Anjuman-i-Islam's
M. H. Saboo Siddik College of Engineering
8, Saboo Siddik Polytechnic Road, Byculla, Mumbai-8

Internal Assessment Examination –I (IAE-I)
September 2023

Blue Print for Students Information

Subject Code : CSC702	Subject Name : Big Data Analysis	SEM : VII

Table 1: Course Objectives and Outcomes

Course Objectives	
CO1	To provide an overview of the big data platforms, its use cases and Hadoop ecosystem.
CO2	To introduce programming skills to build simple solutions using big data technologies such as MapReduce, Scripting for No SQL and R
CO3	To learn the fundamental techniques and principles in achieving big data analytics with scalability and streaming capability.
CO4	To enable students to have skills that will help them to solve complex real-world problems for decision support.
Course Outcomes	
CSC702.01	Understand the building blocks of Big Data Analytics.
CSC702.02	Apply fundamental enabling techniques like Hadoop and MapReduce in solving real world problems.
CSC702.03	Understand different NoSQL systems and how it handles big data.
CSC702.04	Apply advanced techniques for emerging applications like stream analytics.
CSC702.05	Achieve adequate perspectives of big data analytics in various applications like recommender systems, social media applications, etc.
CSC702.06	Apply statistical computing techniques and graphics for analyzing big data.

Table 2: Detail Syllabus Modules / Topic for IAE - I

Module No	Content / Sub Modules	CO	No of Hrs in Syllabus	% Marks Weight-age	% of Syllabus	Marks Range (Including Options) [20 - 32]	
1	Introduction to Big Data and Hadoop	1	2	13	5	3	4
	1.1 **Introduction to Big Data** - **Big Data Characteristics and Types of Big Data** 1.2 **Traditional vs. Big Data Business Approach** 1.3 **Case Study of Big Data Solutions** 1.4 Concept of Hadoop, **Core Hadoop Components; Hadoop Ecosystem**						

2	Hadoop HDFS and MapReduce	2	4	25	10	5	8
	2.1 **Distributed File Systems: Physical Organization of Compute Nodes, Large- Scale File-System Organization.** 2.2 MapReduce: The Map Tasks, Grouping by Key, The Reduce Tasks, Combiners, Details of MapReduce Execution, Coping With Node Failures. 2.3 **Matrix Multiplication or Word Count using Map Reduce**						
3	NoSQL	3	6	15	8	8	12
	3.1 Introduction to NoSQL, **NoSQL Business Drivers** 3.2 **NoSQL Data Architecture Patterns: Key-value stores, Graph stores, Column family (Bigtable)stores, Document stores,** Variations of NoSQL architectural patterns, NoSQL Case Study						
6	Data Analytics with R	6	4	10	5	5	8
	6.1 **Exploring Basic features of R**, Exploring RGUI, Exploring RStudio, **Handling Basic Expressions in R, Variables in R, Working with Vectors, Storing and Calculating Values in R, Creating and using Objects**, Interacting with users, Handling data in R workspace, Executing Scripts, **Creating Plots**, Accessing help and documentation in R 6.2 **Reading datasets and Exporting data from R, Manipulating and Processing Data in R**, Using functions instead of script, built-in functions in R 6.3 **Data Visualization**: Types, Applications						

Er. Ahlam A. Subject In-charge	Er. Fatima Ansari Exam Co-ordinator	Er. Fatima Ansari NBA Co-ordinator	Dr. Mohd. Ahmed Head of Computer Dept.

1. Example creation of plots
2. Write a R script to create histogram or other plot.
3. Explain dataframes and factors with example

Something on this lines will come in the paper.