Loading Data into PostgreSQL

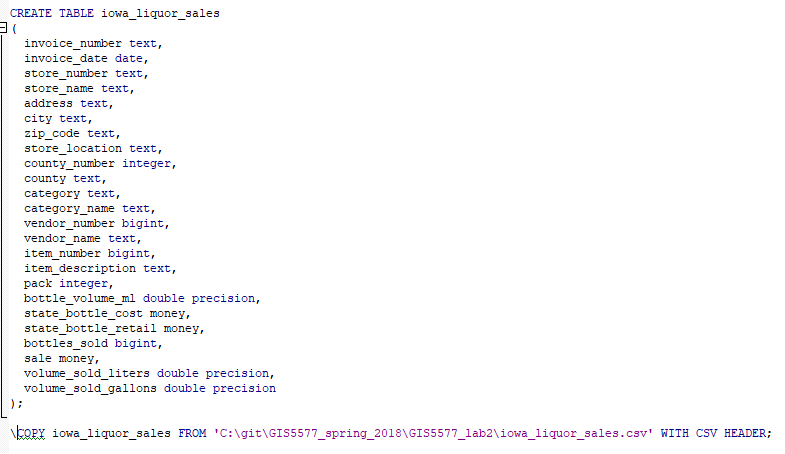
Loading data into PostgreSQL can be a difficult process, but it is extremely important that you learn how to do it. You have already learned how to learn data using the INSERT command with tuples

INSERT INTO <tableName>

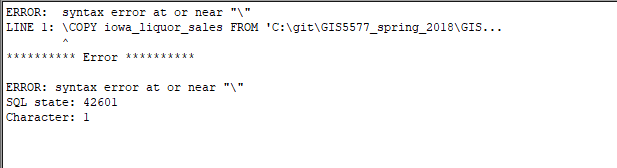
(field1, field2, field3) VALUES (item1, ‘item2’, item3), (item1, ‘item2’, item3), (item1, ‘item2’, item3)…;

However, this is not the most efficient way to load data into the system. The \COPY command is actually the command you wan to use when loading data.

The approach you should use would be to create the table and use the “\COPY” command



Unfortunately this creates an error. This is a pgAdmin Error. It occurs in both pgAdmin3 and pgAdmin4

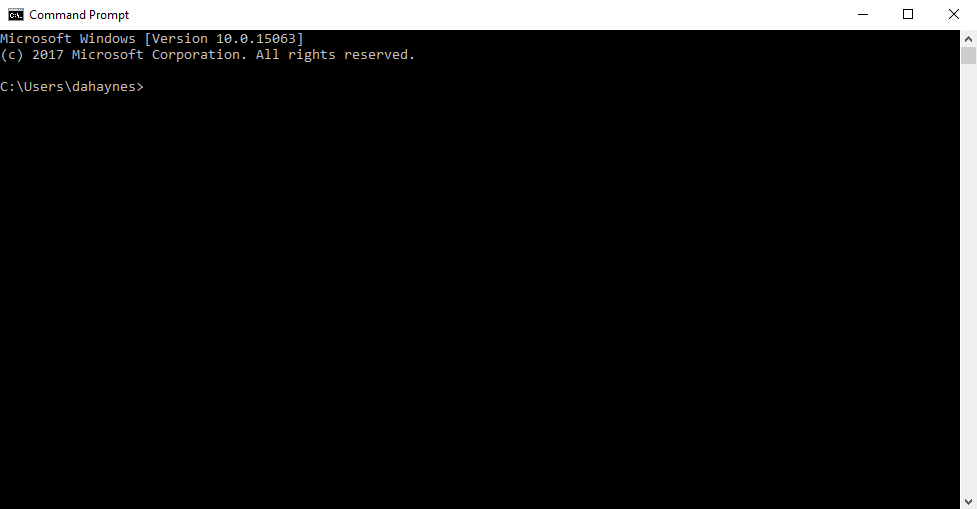


Let’s simply the error to idea that pgAdmin doesn’t care for “\”. A solution might be to drop the “\” and use “COPY” but your accounts (non-superusers) don’t have access to that command.

# Copying Data

What we need to do is connect to the database and issue the “\COPY” command without getting a syntax error. To do that you need to connect to the database. Using the terminal.

