CMSC424 – Database Design

SQL Assignment (Parts A and B)

Spring 2008
Prof. Mihai Pop
TA: Sharath Srinivas

Part A Due Date: Feb 19, before 11:59 PM
Part B Due Date: Feb 26, before 11:59 PM

- The assignment is to be done by yourself.

- To start with, log in to a grace machine, run ”tap oraclient”. The SID is ”dbclass1”. This will initialize the environmental variables. Then run ”sqlplus”, and log in using the user/password emailed to you. You can now start using it.

- **Data**: The dataset contains geographical information about the countries of the world. The information was collected from several web sources, with the CIA World Factbook (https://www.cia.gov/library/publications/the-world-factbook/) being the prominent source.

- **Schema Information:** There are 16 tables in the schema. Most of the table and column names are self-explanatory. The Schema diagram is provided on page 5 and 6 of this document. The ER diagram is provided on page 7. An appendix which lists the meanings of the columns of the tables is on page 8, 9 and 10.

- **Assumptions and Instructions:**

    1. The database has several geographical inaccuracies. But, for the questions in this assignment it is sufficient if your answers are consistent with the supplied dataset, even if they are not geographically correct.

    2. Answer to the questions 1-16 must be a single query. You can make use of temporary tables and nested sub-queries, but you cannot make use of two different queries to solve one question. For questions 17-20, there is no restriction on the number of queries you can use.

    3. The country codes cannot be used for querying any of the tables. E.g., In the “Economy” table the column “Country” is actually the country code and not the name of the country. So, if you have to find the GDP of France, you should make use of the “Country” table to get the code corresponding to France and then use that code to retrieve the GDP of France from the “Economy” table.

4. The string literals are case sensitive. In order to query on them make use of the SQL "LIKE" condition and the string enclosed within **single quotes**. The "%" symbol can be used as a wild card.

   E.g., select * from country where name like 'United States%';

- **Instructions to load the data:** A compressed file containing the dataset and the instructions for loading it into Oracle will be provided. Go through the README file in dataset.tar.gz for instructions to load the data.

- **How to hand in your work:** You should be making two separate submissions, one for Part A and one for Part B. Please included solutions to all the problems in a file named <lastname>-A.sql for Part A, and a file named <lastname>-B.sql for Part B. Use SQL-style comments at the start of each query to specify what question the query will solve. IMPORTANT: If you do not signify what problem a query addresses, you will receive no points for that query.

  Aside from comments made inline with your SQL queries, all other comments or notes to the TA should be made in a file named README. Do not include a spool of the output of your queries, we will run your queries to generate the output. You should probably only need to submit at most two files (the README and the file containing your SQL queries). Compress your files using tar, gzip, etc. Call your submission <lastname>-sqlA.tar.gz (or whatever extension matches the compression method you used) for Part A, and <lastname>-sqlB.tar.gz (or whatever) for Part B. Email your submission to sharath@cs.umd.edu.

Part A (4 pts each = 40 pts)

1. Write a query to change your Oracle password. When submitting the answer, don't write the new password (write "****" instead).

2. Figure out how to find out all the attributes and their types for a give table in SQL*Plus and write down the statements and the result on the "Organization" table. This kind of a statement is usually unique to the client, and is not part of standard SQL.

3. Write a query to display the names, capitals and the population of all the countries in the world.

4. Find the countries with the highest and the lowest population. The result should include the country name and its population.

5. For all the cities in the world that are capitals of some country, display their longitude and latitude.

6. What percentage of Russia lies within Europe?

7. Write a query to find all the countries that border Romania.

8. Find the total number of people living in all the countries that border Romania.

9. Create a View called "Asian_Rivers" that just contains the rivers flowing through any of the countries in Asia and their lengths.

10. Retrieve the list of countries sorted by their continents, and within the continent, arranged by their GDP in ascending order.

PART B (6 pts each = 60 pts)

11. Write a query to add a column called "POP_DENSITY" to the "country" table.

12. Initially the POP_DENSITY column of the "country" table is empty. Write a query to update this column with the population density of the country. (Population density = Population/Area) You can use this table in further queries if you want.

13. Find all the countries whose population density differs by a factor of 10,000 or more.

14. Find the 5th largest economy in the world. (Assume that the world economies are ranked on their GDP.)
    Note: Your result should have only one row: The 5th largest economy in the world.

15. Find all the countries that span across more than one continent. Display the count and names of continents across which the country spans.

