Lab 6: Simple Queries Part II

Due date: Monday, February 18, midnight.

Note: February 18 is an academic holiday, no classes. Lab 7 will be assigned on February 20.

Lab Assignment

Assignment Preparation

This is an individual lab. Each student has to complete all work required in the lab, and submit all required materials exactly as specified in this assignment.

The assignment will involve writing SQL queries for different information needs (questions asked in English) for each of the five course datasets.

The Task

You are to write and debug (to ensure correct output) the SQL queries that return information as requested for each of the information needs outlined below. The information needs can be addressed with either a simple SELECT statement (i.e., a SELECT statement without grouping, aggregation and nested subqueries) and/or with UNION, MINUS and INTERSECT statements, or with a SELECT statement that uses some aggregation. Each information need must be met with a single SQL statement.

For this assignment, you will prepare one SQL script for each database. In addition to SQL statements you may need to include some SQL*plus formatting instructions to ensure that your output looks good. In particular, every row of every resulting table must be printed in a single line. If that means changing the size of the line - do it. Similarly, there should not be awkward pagination of the answers - change page size as needed.
NOTE: Please provide a comment in front of each SQL statement in each of your files. The simplest comment can just state the query number (e.g., "--- Q3.") for this particular database. This is very useful for the situations when for one reason or another you elected not to implement a query.

STUDENTS dataset

For the STUDENTS dataset, write an SQL script STUDENTS-lab6.sql containing SQL statements answering the following information requests.

1. Find all classmates of TOBIE SAADE. Report first and last name. Do not report TOBIE SAADE. Sort output in alphabetical order by last name.
2. Find all first-grade students who are NOT taught by JEROME COVIN. Report their first and last names in alphabetical order by last name.
3. Report the total number of first-graders in the school.
4. Find and report the number of students taught by OTHA MOYER.
5. List all students whose first names start with letter 'R'. Report first and last names and the grade. Sort in alphabetical order by last name.

BAKERY dataset

Write an SQL script BAKERY-lab6.sql containing SQL statements answering the following information requests.

Note: Your queries must match exactly the wording of the information need. For example, if you are asked to find the price of an Apricot Tart, the following query

```
SELECT price
FROM goods
WHERE CODE = '90-APR-PF';
```

is considered to be incorrect because nowhere in the query was the code '90-APR-PF' mentioned. (This is especially important when you are expected to produce a join of two or more tables, but instead look up the foreign key value and use it verbatim in the query. Such queries will be marked as incorrect on the spot).

1. Find all purchases in which a customer bought multiples of a single type of a cake (e.g., two Casino Cakes). Report receipt number, date of purchase, purchased cake (flavor, food) and the name of the customer. Sort the output by purchase date, and by the receipt number for purchases made on the same day. Each item must be reported once.
2. Find all dates in the first half of October of 2007 (October 1 to October 15 inclusive) on which one customer made multiple purchases. Report each date exactly once, output dates sorted in ascending order.

3. Find all customers who purchased, during the same trip to the bakery, two different Croissants. Report first and last names of the customers in alphabetical order by their last name.

4. Find all customers who did not make a purchase between October 14 and October 19, 2007 (inclusive). Report their first and last names sorted alphabetically by last name.

5. Find all days on which at least one purchase was made, but no one purchased a cake.

6. Report the total number of purchases (receipts) that included exactly five items.

7. Report the total amount of money TERRELL ARZT spent at the bakery during the month of October, 2007.

8. Report the total amount of money spent by bakery customers in October 2007 on Cookies.

**CARS dataset**

1. Find all cars that have that are more powerful (more horsepowers) than the 1970 dodge d200. Report the makes, the years, the horsepower of the car and the short name and the country of the maker. Output in ascending order by horsepower.

2. Find all cars made after 1980 with gas mileage better than the 1982 honda civic. Report full name of the car, year it was made and the name of the manufacturer. Sort output in descending order by gas mileage.

3. Find the average, maximum and minimum horsepower for 4-cylinder vehicles manufactured by chrysler between 1971 and 1976 inclusively.

4. Find how many cars produced in 1971 had better acceleration than a 1972 volvo 145e (sw). Report just the number.

5. Find how many different car manufacturers produced a vehicle heavier than 4000 lbs.

**CSU dataset**

Here are the queries for the CSU dataset. Name the SQL script CSU-lab6.sql

1. For each university with over 20,000 students enrolled in 2004, report the difference and the ratio between the student fees in 1996 and 2004. Sort the output in descending order by the ratio.
2. Report all disciplines in which graduate enrollment in 2004 in California State Polytechnic University-Pomona was less than the enrollment in California Polytechnic State University-San Luis Obispo. Report the discipline names and the enrollments in both universities. Output the discipline names in alphabetical order.

3. Find all campuses in which the number of faculty full-time equivalent positions (FTEs) increased on two consecutive years. Report campus, the three years involved (year1→increase→year2 → increase → year3) and the FTE numbers for each year.


5. Find all campuses which in 2004 had more Computer and Info. Sciences than Education students in the graduate program. Report campus names and the number of graduate students in each of the two disciplines.

