Sargam Shah

1001275800 | 4/10/2016

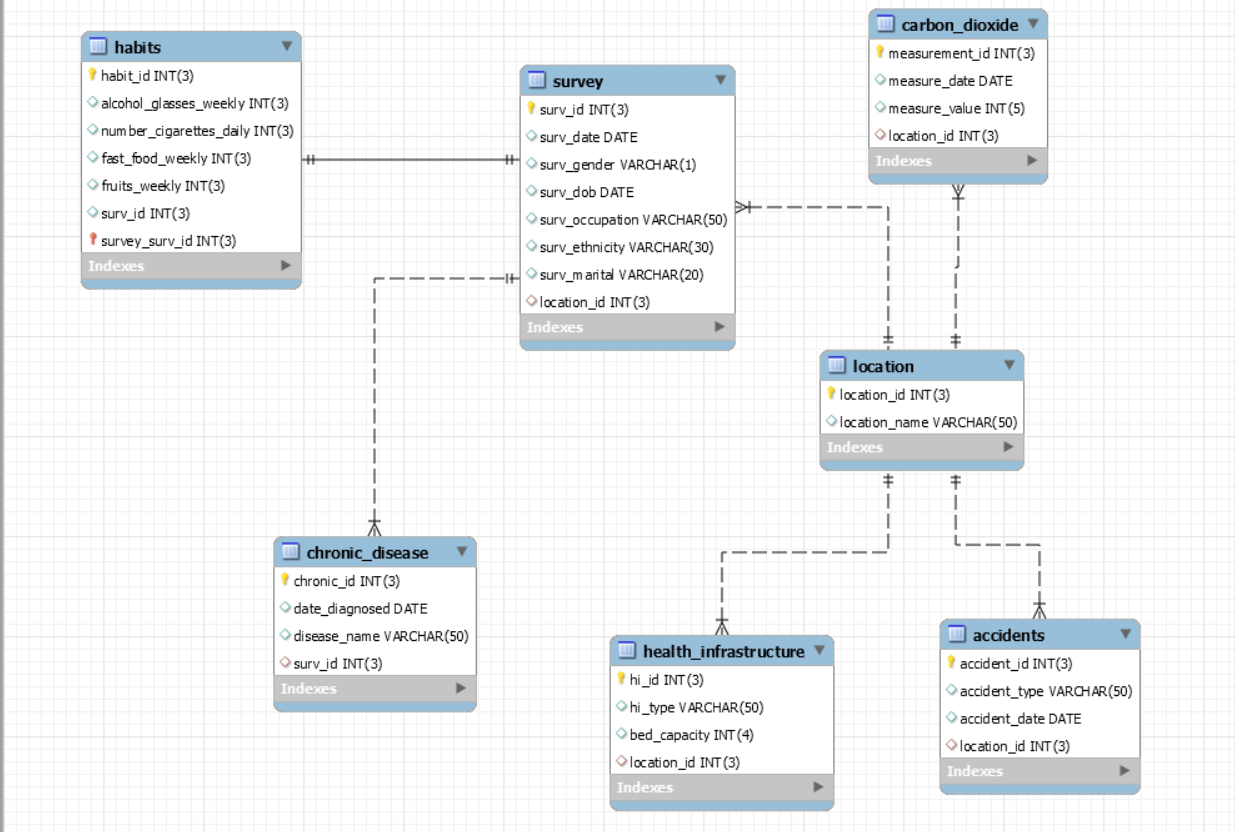
CSE 5339

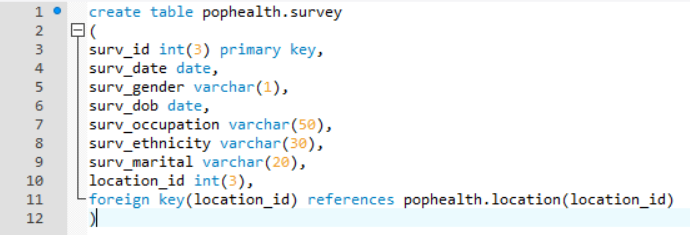
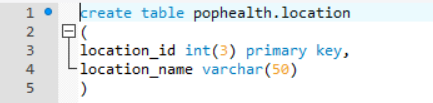
homework 3

