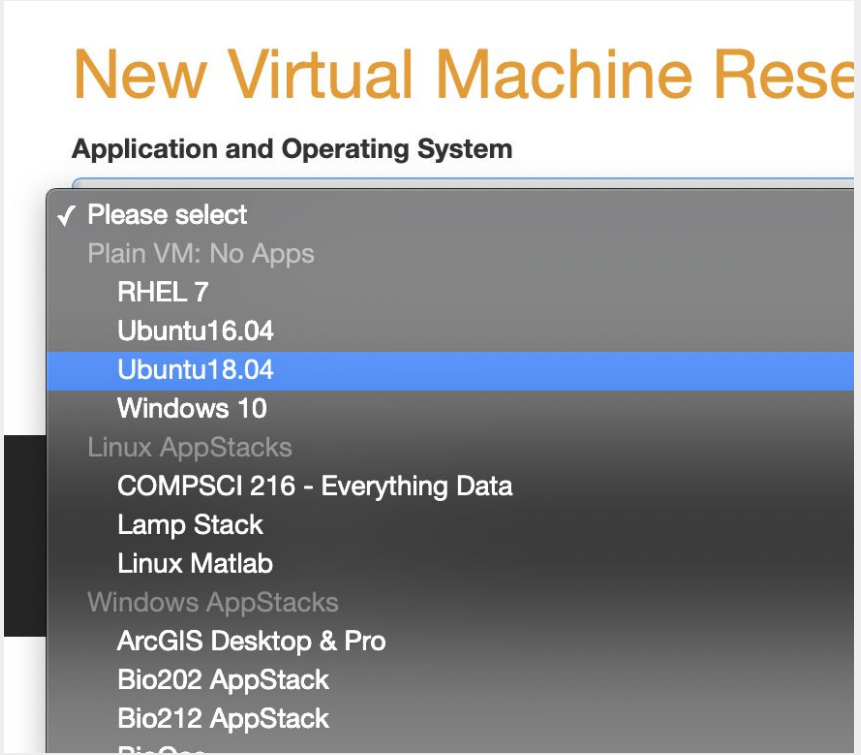
Reserve a virtual machine

1. Go to https://vcm.duke.edu
2. Click “Reserve a VM”
3. Login with netid and password



Reserve a virtual machine (Cont'd)

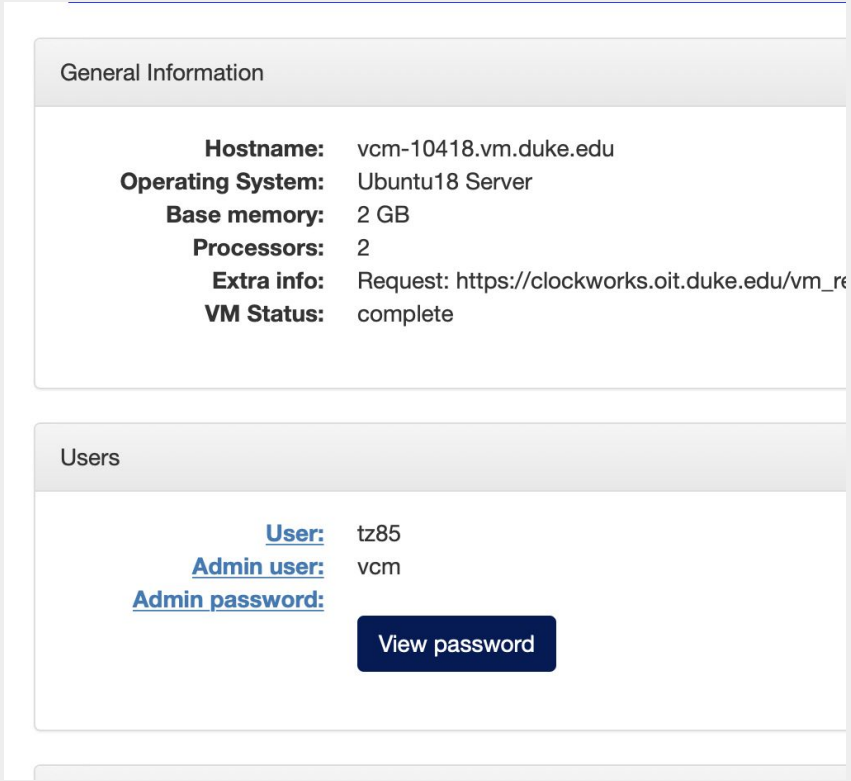
1. Click the drop down and select "Ubuntu 18.04" from the list
2. Read (or not) the agreement, agree and continue



