School of Computer Science & IT, RMIT

COSC2406/2407 Tutorial Exercises

Note: Numbered homework exercises, and exercises marked "*", are from Ramakrishnan and Gehrke.

File organisation and indexes: Week 5

1. We wish to store records about customers in one of: a heap file, a sorted file, or a static hashed file.
   (a) If the most common search is "Find all customers who have a surname beginning with M", which file organisation would you recommend? Why?
   (b) If the most common search is "Show all customers", which file organisation would you recommend? Why?
   (c) If the most common search is to find and retrieve details about one customer, which file organisation would you recommend? Why?
   (d) If all of the above queries are common, which file organisation would you recommend? Why?

2. Explain the difference between the following:
   (a) Primary and secondary indexes
   (b) Dense and s
   (c) Cluster

3. Shown below is a data file consisting of records of the form `(Student_Name, Age, Student_ID, Exam_Mark)`. The data file consists of three pages. _ ary key of the records, and the file is sorted by `Age`.


```
Meier, 25, 0815A, 77
Schultze, 25, 1234B, 33
Klinsmann, 30, 1234C, 54
---------------------------------
Bierhoff, 31, 5321A, 82
Breitner, 43, 1234S, 82
Rummenigge, 44, 3234R, 76
---------------------------------
Hrubesch, 49, 6353L, 42
Beckenbauer, 52, 4113Y, 76
Kaltz, 55, 3456K, 89
```

   (a) Build a dense, secondary index on the `Exam_Mark` field using "Alternative 2" for data entries.
   (b) Build a sparse index on the `Age` field.
   (c) Is the sparse index on the `Age` field a clustered or an unclustered index? Explain.

4. Homework: 8.2, 8.3, 8.8, 8.10