COMS W4111-Introduction to Databases, Fall 2011
Project 1, Part 2
(worth 25% of overall Project 1 grade)

FAQs

Your team and "project mentor"

You will do Part 2 of Project 1 with the same team as for Part 1. If your team partner dropped the class and you did **not** submit a contingency plan for this with your Part 1 submission, then unfortunately you will still have to complete the whole project by yourself. If you team partner dropped the class and you did submit a contingency plan for this with your Part 1 submission, then you are welcome to switch to this reduced version of your project.

For continuity, the TA who grades your Part 1 will grade the two remaining parts of your project. (In many but not all cases, this is the TA with whom you met to discuss Part 1.) This TA will be your **"project mentor"** for the remainder of your project. You are welcome to contact other members of the class staff (including the instructor) about your project, but your project mentor should be your main contact, for continuity, since your mentor will be grading all parts of the project.

As mentioned earlier, both students in a team will receive the same grade for Project 1. Team partners are expected to fully collaborate with each other on solving the project. However, communication about project details with somebody other than your partner is not permitted, and is considered cheating. If in doubt about what kinds of consultations are allowed, check with the instructor, or see the Computer Science Department's policies and procedures on academic honesty. Questions of a general nature that may be of interest to the whole class should be posted to the CourseWorks discussion board.

Grading of Part 1

Teams can pick up their **graded Part 1 in class starting on Wednesday October 12**; if you could not attend the class, please contact the TAs. Once you have received your graded Part 1, please modify your database according to your mentor's instructions. Meanwhile, please start getting familiarized with the Oracle Database Management System.

Computer accounts

You can choose to do this project on your Windows/Mac laptop or on Unix.

If you choose **Windows** or **Mac**, then you do not need a CS account for the project, but you will need to install software on your computer (see "**Software installation instructions,"** below).

In contrast, if you choose **Unix** for the project, then at least one of the team members should have a CS account, but you will not need to install any software for this part of the project. If neither of you have a CS account, please open one (see item above).

Overview of Part 2 of Project 1

For **Part 2** of this project, which you should submit **electronically**, you should implement your relations

of Part 1 over an Oracle Database server and load your tables with some real or realistic data, as you outlined in Part 1.

What you need to do for Part 2

1. Familiarize yourself with the Oracle DBMS by reading the materials available at http://www.oracle.com/technology/documentation/index.html . Our servers are running Oracle 10gr2. The documentation for Oracle 10gr2 can be found at http://www.oracle.com/pls/db102/homepage and http://www.oracle.com/pls/db102/to_toc?pathname=server.102%2Fb14200%2Ftoc.htm&remark=portal+%28Getting+Started%29. Please **check these materials carefully before sending email** to the class staff with questions on syntax or supported features.
2. Set up your Oracle database and change your Oracle account password (see below). You only need to set up one Oracle account/database per team. You don't need to turn anything in for this part.
3. **Make all suggested changes** to the overall design in general, and to the SQL schema in particular, once you receive your graded Part 1. You are likely to have extensive comments from your project mentor in your graded Part 1. Your Part 2 grade will be based in part on how well you have incorporated your project mentor's comments. If you have any questions about these comments once you received them, please contact your project mentor as soon as possible to clarify.
4. **Create all the SQL tables** in your revised SQL schema on your Oracle account, **including all constraints** that you could specify in the table declarations. ***Suggestion:*** *Please use VARCHAR2 as the domain for variable-length string attributes such as names, instead of a fixed-length CHAR domain. VARCHAR2 will simplify your handling of such attributes in Part 3, particularly if you are following the Web Front-End Option.*
5. Add any additional **attribute-based CHECK constraints and tuple-based CHECK constraints** (as discussed in class) that you need so as to express any real-world constraints of your application that are missing from your SQL schema. Note that Oracle does not support general assertions, and that triggers are slightly different from our class description. You do not need to use triggers for this project, and you can ignore any real-world constraints that you could not capture with either **good-style** (as discussed in class) attribute- or tuple-based CHECK constraints, or with PRIMARY KEY, UNIQUE, FOREIGN KEY, and NOT NULL constructs. (Hint: "Good-style" attribute- or tuple-based CHECK constraints tend to refer only to the table in which they are defined, not to other tables.)
6. **Insert into each table in your database, on average, at least 10 tuples of real or "realistic" data**, as you described in your Part 1 project description. This data will help you test and play with your database. Of course, issue queries of your choice to make sure that everything works as you intend.

Software installation instructions for Windows users only

If you choose to do your project on your Windows computer, then you will need to install the **Oracle XE client** software, which is freely available from Oracle, as explained next. (Note that your database—both the schema and the data—will reside on one of the Oracle servers that we will provide and which will be running on two dedicated CS computers. You will connect to the remote Oracle servers from your computer using the client software that you will install on your Windows computer.)

