National University of Computer and Emerging Sciences, Lahore Campus

Course:	Data Warehousing and Data Mining	Course Code:	CS409
Program:	BS(Computer Science)	Semester:	Fall 2018
Out Date:	23-Oct-2018	Total Marks:	
Due Date:	Tue 30-Oct-2018 (start of class)	Weight:	
Section	CS	Page(s):	1
Assignment:	3 (Joining Techniques)		

Instructions: You will have to take assumptions for block size (B), record size (R), number of records (r), available memory (K), indexing column, index type (clustered/non-clustered), and all remaining factors required in calculations.

Question. Consider two tables A and B which have to be joined. Calculate the cost of joining the two tables on their common attribute. You will have to consider the following cases and have to calculate cost (i.e. I/O cost) for all these cases:

1. When the table sizes are almost similar. Let's say 1 row of table A joins with only 1 row of table B.

a. When available main memory (K) is not sufficient to store any operand table.
b. When available main memory (K) is sufficient to store at least one operand table.

2. When the sizes of both tables are significantly different. Let's say 1 row of table A joins with 10 rows of table B.

a. When available main memory (K) is not sufficient to store any operand table.
b. When available main memory (K) is sufficient to store at least one operand table.

Your solutions should include costs of I/Os for:

- **Naïve nested loop join (Block NLJ/Index NLJ/clustered indexed NLJ)**
- **(Sort) merge join**
- **Hash join**