National University of Computer and Emerging Sciences, Lahore Campus

	Course:	**Advance Database Concepts**
	Program:	**BS (Computer Science)**
	Instructor:	**Muhammad Ishaq Raza**
	Practice Problems:	**Indexing Structures**

SOLUTION

Topic: Indexing Structures and Physical DB Design

Q1. Consider a file of customer data that consists of 100,000 records, spread over 5000 disk blocks. There is a secondary index on the key attribute, customer Id with 5 levels. There is another cluster index on total amount spent by customer with 4 levels. For each of the following selection queries, estimate the I/O cost of the best possible solution, making use of the access paths available. Justify your answer.

a. SELECT * FROM customer WHERE customerId IN (111, 222, 333, 444);
b. SELECT COUNT(*) FROM customer WHERE customerId IN (111, 222, 333, 444);
c. SELECT * FROM customer WHERE totalAmountSpent >= 50000; *(Assume 2% of the customers have spent 50,000 or more)*

Answer:
a. 4(X+1) = 4(5+1) = **24**
b. 4(X) = 4(5) = **20**
c. x + s/bfr = 4 + 2000/20 = 4 + 100 = 104

Q2. Assume: A block size is B= 2048 bytes, file has r= 100,000 records, each record is 100 bytes long, a block pointer is P= 10 bytes, a record pointer is Pr= 11 bytes, and a key field for the index is 5 bytes long. A database system uses a *B^+-trees* index on key field. A leaf node and non-leaf node are one block in size and contain as many keys (and appropriate pointers) as will fit in a block. How many blocks will this index use? Show your working.

Answer:
order of p: p * 10 + (p-1) * 5 < 2048, which gives us **order p** = (2048 + 5)/15 **= 136**
order of p_{leaf}: p * (5+11) + 10 < 2048, which gives us **order p** = (2048 - 10)/16 **= 127**

This means the first level i.e. leaves (**b1**) will require ceiling (100,000/127) = **788 blocks**
Thus, our second level (**b2**) above the leaves will require ceiling (788/136) = **6 blocks**
The third level (**b3**) above that will require **1 block**
The **total blocks** this index use, are **795 blocks**

Q3. Repeat Q2 for the *B-tree* index.

Answer:
order of p: p * 10 + (p-1) * (5+11) < 2048, which gives us **order p** = (2048 + 16)/26 **= 79**

This means the first level i.e. leaves (**b1**) will require ceiling (100,000/79) = **1265 blocks**
Thus, our second level (**b2**) above the leaves will require ceiling (1265/79) = **16 blocks**
The third level (**b3**) above that will require **1 block**
The **total blocks** this index use, are **1282 blocks**