



**VIT<sup>®</sup>**

**Vellore Institute of Technology**

(Deemed to be University under section 3 of UGC Act, 1956)

## **School of Computing Science and Engineering**

### **LAB - 8 Exercises**

<b>Course Code</b>	<b>:</b>	<b>CSE3025 - Large Scale Data Processing</b>	<b>Date</b>	<b>:</b>	<b>25/09/2019</b>
<b>Lab Experiment</b>	<b>:</b>	<b>Practice of MapReduce Programming using Counters and Combiner Concept</b>	<b>Slots</b>	<b>:</b>	<b>L15+L16</b>
<b>Instructors</b>	<b>:</b>	<b>Dr. Bharadwaja Kumar and Prof. Ramesh Ragala</b>			

Objective:

1. To understand the detailed processing of MapReduce Framework using Counters and Combiners

Problem- 1:

Consider a student dataset with attributes student\_name, marks, address respectively. Create your own dataset based on the attributed discussed. Use the counters in MapReduce framework to find the missing data in address field of a dataset. Please write individual java files to mapper, reducer and driver program.



**VIT<sup>®</sup>**

**Vellore Institute of Technology**

(Deemed to be University under section 3 of UGC Act, 1956)

**Problem - 2:**

Write a MapReduce application using Combiner to find the number of records for each exchange of a given stock mark dataset. The attributes of the stock market dataset are id, exchange, stockname, sector, country, date, open, high, low, close, volume, adj\_close.

Please use the following link to download dataset:

<https://github.com/protechskills/stock-data>