

Important !!

This assignment will be evaluated by an **AUTOMATED** system. It is important that you follow the guidelines described in this section in preparing the answers. As a result, you either get complete points for a particular question or zero ! There will not be any partial points.

Additionally, if the guidelines are not followed, TAs will not manually edit your submissions to make it work with the evaluation tool, and you will receive zero points.

All assignment questions will have to work on the **POSTGRES** database installed in `winter2021-comp421.cs.mcgill.ca` otherwise it will not be graded.

Your queries should output **ONLY** those columns that are specified in the question in the requested format. Do not **RENAME** columns unless explicitly specified. If an **ORDERING** of rows is requested in the question, make sure that it is performed. All ordering is implied to be **ASCENDING** unless explicitly stated otherwise.

Make sure your queries are always terminated with a **SEMILCOLON** (;)

You may write comments in the SQL file, as long as it is standard SQL comments (i.e. starts with `--`).

Example 1

Q You need to find all the employees who live in Laval and their department. The output rows of the query will consist of first the column NAME from the employee table followed by ADDR_CITY from the department table. Order them based on ADDR_CITY and then on NAME.

Correct SQL 1

```
SELECT NAME, ADDR_CITY
....
ORDER BY 2, 1
;
```

Correct SQL 2

```
SELECT NAME, ADDR_CITY
....
ORDER BY ADDR_CITY, NAME
;
```

Correct SQL 3

```
SELECT E.NAME, E.ADDR_CITY
....
ORDER BY ADDR_CITY, NAME
;
```

In-Correct SQL 1

```
SELECT E.NAME EMPLOYEE_NAME, E.ADDR_CITY
....
ORDER BY ADDR_CITY, EMPLOYEE_NAME
;
```

Incorrect - because the employee name column is renamed.

In-Correct SQL 2

```
SELECT E.NAME, E.ADDR_CITY
....
ORDER BY NAME, ADDR_CITY
;
```

Incorrect - because the ordering should be first by ADDR_CITY and then NAME and not the other way around.

Example 2

Q You need to find the city in which the most number of employees live and the number of employees that live in that city. Your output rows should contain the column ADDR_CITY followed by the number of employees who live in that city, name this second column EMPLOYEE_COUNT . To handle the situation where you have more than one city with the most number of employees in it, make sure the output is ordered by ADDR_CITY.

Correct SQL 1

```
SELECT ADDR_CITY, COUNT(*) EMPLOYEE_COUNT
....
```

```
ORDER BY 1  
;
```

In-Correct SQL 1

```
SELECT ADDR_CITY, COUNT(*)  
.....  
ORDER BY 1  
;
```

Incorrect : COUNT(*) should have been renamed EMPLOYEE_COUNT

Setup Instructions

Download the package PG-XXXXXXXXX.tar.gz

These scripts are meant to be used in the winter2021-comp421.cs.mcgill.ca . You must copy (scp/filezilla/etc..) the package to that server under your home directory.

untar the contents and **RENAME** the folder to replace XXXXXXXXXX with your student id number. Say if your student id is 342567987, this can be accomplished by executing the command

```
$ mv PG-XXXXXXXXX PG-342567987
```

change to the directory

```
$ cd PG-342567987/data
```

and execute the script to create tables.

```
$ ./setup.sh
```

Enter your socs unix password if requested.

There is a script `delNdrop.sh` that is supplied in case you want to drop the tables that are created as part of the assignment.

You can execute it by running the command.

COMP 421/764 Assignment 1 - SQL formatting instructions

```
$ ./delNdrop.sh
```

Remember, you may have to add some records of your own into the tables to test different “corner cases” of some of your queries. You can also do this by adding it to the `setup.sh` for your convenience.

For each question, once you have verified your query is working, save the query into the corresponding sql file in the folder. For example, if you want to save the query for Q1, edit the file `pgsql/q01.sql` and put the contents inside of it. **Make sure to end your query with a semicolon.**

You can test your query output again by executing from your shell prompt. This is very important to ensure that when the automated system executes your query, it works correctly. If your submission SQL files are not executable by the below mechanism, you will get 0 points. It is your responsibility to check this before you turn in your assignment.

For example, to test the query for the first question, from the `pgsql` directory, do the following.

```
$ psql cs421 < q01.sql
```

Enter your SOCS password, if requested. You should see the output result from the query that you can verify.

What to Submit !

Please compress your entire folder and submit it in my courses. DO NOT submit the outputs . We don't need the data folder, but it is ok, if you cannot remove it. Following are some examples of how you can compress what we need. Make sure the parent directory structure with your student id in it is preserved. Following the above examples, you will submit a compressed version of `PG-342567987` (of course you will have your student id here instead).

In the below commands, please replace `342567987` with YOUR student id.

Correct

```
$ tar -zcvf PG-342567987.tar.gz PG-342567987/pgsql/
```

Submit the file `PG-342567987.tar.gz` to mycourses.

Correct

```
$ tar -zcvf PG-342567987.tar.gz PG-342567987/
```

COMP 421/764 Assignment 1 - SQL formatting instructions

This includes the data folder and other things that we do not need, but is still ok as it preserves the parent folder as well as includes the contents of the sql folder which contains your answer queries.

Incorrect

```
$ cd PG-342567987
$ tar -zcvf PG-342567987.tar.gz pgsql/
```

incorrect because now you lost the parent directory information which is PG-342567987 that is needed for us to know your student id.

If your submission failed to capture the directory information to include your student id properly as shown in the above examples, you will get 0 points.