

## LAB 3

### Lab Overview

In this lab we will leave behind the social network and query a professional database with real world data! The queries are more complex and will utilize everything you learned in the previous lab. Take your time deconstructing what is asked for and build your CTE's one step at a time.

### Rules

The installation guide provides students with the proper software needed to complete this lab.

1. You must follow the rules of the [honor code](#). The labs must be done in groups of exactly two people. No larger groups are allowed, and if you have extraordinary extenuating circumstances that force you to do the labs alone, you must obtain permission to do so from the course leader.
2. Presenting P+ assignments are optional for a higher grade if the given tasks are completed and passed.
3. This is a PSQL lab. No other programming languages, either embedded in the database or external to it, are allowed.
4. A number of insert statements are expected to fail when you attempt to execute them.

## LAB 3

### Lab Presentation

Course related terminology is expected during the the lab presentation and make sure you have the following ready to be presented to the TA:

- Queries for the Mondial database and run them to show the results.
- Motivations for how the solution for each requested task is sensible. (Not all tasks have one answer, but they have to be good enough for e.g. a client to accept.)

### Before you start

As usual you should save everything you write so that we can grade you. PostgreSQL uses the “.psql” file format. You can use your preferred text editing software as long as it has color highlighting.

### Part 1: Querying someone else’s database

You have now made a complete database on your own. But working with databases is not only about setting up your own database. A computer engineer is expected to be able to gather necessary information by querying new or unknown databases too, it is in fact the more common type of everyday database interaction. Furthermore, often the information wanted is available only indirectly by combining and calculating different parts of the database. Here, we will query the open source mondial database to learn some cool facts about the world.

Setup the Mondial Database and complete the queries.

#### Setup:

- a. Download the files **mondial-schema.psql** and **mondial-inputs.psql** from Canvas. Do not get them from the official Mondial site as they may have been updated which might change the output.
- b. One file at a time, move the downloaded files from the Mondial DB to a folder in your home catalogue. Suggested methods for doing this remotely are available under “Moving files local → KTH-remote “ in the installation guide.
- c. Run the **psql -h psql-dd1368-ht24.sys.kth.se** command and connect to the database
- d. Drop everything (if you already have data there).
- e. Run the **\i <filename>.psql** command with your moved files and remember to create the schema before the inputs.

**Important note:** *You are forbidden from hardcoding anything in the database*

## LAB 3

*Tip: Here's the [ER diagram](#) for the Mondial database*

### Queries

1. Of the countries that have any land borders with other countries, show a list of those who **ONLY** have the lowest number of bordering land neighbors.

Expected outcome (order may be different):

name	num
Dominican Republic	1
Ireland	1
San Marino	1

2. Write a query for all the languages in the database, that states number of speakers and sorts them from most spoken to least spoken.

Expected outcome:

language	numberspeaker
Mandarin	1008748679
Hindi	456242015
Spanish	444683602
English	381519855

3. Which bordering countries have the greatest contrast in wealth? We define wealth as GDP.

*Tip: Remember to check the difference in wealth in both directions so that you don't end up with just half the results. E.g. Germany and Switzerland but also Switzerland and Germany.*

Expected outcome:

## LAB 3

country1	gdp1	country2	gdp2	ratio
BHT	2133	CN	9330000	4374
BR	2190000	GF	800	2738
CN	9330000	KG	7234	1290
E	1356000	GBZ	1106	1226
I	2068000	RSM	1866	1108
CN	9330000	TJ	8513	1096
CN	9330000	LAO	10100	924

### Part 2: Views

Views are useful tools when using databases. There are two main kinds of views: materialised and non-materialised (virtual).

Questions for the presentation:

1. When would a view be beneficial compared to regular queries? For what reasons?
2. What is stored in the database for non-materialised or materialised views respectively?
3. Name one benefit of using materialised views over non-materialised ones
4. Name one benefit of using non-materialised views over materialised ones

### P+

1. Present a table of the countries that can be reached from Sweden crossing a land border no more than 5 times, provided you never double back over a border you have already crossed. Your answer should show the country code and the name of the country along with the minimal number of crossings required to reach it.

**Hint:** Recursion.

Expected outcome (order may be different for countries with equal min):

code	name	min
SF	Finland	1
N	Norway	1
R	Russia	2
PL	Poland	3
UA	Ukraine	3
EST	Estonia	3

## LAB 3

- Consider a river system as a main river and all the tributary rivers that flow into it, and their respective tributary rivers and so on. Present a table that shows the names of the rivers of the longest branch(es)\* of each river system and the total length of each shown branch, as well as the number of rivers in each and the rank of the number of rivers in the branch. *Start from the rivers the Nile, Amazonas, Yangtze, Rhein, Donau and Mississippi and explore from there.* You are allowed to hard-code the names of those rivers.

\* Longest, in terms of the number of rivers in the displayed branch. For example, the longest branch of Yangtze has two rivers so include all branches of the Yangtze with 2 rivers.

Expected outcome:

rank	path	numrivers	totlength
1	Yangtze-Han	2	7912
1	Yangtze-Gan Jiang	2	7138
3	Mississippi-Missouri-Platte River-North Platte River	4	9501
3	Mississippi-Missouri-Platte River-South Platte River	4	9118
3	Amazonas-Rio Madeira-Beni-Rio Madre de Dios	4	8038

Note:

- Depending on the rank function you use, the rank number can be different. As long as the river connections are correct the solution is fine.
- The rank is on the number of branching rivers.