6. Report the total number of degrees granted by California Polytechnic State University-San Luis Obispo in the period between 1995 and 2000 (inclusively).

7. Find the largest, the smallest and the average fee on a CSU campus in 2005.

8. Find the average fee on a CSU campus in 2003 among the campuses with a better student-to-faculty ratio (use full-time equivalents) than San Diego State University.

INN dataset

For the INN dataset, create a SQL script file INN-lab6.sql with SQL queries for the following information needs. (When no year is supplied in the query descriptions below, assume 2010).

1. Find all rooms that are not occupied on August 20 and 21, 2010. Report room code, full name of the room and the base rate. Sort output in alphabetical order by room name.

2. Find all customers who stayed at the inn more than once during the year, with one of the stays falling completely inside the month of May. Report customer names (first, last) in alphabetical order. Each name shall be reported once.

3. Find all guests who stayed at least twice in the same room. Report the first and last name of each guest and the full name of the room they stayed multiple times at. Sort in alphabetical order by room name, then by last name of the guest.
4. Find the names of all people\textsuperscript{1} staying at the inn at the same time as JESS TORNQUIST. Sort the output in alphabetical order by last name.

5. Find the number of April reservations (both checkin and checkout dates are in April) where two adults are staying with two children.

6. Order all reservations that commenced on a Thursday of the month of May\textsuperscript{2} by the total amount paid for the reservation. For each reservation report the room name, the last name of the person staying, checkin date, number of nights and the total paid.

**MARATHON dataset**

For this dataset, all times must be outputed in the same format as in the original dataset (in the file \textit{marathon.csv}). The information needs are below. Name the file \textit{MARATHON-lab6.sql}.

1. Find all gender-age groups that did not have runners in the top 50.

2. Find all male runners from MA who had a better result than KENDRIK HOLZ. Report first and last name of the runner, sort in descending order by their place in the race.

3. Find all winners in their respective gender-age groups who came in ahead of the fastest male runner in the 50-59 age group. Report the name of the runner, their gender-age group, overall place in the race and the time. Sort in ascending order by the overall place in the race.

4. Find the number of 16-year old runners who participated in the race.

5. Find the best result (place) for a 23-year old female runner.

**AIRLINES dataset**

1. Find all airlines which do not have flights to the LTS airport. Report just the abbreviated names of the airlines.

2. Find all airports that are reachable via a direct flight from both ASY and ANN airports. Report just airport codes.

3. Find all airlines that have multiple flights out of both CVO and CAK. Report abbreviated names of airlines.

4. Report the total number of flights from ASY.

5. Report the total number of JetBlue flights from the airports whose airport code does not start from letter A.

\textsuperscript{1}We only know the names of the people who made the reservations, so only those names are subject to the query.

\textsuperscript{2}Look up the dates in the calendar.
WINE dataset

Create a SQL script WINE-lab6.sql containing SQL statements representing the following information needs.

1. Find all Sonoma county AVA s which do not have a Zinfandel wine in the database. Report the full name of the AVA sorted in alphabetical order.

2. List all 2006 vintage wines from Napa (any appellation within the county) whose total revenue exceeds that of the 2006 ‘Appelation Series’\(^3\) Paso Robles Zinfandel from ‘Rosenblum’ winery. For each wine report grape, winery and name, score and revenue. Order by revenue.

3. Find all wines produced in the same vintage year as the Tor Chardonnay, which have both the higher score and the higher production.

4. Find all 2008 wines with scores higher than 90 and higher revenue (assuming full sales) than the 2008 Merry Edwards Sauvignon Blanc from Russian River Valley. Report grape, winery and the revenue, sorted in descending order by the revenue. Recall, that a single case of wine is 12 bottles.

5. Find the average score of a Paso Robles Zinfandel.

6. Find the total revenue from all red wines made by Kosta Browne.

7. Find the average number of cases of a Pinot Noir produced from grapes sourced from the Central Coast.

8. Report the overall number of different red wines produced in Russian River Valley during the year when this AVA had a wine with a score of 98.

Submission Instructions

You must submit all your files in a single archive. Accepted formats are gzipped tar (\texttt{.tar.gz}) or zip (\texttt{.zip}). The file you are submitting must be named \texttt{lab4.ext} where \texttt{ext} is one of the extensions above. The archive shall contain eight directories: AIRLINES, CARS, CSU, INN BAKERY, STUDENTS, MARATHON and WINE.

Each directory shall contain the following SQL scripts:

- Database creation script. (e.g., CARS-setup.sql). Use the scripts from Lab 2/Lab 4 submissions.

- Table creation script. Use <DATASET>-insert.sql (e.g., CARS-insert.sql) file from Lab 4 submission.

\(^3\)There is a typo there. Let it be for now.
• The cleanup script (e.g., CARS-cleanup.sql). Use the scripts from Lab 2/Lab 4.

• **NEW script.** One script per database, containing all SQL statements and any SQL*plus statements needed for formatting. Name the script (as specified above) <DATASET>-lab6.sql (e.g., CARS-lab6.sql).

Submit using handin:
Section 01:

$ handin dekhtyar lab06-01 <file>

Section 03:

$ handin dekhtyar lab06-03 <file>