Laboratory Manual

ITL801
Big Data Lab
BE - 8th Semester
[Information Technology]

Prepared By
Saurabh Kulkarni
Assistant Professor
Information Technology

FR. Conceicao Rodrigues College of Engineering

January 2020

Course Outcomes

Sr. No.	**Course Outcome Statement**
ITC801.1	Explain characteristics of and trends in big data.
ITC801.2	Use tools like hadoop and NoSQL to solve big data related problems.
ITC801.3	Apply appropriate algorithms for extracting knowledge from given dataset.
ITC801.4	Apply Big data analytics in real life applications.

UNIVERSITY SYLLABUS

Big Data Lab (ITL801)

Perform the following exercises.

1. Study of Hadoop and Hadoop Ecosystem.
2. To count number of words in a file using MapReduce.
3. To find yearly maximum temperature using MapReduce.
4. To perform matrix multiplication using MapReduce.
5. To perform CRUD operations in MongoDB.
6. To perform K-means clustering using MapReeduce.
7. To perform KNN using MapReduce.
8. To find common friends in social network graph using Map Reduce.

9. To develop a recommendation system for a given dataset.
10. Mini Project

CONTENTS

EXPERIMENT 1

Hadoop and ecosystem

1.1 About the Assignment

In this assignment, you are required to study Hadoop installation. The document containing the same is available on moodle with the name *hadoopconf.pdf*. You are supposed to write the steps in your own words. The next task in the assignment is study of eclipse plugin for hadoop. You can get the required document on moodle with the name *hadoopeclipse.pdf*. In the theory you are supposed to explain hadoop components i.e. *hadoop core, YARN, HDFS and MapReduce*. The postlab question includes hadoop ecosystem components which should at least description of *pig, hive, hbase, cassandra, ambari, mahout,*

zookeeper, avro, sqoop. The worksheet for the same is attached after this document.

1.2 Worksheet for the assignment

Name of the student:		**Roll No.**	
Practical Number: 1		**Date of Practical:**	
Relevant CO's: ITC802.2	**At the end of the course students will be able to use tools like hadoop and NoSQL to solve big data related problems.**		
Sign here to indicate that you have read all the relevant material provided before attempting this practical			**Sign:**

Practical grading using Rubrics

Indicator	**Very Poor**	**Poor**	**Average**	**Good**	**Excellent**
Timeline (2)	More than a session late (0)	NA (0.5)	NA(1)	NA (1.5)	Early or on time (2)
Completeness (3)	N/A	N/A	Not Completed (1)	Partially Completed (2)	Completed(3)
Legibility (3)	N/A	N/A	poor(1)	Good(2)	Very Good (3)
Postlab (2)	N/A	N/A	N/A	Partially Correct(1)	All correct answers (2)