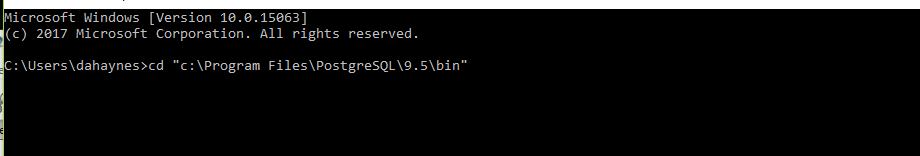
1. First open a windows prompt, by going to the start menu and typing "cmd"



1. Next open find the psql.exe file. For windows users it is usally located in Program Files

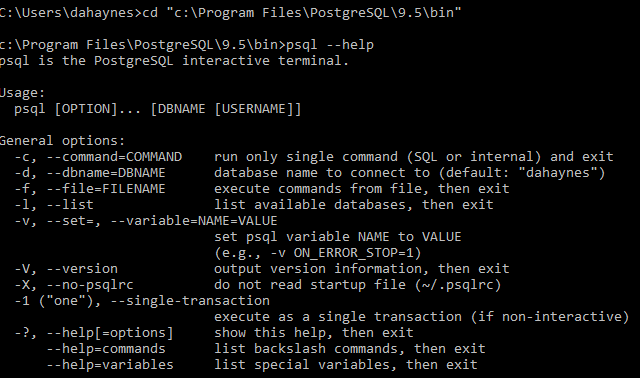
“C:\Program Files\PostgreSQL\9.6\bin” --The version <9.6> depends upon what you installed

In the command prompt change the directory by typing "cd". "cd" is a command that means change directory. After telling the computer to change the directory you need to tell it the directory that is should be looking at. In our case it is "C:\Program Files\PostgreSQL\9.6\bin"



c:\>cd C:\Program Files\PostgreSQL\9.6\bin

1. You should be able to type "psql --help"



Psql is postgresql’s actual terminal for communicating with the database. Throughout the course I will show you some of the options contained within the terminal. You can use this as a way of connecting to a database. Similar to connecting with pgAdmin4.

1. To connect to the server you will need to specify the following
   1. host address
   2. username
   3. database

PSQL uses “flags” to identify these. “-h” is for host “-U” is for username and “-d” is for database

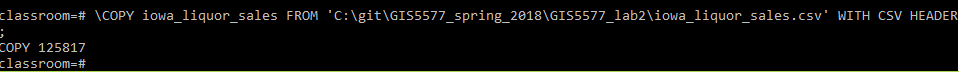


1. Now you can issue your “\COPY” command.

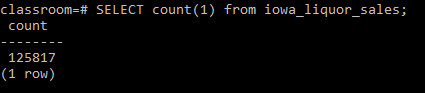
\COPY iowa\_liquor\_sales FROM 'C:\git\GIS5577\_spring\_2018\GIS5577\_lab2\iowa\_liquor\_sales.csv' WITH CSV HEADER;



1. It is important to have the “;” so psql knows when the command ends
2. Hit the enter key to run the command. It takes about 4 minutes for the data to load



1. When the psql is finished loading the data prompt <database># will be available for you to type
2. You can run other SQL commands here to verify the data is loaded.



1. Type “\q” to quit psql. You should see the new table in pgAdmin4 and be able to finish you assignment.

# One-step Connect&Copy

There is a way to combine steps 4 & 5 into a single step. You do this by issuing a 4th flag. I don’t recommend that you try this until you have successfully connected to the database and issued the copy command.

1. To connect to the server you will need to specify the following
   1. “-h” host address
   2. “-U” username
   3. “-d” database
   4. “-c” command

>psql -h 149.165.170.114 -U <username> -d <database> -c “\COPY ……”

1. To load the data simply add the "-c" command for copying the data into the table. The command for copying the data must be enclosed in double quotes “”

"\COPY <tablename> FROM 'csvPath' WITH CSV HEADER;

-- The "\" is important as are the double quotes surrounding the statement “”, single quotes are used inside for specifying the csvpath. csvPath must be the complete file path including the file extension.



C:\Program Files\PostgreSQL\9.6\bin>psql -h 149.165.170.114 -U david -d david -c "COPY iowa\_liquor\_sales FROM 'C:\work\GIS5577\GIS5577\_lab2\iowa\_liquor\_sales.csv' WITH CSV HEADER;"