1. Suppose a database has the following table for suppliers, orders, items ordered and parts combined where (OrderNo, ItemNo, PartName) is the primary key and this is the complete table. (3 points)

<table>
<thead>
<tr>
<th>SupplierNo</th>
<th>SupplierName</th>
<th>OrderNo</th>
<th>OrderDate</th>
<th>ItemNo</th>
<th>PartName</th>
<th>PartPrice</th>
<th>Qty</th>
</tr>
</thead>
<tbody>
<tr>
<td>112</td>
<td>Phillips</td>
<td>100</td>
<td>04/14/00</td>
<td>1</td>
<td>Screw</td>
<td>$2.25</td>
<td>5</td>
</tr>
<tr>
<td>212</td>
<td>Starks</td>
<td>101</td>
<td>04/14/00</td>
<td>2</td>
<td>Bolt</td>
<td>$3.99</td>
<td>5</td>
</tr>
<tr>
<td>212</td>
<td>Starks</td>
<td>101</td>
<td>04/14/00</td>
<td>3</td>
<td>Washer</td>
<td>$1.99</td>
<td>9</td>
</tr>
<tr>
<td>312</td>
<td>Kapur</td>
<td>102</td>
<td>04/15/00</td>
<td>1</td>
<td>Screw</td>
<td>$2.25</td>
<td>5</td>
</tr>
</tbody>
</table>

a. Give an example of an insertion anomaly for Suppliers.

**Answer**

Can’t insert Supplier information without having the Supplier on an order.

b. Using row 1, give an example of a deletion anomaly.

**Answer**

If order 100 is deleted, all information about Supplier 112 is lost and visa versa.

c. Using rows 2 and 3, give an example of an update anomaly.

**Answer**

If the name of a Supplier is changed, it must be changed in both places.

or

If the order date is changed, it must be changed in all three places.
2. Fill in the blank:
   (3 points)
   a. Dependencies based on only a part of a composite primary key are called __________ dependencies.
   
   b. Design relation schemas so that they can be joined with equality conditions on attributes that are either
      primary keys or foreign keys in a way which guarantees that no _________ tuples are generated.
   
   c. \( X \rightarrow Y \) ______ means that whenever two tuples in \( R \) agree on all the attributes in \( X \), they must also
      agree on all attributes of \( Y \), where \( X \) and \( Y \) are sets of attributes in a relation \( R \).

3. Consider the following relation \( R(A, B, C, D, E) \) and \( \text{FDs} = \{C \rightarrow E, B \rightarrow D\} \) (2 points)
   a. Is \( R \) relation in 1NF? Why
      Yes
      Why? All attributes are simple and single
   
   b. Is \( R \) relation in 2NF?
      No
      Why? Because of \( B \rightarrow D \) partial dependency

4. Consider relation \( R \) with five attributes ABCDE, and the following functional dependencies: (1 point)
   \[ A \rightarrow B \quad BC \rightarrow D \quad D \rightarrow E \]
   Give the complete closure for \( AC \), i.e., find \( (AC)^+ \).
   ABCDE

5. Consider a relation \( R(A, B, C, D) \) with FD’s \( AB \rightarrow C \), \( C \rightarrow D \), and \( D \rightarrow A \).
   List the two keys for \( R \). (2 points)
   \[
   \{A, B\}^+ = \{A, B, C, D\} \quad \{B, D\}^+ = \{B, C, D, A\} \\
   \{B, C\}^+ = \{B, C, D, A\}
   \]
6. For relation \( R(A, B, C) \), suppose \( AB \rightarrow C \) and \( C \rightarrow B \). List all normal forms (if any) that \( R \) is in. You only need to consider 3NF, 4NF, and BCNF. (1 point)

3NF

7. We perform decomposition to normalize an original schema to be of certain normal forms. For such a decomposition to be “equivalent” to the original schema, it is desirable to be lossless.

To study this concept, let’s consider an original schema \( R(A, B, C) \). Suppose we decompose \( R \) into \( R1(A, B) \) and \( R2(A, C) \).

a. Is this decomposition always lossless? Answer yes or no and briefly explain why. (1 point)

No. Unless \( A \rightarrow B \) or \( A \rightarrow C \) is satisfied in \( R \), \( R1 \) \( \propto \) \( R2 \) may generate entries that do not belong to \( R \).

b. Give an example instance of \( R \) (i.e., an example table with 2 tuples) and demonstrate its decomposition, to support your answer in (a). (1 point)

<table>
<thead>
<tr>
<th>( R1 )</th>
<th>( R1 )</th>
<th>( R2 )</th>
<th>( R1 \propto R2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. Consider a relation \( R \) with five attributes ABCDE. The following dependencies are given:

\[
A \rightarrow B \\
BC \rightarrow E \\
ED \rightarrow A.
\]

The keys for \( R \) are CDE, ACD, BCD

a. Is \( R \) in 3NF? Briefly explain why. (1 point)

YES, it is in 3NF because B, E, and A all parts of keys.

b. Is \( R \) in BCNF? If yes, please explain why. Otherwise, decompose \( R \) into relations that are in BCNF. (3 points)

No. For example, \( BC \rightarrow E \) violates BCNF. If we decompose using this FD (\( BC \rightarrow E \)), we get BCE, and ABCD.

