CS3543 – Hands-on 5

Fall, 2023, UNB, Fredericton

Due: December 7 at 23:00 pm

In this hands-on, you'll become familiar with OLAP queries in PostgreSQL.

Setup

1. Log into the lab Desktop with ITS credentials and then open a browser and go to the URL:
   https://remotelab.cs.unb.ca/
   Log in with your FCS credentials.

2. Launch the Data Science Lab VM as follows:
   from the *Applications* (top left corner) drop-down menu select *Applications* > *FCS VMs* > *Data Science Lab VM*
   VM login and password are both *bigdata*.

3. Start PostgreSQL database server by running the command on a Terminal: `pg_ctl start`

Task 1

The dataset for this hands-on is about climate change, specifically, annual CO2 emissions (tons) and the contribution of the countries. Create two tables and load data into them.

1. Get the data files:

```
wget    https://www.cs.unb.ca/~sray/teaching/dbsysadm/handson5/annual-co2-emissions-per-country.csv

wget https://www.cs.unb.ca/~sray/teaching/dbsysadm/handson5/country_info.csv
```

2. Start a psql SQL client by running the command: `psql testdb -U dbuser`

3. Create two tables:

```
CREATE TABLE co2_emissions (
  Country VARCHAR(50),
  Code CHAR(3),
  Year integer,
  Annual_CO2_emissions real,
  PRIMARY KEY (Code, Year)
);

CREATE TABLE country_info (
  Country VARCHAR(50),
  Capital VARCHAR(50),
  Code CHAR(3),
  Continent VARCHAR(50)
);
```

4. Load data from the data files into the two tables:

```
COPY co2_emissions(Country, Code, Year, Annual_CO2_emissions)
FROM '<FILEPATH>/annual-co2-emissions-per-country.csv'
DELIMITER ','
CSV HEADER;
```

```
COPY country_info(Country, Capital, Code, Continent)
FROM '<FILEPATH>/country_info.csv'
DELIMITER ';'
CSV HEADER;
```

`<FILEPATH>` is the folder where you downloaded the csv files

Task 2

Q1. Write an Online Analytical Processing (OLAP) query that can be used to create the following cross-tabulation report showing the total CO2 emissions (tons) for all countries during 2012 and 2022. Note that in the cross-tabulation, data is presented for the total CO2 emissions for each country for each year during 2012 and 2022, the sub-total emissions for each country, the sub-total emissions for each year during 2012 and 2022, and the grand total emissions. Data is sorted by country name and year.

Year **Country**	2012	2013	. . .	2022	*Sub-total* *(by country)*
Afghanistan					
Albania					
. . .					
Zimbabwe					
Sub-total *(by year)*					*Grand total*

Q2. Write an Online Analytical Processing (OLAP) query that can be used to produce a roll-up report showing the total CO2 emissions (tons) for all continents during 2012 and 2022. Note that in the roll-up, data is presented for the total CO2 emissions for each continent for each year during 2012 and 2022, the sub-total emissions for each continent and year combination during 2012 and 2022, the sub-total emissions for each continent, and the grand total emissions. Data is sorted by continent name and year.

Continent	**Year**	*Sub-total*
Africa	2012	
Africa	2013	
. . .		
Africa	2022	
Africa		
Americas	2012	
Americas	2013	
….		
Americas	2022	
Americas		
…	…	
…		
		Grand total

Submission instructions

1. Submit one pdf file with the following in a pdf file:

Task 1
No submission is necessary for Part 1.

Task 2
i) The OLAP query Q1 and a screenshot of your output.
ii) The OLAP query Q2 and a screenshot of your output.

2. Submit through Desire To Learn (D2L) course drop box Handson1. Mention the following on the top of your submitted file: your name and hands-on#

3. Hands-on not submitted electronically via D2L or submitted after the due date will NOT be marked.

Notes about plagiarism

Please note that the handsons are meant to be done individually. Any submission that appears to be in violation of an academic offence (plagiarism) may be reported to the Registrar's Office as per UNB regulations (See section VII of UNB Undergraduate Calendar).