Access your virtual machine

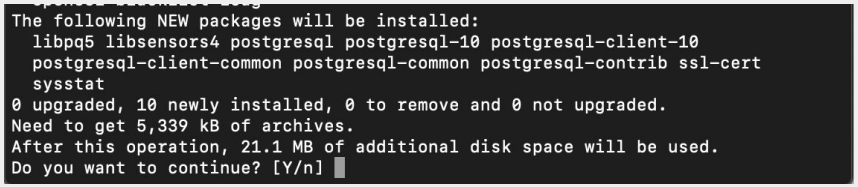
SSH into your new virtual machine using the hostname and your netid

$ ssh <netid>@<hostname>



Installing Postgres on Ubuntu

1. Update your machine
   - `sudo apt update`
2. Install latest version of psql
   - `sudo apt install postgresql postgresql-contrib`



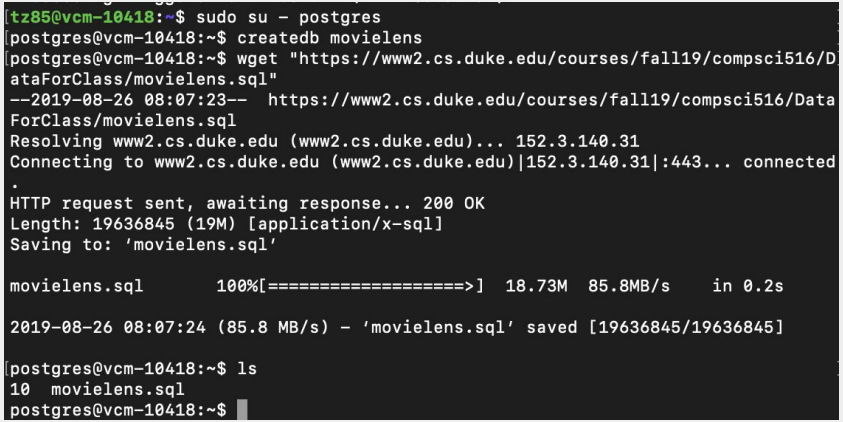
Create database and download dataset

1. Switch user to postgres

   `sudo su - postgres`
2. Create database “movielens”

   `createdb movielens`
3. Download dataset

   `wget https://www2.cs.duke.edu/courses/fall19/compsci516/DataForClass/movielens.sql`



Load dataset into the new database

1. Create tables from sql file

   `psql -U postgres movielens < movielens.sql`
2. Connect to database

   `psql movielens`
3. List all tables

   `\dt`

```
postgres@vcm-10418:~$ psql movielens
psql (10.10 (Ubuntu 10.10-0ubuntu0.18.04.1))
Type "help" for help.

movielens=# \dt
              List of relations
 Schema |      Name       | Type  |  Owner
--------+-----------------+-------+----------
 public | movies          | table | postgres
 public | ratings         | table | postgres
 public | users           | table | postgres
```

Try queries similar to lecture slides on the MovieLens dataset

Schema

movie

- movie_id
- movie_title
- release_date
- Action
- Adventure
- Animation
- ….

ratings

- movie_id
- user_id
- rating
- timestamp

users

- user_id
- age
- gender
- occupation
- zipcode

Sample query:

```
SELECT AVG(ratings)
FROM movie M, ratings R
WHERE M.movie_id = R.movie_id
```

Note: for the genre columns in movies, like Action, Adventure, etc., use “..” in queries, e.g.,:

```
SELECT “Action”
FROM movies
```

Postgres converts all unquoted identifiers to lower case:
https://stackoverflow.com/questions/37910287/sql-hint-to-reference-a-column