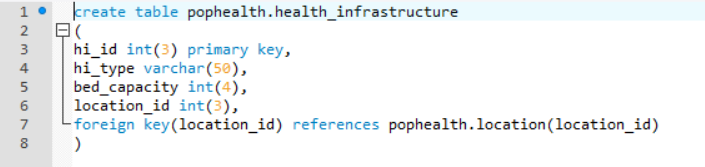
Homework 3

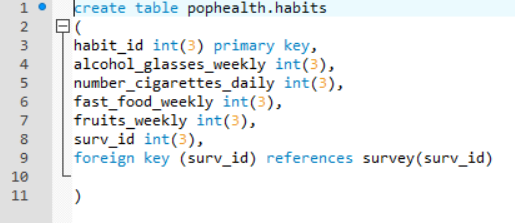
Task 1 DATABASE DEVELOPMENT

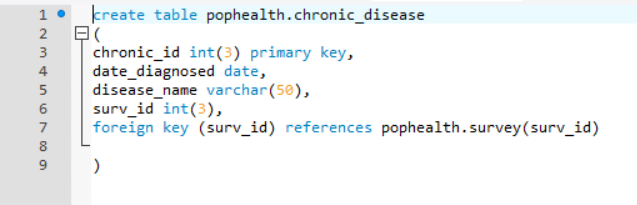


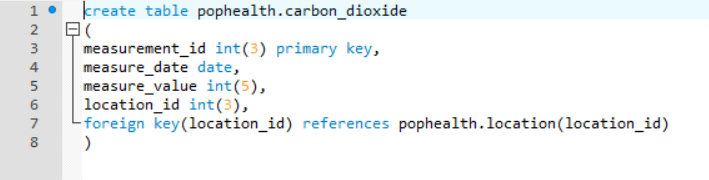


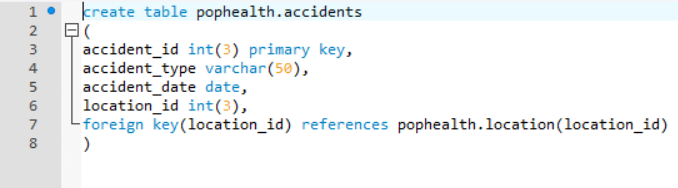
 











Task 2 INSERTION

The data has been inserted as per the instructions.

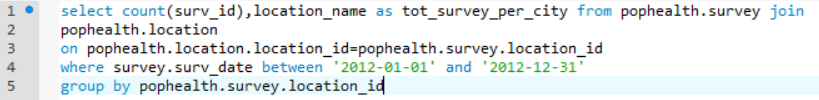
Task 3 HEALTH SURVELLANCE

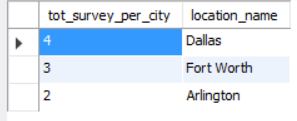
**1. Total number of surveys performed all time.**



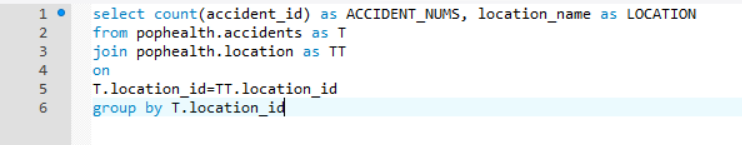


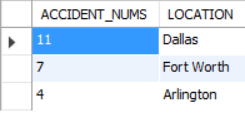
**2. Surveyed population of each city during the year 2012**