Further decomposing ABCD using \( A \rightarrow B \) we get AB, ACD.

Therefore, it could be decomposed into BCE, AB, ACD.
9. Consider the relation for an appliance dealer \( S(A_1, A_2, A_3, A_4, A_5) \) (2 points)

Assume that \( A_1 \rightarrow A_4, A_1A_2 \rightarrow A_3, A_4 \rightarrow A_5 \)

Consider the decomposition \( D \) \( S_1(A_1, A_2, A_4) \) \( S_2(A_1, A_3, A_5) \).

Show that the decomposition \( D \) is not lossless join?

\[
S_1 \cap S_2 = A_1 \quad S_1 - S_2 = (A_2, A_4) \quad S_2 - S_1 = (A_3, A_5)
\]

Because \( A_1 \) doesn’t functionally define either \( (S_1 - S_2) \) or \( (S_2 - S_1) \), this decomposition is not lossless join.

10. Consider the relation \( R(A, B, C) \) with MVD \( C \rightarrow\rightarrow B \) and \( C \rightarrow\rightarrow A \). \( R \) currently has the following tuples: (2 points)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>a</td>
<td>x</td>
</tr>
<tr>
<td>1</td>
<td>a</td>
<td>y</td>
</tr>
<tr>
<td>1</td>
<td>b</td>
<td>x</td>
</tr>
<tr>
<td>1</td>
<td>b</td>
<td>y</td>
</tr>
<tr>
<td>2</td>
<td>c</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>c</td>
<td>z</td>
</tr>
<tr>
<td>2</td>
<td>d</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>d</td>
<td>z</td>
</tr>
</tbody>
</table>

a. The above relation is not in 4NF. Normalize it into 4NF.

\[ R_1(C, B) \quad \text{and} \quad R_2(C, A) \]

b. What is the table look like after decomposition?

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>y</td>
<td></td>
</tr>
</tbody>
</table>

11. Select the correct sentence. (1 point)

a. The second and higher levels must be sparse in an index of multiple levels on sequential files.

b. The order of insertions into a B+ tree will not affect the tree's final structure at the end.

c. The time needed to access a page on disk is comprised of two components: disk seek time and data transfer time.

d. For disk latency, seek time refers to the time for the head to find the desired sector—i.e., for the disk to rotate so the first of the sectors containing the desired data reaches the head.
The questions pertain to the relation and indices described below:

### Car Relation

**Car Relation**

Car(Vin, Year, Model, Price)

The *car* relation belongs to an automobile database. It contains a total of 10,000 tuples. The data includes the vehicle identification numbers, year, model, and base price for 50 different models over a 40 year period from 1960-1999 (inclusive). Each block of the file where the car relation is stored contains 20 tuples. The records in the file are ordered sequentially according to model.

12. According to the information given, it would be possible (without reorganizing the file) to create a primary index on attribute ___________.  
   (1 point)
   a. Vin       b. Year       c. Model       d. Price       e. None of the above

13. According to the information given, it would be possible (without reorganizing the file) to create a clustering index on attribute ___________.  
   (1 point)
   a. Vin       b. Year       c. Model       d. Price       e. None of the above

14. A secondary index:  
   (1 point)
   a. Must use a tree structure such as a B+-tree or B-tree.  
   b. Must be a dense index.  
   c. Cannot be created on the primary key of a relation.  
   d. All of the above.

15. If each node in a B+-tree index file contains exactly p pointers, and the number of search key values to be indexed is n, then the height of the tree is proportional to:  
   (1 point)
   a. n       b. pn       c. log₂(n)       d. log₂(pn)       e. logₚ(n)

16. How many blocks are on a disk with the following characteristics? Sector size of 512 bytes, 16 sectors per track, 16384 tracks per surface, 4 double sided platter, and 4096 bytes per block.  
   (1 point)
   \[512 \times 16 \times 16384 \times 4 \times 2 / 4096 = 262144 \] blocks
17. Consider constructing a B*-tree of order 4 (i.e., \( f = 4 \), each index node can hold \( f \) pointers and \( f-1 \) keys). Assume the following structure for a non-leaf node.

(a) Show the resulting tree after inserting keys 40; 10; 50; 30; 90; 80; 70; 20; 60; 100, in this order. (2)

(b) Based on the tree in (a), how many nodes are accessed to find all records in the range 20 to 60. (1)

(c) Show the resulting tree after deleting key 60 from the tree in (a). (1)