1. Download Oracle XE for Windows at http://www.oracle.com/technology/software/products/database/xe/htdocs/102xewinsoft.html: after clicking on “accept,” choose the **Oracle Database 10*g* Express Edition (Western European)** version. You might need to create an account to successfully complete the Oracle XE download process.
2. Click on the downloaded executable to install Oracle XE.
3. Configure your Oracle client by going to the directory in your Oracle installation that contains file **tnsnames.ora.** For example, if you installed Oracle XE in your **C:** drive, this file will be in directory **C:\oraclexe\app\oracle\product\10.2.0\server\NETWORK\ADMIN**. Edit file **tnsnames.ora** by

adding the following lines at the end of the file. These lines specify the two Oracle servers for the class with which your client will be able to communicate:

```
ADB =
 (DESCRIPTION =
  (ADDRESS = (PROTOCOL = TCP)(HOST = w4111c.cs.columbia.edu)(PORT = 1521))
  (CONNECT_DATA =
   (SERVER = DEDICATED)
   (SERVICE_NAME = ADB)
  )
 )

ADB2 =
 (DESCRIPTION =
  (ADDRESS = (PROTOCOL = TCP)(HOST = w4111b.cs.columbia.edu)(PORT = 1521))
  (CONNECT_DATA =
   (SERVER = DEDICATED)
   (SERVICE_NAME = ADB)
  )
 )
```

To connect to one of our Oracle servers (see below), you will use the Oracle **SQL*Plus** client. (Check our useful SQL*Plus tips.) This client will let you create tables and populate them. Please check the section "**Setting up your Oracle database**" below.

Software installation instructions for Mac users only

Follow the instructions below for installing SQLPlus on a Mac which will let you connect to the class Oracle servers from your OS X terminal:

1. Download Oracle Instant Client here (make sure you get the right version for you mac: 32 or 64 bit): http://www.oracle.com/technetwork/topics/intel-macsoft-096467.html
   To get SQLPlus working you need to download "Instant Client Package - Basic" and "Instant Client Package - SQLPlus". You'll have to register a free account with Oracle before downloading.
2. Expand these downloaded files and put the contents of both into the same folder. Let's say you put yours in: /Applications/instantclient_10_2. Put the following in your .bash_profile file:

   ```
   export ORACLE_BASE=/Applications/instantclient_10_2
   export ORACLE_HOME=$ORACLE_BASE
   export ORACLE_SID=ADB
   export EDITOR=vi

   export CLASSPATH=$ORACLE_HOME/jdbc/lib/ojdbc14.jar:$ORACLE_HOME
   /jlib/orai18n.jar:$CLASSPATH
   export DYLD_LIBRARY_PATH=$ORACLE_HOME
   export LD_LIBRARY_PATH=$ORACLE_HOME/lib:$LD_LIBRARY_PATH
   export PATH=$ORACLE_HOME:$ORACLE_HOME/bin:$PATH
   ```
3. Create the file tnsnames.ora in /etc and add the following to the file:

   ```
   ADB =
    (DESCRIPTION =
   ```

```
  (ADDRESS = (PROTOCOL = TCP)(HOST = w4111c.cs.columbia.edu)(PORT = 1521))
  (CONNECT_DATA =
    (SERVER = DEDICATED)
    (SERVICE_NAME = ADB)
  )
 )

ADB2 =
 (DESCRIPTION =
  (ADDRESS = (PROTOCOL = TCP)(HOST = w4111b.cs.columbia.edu)(PORT = 1521))
  (CONNECT_DATA =
    (SERVER = DEDICATED)
    (SERVICE_NAME = ADB)
  )
 )
```

4. You should now be able to connect to the oracle server like this:
   sqlplus username@ADB or
   sqlplus username@ADB2

Unix users only

1. Get a **CS account** if neither you nor your team-mate has one.
2. Use **ssh** to **clic.cs.columbia.edu** to access Linux machines in the CLIC cluster.
3. You should have received an email from Pranjal with your Oracle database username and password. This email was sent to your @columbia.edu email address. If you have not received such an email, please contact Pranjal at pg2354@columbia.edu immediately.
   Note: You only need to set up **one single Oracle account/database per team**, even though your team was assigned one Oracle account per team member. Pick just one of the two accounts arbitrarily and work on it together with your team-mate.
4. Add the following lines to your `.bashrc` file or equivalent:

```
export ORACLE_BASE=/usr/lib/oracle/xe/app/oracle
export ORACLE_HOME=$ORACLE_BASE/product/10.2.0/server
export ORACLE_SID=ADB
export EDITOR=vi

export CLASSPATH=$ORACLE_HOME/jdbc/lib/ojdbc14.jar:$ORACLE_HOME
/jlib/orai18n.jar:$CLASSPATH
export LD_LIBRARY_PATH=$ORACLE_HOME/lib:$LD_LIBRARY_PATH
export PATH=$ORACLE_HOME/bin:$PATH
```

   Set `ORACLE_SID` to `ADB2` if you are connecting to the Oracle server on w4111b.cs.columbia.edu.
   After editing your `.bashrc` file, please logout and login again, so that the new version of the file is loaded.