16. When a new Organization is established, the country in which the headquarter of the organization is located, by default becomes a member of that Organization. Write a trigger to insert a new record in the "is_member" table whenever a row is added to the "organization" table. Set the "type" column of "is_member" table to the string value: "host".

17. If all the statistics in the dataset are for the year 2007, write a query to predict the population of Australia in the year 2010. You can assume that the population growth rate of Australia remains constant.
HINT: Population in 2008 = population in 2007 + (growth_rate*population in 2007)

18. Write a query to create a table "EURO_PERCENTAGE" with the following schema:

Name	Percentage

Meaning of the columns:

Name: Name of the country that lies within Europe
Percentage: Percentage of Europe's population living in that country. E.g., 10.55% of Europe's population lives in Germany.

This table lists only those countries that lie entirely or have at least some part of their territory within Europe. Further, for those countries that lie partially in Europe, you might assume that their entire population resides within the European border of that country.

19. Find the countries that are 3 hops away from Turkey. (A country is one hop away from another country if they share a common border)

20. Write a query to find the total number of people living in Asia. You can assume that the population of the countries is uniformly distributed. E.g., If 80% of Russian territory is inside Asia, you can assume that 80% of Russians live in Asia.

Schema Diagram

Country:

Name	Code	Capital	Province	Area	Population

City:

Name	Country	Province	Population	Longitude	Latitude

Province:

Name	Country	Population	Area	Capital	CapProv

Economy:

Country	GDP	Agriculture	Service	Industry	Inflation

Population:

Country	Population	Infant_Mortality

Continent:

Name	Area

borders:

Country1	Country2	Length

encompasses:

Country	Continent	Percetage

Organization:

Abbreviation	Name	City	Country	Province	Established

is_member:

Country	Organization	Type

Sea:

