

Marmara University – Faculty of Engineering – Department of Computer Engineering

**Fall 2020 – CSE3055 Database Systems
Homework #8****Due: 22.01.2021.Fri 23:59**

- 1) [35 pts] Consider the unnormalized relation R with six attributes ABCDEF and the following functional dependencies:

 $AB \rightarrow CDE$ $B \rightarrow F$ $C \rightarrow D$

- a) [5 pts] What is the key(s) for the relation?
- b) [10 pts] What is the normal form of this relation? Explain it.
- c) [20 pts] Decompose R into 3NF relations step by step if it is not in 3NF.
- 2) [15 pts] Consider the following normalized relations from a database in a large retail chain:

STORE (StoreID, Region, ManagerID, SquareFeet)**EMPLOYEE** (EmployeeID, WhereWork, EmployeeName, EmployeeAddress)**DEPARTMENT** (DepartmentID, ManagerID, SalesGoal)**SCHEDULE** (DepartmentID, EmployeeID, Date)

What opportunities might exist for denormalizing these relations when defining the physical records for this database? Under what circumstances would you consider creating such denormalized records?

- 3) [14 pts] Consider the following two relations for Millennium College:

STUDENT (StudentID, StudentName, CampusAddress, GPA)**REGISTRATION** (StudentID, CourseID, Grade)

Following is a typical query against these relations:

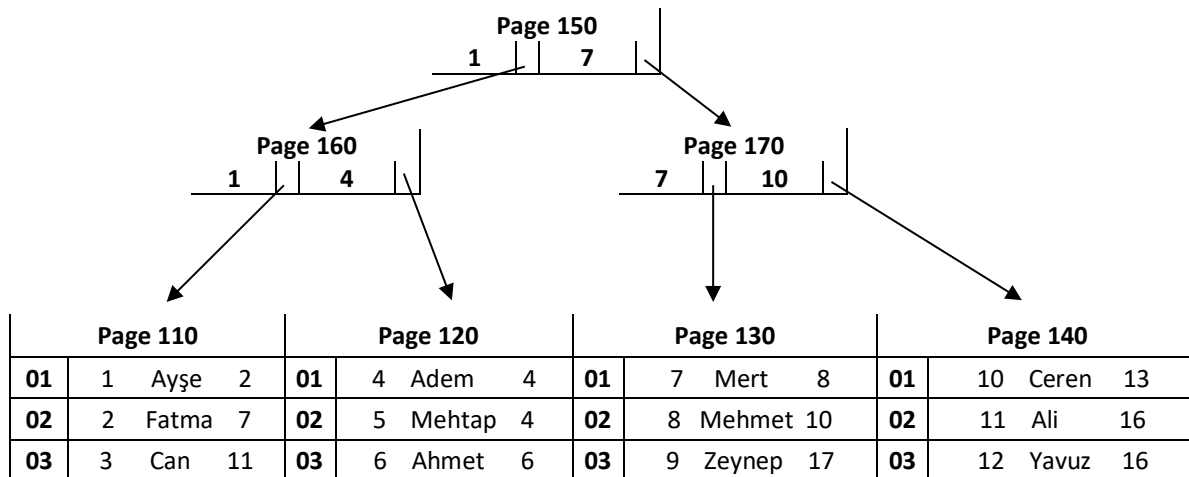
```
SELECT Student.StudentID, StudentName, CourseID, Grade
FROM Student, Registration
WHERE Student.StudentID = Registration.StudentID AND GPA > 3.0
ORDER BY StudentName;
```

- a) [6 pts] On what attributes should indexes be defined to speed up this query? Give the reasons for each attribute selected.
- b) [8 pts] Write SQL commands to create indexes for each attribute you identified in part a.

- 4) [36 pts] You have a STUDENT table that has SID, Name, and Age columns. Which data pages are accessed to execute the queries below, under situations given at (a) and (b)? (Assume that index seek is used whenever possible)

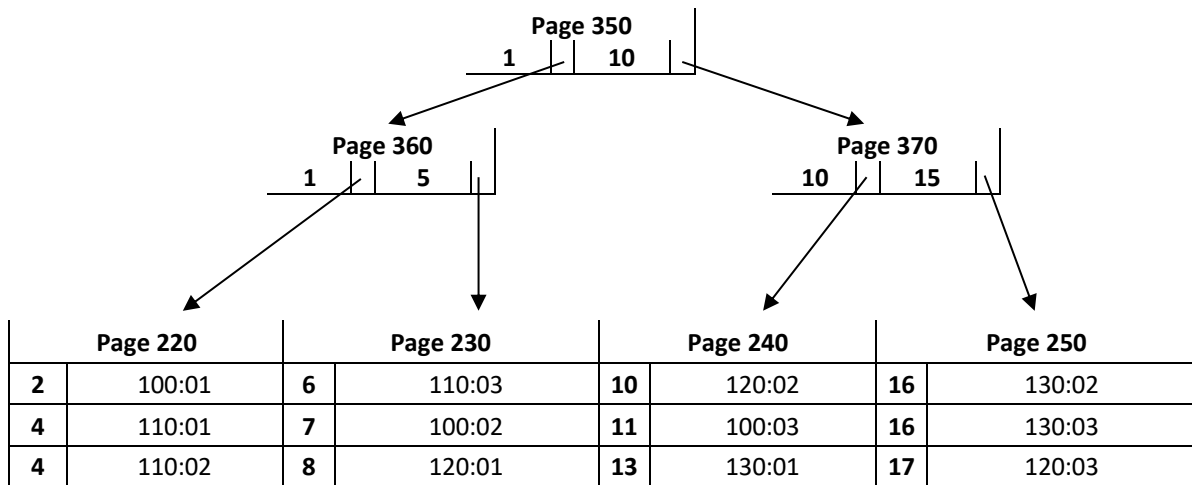
STUDENT		
SID	Name	Age

- a) [18 pts] The table has a *clustered index* on **SID** column, and no other indexes. The index structure and data is stored on data pages as the following:



- i) **Query 1:** select Name from STUDENT where SID < 11
- ii) **Query 2:** select * from STUDENT where Age = 16
- iii) **Query 3:** select * from STUDENT where SID = 7

- b) [18 pts] The table has a *non-clustered index* on **Age** column, and no other indexes. The index structure and data is stored on data pages as the following:



Page 100				Page 110				Page 120				Page 130			
01	1	Ayşe	2	01	4	Adem	4	01	7	Mert	8	01	10	Ceren	13
02	2	Fatma	7	02	5	Mehtap	4	02	8	Mehmet	10	02	11	Ali	16
03	3	Can	11	03	6	Ahmet	6	03	9	Zeynep	17	03	12	Yavuz	16

- i) **Query 1:** select Age, Name from STUDENT where SID < 9
- ii) **Query 2:** select Age from STUDENT where Age < 8
- iii) **Query 3:** select * from STUDENT where Age = 8