

Big Data Hadoop, Module 2: Assignment - 2

Problem statement:

ABC Telephone Company keeps track of every call made using its services. You work as a Big Data Analyst at this organization. You have been given the task to analyze the data that has been accumulated.

Link for the dataset and resources: <https://intellipaate-course-attachments.s3.ap-south-1.amazonaws.com/Hadoop/phone.txt>

Dataset Description:

The data format is comma separated values, and the dataset is

'phone.txt'. Following are the attributes in sequential order:

- Caller ID
- Receiver ID
- Call Start Time
- Call End Time
- Flag (0 -> Local Call, 1 -> Long Distance Call)

Tasks to be completed:

Find out the list of customers who have the total call time in long-distance calls of more than 60 minutes
