1. Answer the following questions about files and indexes:
    a. What alternatives are available for the data entries in an index?

        - A data entry h is an actual data record (with search key value k).

        - A data entry is a (k, rid) pair, where rid is the record id of a data record with search key value k.

        - A data entry is a (k. rid-list) pair, where rid-list is a list of record ids of data records with search key value k.

    b. What is the difference between a clustered index and an unclustered index? If an index contains data records as 'data entries', can it be unclustered?

        In a clustered index, the order of the index corresponds to the physical order of data records in the table while unclustered indexing is more random and the indexes are seemingly out of order

    c. How many clustered indexes can you create on a file? Would you always create at least one clustered index for a file?

        There should only be one clustered index per file due to it being the way of ordering indexes in a set order but there is not a need to create a clustered index for a file as there are unclustered ways to order a file

2. Explain the terms seek time, rotational delay, and transfer time.

    Seek time – the time taken to move the disk heads to the track on which a desired block is located

    Rotational Delay – the waiting time for the desired block to rotate under the disk head

    Transfer Time – The time to actually read or write the data in the block once the head is positioned

3. What is sequential flooding of the buffer pool?

    Sequential flooding is when the file is scattered in the buffer pool so everything is scanned so that the result always includes every page of the file

4. Describe two possible record formats. What are the trade-offs between them?

    The two possible formats are Fixed-Length Records and Variable-Length Records, with fixed-length there are set lengths that each field has and the number of fields are fixed this can help with easy consecutive storage and the lengths and addresses can be calculated due to this fixed length. The other format is variable-length which lets the fields be any length resulting in more dynamic storage and can help when dealing with null values.

5. Why do frames in the buffer pool have a pin count instead of a pin flag

Pin counts are used instead to keep better counts of the number of current users of the page.