Total Marks (10)	**Sign of instructor**

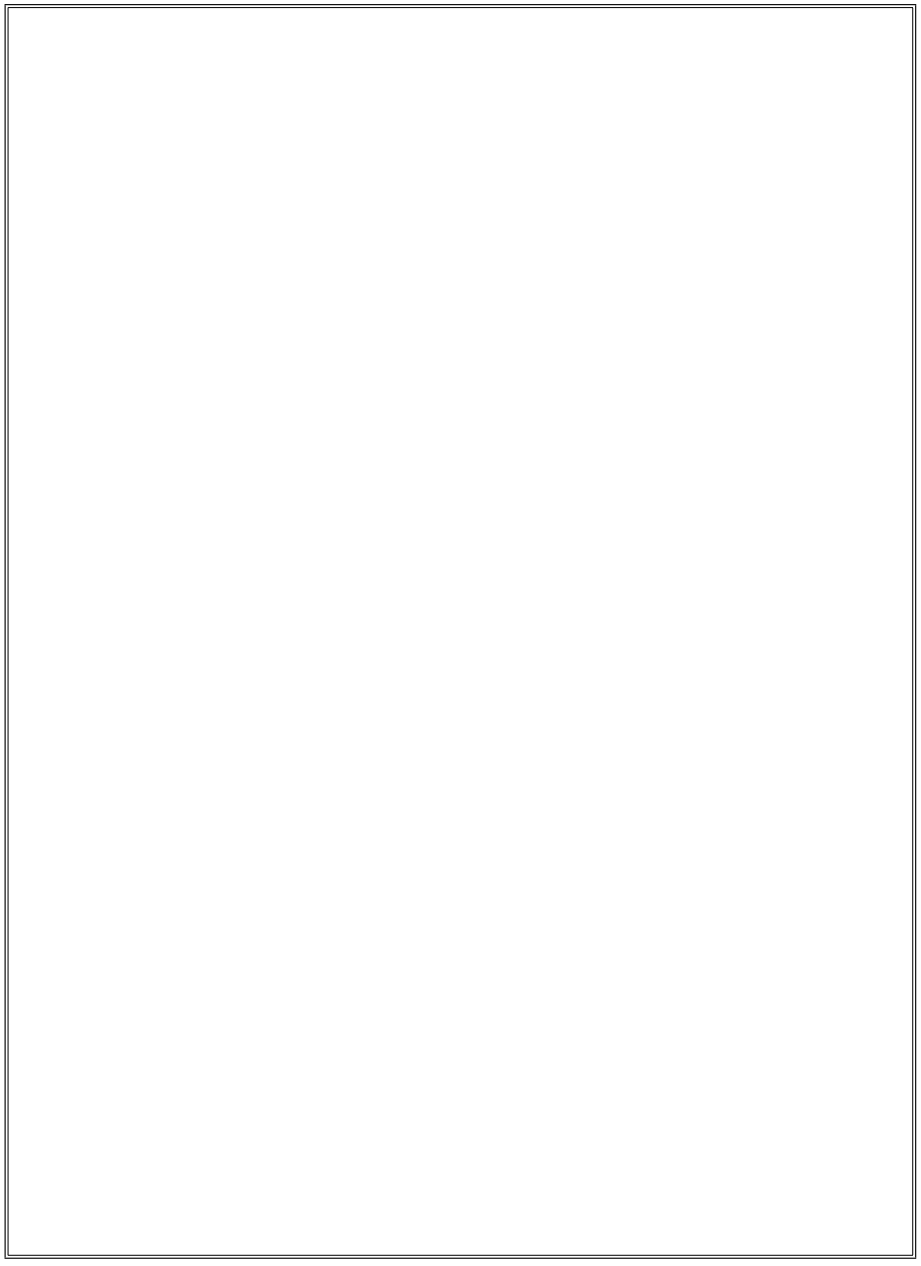
Course title: Big Data Analytics

Practical

Course title: Big Data Analytics
Course term: 2019-2020

Problem Statement: Study of hadoop installation, eclipse plugin configuration for hadoop and hadoop ecosystem.

