



# ವಿಶ್ವೇಶ್ವರಯ್ಯ ತಾಂತ್ರಿಕ ವಿಶ್ವವಿದ್ಯಾಲಯ

("ವಿ ಟಿ ಯು ಅಧಿನಿಯಮ ೧೯೯೪" ರ ಅಡಿಯಲ್ಲಿ ಕರ್ನಾಟಕ ಸರ್ಕಾರದಿಂದ ಸ್ಥಾಪಿತವಾದ ರಾಜ್ಯ ವಿಶ್ವವಿದ್ಯಾಲಯ)



## VISVESVARAYA TECHNOLOGICAL UNIVERSITY

(State University of Government of Karnataka Established as per the VTU Act, 1994) "Jnana Sangama" Belagavi-590018, Karnataka, India)

**Prof. B. E. Rangaswamy, Ph.D.**  
REGISTRAR

Phone: (0831) 2498100

Fax: (0831) 2405467

REF: VTU/BGM/BOS/New UG-PG Prog/2023-24/4591

DATE: 4 DEC 2023

### CIRCULAR

**Subject:** Shifting of the 'Big Data' course code: 22MBABA405 from the 4<sup>th</sup> semester to the third semester in place of 'Marketing, Web and Social Media Analytics' course code: 22MBABA306 for Business Analytics specialization in MBA.

**Reference:** Chairperson & Members Board of Studies in MBA VTU Belagavi opinion dated 01.12.2023, VC sir approval dated : 04.12.2023

Returning to the subject above, this is to notify all MBA-affiliated colleges that the Board of Studies has added **Business Analytics** as one of the MBA specializations, expanding the options under CBCS. However, the Board of Studies in Management Studies believes it is wise to move the course (subject) "**Big Data**" code 22MBABA405, from the **fourth** semester to the **third** semester instead of the "Marketing, Web, and Social Media Analytics course" code 22MBABA306, due to some technical difficulties. However, this shall apply to the students taking Business Analytics as a core specialization.

**We kindly suggest that the section heads and teachers in question take note of the above adjustments and apply them as appropriate.**

Sd/-  
Registrar

To,

All the Principals of Affiliated /Constituent Engineering Colleges, where MBA program being offered.

The Chairperson / Program Coordinator, Department of Business Studies at Kalaburagai, Belagavi, Bengaluru and Mysuru

#### Copy to:

1. The Hon'ble Vice-Chancellor through the secretary to VC VTU Belagavi for information
2. The Registrar (Evaluation) VTU Belagavi for information and needful
3. The Director, ITI SMU, VTU Belagavi for information and request to make arrangements for uploading of this circular on the VTU web portal in the section of Circular/Notification @ [https:// vtu.ac.in/ en/ category / administration/](https://vtu.ac.in/en/category/administration/)
4. The Special Officer, QPDS Examination Section VTU Belagavi for needful.
5. Office Copy

R  
04/12/23 BE

REGISTRAR

7-

## BIG DATA

<b>Course Code</b>	<b>22MBABA306</b>	<b>CIE Marks</b>	<b>50</b>
<b>Teaching Hours/Week (L:P:SDA)</b>	<b>2:2:0</b>	<b>SEE Marks</b>	<b>50</b>
<b>Total Hours of Pedagogy</b>	<b>40</b>	<b>Total Marks</b>	<b>100</b>
<b>Credits</b>	<b>03</b>	<b>Exam Hours</b>	<b>03</b>
<b>Course Learning objectives:</b>			
<ul style="list-style-type: none"> <li>• Understand the Big Data Platform and its Use cases</li> <li>• Provide an overview of Apache Hadoop</li> <li>• Provide HDFS Concepts and Interfacing with HDFS</li> <li>• Understand Map Reduce Jobs</li> <li>• Exposure to Data Analytics with R.</li> <li>• Apply analytics on Structured, Unstructured Data.</li> </ul>			
<b>Module-1                    6 Hours</b>			
<b>INTRODUCTION TO BIG DATA AND HADOOP</b> :Types of Digital Data, Introduction to Big Data, Big Data Analytics, History of Hadoop, Apache Hadoop, Analysing Data with Unix tools, Analysing Data with Hadoop, Hadoop Streaming, Hadoop Echo System, IBM Big Data Strategy, Introduction to Info sphere Big Insights and Big Sheets.			
<b>Module-2                    7 Hours</b>			
<b>HDFS (Hadoop Distributed File System)</b> :The Design of HDFS, HDFS Concepts, Command Line Interface, Hadoop file system interfaces, Data flow, Data Ingest with Flume and Scoop and Hadoop archives, Hadoop I/O: Compression, Serialization, Avro and File-Based Data structures.			
<b>Module-3                    6 Hours</b>			
<b>Map Reduce</b> : Anatomy of a Map Reduce Job Run, Failures, Job Scheduling, Shuffle and Sort, Task Execution, Map Reduce Types and Formats, Map Reduce Features.			
<b>Module-4                    7 Hours</b>			
<b>Hadoop Eco System</b> :Introduction to PIG, Execution Modes of Pig, Comparison of Pig with Databases, Grunt, Pig Latin, User Defined Functions, Data Processing operators.  Hive: Hive Shell, Hive Services, Hive Metastore, Comparison with Traditional Databases, HiveQL, Tables, Querying Data and User Defined Functions. HBase :HBasics, Concepts, Clients,			
<b>Module-5                    7 Hours</b>			
<b>Data Analytics with R</b> : Introduction, Supervised Learning, Unsupervised Learning, Collaborative Filtering.Big Data Analytics with Big R.			
<b>Module-6                    7 Hours</b>			
<b>NoSQL , Searching and Indexing Big Data</b> : Structured and Unstructured Data, Taxonomy and NoSQL Implementation, Discussion of basic architecture of Hbase, Cassandra and MongoDB			