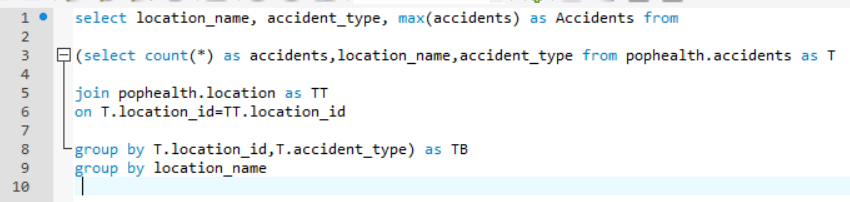


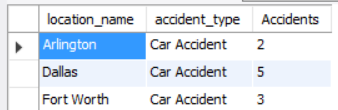
**3. Number of accidents in each city**



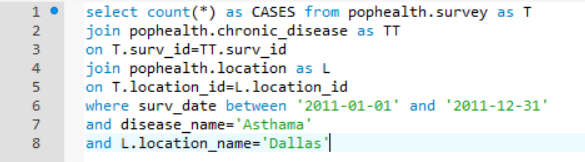


**4. Most common accident type for each city**



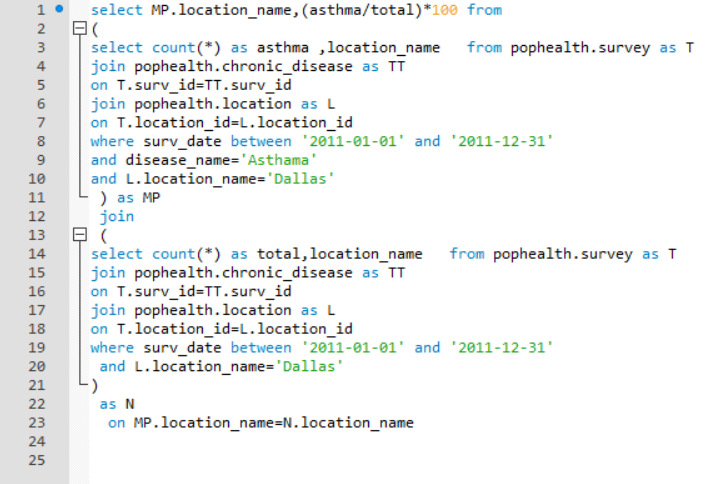


**5. Cases of asthma in Dallas for the year 2011**





**6. Prevalence of asthma in Dallas for 2011**





Prevalence of Asthma in year 2011 in the city of Dallas = (1/4) \* 100 = 25%..

**7. City with the lowest asthma prevalence for the year 2012**

This can also be made into a SINGLE QUERY as follows:

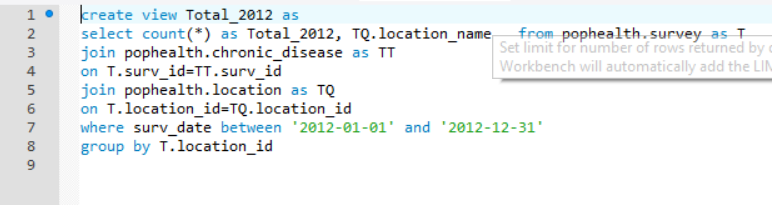


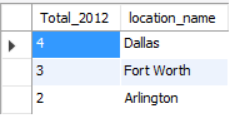


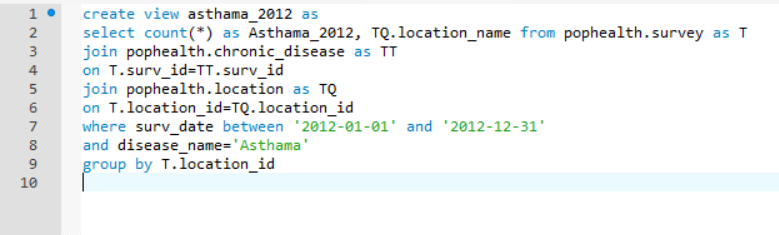
OUTPUT:

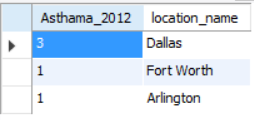


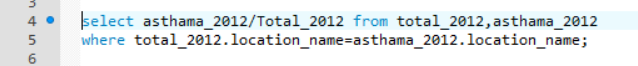
Step by step query with explanation is as follows:

