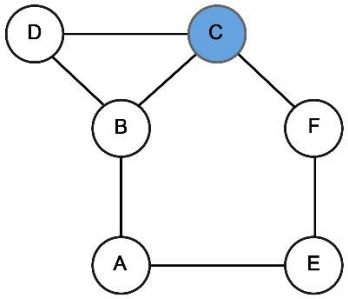
# CSE 4/587: Exam-3

**Due May 13, 2022 (Friday) 11.00 PM**

**Notes:** The exam must be done individually, with no discussion or help with others. Breaking this rule will result in an automatic 0 grade.

## PART A (25 Points)

1. Discuss the Resilient Distributed Datasets in Spark and their significances. What is lazy evaluation in Spark? **(5 points)**
2. Differentiate Narrow Transformation and Wide Transformation? Explain with example. **(5 points)**
3. Define k-core subgraph. Discuss the interpretation of a k-core subgraph in the context of Community Analysis. **(5 points)**
4. Calculate Betweenness Centrality for **Node C** in the following graph. Show your work. **(10 points)**




## PART B (15 Points)

5. A data set named "GRAPH_DATA.csv" is uploaded to the Exam-3 folder. Implement Dijkstra's algorithm to find the shortest path distance from Node “N” to node “O” using Neo4j. Show your work, including code, Initial Graph, and Shortest Distance Path.

## PART C (10 Points)

6. Write an abstract/summary of your group project in your words. Discuss your specific contribution to this team effort.