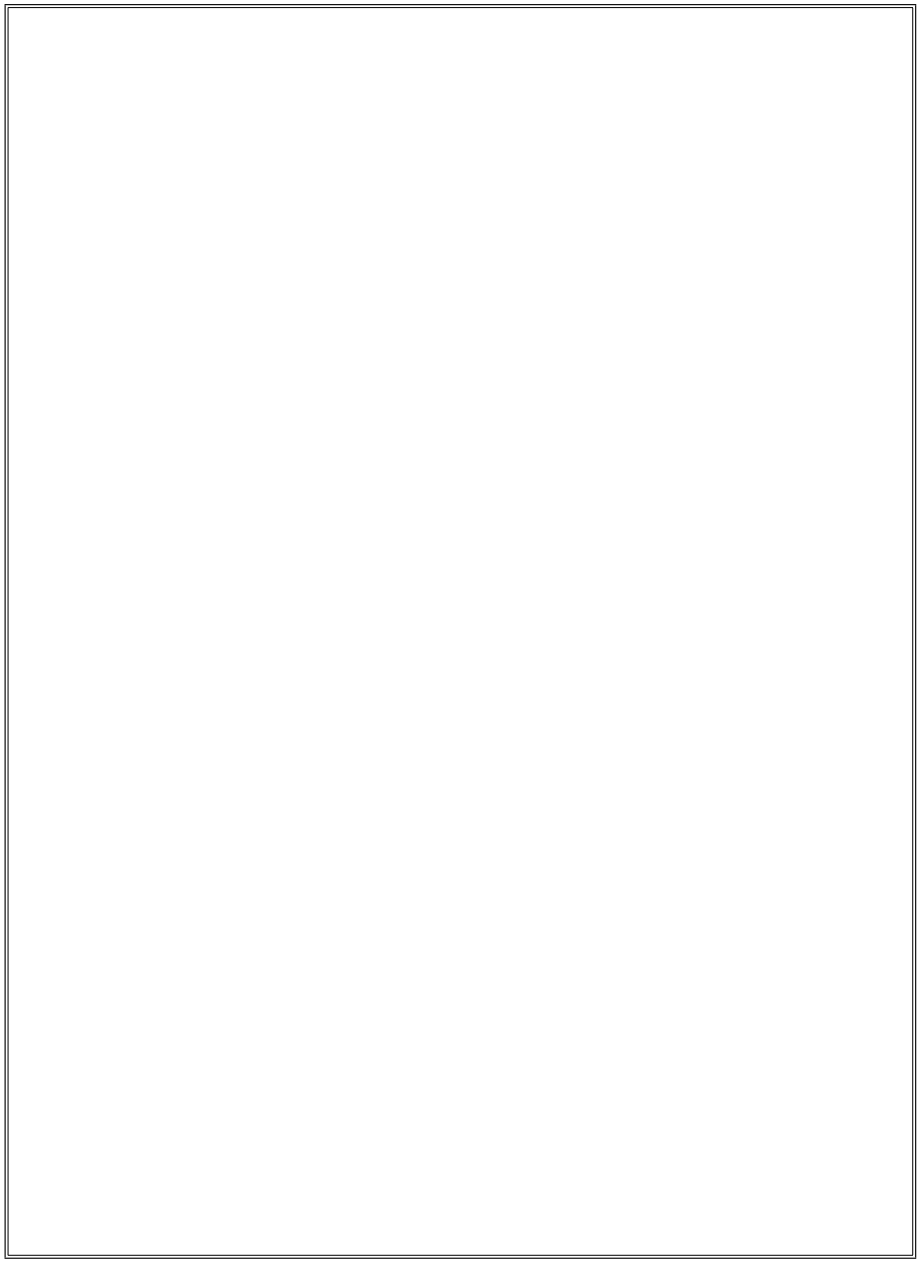
Theory:Write about components of hadoop



Course title: Big Data Analytics

Explain the step by step procedure of hadoop installation.

