**Method to restore data on localhost Postgresql:**

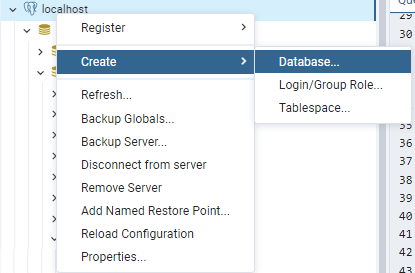
1. **By terminal (psql)**

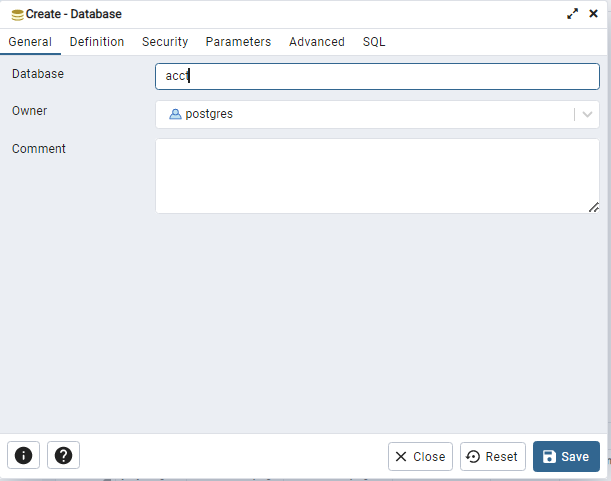
Once entered into psql using the command line/terminal you can run the followinf to commands to restore the database ( downloaded file )

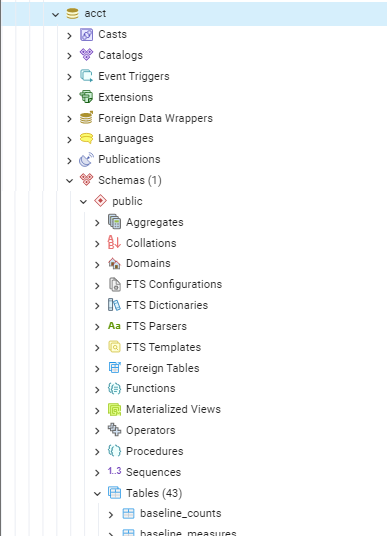
1. Createdb acct;
2. pg\_restore -e -v -O -x -d aact --no-owner /path/to/data/postgres\_data.dmp
3. **By PgAdmin:**

Follow the steps in order of screenshots attached after opening localhost on PgAdmin

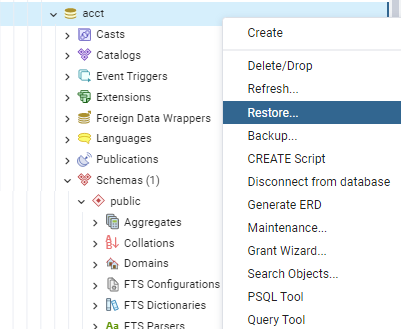
1. Right Click on localhost and create custom database

