A document with writing on it

Description automatically generated

1)

A)

B)

C)

D)

E)

2)

A)

B)

C)

D)

E)

3)

A)

This expression runs in O(r), evaluating each tuple in R only once. It is optimal as each tuple in a select operation needs to be read, evaluated, and possibly written at least once, and this query evaluate each of the 4 filter functions on the same tuple without needs to iterate through the data more than once.

B)

This condition ensures that subsequent project operations do not attempt to store previously eliminated attributes. The expression could be optimized by eliminating the L2 and L3 projections, as either way, we still will only end up with the L1 projection.

C) Possibly but not always. In order for to be equivalent to , all of the attributes that are being used in F must be present in A. This is not the case for .

D) This is not a valid operation. For difference operations, R1 and R2 must be schema compliant.

4)

A) Minimum 0, occurs where no tuples are similar between R and T (R intersect T yields null set). Maximum is the smaller of r and t (min(r, t)), occurs when one (or both) of r or t is a subset of the other.

B) Minimum 0, occurs where no tuples are similar between R and T (R intersect T yields null set). Maximum is the smaller of r and t (min(r, t)), occurs when one (or both) of r or t is a subset of the other.

C) The minimum is max(r,s), occurs when one of R or S is a subset of the other. The maximum is r + s, occurs when R and S are disjoint (R union S yields null set).

D) Minimum 0, occurs where no tuples are similar between R and T (R intersect T yields null set). Maximum r, occurs when all tuples in R have a match in T.

5)

Task 1

A screenshot of a computer program

Description automatically generated

A screenshot of a computer

Description automatically generated

Task 3

I utilized a python script to transform the data as specified. Code is shown here:

A screen shot of a computer program

Description automatically generated

Starting from the top, I first open the write file, and before anything else, add the command “SET DEFINE OFF”. This command prevents SQL Developer from asking for an input variable whenever a “&” is discovered in a string. Next the read file is opened. Starting from the first line of data, each line is read, any random quotation marks are removed, single quotes are escaped with double single quotes (to prevent the program from thinking the string has ended prematurely), ‘\n’’s are stripped from the end of the line, and finally split into a list by commas. Then the script checks whether the address block had any commas which caused an undesired split, and if this occurred, reconsolidates the address data. Then, the data from non-desired columns is eliminated. After that, the date is converted to the desired SQL format using f-strings and single quotes are added around the remaining string data. Finally, the list of data is recombined with commas, wrapped with the necessary SQL INSERT commands, and written into the desired file.

Task 4

A screenshot of a computer

Description automatically generated

Task 5

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated