



BIG DATA
DEVELOPMENT

ACADGILD

Session 5: MapReduce

Assignment 3

Session 5: MapReduce

Assignment

Table of Contents

1. Introduction	3
2. Objective	3
3. Prerequisites	3
4. Associated Data Files	3
5. Problem Statement.....	3
6. Approximate Time to Complete Task	3

Big Data and Hadoop Development

1. Introduction

In this assignment you need to understand basic operations on Java string.

2. Objective

Hadoop is majorly used for Text data analysis. This assignment will help you to understand how to handle text data.

3. Prerequisites

Acadgild's VM , or Linux operating system with Hadoop installed in it.

4. Associated Data Files

No data files

5. Problem Statement

Note: consider this string line for below assignment problems.

String line="Hadoop is open source frame work ,Hadoop is good framework for handling big data";

1. Class a sample class with name String_demo and inside the main method perform the following operation
 - Split the line by space and store in a new string array
 - Count number of occurrence of word Hadoop in the given string
 - Replace the ',' with space in the given string
 - Create a array list with generic string and store the splitted string inside it

6. Approximate Time to Complete Task

45 min