Data Systems Group Assignment 3 – Due Monday 5th October 11 PM

1. Consider the following information

Relation/number of rows	Column/Attribute	Length (in bytes)	Number of distinct Values/Distribution
R/ 2^{20} rows.	KR	32	key
	A	24	10 - uniform
	B	28	2^{10} - uniform
	C	4	2 values - uniform
	KS	16	Foreign Key – (see relation S) uniform
	Other columns	192	-
S/ 2^8 rows	KS	16	key
	A	24	10 - uniform
	B	28	2^8 - uniform
	C	4	2 values - uniform
	Other columns	188	-

Consider the

a) Disk block size (B) is 512 bytes, 4096 bytes, and 32,768 bytes. Three possibilities.
b) Block Pointer length is 8 bytes.
c) The record pointer length is 16 bytes.

Now compute the number of block accesses needed **(for each of the above block sizes)** to execute the following **retrievals**. Provide detailed answers with all intermediate calculations, formulae used, and the rationale. Use floor and ceil appropriately for the computation of the number of block accesses.

a) R and S are unordered files, determine the number of block accesses for the retrieval
    i. KR=val
    ii. KR>val
    iii. S.A=val
    iv. S.A>val
b) R and S are ordered files, determine the number of block accesses for the retrieval
    i. S.KS=val
    ii. S.KS>val
    iii. R.B=val
    iv. R.B>val
c) Size in the number of blocks for Primary Index
    i. On R.KR and the number of block accesses for R.KR=val.
    ii. On S.KS and the number of block accesses for S.KS =val.

d) Size in the number of blocks for the Clustered Index
    i. On R.B and the number of block accesses for R.B <val.
    ii. On S.A and the number of block accesses for R.A=val
e) Size in the number of blocks for the secondary index on key
    i. On S.KS and the number of block accesses for S.KS=val
f) Size in the number of blocks for the secondary index on non-key
    i. On R.C and the number of block accesses for R.C =val
    ii. On S.B and the number of block accesses for S.B<val
    iii. On R.KS and the number of block accesses for R.KS =val
g) Consider a B-Tree on R.KR with each block 2/3rd full. Determine the number of block accesses for R.KR=val retrieval.
h) Consider a B+ Tree on R.KS with each block 2/3rd full. Determine the number of block accesses for R.KS=val retrieval.
i) Consider the main memory buffer of (2^10+2) blocks. Compute the number of block accesses for the block nested loop equijoin between R and S on attributes R.KS=S.KS.
j) Consider an equijoin of R and S, on attributes R.KS=S.KS, compute the number of block accesses for nested loop join using B+tree index on S.KS.
k) Consider B+tree secondary indices on R.C and S.C are available. Which join algorithm will you use, and how many block accesses will you need to perform R and S's equijoin on attributes R.C=S.C.
l) There is a retrieval on a single relation using an indexing method that will take the maximum number of block accesses. What is that retrieval, and how many block accesses will it take?
m) Change the block size so that the complete primary index on S.KS will fit in a single block. Recompute the number of block accesses for the above query (l).