National University of Computer and Emerging Sciences, Lahore Campus

	Course:	Data Warehousing & Data Mining	Course Code:	CS409
	Program:	BS(Computer Science)	Semester:	Fall 2017
	Date:	31-Oct-2017	Total Marks:	10
	Roll No:		Weight:	
	Section:		Page(s):	1
	Quiz:	4 (Indexing Techniques)		

Instruction/Notes:

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

ACCOUNT (accId, title, accType, rating, openingDate, …);

Block Size= 4 KB; Available Memory= 100 Blocks; Rows= 250,000; Row Width= 500 bytes; Index entry size (i.e. RID Width)= 8 bytes. Assume accounts with 'SAVING' accType are 4%, accounts with 'CHECKING' accType are 10%, and accounts with '1' rating are 6%.

Query: SELECT COUNT(*) FROM account WHERE (accType= 'SAVING' OR accType= 'CHECKING') AND Rating= 1

Calculate the I/O cost for the above query using

a) Composite index access (Assume a composite index exist on accType and rating columns)
b) Dynamic Bitmap index access (Assume indexes exist on accType and rating columns separately)
c) Clustered index access (Assume only clustered index exist on accType column)

Ans:
B=4K, bfr= 8, bfr_i= 512, b= 31,250

a) Composite Index:
4% of (6% of 250000)= 600
10% of (6% of 250000)= 1500
Total I/Os (Index access cost only) = 600/512 + 1500/512 = 2 + 3= **5 blocks**

b) Dynamic Bitmap Indexes:
Combined selectivity for combination ('SAVING' or 'CHECKING' and '1') = (4+10)% of (6% of 250000) = **2100 rows**
I/Os to access index for accType ('SAVING' or 'CHECKING') i.e. (4%+10%) = (10,000/512) + (25000/512) = 20+49= **69 blocks**
I/Os to access index for rating ('1') i.e. 6% = 15000/512 = **30 blocks**
Total I/Os (Index access + Base table access) = (30 + 69) + 2100 = **2199 blocks**

c) Clustered Index Access
I/Os to access base table = (10,000/8) + (25,000/8)= 1250 + 3125= **4375 blocks**
I/Os to access index for accType ('SAVING' or 'CHECKING') = (10,000/512) + (25000/512) = 20+49= **69 blocks (for dense index)**
Total I/Os (Index access + Base table access) = 69 + 4375 = **4444 blocks**