

1. Create a MR job with the following features
  - a. It should be able to get a word-count
  - b. Find how many "and" ( case insensitive ) occurred in the input text.
  - c. Find the following stats from your job
    - i. io.buffer were used during the entire job execution
    - ii. Total mappers
    - iii. Total reducers
  - d. Find the input and output formats that were used in your job while processing.
  
2. Create a MR job with the following features:
  - a. You receive data from all the trains that are coming and going via a junction.
  - b. Input Data is something like the following:
    - i. TRAIN\_NO | TRAIN\_TIME | TRAIN\_SPEED | TRAIN\_DIRECTION  
 1000 | 1000 | 60 | E  
 1001 | 1010 | 80 | W  
 1002 | 1015 | 40 | N  
 1003 | 1020 | 60 | S  
 1002 | 1015 | 40 | E  
 0110 | 1000 | 60 | W  
 Note : Refer attached train\_data.txt for the input
    - ii. Using the above data, answer the following:
      1. Total how many train data in present in the input data.
      2. How many trains are travelling from this station in each hour
      3. How many trains are travelling in each direction in each hour
      4. What is the avg speed of each train travelling in each direction in each hour.
      5. If you receive some data that is incorrect in input file, add additional code to handle that inside your mapper / reducer