```
-- This is a list of the SQL queries we executed in class

---------------------------------------------------------------
-- SQL queries using one relation

-- What is the population of the US?
SELECT Population
FROM Country
WHERE Code = 'USA';

-- Which countries gained independence after 1989?
SELECT Name, IndepYear
FROM Country
WHERE IndepYear >= '1990';

-- Return all the attributes for cities with population over 1 million in the US
SELECT *
FROM City
WHERE Population >= '1000000' AND CountryCode = 'USA';

-- This query returns only the name of these cities, and renames the attribute
-- Name to LargeUSACity
SELECT Name AS LargeUSACity
FROM City
WHERE Population >= '1000000' AND CountryCode = 'USA';

-- Cities with population over 5 million (changing the unit to millions)
SELECT Name, (Population / 1000000) AS PopulationInMillion
FROM City
WHERE Population >= '5000000';

-- If we want to get a float value, and not the rounded integer value
SELECT Name, ((Population * 1.0)/ 1000000) AS PopulationInMillion
FROM City
WHERE Population >= '5000000';

-- Find the countries that have a form of goverment related to monarchy
SELECT Name, GovernmentForm
FROM Country
WHERE GovernmentForm LIKE '%Monarchy%';

-- Find all forms of goverment (the below query IS NOT CORRECT !!)
SELECT GovernmentForm
FROM Country ;

-- Hmmm, the above query keeps multiple copies of the same value.
-- Add the DISTINCT keyword to remove all the duplicates
SELECT DISTINCT GovernmentForm
FROM Country ;

-- We can use ORDER BY to order the city population by decreasing population
SELECT Name, (Population / 1000000) AS PopulationInMillion
FROM City
WHERE Population >= '5000000'
ORDER BY PopulationInMillion DESC;

-- Find the top two most populated cities!
-- Here we can use LIMIT to limit the output
SELECT Name, (Population / 1000000) AS PopulationInMillion
FROM City
ORDER BY PopulationInMillion DESC
LIMIT 2;
```

```
-- Which is the first country that became independent?
-- The query below is NOT CORRECT! NULL values are present
SELECT Name, IndepYear
FROM Country
ORDER BY IndepYear ASC
LIMIT 1;

-- Instead the correct way is to filter out the NULL values by using NOT NULL
SELECT Name, IndepYear
FROM Country
WHERE IndepYear NOT NULL
ORDER BY IndepYear ASC
LIMIT 1;

----------------------------------------------------------------
-- Multi-relational queries

-- What are the names of all countries that speak Greek?
-- Here we need to join two relations
SELECT Name
FROM Country, CountryLanguage
WHERE Code = CountryCode AND Language = 'Greek';

-- Better style for the above query
SELECT Country.Name
FROM Country, CountryLanguage
WHERE Country.Code = CountryLanguage.CountryCode
AND CountryLanguage.Language = 'Greek';

-- or even better (prefer this one!)
SELECT C.Name
FROM Country C, CountryLanguage L
WHERE C.Code = L.CountryCode
AND L.Language = 'Greek';

-- Which countries speaks at least 50% Greek?
SELECT C.Name
FROM Country C, CountryLanguage L
WHERE C.Code = L.CountryCode
AND L.Language = 'Greek'
AND L.Percentage >= 50 ;

-- What is the district of the capital of USA?
SELECT T.district
FROM Country C, City T
WHERE C.code = 'USA'
AND C.capital = T.id ;

-- Which countries speak Greek and English?
SELECT C.Name
FROM Country C, CountryLanguage L1, CountryLanguage L2
WHERE C.Code = L1.CountryCode
AND C.Code = L2.CountryCode
AND L1.Language = 'Greek'
AND L2.Language = 'English';

-- Alternatively we could also use the set operator INTERSECT
SELECT C.Name
FROM Country C, CountryLanguage L
WHERE C.Code = L.CountryCode
AND L.Language = 'Greek'
INTERSECT
SELECT C.Name
FROM Country C, CountryLanguage L
WHERE C.Code = L.CountryCode
```

```
AND L.Language = 'English' ;


-- Countries that have population more than 100 million and do not
-- speak English: here we use the set operator EXCEPT
SELECT C.Name
FROM Country C
WHERE C.Population >= 100000000
EXCEPT
SELECT C.Name
FROM Country C, CountryLanguage L
WHERE C.Code = L.CountryCode
AND L.Language = 'English' ;


-------------------------------------------------------
-- Nested queries

-- In which country is Berlin?
SELECT C.Name
FROM Country C
WHERE C.code =
  (SELECT C.CountryCode
   FROM City C
   WHERE C.name = 'Berlin');

-- Find all countries in Europe with population more than 50 million
SELECT C.Name
FROM (SELECT Name, Continent
      FROM Country
      WHERE Population >50000000) AS C
WHERE C.Continent = 'Europe' ;

-- Find all countries in Europe that have *some* city with population
-- more than 5 million
SELECT C.Name
FROM Country C
WHERE C.Continent = 'Europe'
AND C.Code IN (SELECT CountryCode
               FROM City
               WHERE Population > 5000000);

-- alternatively:
SELECT C.Name
FROM Country C
WHERE C.Continent = 'Europe'
AND EXISTS (SELECT *
            FROM City T
            WHERE T.Population > 5000000
            AND T.CountryCode = C.Code) ;

-- alternatively (the query below is NOT supported in SQLite!)