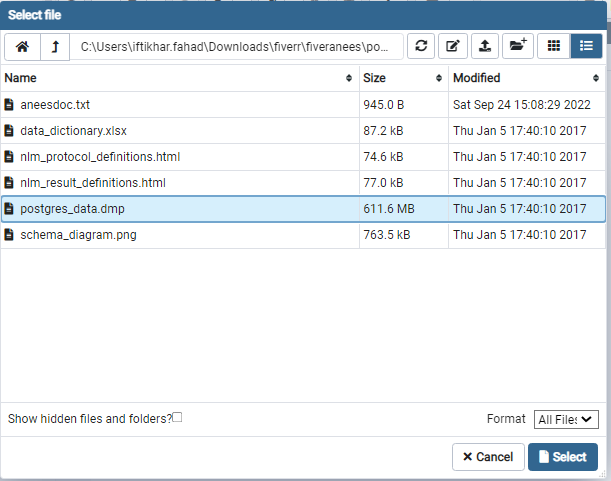


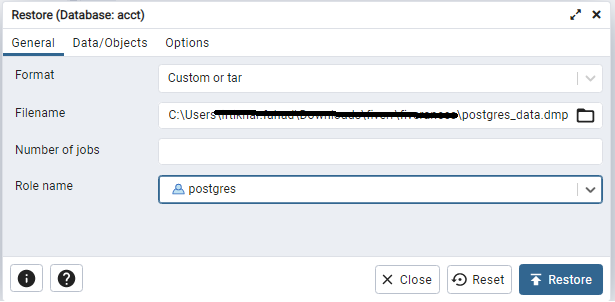




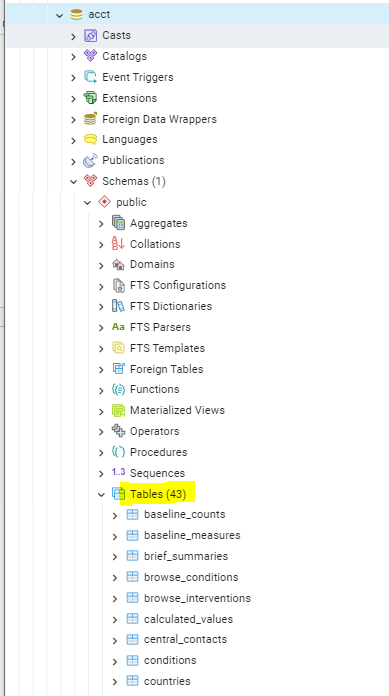
1. Restore your backup file. First right click on the database acct click on Restore. E







Once done the table section of your restored database will look like this instead of empty. After this we can start querying the data using query builder.



**Queries:**

**1.**

create view trial\_details as

select s.nct\_id,c.name,e.criteria,f.city || ', ' || f.country as location,i.intervention\_type,count(p.id) as total\_participants

from studies s join conditions c on s.nct\_id=c.nct\_id

join eligibilities e on s.nct\_id=e.nct\_id

join facilities f on s.nct\_id=f.nct\_id

join interventions i on s.nct\_id=i.nct\_id

join participant\_flows p on s.nct\_id=p.nct\_id

where s.overall\_status='Completed'

group by s.nct\_id,c.name,e.criteria,location,i.intervention\_type;

**Explanation:**

This query makes a view with name trial\_details while joining multiple tables to get desired colums we concat columns city and country of facilities table to give a meaningful location column.We use the participants\_flow table to get no of participants per trail in this approach. To access this view you will use another query:

Select \* from trial\_details;

**2.**

Create view trial\_effects as

select s.nct\_id,o.outcome\_type,r.adverse\_event\_term from studies s

join outcomes o on s.nct\_id=o.nct\_id

join reported\_events r on s.nct\_id=r.nct\_id

where r.adverse\_event\_term is not null;

**Explanation:**

This query makes a view with name trial\_effects while joining multiple tables to get desired . To access this view you will use another query:

Select \* from trial\_effects;

**3.**

select s.nct\_id,count(om.id) as patient\_count

from studies s join outcomes o on s.nct\_id=o.nct\_id

join outcome\_measures e on o.id=e.outcome\_id

join outcome\_measurements om on om.outcome\_measure\_id=o.id

where s.overall\_status='Completed'

group by s.nct\_id

order by patient\_count desc limit 1;

**Explanation:**

This query displays the trial with the most patient\_counts ( taken as count from outcome\_measurements table. Order by orders the first row to be that with largest patient\_count and limit 1 makes the result limited to that first row only giving us our required results.

**4.**

select count(nct\_id) from studies where

TO\_DATE(start\_month\_year,'Month YYYY dd')>='2006-01-01'

and TO\_DATE(completion\_month\_year,'Month YYYY dd')<='2009-12-31';

**Explanation:**

In this query we use columns start\_month\_year and completion month year and firstly convert them to date format using to\_date function. Then we apply necessary filter ( after 2005- before 2010 ) to get required results.

**5.**

1. Distribution For states for all countries available in dataset :

select count(s.nct\_id),f.city,f.country from studies s

join facilities f on s.nct\_id=f.nct\_id

group by f.city,f.country

order by f.country,f.city;

1. Distribution For states of United States of America :

select count(s.nct\_id),f.city,f.country from studies s

join facilities f on s.nct\_id=f.nct\_id

where f.country='United States'

group by f.city,f.country

order by f.country,f.city;