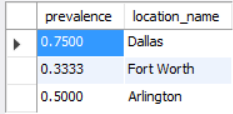


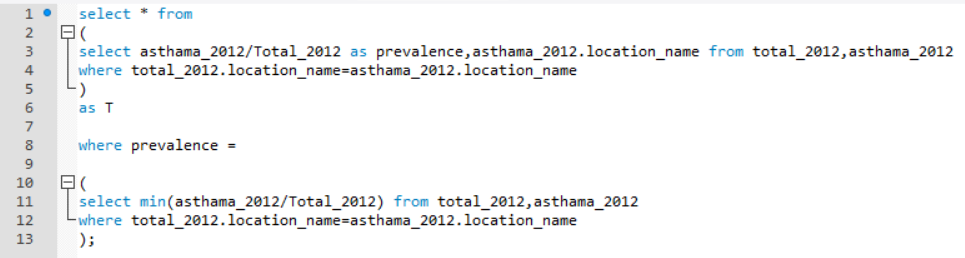


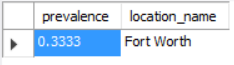




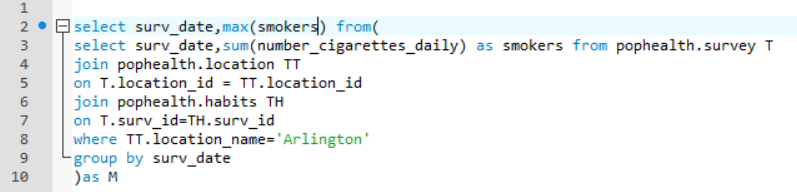








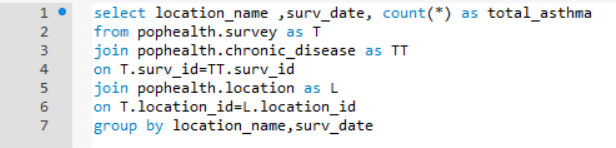
**8. Year with the highest smoking ratio for Arlington**

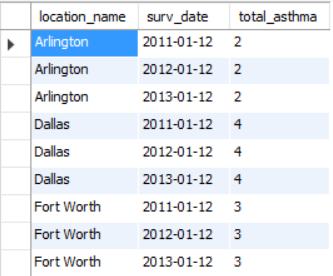


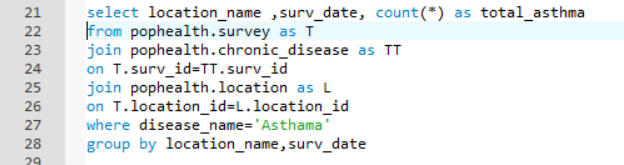


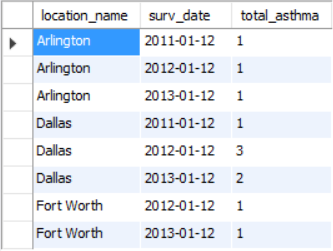
**9. Yearly change in asthma prevalence for the three consecutive years (2011-2012**

