1. Read the data available in Data tab in R

2. Check missing value in the data

3. Activate rpart and partykit library

4. set seed 5 - this will ensure that you get the same output as mine

5. Find percentage distribution of Term\_Dep subscription in complete data set

6. develop a classification tree

7. Print node statistics

8. Which is the leaf node which has least term deposit subscription %?

9. What is the % of Yes - in the leaf node, which has least term deposit subscription?

10. Which is the leaf node which has maximum term deposit subscription %?

11. What is the % of Yes - in the leaf node, which has maximum term deposit subscription?

12. Find KS of this decision tree, to preeict Term\_dep= yes

13. If the budget allows you to call only 6000 people, which are the leaf node, whom you won't like to call

14. Which is the leaf node, which has highest number of term deposit but difficult to contact