DS5003 Data Engineering Lab

Final, Date: 08/11/ 2024

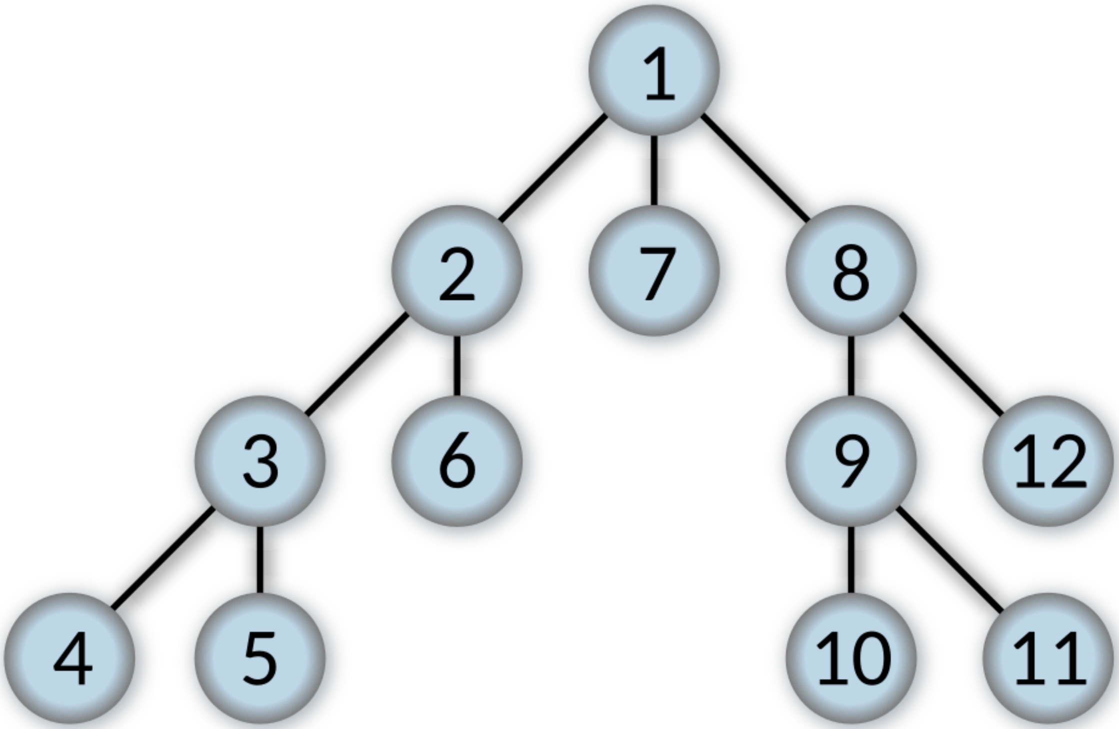
Timing: 2:00 to 4:30 PM Max Marks: 20

Instructions

1. Submit one .ipynb file containing all answers. The name should be **[student name]_final.pdf**
2. Write questions in separate text blocks in Jupyter Notebook before the code blocks containing answers.
3. Read the questions carefully before answering. If a question asks to follow a particular approach or use a specific data structure, it must be followed.

---

1. Use the graph below to answer questions a and b



   a. Use BFS to find the path between vertex 3 and 11. List the nodes in the path. **2**
   b. Find the shortest path from vertex 1 to vertex 12 using BFS. List the nodes in the order they appear on this path. **2**

2. Given an integer n, find how many digits to remove from the number to make it a perfect square. **4**

Example- Input: 8314, Output: 81, 2

(If we remove 3 and 4, the number becomes 81 which is a perfect square.)

Input : 57, Output : -1

3. Create a database named IIT Palakkad and the following tables with at least 10 entries each:

   student (id, name, dept_name, gate_score, cgpa)

   assistant (id, course_id, sec_id, sem, year)

   course (course_name, course_id, credit, dept_name)

   Sample Data: dept_name = ("DS", "PEPS", "SOCD"), sem = (1, 2, 3, 4), year = (2023, 2024), credit = (3, 4, 5), gate_score is out of 1000

   a. Retrieve the names and CGPA of all students from the "DS" department whose CGPA falls between 7.5 and 8.5 **1**
   b. For each department, retrieve the name of the student(s) with the highest GATE score. Additionally, include the department name and the average CGPA of all students within that department. Ensure to handle cases where multiple students share the highest GATE score. **3**
4. Use **Northwind database**
   a. Create a function that, given a category name, returns the average price of all products in that category using sqlAlchemy **2**
   b. Create a view that summarizes total sales by month for each product, including product names and total sales amounts for each month in 1997. Provide screenshots of view creation and verification of its contents. **2**
   c. Design a trigger that automatically updates the discount in the order_details table based on the total quantity ordered. For example, if the total quantity ordered for a product_id across multiple orders exceeds 10 units, apply a 5% discount on each order line for that product.. **4**

      Note: Create the SQL trigger and demonstrate its functionality. After creating the trigger, insert or update records in the table(s) to test it. Provide screenshots of your SQL code, the trigger creation, and the output showing the discount update.