INST 123 – Databases for All

Assignment 2 - Questions

(See ELMS for the assignment deadline)

There are two questions in this assignment. Be sure to complete both questions.

You will use the ***us_counties_2010*** table for this assignment. If you haven't created and populated the ***us_counties_2010*** table on your server yet, use the CREATE TABLE query and the csv data file available at: https://github.com/anthonydb/practical-sql/tree/master/Chapter_04. Contact a TA or AMP, if you encounter any issues creating or populating the ***us_counties_2010*** table.

Q.1) (20 points) Write a query that replicates the result set below, matching all data, including row order. HINT: You may need to transform some data types to achieve the correct results.

NOTE: Focus on counties in these three states: California, Florida, and Pennsylvania

NOTE: Column name "p0030001" refers to Total Adult Population

	County character varying (90)	State character varying (2)	Latitude numeric (10,7)	Longitude numeric (10,7)	Area (Water) bigint	Population (Adults) integer	pct_adult numeric
1	San Francisco County	CA	37.7272391	-123.0322294	479190317	697711	86.65
2	Amador County	CA	38.4435501	-120.6538563	29456575	31698	83.22
3	Sierra County	CA	39.5769252	-120.5219926	23299108	2688	82.96
4	Tuolumne County	CA	38.0214510	-119.9647084	138683578	45683	82.51
5	Mariposa County	CA	37.5700335	-119.9128599	36271400	15009	82.24
6	Plumas County	CA	39.9951699	-120.8295155	156395470	16406	82.00
7	Lassen County	CA	40.7210889	-120.6299314	463440721	28602	81.97
8	Trinity County	CA	40.6477241	-123.1144043	73407861	11266	81.72
9	San Luis Obispo County	CA	35.3852268	-120.4475404	820974646	218796	81.14
10	Nevada County	CA	39.2951907	-120.7734456	41514942	79658	80.65
11	Calaveras County	CA	38.1878437	-120.5551154	43810420	36635	80.38
12	Humboldt County	CA	40.7066731	-123.9258181	1254256396	107562	79.90
13	Marin County	CA	38.0518169	-122.7459738	797420433	200195	79.31
14	Siskiyou County	CA	41.5879861	-122.5332868	179889482	35575	79.23
15	Mono County	CA	37.9158363	-118.8751668	214695487	11223	79.02
16	Butte County	CA	39.6659588	-121.6019188	105325758	173832	79.01
17	Inyo County	CA	36.5619770	-117.4039269	119208057	14646	78.97
18	Santa Cruz County	CA	37.0124883	-122.0072050	419568865	206964	78.88
19	Lake County	CA	39.0948019	-122.7467569	188962335	50993	78.86
20	Del Norte County	CA	41.7499033	-123.9809983	578525663	22472	78.55
21	Alpine County	CA	38.6176096	-119.7989986	12557251	919	78.21
22	Modoc County	CA	41.5929185	-120.7183704	739774453	7562	78.07
23	Sonoma County	CA	38.5325740	-122.9451943	497530353	377407	78.00

There are 192 result rows returned.

Q.2) (30 points) The solution to this question consists of 3 steps. **Paste in your submission file all three queries you use for this question, as well as the csv file you create in Step 1.**

1. Export data from the us_counties_2010 table for DC, Maryland, and Virginia to create a CSV file named ***dmv_states.csv***. The columns to be exported are as follows. Paste the COPY query into the submission file (*yourlastname_A2_queries.txt*), and include the ***dmv_states.csv*** file in your submission.
    a. geo_name
    b. state_us_abbreviation
    c. internal_point_lat
    d. internal_point_lon
    e. housing_unit_count_100_percent as `numeric(12,2)`
    f. population_count_100_percent as `numeric(12,2)`

2. Create the ***dmv_states*** table in PgAdmin with the following columns and data types. Paste the CREATE query into the submission file (*yourlastname_A2_queries.txt*).
    a. county `varchar(90)`
    b. state `varchar(2)`
    c. latitude `numeric(10,7)`
    d. longitude `numeric(10,7)`
    e. household `numeric(12,2)`
    f. population `numeric(12,2)`

3. Import the data from the ***dmv_states.csv*** file you created in Step 1, into the ***dmv_states*** table you created in Step 2. Paste the COPY query into the submission file (*yourlastname_A2_queries.txt*).

The ***dmv_states*** table should look like this when you have successfully completed the 3 steps.

	county character varying (90)	state character varying (2)	latitude numeric (10,7)	longitude numeric (10,7)	household numeric (12,2)	population numeric (12,2)
1	District of Columbia	DC	38.9041485	-77.0170942	296719.00	601723.00
2	Allegany County	MD	39.6123087	-78.7031084	33311.00	75087.00
3	Anne Arundel County	MD	38.9933743	-76.5605111	212562.00	537656.00
4	Baltimore County	MD	39.4431666	-76.6165693	335622.00	805029.00
5	Calvert County	MD	38.5213575	-76.5258639	33780.00	88737.00
6	Caroline County	MD	38.8715312	-75.8316311	13482.00	33066.00
7	Carroll County	MD	39.5631892	-77.0155121	62406.00	167134.00
8	Cecil County	MD	39.5623524	-75.9415843	41103.00	101108.00
9	Charles County	MD	38.4728532	-77.0154272	54963.00	146551.00
10	Dorchester County	MD	38.4291957	-76.0474333	16554.00	32618.00
11	Frederick County	MD	39.4704273	-77.3976265	90136.00	233385.00
12	Garrett County	MD	39.5298710	-79.2694163	18854.00	30097.00
13	Harford County	MD	39.5374292	-76.2997894	95554.00	244826.00
14	Howard County	MD	39.2522639	-76.9244057	109282.00	287085.00
15	Kent County	MD	39.2391765	-76.1242003	10549.00	20197.00
16	Montgomery County	MD	39.1373815	-77.2030633	375905.00	971777.00
17	Prince George's County	MD	38.8258796	-76.8472724	328182.00	863420.00
18	Queen Anne's County	MD	39.0406929	-76.0824053	20140.00	47798.00
19	St. Mary's County	MD	38.2226660	-76.5342705	41282.00	105151.00
20	Somerset County	MD	38.0744501	-75.8533228	11130.00	26470.00
21	Talbot County	MD	38.7494273	-76.1791304	19577.00	37782.00
22	Washington County	MD	39.6036207	-77.8146709	60814.00	147430.00
23	Wicomico County	MD	38.3673888	-75.6322064	41192.00	98733.00
24	Worcester County	MD	38.2221332	-75.3099315	55749.00	51454.00

There should be 159 rows in the ***dmv_states*** table.