

AHSANULLAH UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Course No: CSE4125

Course Title: Distributed Database Systems

Fall 2019 | Quiz – 4 | Marks 20 | Time: 60 Minutes (50+10)

1. Four fragments R, S, M and T of a relation are given. We want to perform the following query Q. Where, $Q = (PJ_a(R \cup S)) \Join_{a=a} (SL_{m=value}(M \cup T))$

Database profiles are provided below.

Card (R) = 300, Site (R) = 1

	a	b	c	d
Size	6	7	2	10
Val	300	1000	30	50

Site (S) = 4

	a	b	c	d
Size	6	7	2	10
Val	100	10	20	15

Card (T) = 2000, Site (T) = 3

	a	m	n
Size	6	5	4
Val	2000	5	5

Card (M) = 2000, Site (M) = 2

	a	m	n
Size	6	5	4
Val	2000	5	5

Assume that, the result of $(R \cup S)$ has no duplicate values in attribute **a**, and the same property stands for $(M \cup T)$.

Now answer the following questions.

- a. If attribute **a** is the primary key of **S**, then **Card (S) = ?** 1
- b. **size(R SJ_{a=a} M) = ?** 2
- c. For the simple selection **SL_{m=value}(M UN T)**, estimate the selectivity **ρ** 2
- d. Estimate the cardinality of the result of **Q**. Indicate the formulas applied. 3
- e. Estimate the total size of data in the result of **Q**. Indicate the formulas applied. 2
- f. Suppose, we want to apply **strategy – 2** on query **Q** at site **(last 3 digits of your ID % 4) + 1**. Determine the transmission delay for the strategy when network wide transfer rate is 10000 bits/second. Note that, the system provides benefits of parallel processing and the initial delay is 0. 4
- g. Write down the steps to perform the join using semi-join program of query **Q** at site **(last 3 digits of your ID % 4) + 1**. [only steps, no figure, no calculations] 3
- h. Write a reducer program for the query **Q** to optimize the corresponding operator tree. Draw the obtained optimization graph. 3