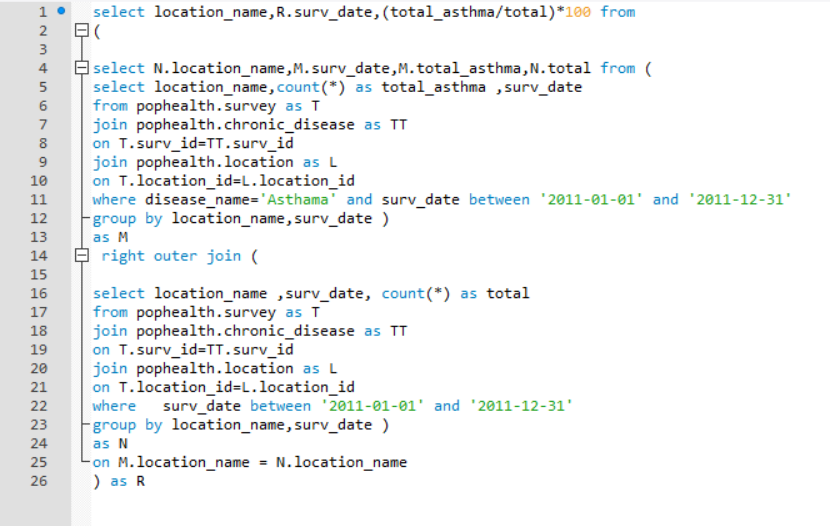
and 2012-2013) in each city



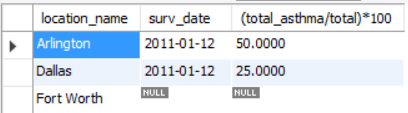




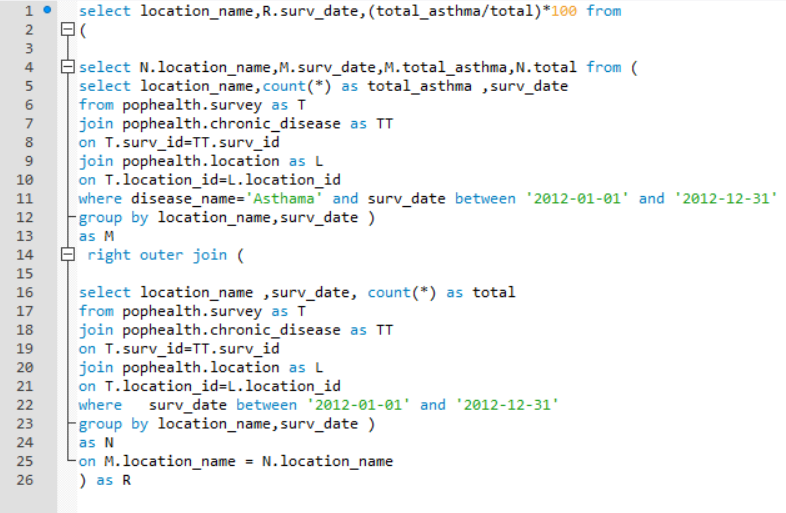




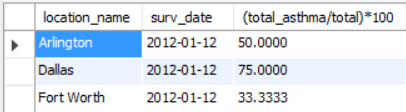
Asthma prevalence in the year 2011:

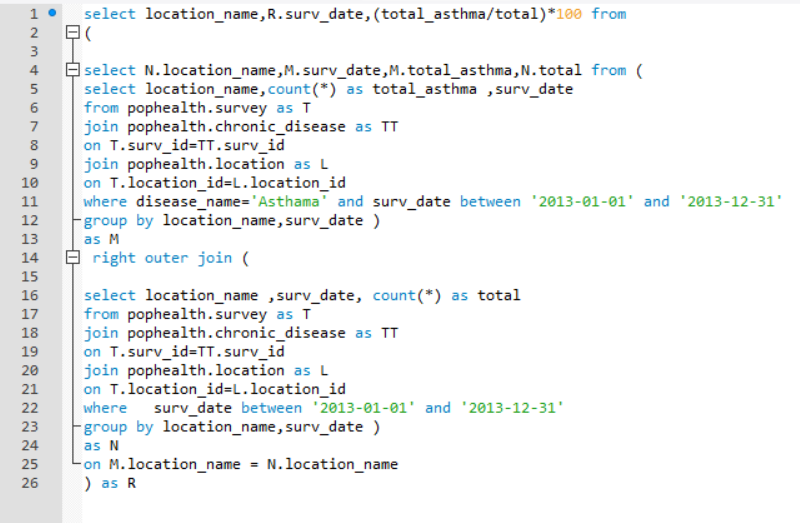


Note for Fort Worth it is 0.

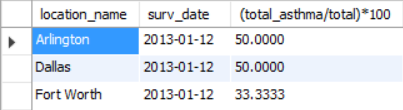


Asthma prevalence for the year 2012:



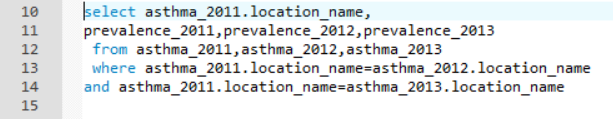


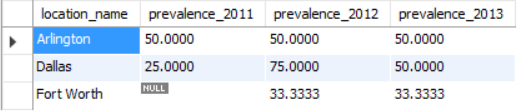
Asthma prevalence for the year 2013:

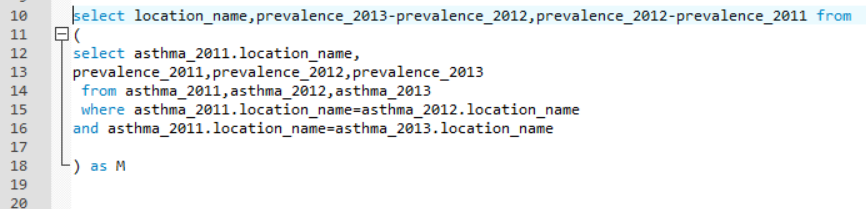


Three views for 2011,2012,2012 were created for the above three queries.

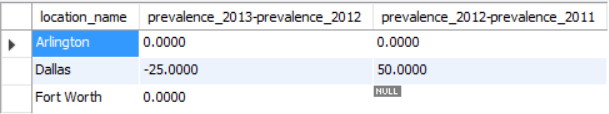
The tables are combined into one as follows







This table gives the difference in the percentages of the prevalence.

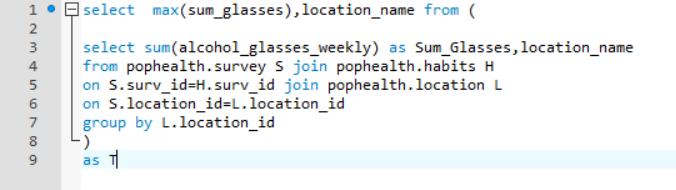


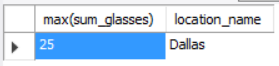
It can be seen in the table, for Arlington there is no change in prevalence percentage over the years.

For dallas, it first increases and then decreases.

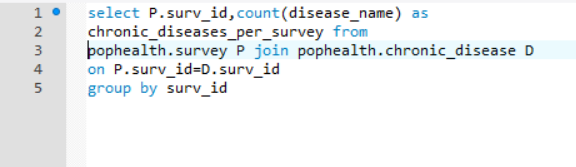
For Fort Worth, it first rises from 0 case of Asthma to 33.33 percent and then no change in prevalence for the next year.

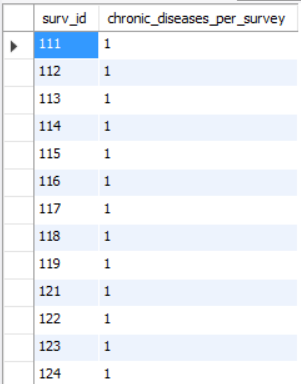
**10. Most “alcoholic” city**





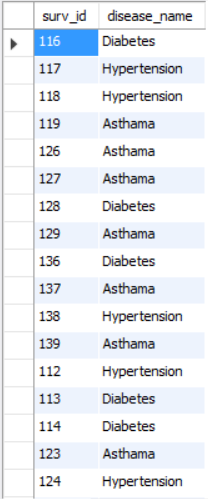
**11. The chronic diseases reported for each survey**