Name	Depth

geo_sea

Sea	Country	Province

River

Name	River	Lake	Sea	Length

geo_River

River	Country	Province

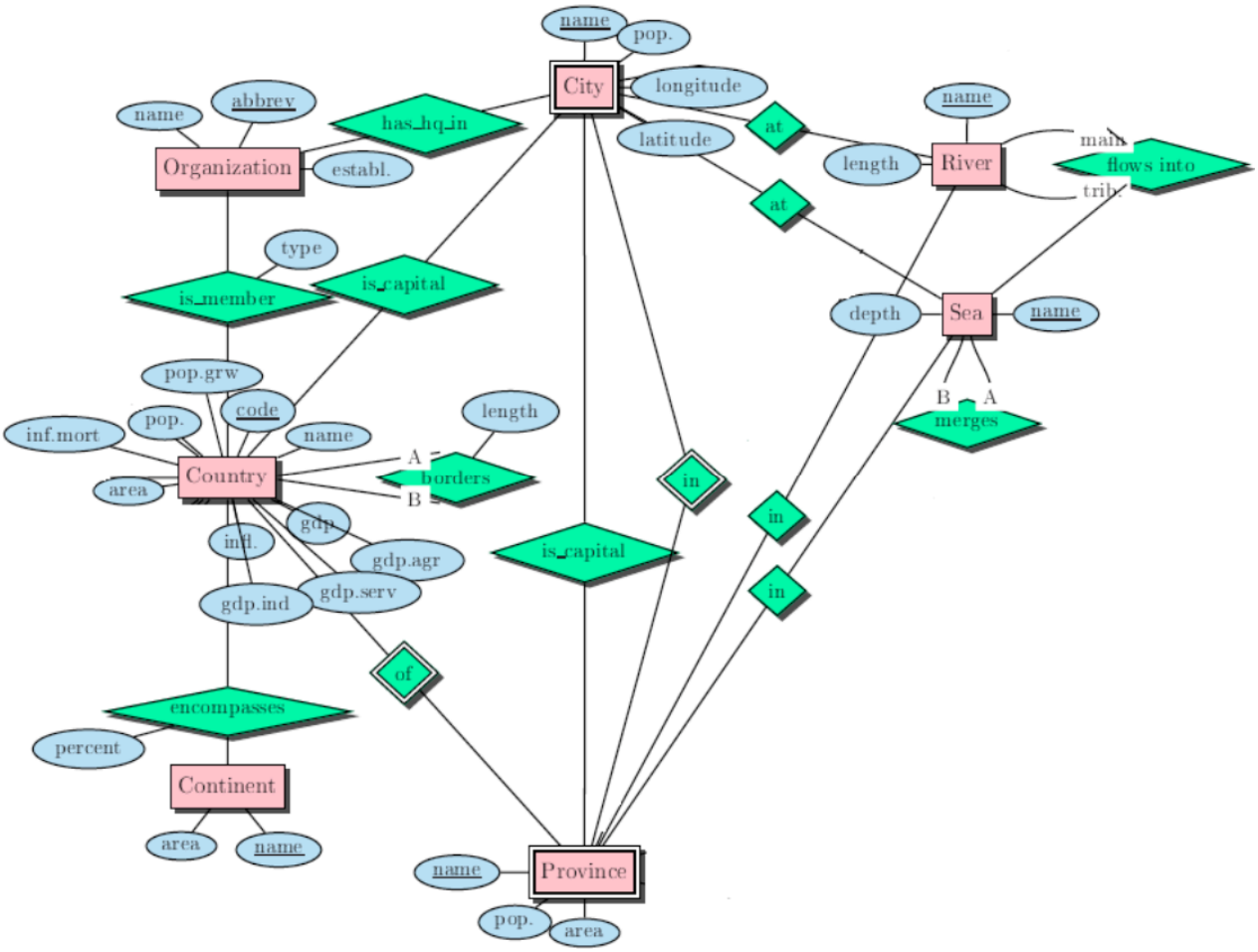
merges_with

Sea1	Sea2

located

City	Province	Country	River	Lake	Sea

E-R diagram


name
pop.
City
longitude
latitude
at
name
abbrev
has_hq_in
Organization
establ.
length
River
main
flows into
trib.
at
type
is_member
is_capital
depth
Sea
name
pop.grw
code
length
inf.mort
pop.
name
A
borders
B
area
Country
gdp
infl.
gdp.agr
gdp.serv
gdp.ind
is_capital
in
in
in
B
A
merges
of
encompasses
percent
Continent
area
name
name
Province
pop.
area

Appendix: Meanings of the database columns

Country: The countries of the world with some data.
Name: The country name.
Code: The country code
Capital: The name of the capital
Province: The province where the capital belongs to
Area: The total area
Population: The total population

City: information about cities.
Name: the name of the city
Country: the country code where it belongs to
Province: the name of the province where it belongs to
Population: the population of the city
Longitude: geographic longitude
Latitude: geographic latitude

Province: information about administrative divisions
Name: the name of the administrative division
Country: the country code where it belongs to
Area: the total area of the province
Population: population of the province
Capital: the name of the capital
Capprov: the name of the province where the capital belongs to
Note that *capprov* is not necessarily the same as *name*. E.g., the municipality of *Bagota* (*Columbia)* is a province of its own, and *Bagota* is the capital of the surrounding province *Cudinamarca*.

Economy: economical information about the countries
Country: the country code
GDP: gross domestic product
Agriculture: % of agriculture of the GDP
Service: % of services of the GDP
Industry: % of industry of the GDP
Inflation: inflation rate (per annum)

Population: information about the population of the countries
Country: The country code
Population_growth: population growth rate per annum (%)
Infant_mortality: infant mortality (per 1000)

Borders: information about neighboring countries. Note that in this relation, for every pair of neighboring countries (A,B), only one tuple is given- thus the relation is *not* symmetric.

Country1: a country code
Country2: a country code
Length: length of the border between country1 and country2

Continent: Information about continents.
Name: name of the continent
Area: total area of the continent

Encompasses: information to which continents a country belongs
Country: the country code
Continent: the continent name
Percentage: percentage, how much of area of a country belongs to the continent

Organization: information about political and economical organizations.
Name: the full name of the organization
Abbreviation: its abbreviation
City: the city where the headquarter is located.
Country: the country code where the headquarter is located
Province: the province where the headquarter is located
Established: the date of establishment

Is_member: memberships in political and economical organizations
Organization: the abbreviation of the organization
Country: the code of the member country
Type: the type of membership

River: information about the rivers
Name: name of the river
Length: the length of the river
River: the river where it flows to
Lake: the lake where it flows to
Sea: the sea where it flows to

Sea: information about the seas
Name: the name of the sea
Depth: the maximum depth of the sea

Geo_sea: geographical information about the seas
Sea: name of the sea
Country: the country code where it is located
Province: the province of this country

Note for a country there can be several provinces where the sea is located in.

Analogous for *geo_river*
merges_with: information about neighboring seas
sea1: a sea
sea2: a sea

located: information about cities located at rivers, lakes and seas
city: the name of the city
country: the country code where the city belongs to
province: the province where the city belongs to
river: the river where it is located at
lake: the lake where it is located at
sea: the sea where it is located at

Note that for a given city, there can be several lakes/seas/rivers where it is located at.