DSAA 5012: ADVANCED Database Management FOR DATA SCIENCE

Lecture 11 Exercises
Storage and File Structure

Exercise 1: A Student file has 20,000 records of fixed-length. Assume the page size is 512 bytes and each record has the following fields: name (30 bytes), studentId (8 bytes), address (40 bytes), phone (8 bytes), birthdate (8 bytes), gender (1 byte), majorDeptCode (4 bytes), minorDeptCode (4 bytes), classCode (4 bytes), and degreeProgram (3 bytes). An additional byte is used as a deletion marker.

a) What is the record size in bytes?

b) What is the blocking factor $bf_{Student}$?

c) How many pages are needed to store the file?

Exercise 2: How many page I/Os are needed to search for a record given its studentId value if the file of Exercise 1 is organized as

a) a heap file?

b) a sequential file sorted on studentId?

Do not upload this exercise worksheet to Canvas.

Name: _____/_____ **Student#:** _____ **Date:** _____
Family/Last (PRINT) Given/First (PRINT)

DSAA 5012: ADVANCED Database Management FOR DATA SCIENCE

Lecture 11 Exercises
Storage and File Structure

Exercise 3: An Employee file has 30,000 records of fixed-length. Assume the page size is 1,000 bytes and each record has the following fields: name (25 bytes), hkid (8 bytes), address (35 bytes), deptCode (8 bytes), phone (8 bytes), birthdate (8 bytes), gender (1 byte), jobCode (3 bytes), salary (4 bytes). An additional byte is used as a deletion marker.

a) What is the record size in bytes?

b) What is the blocking factor $bf_{Employee}$?

c) How many pages are needed to store the file?

Exercise 4: For the file of Exercise 3, how many page I/Os are needed to search for

a) a record given its hkid value if the file is organized as a sequential file sorted on hkid?

b) all the records with a given jobCode value if the file is organized as a sequential file sorted on hkid?

c) a record given its hkid value if the file is organized as a hash file hashed on hkid and there are no overflow pages?

Upload this completed exercise worksheet to Canvas by **11 p.m. on March 10th.**