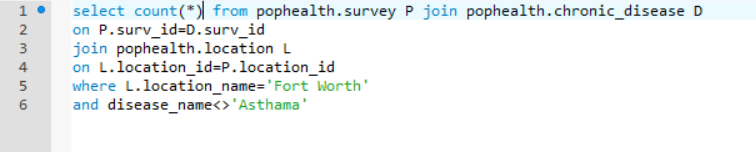


**12. All the survey id’s and their reported chronic diseases**



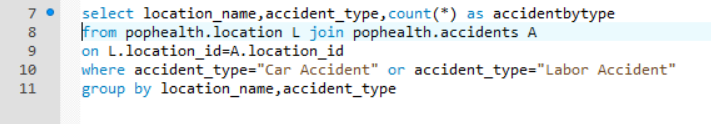


**13. The number of surveys occurred in Fort Worth which had no reporting of asthma**

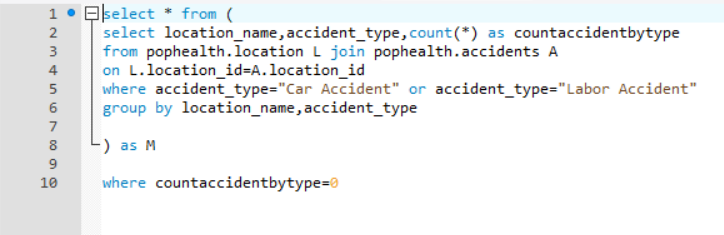


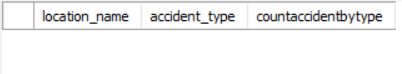


**14. The city(ies) where no accidents have been reported**





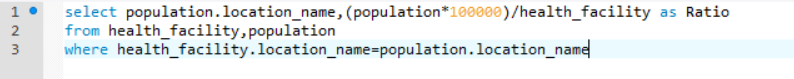


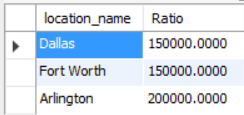


This query gives the city where no accidents have been reported which is none in our scenario.

**15. The availability of health resources for Waco for the whole observation period.**

This is the ratio of citizens per hospital.

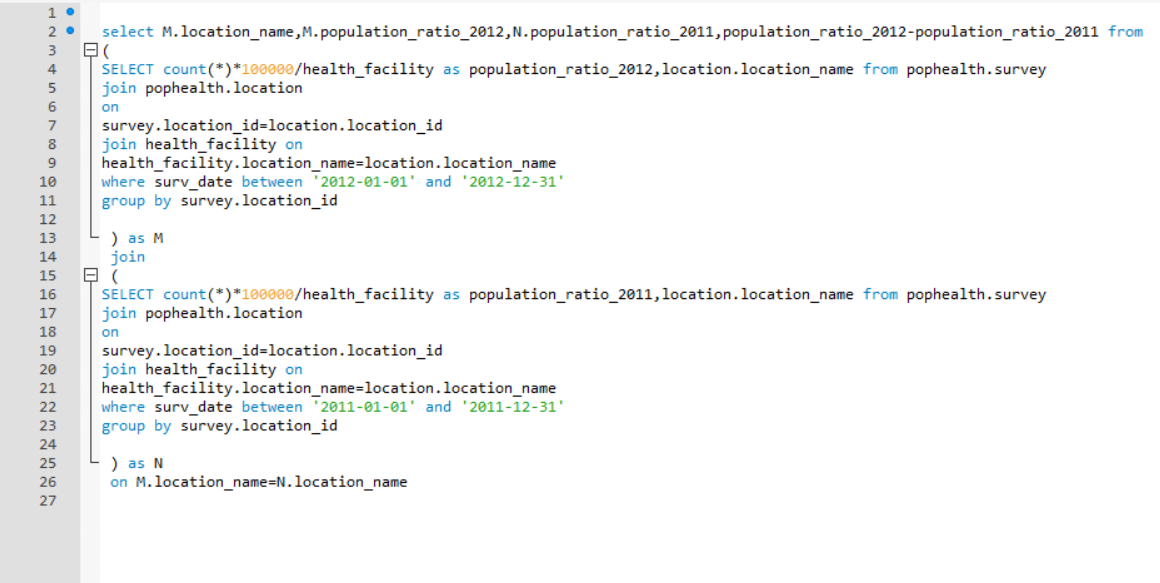




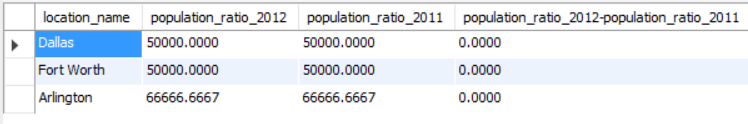
**16. Has there been any change to the above indicator in the year 2012 compared to**