SELECT C.Name
FROM Country C
WHERE C.Continent = 'Europe'
AND 5000000 > ANY (SELECT T.Population
                   FROM City T
                   WHERE T.CountryCode = C.Code) ;

-- Find all countries in Europe that have all cities with population
-- less than 1 million
SELECT C.Name
FROM Country C
WHERE C.Continent = 'Europe'
AND NOT EXISTS (SELECT * FROM City T
```

```
                WHERE T.Population > 1000000
                AND T.CountryCode = C.Code) ;

-- alternatively (the query below is NOT supported by SQLite)
SELECT C.Name
FROM Country C
WHERE C.Continent = 'Europe'
AND 1000000 > ALL (SELECT T.Population
                   FROM City T
                   WHERE T.CountryCode = C.Code) ;



-----------------------------------------------------
-- Aggregate queries

-- Find the average population of countries in Europe
SELECT AVG(Population)
FROM Country
WHERE Continent = 'Europe';

-- Or max population:
SELECT MAX(Population)
FROM Country
WHERE Continent = 'Europe';

-- How many countries are in Europe?
SELECT COUNT(*)
FROM Country
WHERE Continent = 'Europe';

-- How many languages are spoken in the USA?
SELECT COUNT(*)
FROM CountryLanguage
WHERE CountryCode = 'USA';

-- Count the number of languages (this will NOT give the correct answer!!!)
SELECT COUNT(Language)
FROM CountryLanguage ;

-- Instead use DISTINCT to eliminate duplicates:
SELECT COUNT(DISTINCT Language)
FROM CountryLanguage ;

-- Find the name and population of the country with the max population in Europe!
-- The example below is not correct! Why?
SELECT Name, MAX(Population)
FROM Country
WHERE Continent = 'Europe';

-- Instead we have two alternatives:
-- # 1: Use subqueries to do this!
SELECT Name, Population
FROM Country
WHERE Population =
  (SELECT MAX(Population)
   FROM Country
   WHERE Continent = 'Europe');

-- # 1: Use order by and limit!
SELECT Name, Population
FROM Country
WHERE Continent = 'Europe'
ORDER BY Population DESC
LIMIT 1 ;
```

```
-- Find how many countries have each form of goverment:
SELECT GovernmentForm, COUNT(Code)
FROM Country
GROUP BY GovernmentForm ;

-- Find how many countries speak each language (with percentage > 50%) and
-- output them in decreasing order!
SELECT Language, COUNT(CountryCode) AS N
FROM CountryLanguage
WHERE Percentage >= 50
GROUP BY Language
ORDER BY N DESC ;

-- Find languages that are spoken in at least 3 different countries with
-- percentage at least 50
SELECT Language, COUNT(CountryCode) AS N
FROM CountryLanguage
WHERE Percentage >= 50
GROUP BY Language
HAVING N > 2
ORDER BY N DESC ;

-- The query below is NOT semantically correct!
-- HAVING clause is not correctly applied
SELECT Language, COUNT(CountryCode) AS N
FROM CountryLanguage
GROUP BY Language
HAVING Percentage > 50 ;

-- Output for each country the population most populated city,
-- for countries with at least 10 cities!
SELECT C.NAME AS Country, MAX(T.Population) AS N
FROM City T, Country C
WHERE C.Code = T.CountryCode
GROUP BY C.Name
HAVING COUNT(T.ID) > 9
ORDER BY N DESC ;

-- How do we get the name of the most populated city as well?
SELECT Temp.Country, T.Name
FROM (SELECT C.NAME AS Country, C.Code AS Code, MAX(T.Population) AS N
      FROM City T, Country C
      WHERE C.Code = T.CountryCode
      GROUP BY C.Name, C.Code
      HAVING COUNT(T.ID) > 9) AS Temp, City T
WHERE T.Population = Temp.N
    AND T.CountryCode = Temp.Code;

---------------------------------------------------------------
-- NULL behavior

SELECT COUNT(*)
FROM Country
WHERE IndepYear > 1990 OR IndepYear <= 1990 ;

-- Some countries are missing! What can we do?
SELECT COUNT(*)
FROM Country
WHERE IndepYear > 1990 OR IndepYear <= 1990 OR IndepYear IS NULL;


-- Returns max population of a city for each country, and for countries
-- without any city returns NULL
SELECT C.Name AS Country, MAX(T.Population) AS N
```

```
FROM Country C LEFT OUTER JOIN City T ON C.Code = T.CountryCode
GROUP BY C.Name
ORDER BY N DESC ;

-------------------------------------------------
-- DB Modifications

-- Inserting a new tuple!
INSERT INTO CountryLanguage VALUES('USA', 'C++', 'F', 0.5);

-- Let's delete it now.
DELETE FROM CountryLanguage WHERE LAnguage = 'C++';

-- Let's update something
UPDATE CountryLanguage
SET IsOfficial = 'T'
WHERE CountryCode = 'USA' AND Language = 'Spanish';

-------------------------------------------------
-- Views

-- Let's create a view that stores the name, and the official languages for each country
CREATE VIEW OfficialCountryLanguage AS
SELECT C.Name AS CountryName, L.Language AS Language
FROM CountryLanguage L, Country C
WHERE L.CountryCode = C.Code
AND L.IsOfficial = 'T' ;

-- Now we can use this view! Find the countries with more than one offficial language!
SELECT CountryName, COUNT(Language)
FROM OfficialCountryLanguage
GROUP BY CountryName
HAVING COUNT(Language) > 1 ;
```