Full text Indexing and Searching, Indexing with Lucene, Distributed Searching with Elastic search.

### **Assessment Details (both CIE and SEE)**

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing marks for the CIE is 50% of the maximum marks. Minimum passing marks in SEE is 40% of the maximum marks of SEE. A student shall be deemed to have satisfied the academic requirements (passed) and earned the credits allotted to each course if the student secures not less than 50% in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together.

#### **Continuous Internal Evaluation:**

There shall be a maximum of 50 CIE Marks. A candidate shall obtain not less than 50% of the maximum marks prescribed for the CIE.

#### **CIE Marks shall be based on:**

- a) Tests (for 25Marks) and
- b) Assignments, presentations, Quiz, Simulation, Experimentation, Mini project, oral examination, field work and class participation etc., (for 25 Marks) conducted in the respective course. Course instructors are given autonomy in choosing a few of the above based on the subject relevance and should maintain necessary supporting documents for same.

#### **Semester End Examination:**

The SEE question paper will be set for 100 marks and the marks scored will be proportionately reduced to 50.

- The question paper will have 8 full questions carrying equal marks.
- Each full question is for 20 marks with 3 sub questions.
- Each full question will have sub question covering all the topics.
- The students will have to answer five full questions; selecting four full questions from question

#### **Suggested Learning Resources:Books**

1. Tom White "Hadoop: The Definitive Guide" Third Edit on, O'reily Media, 2012.
2. Seema Acharya, SubhasiniChellappan, "Big Data Analytics" Wiley 2015.
3. Michael Berthold, David J. Hand, "Intelligent Data Analysis", Springer, 2007.
4. Jay Liebowitz, "Big Data and Business Analytics" Auerbach Publications, CRC press (2013)

**Web links and Video Lectures (e-Resources):**

- [https://r.search.yahoo.com/\\_ylt=AwrKC2qQjvdij9UV\\_2.7HAX.; ylu=Y29sbwNzZzMEcG9zAzMEdnRpZAMEc2VjA3Ny/RV=2/RE=1660419857/RO=10/RU=https%3a%2f%2fwww.crayondata.com%2fdownload-12-free-ebooks-on-big-](https://r.search.yahoo.com/_ylt=AwrKC2qQjvdij9UV_2.7HAX.; ylu=Y29sbwNzZzMEcG9zAzMEdnRpZAMEc2VjA3Ny/RV=2/RE=1660419857/RO=10/RU=https%3a%2f%2fwww.crayondata.com%2fdownload-12-free-ebooks-on-big-)
- <https://www.youtube.com/watch?v=rvJgArru8dI>
- <https://www.coursera.org/courses?query=big%20data>
- <https://www.pdfdrive.com/big-data-books.html>

**Note:** The aforesaid links and study materials are suggestive in nature, they may be used with due regards to copy rights, patenting and other IPR rules.

**Skill Development Activities Suggested**

- Visit amazon website and do observe how often change the design the website and update information
- Understand Walmart strategies for competitive advantage using big data analytics

**Course outcome :At the end of the course the student will be able to :**

Sl. No.	Description	Blooms Level
C01	Understand Big Data and its Business Implications	L2
C02	Apply the knowledge of Hadoop and Hadoop Eco-System in big data analysis	L3
C03	Analyse the big data and provide data visualization and helps in decisions	L4
C04	Develop Big Data Solutions using Hadoop Eco System	L5
C05	Apply Machine Learning Techniques using R	L3

**Mapping of COS and Pos**

	PO1	PO2	PO3	PO4	PO5	PSO 1	PSO 2	PSO 3	PSO 4
C01	1				2	3			
C02		2	2				2		
C03				3		3		2	
C04		2		2			1		2
C05	1				2	3			
C06		2	2				2		