Setting up your Oracle database

1. For this class, we have two identical Oracle database servers:

   w4111b.cs.columbia.edu
   w4111c.cs.columbia.edu

   You may use either server, as we have created **independent** accounts for you on both of them. (Your two accounts are not linked in any way.) There are two servers to help balance load and provide some redundancy. If one server goes down, you may continue development on the other

server. It is important to note that the two machines, though identical, are **independent**. That is, none of your data on one machine will be replicated on the other machine. Therefore, use the **exp** command frequently so you can restore your data on the other machine if your development machine crashes; you should use the **imp** command to automatically recreate your schema and data on the other machine. **See the "Important note on backups" section below** for details about using both the **exp** and **imp** commands.

CRF **has asked us to request that students in cs4111 not call or send emergency tickets to CRF regarding a crashed database server (or any other such problem) outside of CRF's working hours. If you have any server status related issues, please contact Pranjal (pg2354@columbia.edu).**

6. Connect to the Oracle server using SQL*Plus:

   To connect to the Oracle server on w4111b.cs.columbia.edu: `sqlplus <username>@ADB2`

   To connect to the Oracle server on w4111c.cs.columbia.edu: `sqlplus <username>@ADB`

   At the `Password:` prompt, enter the password that was emailed to you.

7. Change your password at the Oracle prompt by typing:

   ```
   ALTER USER <username> IDENTIFIED BY <your new password>;
   ```

   Note that all SQL commands issued from within the SQL*Plus client **need to end in ";"** (or carriage-return followed by "/") to be sent to the server. Pressing <enter> without a semicolon will continue the same command on the next line.

8. Exit your SQL*Plus session by typing `exit;`.

What to submit and when

You will submit this part of the project **electronically on CourseWorks** directly. The deadline is **Monday October 31 at 5 p.m. EST.** Here is how you should proceed:

1. Create a directory named **<your-UNI>-proj1part2**, where you should replace **<your-UNI>** with the Columbia UNI of one teammate (for example, if the teammate's UNI is **abc123**, then the directory should be named **abc123-proj1part2**).
2. Generate the dump file **expdat.dmp** using Oracle's export utility **exp** and place it inside the **<your-UNI>-proj1part2** directory. The export utility lets you store your database objects in an external file. Note that the **exp** command must be executed from your shell if you are using Unix, or from the command prompt if you are using Windows, **not** from within SQL*Plus. The **exp** command creates a file called **expdat.dmp** in the directory where it is executed. Here is how you should run **exp**:
   - For the Oracle server on w4111b.cs.columbia.edu:
     exp userid=<userid>/<password>@ADB2 TABLES=table-1,table-2,…,table-n ROWS=Y
     where **table-1**, **table-2**, …, **table-n** are the names of all the tables that you defined in your database.
   - For the Oracle server on w4111c.cs.columbia.edu:
     exp userid=<userid>/<password>@ADB TABLES=table-1,table-2,….,table-n ROWS=Y
     where **table-1**, **table-2**, …, **table-n** are the names of all the tables that you defined in your database.
3. Generate the following **plain-text file** inside the **<your-UNI>-proj1part2** directory:
   File **queries.txt**: This plain-text file has three "interesting" SQL queries over your database, with a sentence or two per query explaining what the query is supposed to compute. The goal of these queries is to help us better understand your application. You will not be graded on these queries, but we strongly suggest that you submit well-formed queries that run without problems, so please make sure that you have tested your queries by running them on your

database exactly as submitted (use copy and paste).

4. *(Unix option only)* **Tar** and **gzip** the **<your-UNI>-proj1part2** directory, to generate **a single file <your-UNI>-proj1part2.tar.gz** (containing the expdat.dmp and queries.txt files), which is the file that you will submit.
5. *(Windows option only)* Zip the **<your-UNI>-proj1part2** directory, to generate **a single file <your-UNI>-proj1part2.zip** (containing the expdat.dmp and queries.txt files), which is the file that you will submit.
6. Login to Courseworks at https://courseworks.columbia.edu/ and select the site for our class.
7. Select "Class Files" and then "Post File."
8. Enter **"Project 1, Part 2"** in the "Title" field and select your **<your-UNI>-proj1part2.tar.gz** file (from Unix) or your **<your-UNI>-proj1part2.zip** file (from Windows) in the "File" field using the "Browse" button.
9. Select **folder "Project 1, Part 2" (under "Shared Folders")** in the "Post File To" field.
10. Enter the names of the members of the team and their UNI in the "Comments" field.
11. Hit the "Submit" button.

