|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **National University of Computer and Emerging Sciences, Lahore Campus** | | | | |
| final design | **Course:** | **Advance Database Systems** | **Course Code:** | **CS451** |
| **Program:** | **BS(Computer Science)** | **Semester:** | **Spring 2019** |
| **Out Date:** | **11-Mar-2019** | **Total Marks:** |  |
| **Due Date:** | **Tue 19-Mar-2019 *(Start of Class)*** | **Weight:** |  |
| **Section** | **CS** | **Page(s):** | **1** |
| **Assignment:** | **2 (Disk Storage & File Structures)** |  |  |
|  |  | | | |

*Instructions:*

* **This assignment is an individual assignment.**
* **Use proper assignment papers for solving your assignment questions. Assignment done on diary pages, register pages, rough pages will not be credited.**
* **Use any valid assumption where needed.**

**Q1.**

Consider a file system on a disk with block size *B=2048 bytes*. A file has *r=1,000,000* CUSTOMER records of fixed-length (un-spanned). Each record length is *100 bytes*. Assume there are 25 cities and 40,000 customers per city. Primary key of customer file is CustomerID.

Estimate the number of block fetches needed to compute the following queries:

1. SELECT \* FROM customer WHERE CustomerID =12345678; (Assume file is not ordered)
2. SELECT \* FROM customer WHERE CustomerID =12345678; (Assume file is ordered on CustomerID)
3. SELECT \* FROM customer WHERE city = ‘Lahore’; (Assume file is not ordered)
4. SELECT \* FROM customer WHERE city = ‘Lahore’; (Assume file is ordered on City)

**Q2.** A customer file has following customer ID values:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 8 | 16 | 24 | 32 | 40 | 48 | 56 | 64 | 7 | 15 | 31 | 63 | 127 | 1 | 10 | 4 | 108 | 126 |

1. Consider these customer IDs as hash key values. The file uses 6 buckets named 0 to 5. One bucket cannot hold more than 3 records, means at max a bucket can hold 3 records. Load these records in file using hash function h(k) = k mod 6, in the given order.
2. Calculate the average number of block accesses for random retrieval on customer ID.
3. Load the given values in expandable hash files based on extendible hashing, show structure on each step, use hash function h(k) = K mod 16, max 4 records can be kept in one bucket.
4. Load the given values in expandable hash files based on dynamic hashing, show structure on each step, use hash function h(k) = K mod 16, max 4 records can be kept in one bucket.
5. Load the given values in expandable hash files based on linear hashing, show structure on each step, use hash function h(k) = K mod 16, max 4 records can be kept in one bucket.
6. Show the structure after inserting an entry with customer ID value 73 into the final structure of part (c).
7. Show the structure after inserting an entry with customer ID value 72 into the structure of part (f).