**BIG DATA ANALYTICS LAB QUESTIONS**

**Question -1**

1. Setting up hadoop on Single machine (standalone mode)
2. Setting up hadoop on multi-node (distributed cluster environment)

**Question-2:**

1. Basic hadoop file system commands.
2. Write a map-reduce program for word count in give text dataset.
3. Write a map-reduce program for finding maximum temperature in weather dataset.
4. Write a map-reduce program for join of two records.
5. Write a map reduce program to find duplicate record in csv file.
6. Write a map reduce program to find patent citations in patent dataset.
7. Write a Map reduce program for total retail collection.(retail dataset)
8. Write a Map reduce program for store wise collection.(retail dataset)
9. Write a Map reduce program for product wise collection.(retail dataset)
10. Write a map reduce program for display highest ctc in each dept in employ ctc dataset. Note: Cost to company (**CTC**) is a term for the total salary package of an **employee.**

**Extra Programs:**

1. Write a map-reduce program for word count in give text dataset.(with combiner)
2. Write a map-reduce program for word count in give text dataset.(without case sensitive)
3. Write a map reduce program to sum of numbers in text file.
4. Write a map-reduce program for finding minimum temperature in weather dataset.