

National University of Computer and Emerging Sciences, Lahore Campus



Course: Data Warehousing & Data Mining
Program: BS(Computer Science)
Duration: 20 Minutes
Paper Date: 20-Oct-16
Section: B
Exam: Quiz (Joining Techniques)

Course Code: CS409
Semester: Fall 2016
Total Marks: 10
Weight
Page(s): 1
Roll No:

Instruction/Notes:

Question:

Consider the following tables and statistics which are part of a student system:

Student (RollNo, Name, DegreeID, BatchID,); **Attendance** (RollNo, CourseCode, Semester, AttFlag,);

Block Size = 16 KB; **Available Memory** = 100 Blocks;

<u>Table Name</u>	<u>Row Count</u>	<u>Row Width</u> (in bytes)	<u>Table Size</u> (in Blocks)
Student →	128,000	256	2,000
Attendance →	1,280,000	256	20,000

Assume average attendance table rows retrieved per student table qualifying row = 10;

Assume: 3% of students from '2013' Batch (i.e. 3840 students).

Query: *SELECT * FROM student JOIN attendance ON student.rollno=attendance.rollno WHERE BatchId = '2013'*

Calculate the total I/O cost (including the I/O cost to filter the condition on student table) for the above Query using the following joining techniques. You are supposed to filter the condition first and then join. Show all steps.

- 1) Hash Join
- 2) Sort Merge Join
- 3) Block Nested Loop Join
- 4) Naïve Nested Loop Join
- 5) Index Nested Loop Join (Assume there is a non-clustered index on RollNo column of Attendance table and three I/Os are required to read index for each qualifying student)