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. A rotated sorted array is an array that has been sorted and then rotated around an index. For example a sorted array [1, 2, 3, 4, 5], rotated around the index 2 becomes [3, 4, 5, 1, 2].   Write a program that performs **binary search** in a rotated sorted array. Print whether the target integer is found and its position.
   Input: rotated_array = [4, 5, 6, 7, 0, 1, 2], target = 0
   Output: Found 0 at index 4   **3**

2. Use attached **sales dataset** to answer the following questions
   a. Which categories of items in the sales dataset have been purchased the highest and the lowest? Visualize the results using a pie chart.   **1.5**
   b. Use pie charts to show which region generated the maximum profit and which region generated the minimum profit.   **1.5**

3. Use **Northwind database**
   a. For each product category in the database, find the top three products based on the total quantity ordered.   **1**
   b. Create a view that lists products and their categories using sqlAlchemy. Provide screenshots of view creation and verification of its contents. **2**
   c. Write a trigger that automatically updates the reorder_level in the products table when the units_in_stock falls below a specific threshold (100 units) after an update on the order_details table.   **2**
   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.

4. Consider the given graph:
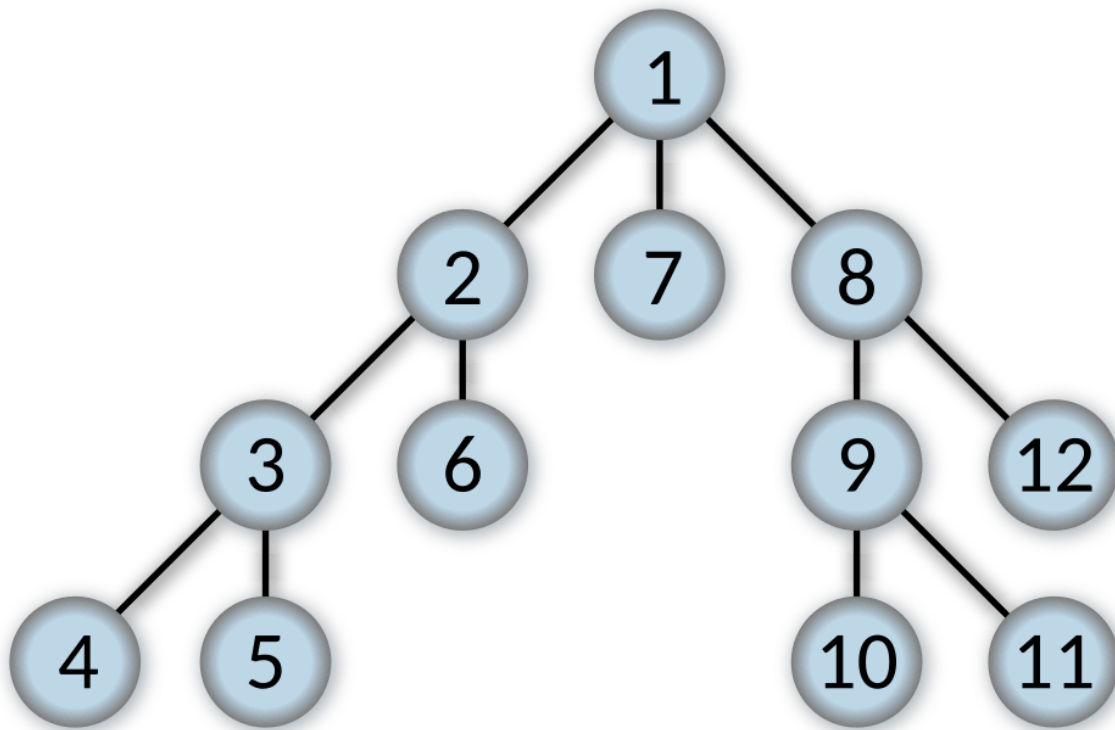
Write an algorithm to find the shortest path from any two graph nodes where each edge has a uniform weight of 1. Provide the path taken and the total edge distance for that path. Verify the algorithm for finding out the path from node 1 to node 11.   **4**

5. Use **Northwind database**
   a. Write a function that retrieves all products within a specific price range, and category including their names and supplier information.   **2**
   b. Show the summary of total sales per territory for orders placed on or after January 1, 1997   **3**