Homework Due 2021-11-05 by 22:55 New York Time

Contents

1 General Instructions

1. You need to follow carefully the instructions for the assignment as written below.

   It is advisable to print out this document and check off various points as they are addressed. It is easy to miss something when switching between the assignment and the solution on a single screen, especially on a laptop with a relatively small screen.

   If you do not have access to a printer, at least review your solution before the submission to make sure that you did all that you were requested to do and only what you were requested to do.

2. If you want to refer to a specific line in this document, refer to the small numbers in the left margin.

3. If you have questions concerning this homework email Shubham Srivastava, `mailto:ss14687@nyu.edu`, *in the way specified in the course description.*

   To be sure that you get an answer to your question before the submission deadline, *do not delay your question to the date on which the assignment is due.*

   If you still have unresolved questions, email Zvi Kedem, `mailto:zk1@nyu.edu`, including all relevant correspondence with the assistant(s) listed above, *in the way specified in the course description.*

4. Submit your homework in an electronic form by uploading it to NYU Classes by the due date and time. Use only permitted software and format. E.g., if you are asked for a relational database specification using SQL Power Architect than that's what you must submit.

5. If you submit a scanned, handwritten assignment when permitted, it has to be written neatly, that is, it should be neatly divided into lines just as a typeset document, etc. You may submit a handwritten assignment only when that is explicitly allowed. And, unless stated otherwise, you must submit such a handwritten assignment as a file in PDF format only.

6. It is important that you follow the directions precisely. Also, please *check* that you submitted what you intended to submit, as you are responsible for making sure of that. The best way to

do is to download what you submitted to check that.

And the best way to manage your work is to dedicate a folder/directory to each assignment.

7. Until the deadline that the system imposes you can resubmit your homework as many times as you like and you may want to submit it relatively frequently in case something happens to your partial work on your machine.

8. Do not email your submission to any of the assistants. If you did not submit your solution on time, please email Zvi Kedem, `mailto:zk1@nyu.edu`, *in the way specified in the course description* with an explanation of what has happened, and if you have a solution (possibly partial), email the solution also.

   If you do need to submit the solution by email, and *only* if you need to submit by email because you are late or for other reasons, please follow the format as described next. Assuming that you are submitting your solution to Homework due 2034-02-15 and your Net ID is abc123, all the files of your homework should be emailed as a zip file named `20340215abc123.zip`. Of course you need to specify the correct date and the correct Net ID.

   Do not communicate with any of the graders concerning a late submission.

9. Be sure to follow the academic integrity rules listed in the posted syllabus. The department, the GSAS, and NYU treat academic integrity very seriously and we are required to report all possible violations.

10. **Note:** Due to the unusual circumstances, we will be more able to extend deadlines, but generally only on a one-by-one case. All such requests need to be addressed to Zvi Kedem, `mailto:zk1@nyu.edu` *in the way specified in the course description*, with a reason for such a request.

2 Homework

Reminder: If you are not officially registered in the class and the class does not show on Albert for you, do not submit any assignments.

Please read and follow carefully the instructions in Section 1.

2.1 Description

This is the first of two assignments dealing with SQL. Both assignments will use the same (or almost the same) small database. This first assignment is simpler than the second one as this is the first time you are actually writing and executing Oracle queries.

2.2 Assignments

1. (a) You are already supposed to know how to run SQL queries/commands on the Oracle systems at CIMS. You were asked to familiarize yourself with and follow the instructions in `How_To_Use_Oracle_At_CIMS.pdf`.

   You were asked to do that earlier in order to save you time while working on this homework.

   So it is assumed that you know how to do that.

(b) Look at the files `ER06.drawio` and `relational06.architect`. They will help you understand the database schema defined in the files `script06.sql` and `dataSetupScript06.sql`. These files fully specify the application.

(c) Read `script06.sql` and `dataSetupScript06.sql` carefully. These scripts both define and create the sample database and serves as the placeholder for putting in your solutions.

Look carefully over ANSWER0 there. It shows you how to insert a result of a query into an empty table. It also uses the temp table TEMP0, just to demonstrate the usage of temp tables.

(d) Input your queries into `script06.sql` after doing what is requested in Item 1e of Section 2.2.

Please note which operations to use. Use only the operations that were introduced in Unit 4 and the CREATE TABLE ... AS ..., used in the example of ANSWER0 in the script. So base your queries on the operations of selection, projection, cartesian product, minus, union, and intersection, with renaming and creation of new tables: creation, as shown in the script. Do not, e.g., use JOINs of any kind. *To reiterate: A solution that used other operations (including those covered in Unit 5) will not be acceptable*

In addition, use DISTINCT and ORDER BY as described below. If the output is to be sorted in a different way, use an appropriate variant of ASC and DESC (ascending and descending) and list the sorting instructions in the appropriate order,

For each query, *unless something else is required by the query* make sure to

i. Remove duplicates from the answer (unless requested otherwise); that's what DISTINCT does

ii. Sort the result in ascending order (unless requested otherwise); that's what ORDER BY does

This is *extremely important* to make the grading more manageable.

So, for example, assuming that you are going to select `a` and `b` and rename `b` as `c`, you should actually *explicitly* use:

```
SELECT DISTINCT a, b AS c
...
ORDER BY a ASC, c ASC;
```

Do not rely the on default removal of duplicates and sorting order. Add the `DISTINCT` *and* `ORDER BY` *instructions even if you think that they are not necessary.*

You may use temporary tables. If you choose to do that, use tables TEMP1, TEMP2, ..., TEMP20.

(e) Replace "zk1" in `script06.sql` with your NetID.

(f) Do not remove the existing sample query.

(g) Execute `script06.sql`, which internally calls `dataSetupScript06.sql` and produces a spool file `spool06.txt`. The spool file must only include the details of the queries from `script06.sql`. The spool file created will be a part of the submission.

Do not be concerned that there are more placeholder ANSWERs, than the queries that you are supposed to produce.

For reference, `spool06.txt` corresponding to the given `script06.sql` (with one sample query) is enclosed.

The requested queries are listed below. Your answers must work for every instance of the database and not only for the specific instance provided. The tables are named AnswerX, where "X" stands for the item number below. So, as the first item is item number 1, the first table is Answer1.

1. Produce table AnswerX(RegNumber, TIN) that lists all the companies, sorted ascending by RegNumber first and descending by TIN. Note that the order of the attributes is not the same in both places.
2. Produce table AnswerX(RegNumber, TIN) that lists all the big companies that work in the 'IT' domain and are managed by people with familyName 'Garcia'.
3. Produce table AnswerX(helperTIN, helpedTIN) that lists all the companies that help each other. Note that (TIN1, TIN2) and (TIN2, TIN1) are not duplicates if TIN1 and TIN2 are not the same.
4. Produce table AnswerX(RegNumber, TIN) that lists all the companies that work in the 'IT' domain, and are helped by both companyOne with TIN '59515298', and companyTwo with TIN '51251930'.
5. Produce table AnswerX(RegNumber, TIN) that lists all the companies (big or small) that do not help any company.
6. Produce table AnswerX(RegNumber, TIN) that lists all the companies, sorted ascending by RegNumber first and ascending by TIN, that do not work in the 'IT' industry.
7. Produce table AnswerX(RegNumber, TIN) that lists all the small companies that work in 'banking' domain that work in the 'IT' domain, and are helped by small companies in 'IT' domain.

2.3 What to submit

Please upload 2 files, named *exactly* as specified and in the format *exactly* as specified.

1. `script06.sql`, the script with your answers
2. `spool06.txt`, the resulting spool file