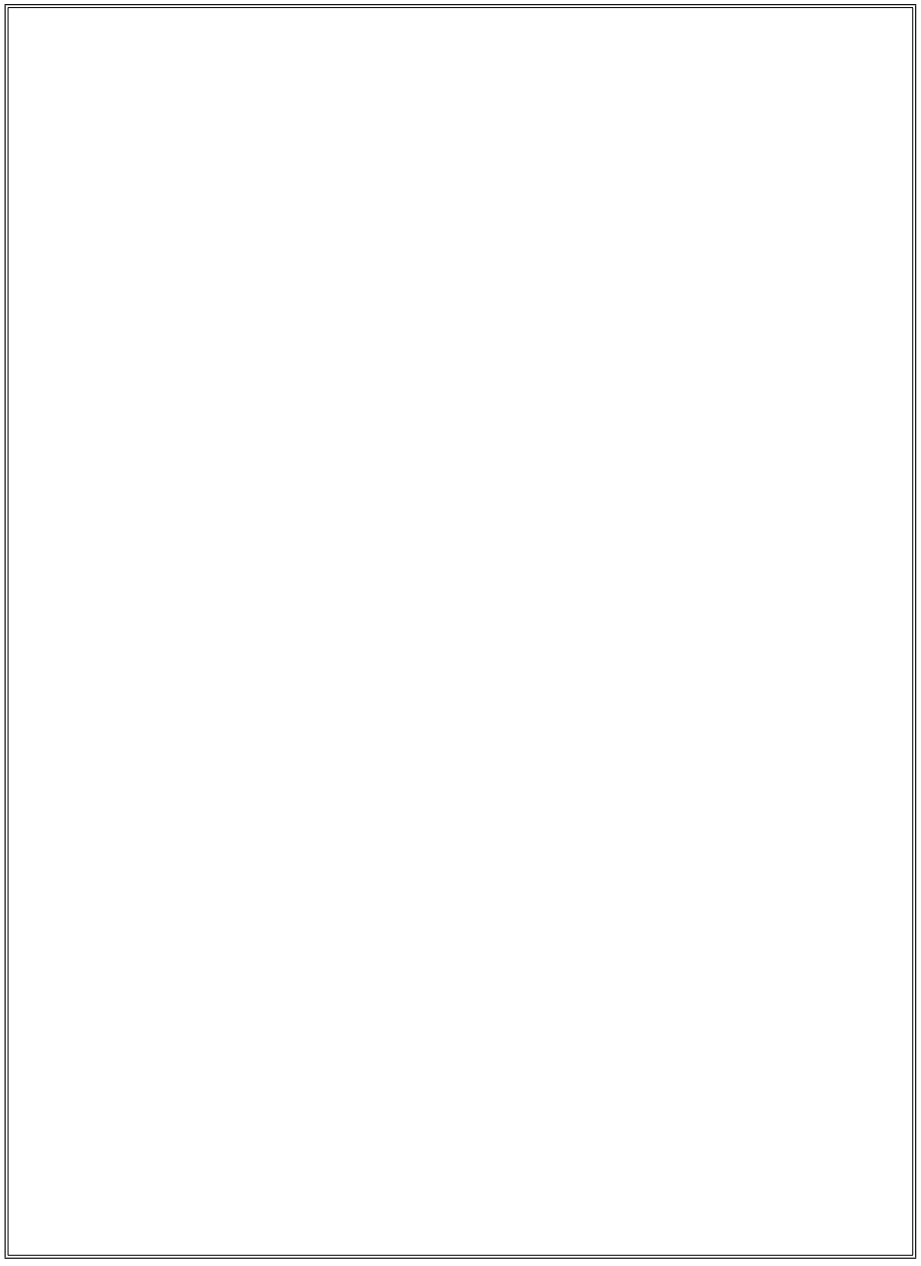
Course title: Big Data Analytics



Course title: Big Data Analytics

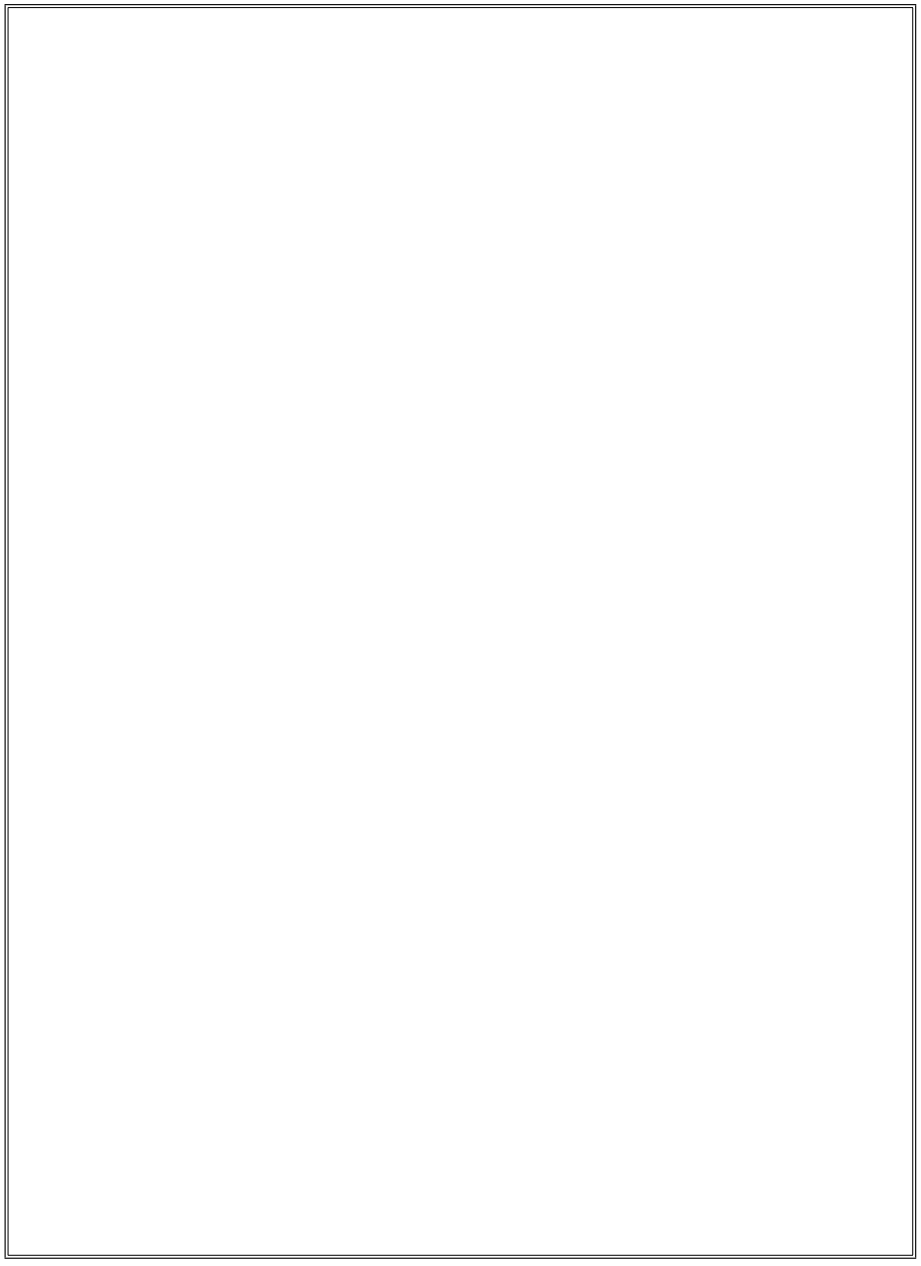
Explain the step by step procedure of installing and configuring eclipse plugin for hadoop.

