Homework #3 Help
Commuter Statistics Problem

Background Information

Every day, millions of workers commute across county lines for their jobs. Commuting patterns vary widely across the United States but are influenced by a number of factors including community development patterns, labor market shifts, and technological changes. Information on commutes is tracked annually by the U.S. Census Bureau as part of its American Community Survey.

Problem Statement

In this assignment, students will analyze commuter statistics for counties across the United States from 2006 to 2010.

Instructions

IMPORTANT: This is not the actual Homework for your section. You will not receive any credit for completing this project.

IMPORTANT: This assignment requires the Windows version of Microsoft 365. macOS users can access a ready-to-use version through Windows Virtual Desktop by following the instructions at https://cs101.wvu.edu/wvd.

IMPORTANT: Complete the steps below in the order they are given. Completing the steps out of order may complicate the assignment or result in an incorrect result.

1. Download and extract the provided Data Files ZIP file. It contains the following files for use in this assignment:

   a. **states.xml** – Information on states.

Table: *States*		
Field Name	**Type**	**Description**
StateAbbrv	Short Text	Primary key. Abbreviation for the state.
StateName	Short Text	Full name of the state.
RegionAbbrv	Short Text	Bureau of Economic Analysis region where this state is located.

b. **counties.xml** – Information on the population and workers for each county in the United States [1].

Table: *Counties*		
Field Name	**Type**	**Description**
StateAbbrv	Short Text	Part of composite key. Abbreviation for the state.
CountyName	Short Text	Part of composite key. Name of the county.
Residents	Number	Total people living in the county.
EmployedResidents	Number	People living in the county who are employed, whether in the county or somewhere else.
PeopleWorkingInCounty	Number	Number of people working in the county, whether they live there or somewhere else.
ResidentsWorkingInCounty	Number	Number of residents of the county who work in the county.

2. Create a new Microsoft Access database named **hwhelp3_csp_*lastname_firstname*.accdb**.

3. Import the following items into the database:

   a. **states.xml** file – Import structure and data into a new table.

   b. **counties.xml** file – Import structure and data into a new table.

4. We need to create a table to store data on the Bureau of Economic Analysis (BEA) regions.

   a. Create a table named *Regions* with the below fields and characteristics. Use appropriate field types and designate a good primary key.

Table: *Regions*	
Field Name	**Description**
RegionAbbrv	Abbreviation for the BEA region.
RegionName	Full name of the BEA region.

   b. Enter records for all regions below [2].

   HINT: The *Regions* table will contain 8 records.

RegionAbbrv	RegionName
FW	Far West
GL	Great Lakes
ME	Mideast (Mid-Atlantic)
NE	New England
PL	Plains
RM	Rocky Mountains
SE	Southeast
SW	Southwest

5. Create the appropriate relationships for the following tables. Do not enforce referential integrity.

    a. *Regions* and *States*

    b. *States* and *Counties*

6. Create separate queries to provide the information requested below. Name each query after the step in which it appears (e.g., name the query in Step 6a as *Query6A*).

    HINT: Run your queries to test them. Make sure that they display all and only the records that you would expect to appear.

    a. Create a query to display information on each county. List the county name, state name, region name, and number of residents.

       Sort by county name in ascending order.

       HINT: This query will show 3,143 records and 4 fields.

    b. We wish to explore bedroom counties where many of their residents travel outside of the county for work. Create a new query listing the state name, county name, number of employed residents, and number of people working in the county.

       Only display records where the number of people working in the county is less than the number of employed residents.

       Sort by state name and then by county name, both in ascending order.

       HINT: This query will show 2,177 records and 4 fields.

    c. Create a query to display information on small Plains counties. List the county name, state name, region name, number of residents, number of employed residents, and number of people working in the county.

       Only display records for the Plains region where the population was less than 10,000 (< 10000).

       HINT: This query will show 287 records and 6 fields.

d. We wish to calculate percentage of residents who are employed for each county. Create a query listing the state name, county name, number of residents, and number of employed residents. Also, include a field to calculate the percentage of residents who are employed.

You can calculate the percentage of residents who are employed using the formula:

$$\frac{[Counties.EmployedResidents]}{[Counties.Residents]}$$

Format the calculated field as a percentage with 1 decimal place.

HINT: This query will show 3,143 records and 5 fields.

e. Create a query to find counties where a majority of workers are in-commuters who reside in another county. List the county name, state name, number of people working in the county, and number of residents working in the county. Also, include a field to calculate the percentage of people working in the county who were in-commuters.

You can calculate the percentage of in-commuters using the formula:

$$\frac{([Counties.PeopleWorkingInCounty] - [Counties.ResidentsWorkingInCounty])}{[Counties.PeopleWorkingInCounty]}$$

Only display records where at least 50% (≥ 0.5) of people working in the county are in-commuters.

Format the calculated field as a percentage with 1 decimal place.

HINT: This query will show 140 records and 5 fields.

f. We wish to calculate the number of people residing and working in each state. Create a query listing, for each state name, the sum of residents, the sum of employed residents, and the sum of people working in its counties.

Format the sums as standard-type numbers with no decimal places.

HINT: This query will show 51 records and 4 fields.

g. We want to explore the daytime population of large counties. Copy-and-paste this SQL code into a new query:

```
SELECT States.StateName, Counties.CountyName,
Counties.Residents,
Counties![Residents]+Counties![PeopleWorkingInCounty]-
Counties![EmployedResidents] AS DaytimePopulation,
Counties![PeopleWorkingInCounty]-Counties![EmployedResidents]
AS DaytimePopulationChange
FROM States INNER JOIN Counties ON States.StateAbbrv =
Counties.StateAbbrv
WHERE (Counties.Residents >=750000)
ORDER BY States.StateName, Counties.CountyName;
```

IMPORTANT: Do not make any modifications to this query other than entering the above SQL code.