Late policy for projects

Please check the Project Lateness Policy.

Grading for Part 2

Your grade for Part 2 of Project 1 will be split as follows:

1. **Quality of final SQL schema and implementation on Oracle: 10 points.**
   We will evaluate the overall quality of your final SQL schema on Oracle, especially in terms of how thoroughly you incorporated any revisions suggested by your project mentor in the grading of Part 1 of your project.
2. **Quality of constraint handling: 10 points.**
   We will evaluate how well you managed to capture real-world constraints through primary key, foreign key, unique, and attribute- and tuple-based CHECK constraints.
3. **Quality of the real-world (or at least realistic) data that you loaded into the database: 5 points.**

Important note on backups

To avoid losing any data, we strongly suggest that you periodically use the Oracle export utility, **exp**, as explained above to generate a file with all your schema definitions and with the data that you uploaded to your database. You should then save this file in a safe place. For a variety of technical reasons that are not under our control, **we cannot provide backups of the contents of your Oracle account, so do back up your data frequently**.

- **Oracle Export Utility exp**: Please refer to Step 2 of the “What to submit and when” section above.
- **Oracle Import Utility imp**: The **imp** import utility allows you to import database objects that were previously exported using the **exp** command. In case of a server crash, you can reconstruct your database (both the schema and the data) from the output of the **exp** command. So we strongly suggest that you run the **exp** command frequently and save its output in a safe place!

  To use the import utility **imp** to import your table schemas and data, as recorded in an expdat.dmp file produced by **exp**, proceed as follows:
  - For the Oracle server on w4111b.cs.columbia.edu: **imp userid=<userid>/<password>@ADB2 file=expdat.dmp full=yes**
  - For the Oracle server on w4111c.cs.columbia.edu: **imp userid=**

<userid>/<password>@ADB file=expdat.dmp full=yes

Important instructions on Oracle timeouts and limit on the number of connections

At any given time, each user can have a **maximum of 5 open connections to each Oracle server**. Also, a connection will be **automatically closed by Oracle if the connection has been idle for 20 minutes.** (In this case, you will receive the following error message: "ERROR at line 1: ORA-02396: exceeded maximum idle time, please connect again.") These limits are in place to make the Oracle servers more stable and avoid overload-originated server crashes as much as possible. **If you find out that your SQL*Plus session has ended "automatically" and/or you are unable to connect to the Oracle servers using SQL*Plus, please check the following before emailing the TA's or posting a question on CourseWorks**:

- **Make sure that you don't have open connections to the server in other windows** (or machines!). Recall that you cannot open more than 5 simultaneous connections to your Oracle account on a server.
- **Wait for 20 minutes and try again**: after 20 minutes, all your idle connections will be closed automatically by Oracle. So even if you forgot to close some connections from a computer that you don't remember at all, after 20 minutes your idle connections will be closed and you will go below the 5-connection limit, so you should be able to connect at that point.

Frequently-asked questions

- **Q: Can you give us some ideas on how to use the Oracle SQL*Plus client?**
  A: Please check our useful SQL*Plus tips.
- **Q: Why use Oracle for Parts 2 and 3 of Project 1? Can I use my favorite DBMS instead?**
  A: As much as we would like to be more flexible, we just don't have the staff to handle several diverse systems and platforms. Unfortunately, you cannot use any other DBMS.
- **Q: Why are we using Oracle instead of MySQL or PostgreSQL?**
  A: Oracle generally supports more advanced functionality than MySQL and PostgreSQL do.
- **Q: How do I learn more about Oracle? Is <feature X> supported on Oracle?**
  A: Please check Oracle documentation at http://www.oracle.com/pls/db102/homepage and http://www.oracle.com/pls/db102/to_toc?pathname=server.102%2Fb14200%2Ftoc.htm&remark=portal+%28Getting+Started%29.
- **Q: I can't log in on the machines running Oracle servers using the userid and password as specified above. How do I fix this?**
  A: You do **not** have accounts on the **machines** running the Oracle servers; your accounts are for Oracle only. This means that you cannot **ssh** to the machines that run the Oracle servers. Rather, you should login to one of the CLIC machines using your regular CS account userid, and follow the instructions above on starting a client locally to connect to the **remote** Oracle server, assuming you are following the Unix Option.
- **Q: I can't find my .bashrc file. How do I proceed?**
  A: You should simply create a plain-text file called **.bashrc** in your home directory and paste the commands on the project description into the file.