

COMP 543 Assignment #1

Due February 3rd at 11:55PM.

1 Description

The goal of this assignment is to write several SQL queries that will answer questions over the database used by the imaginary Southern Sierra Wildflower Club (SSWC), an organization whose members are interested in observing wildflowers in their native habitat in the southern part of the Sierra Nevada mountains of California.

The database maintained by the club has four tables:

```
SIGHTINGS (SIGHT_ID, NAME, PERSON, LOCATION, SIGHTED)
FEATURES (LOC_ID, LOCATION, CLASS, LATITUDE, LONGITUDE, MAP, ELEV)
FLOWERS (FLOW_ID, GENUS, SPECIES, COMNAME)
PEOPLE (PERSON_ID, PERSON)
```

The database tables have the following semantics:

- SIGHTINGS gives information that describes every time that a member of the club observes one of the wildflowers described in the table FLOWERS. NAME tells the name of the flower observed, PERSON describes who saw the flower, LOCATION tells the name of a nearby geographical feature where the flower was seen, and SIGHTED tells the day when the flower was seen.
- FEATURES lists the various locations where flowers have been observed. LOCATION is the name of the place, CLASS is the type of place (there are several types, such as Summit, Mine, Locale, etc.), LATITUDE and LONGITUDE describe where on the surface of the earth the locations are found (if you are not familiar with the concepts of latitude and longitude, you might want to do a web search on them; the first is like an x-coordinate on the Earth's surface, and the second is like a y-coordinate). MAP tells the name of the topographic map where the feature can be found, and ELEV tells the height of the feature.
- FLOWERS lists all of the flowers that the members of the SSWC try to find. GENUS and SPECIES give the scientific name for the flower, and COMNAME gives the non-scientific name (SIGHTING.NAME is a foreign key into FLOWER.COMNAME).
- Finally, PEOPLE lists all of the people in the club.

Your assignment is to write SQL queries that answer the following questions, sorted in general on order of difficulty from easiest to most difficult. List all the results alphabetically, unless otherwise specified.

1. Who has seen a flower at Alaska Flat?
2. Who has seen the same flower at both Moreland Mill and at Steve Spring?
3. What is the scientific name for each of the different flowers that have been sighted by either Michael or Robert above 8250 feet in elevation?

4. Which maps hold a location where someone has seen `Alpine penstemon` in August?
5. Which genus have more than one species recorded in the SSWC database?
6. What is the common name of the most commonly sighted flower (in terms of number of sightings)?
7. Who has not seen a flower at a location of class `Tower`?
8. For each feature class, compute the total number of flower sightings.
9. For each month, compute the fraction of the various flower species that were observed. For example, say that all of the sightings were in May and June. If 56% of the different flowers were observed in May and 74% in June, your query should return `{(May, .56), (June, .74)}`. Sort by month number (e.g. January, February, March, ...)
10. Who has seen a flower on every summit on the `Sawmill Mountain` map, except for `Cerro Noroeste`?
11. For those people who have seen all of the flowers in the SSWC database, what was the date at which they saw their last unseen flower? In other words, at which date did they finish observing all of the flowers in the database?
12. Which latitude range (defined by a lower latitude and an upper latitude) having no more than 20 different locations inside of it had the most flower sightings, and how many sightings were there?

2 Turnin

For each question, you need to do the following:

1. Execute the query using SQL Server
2. Copy and paste the query and the answer given into a word/text processor document
3. Then, turn in the document that you create to Canvas as a .txt file

3 Grading

Each question is worth 10 points. Points will be assigned as follows:

- 0 points: query not attempted, query does not give any results, or it does not compile.
- 5 points: query compiles, runs, and is most of the way towards a correct answer, but may have a major bug.
- 8 points: the query and answer given are almost correct, but there is a slight or subtle bug in the query.
- 10 points: the query is correct and gives the right answer.

4 Getting Started

First, install MySQL on your computer. Some directions for MacOS X and Windows 10 are posted on the course piazza site. Your system may be slightly different, so you might want to install the software earlier rather than later so you have time to get help with the install, if you need it.

To begin on the assignment, download and unzip the SSWmysql.zip archive. You will then want to connect to the database server, create your database using the contents of the zip archive, and then you can start writing your queries. First run createtables.sql, and then run sightings.sql, flowers.sql, people.sql, and features.sql. If you ever want to destroy those tables, you can just run droptables.sql. You will then need to run those scripts again to recreate them. Note that in MySQL, all queries are followed by a “;”. The query is compiled and executed when a ‘;’ is seen. So, once you create the SIGHTINGS table, you might type:

```
SELECT *  
FROM SIGHTINGS;
```

5 A Word of Caution

Start early! The average student might need 6 to 10 hours to get all 12 queries to work.

6 Two Final Notes

Several of these queries will be nearly impossible to write without breaking them up into pieces. The way to do this is with the SQL CREATE VIEW command.

Here’s a big hint. If you are not using an IDE (such as MySQL Workbench) and are instead using some other software to connect and interact with the database via a command prompt, do not type your queries directly into the database prompt. Instead, type them into your favorite text editor or word processor, and then copy and paste your query into the prompt. That way, when your query does not work (it never will the first time!), you don’t have to re-type it. You can simply edit your query and then re-paste it. This may seem obvious, but many people would not do this without first being advised to do it!