HINT: This query will show 73 records and 5 fields.

7. We need to create a new table to store analysis questions responses.

a. Create a table named *AnalysisQuestions* with the fields below. Use appropriate field types and designate a good primary key.

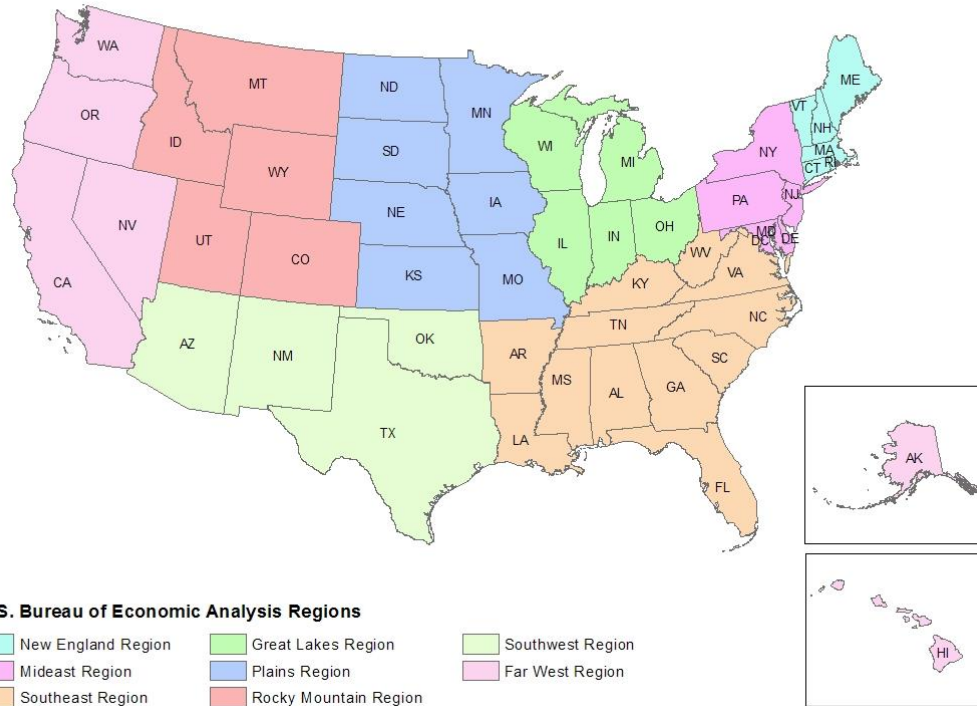| Table: *AnalysisQuestions*	
Field Name	**Description**
QuestionNumber	Question being answered.
Response	Response to the analysis question prompt.

8. In the *AnalysisQuestions* table, answer three of the five analysis questions below. Respond to one question per record.

a. What might the percentage of residents who are employed in a state say about its economy?

b. Considering that commuting costs time and money, why might someone choose to live in a different county than where they work? Name at least 2 reasons.

c. In Virginia, incorporated cities are not considered to be part of any county even if they are the seat of a county's government. A number of Virginia's cities have extremely high in-commuter rates. Why might this be the case?

d. The number of people working and living in the same county varies from 29.6% in the Mideast (Mid-Atlantic) region to 37.7% in the Far West region. What might explain the difference in commuting patterns between the two regions?

U.S. Bureau of Economic Analysis Regions

New England Region	Great Lakes Region	Southwest Region
Mideast Region	Plains Region	Far West Region
Southeast Region	Rocky Mountain Region	

e. Look at the daytime population changes for the largest counties. Are there any patterns you can discern about which counties have negative versus positive changes? Why might this be the case?

9. Run the Compact and Repair Database utility on your database. Ignore any errors you receive when running the utility.

Grading Rubric

This assignment is worth 60 points. It will be graded by your instructor using this rubric, with partial credit awarded as appropriate:

Step 3	3 points	Steps 6a-g	5 points each
Steps 4a-b	5 points total	Step 7a	4 points
Step 5	4 points	Steps 8a-e (pick 3 of 5)	3 points each

Homework #3 Help
Commuter Statistics Problem

The analysis questions in Steps 8a-e will be evaluated using this rubric:

Standard	Meets Requirements (1.5 points)	Does Not Meet Requirements (0 points)
Answer is reasonable.	Answer addresses the question prompt and is factually correct or a reasonable interpretation of available data.	Answer does not address the question prompt, is factually incorrect, or is an unreasonable interpretation of available data.
Answer is supported.	Logical rationale is provided to support the given answer.	Logical rationale is not provided to support the given answer.

Acknowledgments

The image in the introduction appears courtesy of Steve Way [3].

The Bureau of Economic Analysis regions map is courtesy of Iowa State University [2].

References

[1] "Commuter Adjusted Daytime Population: 2006-2010 5-year ACS," U.S. Census Bureau, Jun. 2014. Available: *http://www.census.gov/hhes/commuting/data/acs2006_2010.html*.
[2] Iowa State University, *U.S. Bureau of Economic Analysis Regions*. 2014. Available: *http://www.icip.iastate.edu/sites/default/files/uploads/images/region_maps/BEA_regions.jpg*.
[3] S. Way, *Commuter Quadrille 2*. 2005. Available: *http://www.flickr.com/photos/steve_way/29300038/*.