

PostgreSQL Hands On Labs – CSV Export

In this week's folder, you are given instructions for restoring the tb database into PostgreSQL. After we've loaded our tb database into our database server, we'll want to create a .CSV file to load into R (or another "downstream" program).

Specifically, we want to **export a .CSV file that contains total tuberculosis cases by country by year**. So we need to do some data transformation work in SQL to get our SELECT statement how we want it, then wrap the SELECT statement into a statement that exports the data to a .CSV file. Then we **load the .CSV file into Excel**.

In the figure on the right are the first few rows from the resultset that we want to get from the SELECT statement that we'll run against the tb table in our tb database. This will require two transformations on the single table.

Then you'll wrap your SELECT statement in code like that shown below. Here, we're using the default PC public folders; Mac users should substitute a Mac public folder. In either case, *the designated postgres user account* needs access to where the file will be written.

```
COPY (  
Your-SELECT-statement-here  
) TO '/Users/Public/tb.csv' With CSV Header;
```

	country character varying	year integer	cases bigint
1	Afghanistan	1995	
2	Afghanistan	1996	
3	Afghanistan	1997	128
4	Afghanistan	1998	1778
5	Afghanistan	1999	745
6	Afghanistan	2000	2666
7	Afghanistan	2001	4639
8	Afghanistan	2002	6509
9	Afghanistan	2003	6528
10	Afghanistan	2004	8245
11	Afghanistan	2005	9949
12	Afghanistan	2006	12469
13	Afghanistan	2007	13213
14	Afghanistan	2008	13136
15	Afghanistan	2009	12503
16	Afghanistan	2010	12947
17	Afghanistan	2011	15542
18	Afghanistan	2012	15774
19	Afghanistan	2013	
20	Algeria	1995	
21	Algeria	1996	
22	Algeria	1997	8902
23	Algeria	1998	
24	Algeria	1999	7088
25	Algeria	2000	8050

Possible Solution

Here is some code that may give you some insight into a solution path: select * from tb limit 20;

```
SELECT country, year, sex, child+adult+elderly AS cases
FROM tb;
```

```
DROP TABLE IF EXISTS tb_cases;
```

```
CREATE TABLE tb_cases AS
    SELECT country, year, sex, child+adult+elderly AS cases
FROM tb;
```

```
SELECT *
FROM tb_cases
WHERE country = 'Afghanistan'
ORDER BY country, year
LIMIT 50;
```

```
SELECT country, year, SUM(cases) AS cases
    FROM tb_cases
GROUP BY country, year
ORDER BY country, year
```

```
COPY (
SELECT country, year, SUM(cases) AS cases
    FROM tb_cases
GROUP BY country, year
ORDER BY country, year
) TO '/Users/Public/tb.csv' With CSV Header;
```