

SSN College of Engineering

Department of Computer Science and Engineering
ME(CSE) II Semester

PCP121 – Data Analytics Lab

Academic Year: 2018-19 EVEN

Batch: 2018-2020

Instructor: Dr. J. Suresh

Index

Ex. No	Date	Name of the Exercise	Completion Date
Part I – Hadoop			
1	13.02.2019	Learning Linux – Shell Programming – Network Commands	20.02.2019
2	20.02.2019	Installing & Working - Hadoop – Single Node	06.03.2019
3	06.03.2019	Running MR Programs – Wordcount / Word frequency	20.03.2019
4	20.03.2019	Installing & Working - Hadoop – Multinode Node	27.03.2019
5	27.03.2019	Running MR Programs – Wordcount / Word frequency	03.04.2019
6	03.04.2019	Running MR Programs – Process another dataset (E.g. Weather dataset)	10.04.2019
Part II – R Programming (5)			
1	10.04.2019	Warmup exercises – R programming	17.04.2019
2	24.04.2019	Implement Linear and Logistic Regression (Packages -glmnet, earth)	03.05.2019
3	08.05.2019	Implement SVM / Decision tree classification techniques - Packages (e1071, tree)	15.05.2019
4	15.05.2019	Implement Clustering techniques (Packages - cluster, factoextra, magrittr)	21.05.2019
5	22.05.2019	Visualize data using any plotting framework (Packages – ggplot)	27.05.2019
Part III – Application using NoSQL			
1	22.05.2019	Implement an application that stores big data in HBase / MongoDB using Hadoop.	27.05.2019