**2011? Show your answer with an appropriate query.**

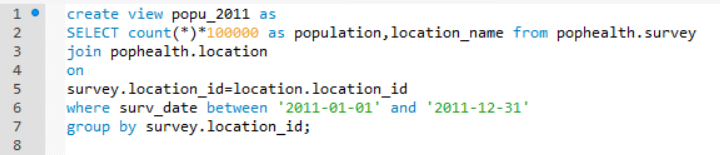
The single query is as follows:



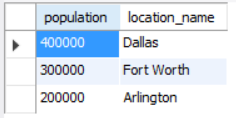
Output:

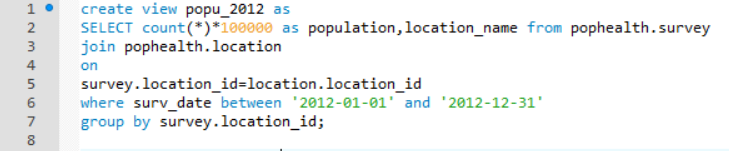


Step by step explanation:

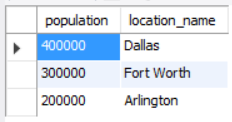


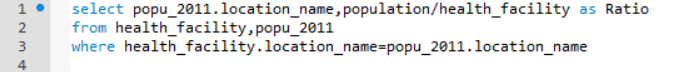


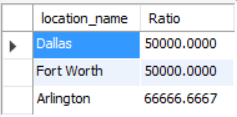


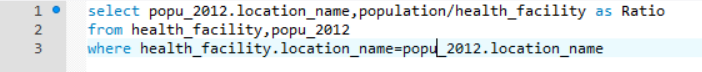


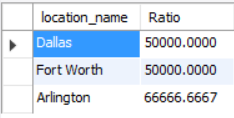












For the questions 15 and 16 we assume that the population of each city is equal to the number

of surveys multiplied by 100,000, since the survey method is supposed to sample a very small

number of each city’s populations.

Task 4

1. Distributed databases are suitable for healthcare. Can you briefly explain why? 2. How can (i) health professionals (ii) researchers (iii) public health officials, benefit from big data?

Answer:

Data in healthcare comes from various sources and also the data keeps on updating, such as latest blood RBC counts etc. Thus, consistency of the latest information needs to be maintained in order to carry out proper medication of the patients. The sources of data could be the lab which provides the test reports, observation and medication records generated by the nurses, diagnosis reports by the doctors, patient survey and feedbacks. Also, various hospitals share their treatment database with other hospitals for their betterment and feedback. This explains the spatial significance of databases in healthcare. Also, the data collected depends on the time right from the patient enters the hospital wherein patient demographics are collected, diagnosis, treatment followed by prognosis. Thus, database in healthcare is also of temporal significance.

Owing to this, distributed databases are the most suitable for healthcare as it allows for the data to be collected at different physical locations. Also, distributed databases allows display of the same data but as per the requirements such as financial department might be interested in knowing the costs, managers might be interested in knowing mortality ratios, doctor might be knowing in number of admitted cases of a particular disease. If it is a distributed database, then all such data presentations could be created seamlessly at the same time. Thus distributed database enables multiple use of the same database.

Benefit from big data:

1. Health Professionals

Big data helps do proper diagnosis for the patients by the doctors. The doctors mine similar cases admitted to the hospital earlier and then do their analysis accordingly. The management can be benefited using big data by improving the quality of service and reducing the costs.

1. Research

Big data analysis could help researchers find out new ailments and medications for the diseases. They could find out the links between a disease and a certain habit. For example, those who smoke 4 cigarettes a day have 40 % chances of getting Lung cancer.

1. Public health officials

Big data could help the officials carry out public health surveillance. They could monitor the health of the people in certain location. By big data analysis, they could also conclude that there are more number of cases of a certain disease in a particular area and thus build specialty hospitals in that area.