Course title: Big Data Analytics



Course title: Big Data Analytics

PostLab:Explain hadoop ecosystem in detail.



Course title: Big Data Analytics

∗ ∗ ∗ ∗ ∗

EXPERIMENT 2

Fourier Transform

2.1 Introduction

Start...

2.1.1 Theory

Explain...

2.1.1.1 Example

Examples...

* * * * *

EXPERIMENT 3

Z-Transform

3.1 Introduction

Start...and & See Fig. 3.1 Refer (3.1)

$$E = mc^2 \tag{3.1}$$

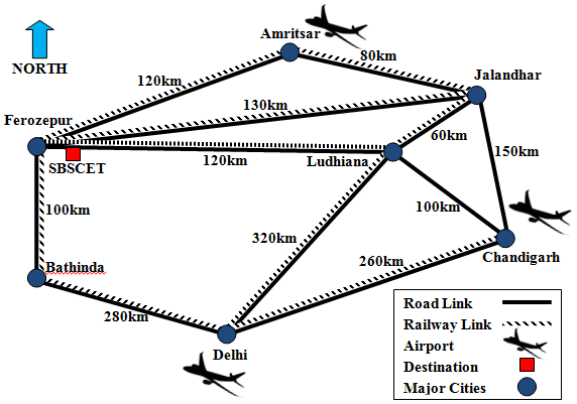


FIGURE 3.1: How to Reach at Ferozepur

My experimented problem gives $V = R_1 I_1 + R_2 I_2$ and $\mu\varphi\psi\omega \sqcap \sqcap \sqcap \sqcap \sqcap \bigstar$

3.1.1 Theory

Explain...

3.1.1.1 Example

Examples...

* * * * *

∗ ∗ ∗ ∗ ∗

BIOGRAPHY

Dr Satvir Singh Sidhu was born on Dec 7, 1975. He received his degree of Bachelor of Technology from Dr. B. R. Ambedkar National Institute of Technology, Jalandhar (PB) with specialization in Electronics & Communication Engineering in 1998, degree of Master of Engineering from Delhi College of Engineering with distinction in Electronics & Communication Engineering in 2000. He obtained his Doctoral degree from Maharshi Dayanand University, Rohtak (HR) in 2010. His professional experience includes the teaching as Assistant Professor and Head, Department of Electronics Engineering at BRCM College of Engineering & Technology, Bahal (HR). Presently, he is Head, Department of Electronics Engineering at Shaheed Bhagat Singh College of Engineering & Technology, Ferozepur (PB).

His fields of special interest include Nature Inspired Computational Intelligence, MATLAB Programming, LaTeX Typesetting, Analog and Digital Circuits and Systems, etc. He has published/communicated research papers in many international conferences and journals. He has conducted a number expert talks in AICTE sponsored short term courses, workshops, and international conferences. At present, he is actively involved in research activities and guiding Ph.D. and M.Tech. scholars. He has also completed AICTE sponsored: (a) MODROB project for Microcontroller and Embedded Systems Lab worth Rs. 5 Lac and (b) two-week Staff Development Programme on Intelligent Computational Techniques at SBS CET Ferozepur.