



LSSBM's  
**PADMASHRI DR. V.B. KOLTE COLLEGE OF ENGINEERING, MALKAPUR,  
DIST-BULDANA (M.S) - 443101**

(NAAC Accredited, Approved by AICTE-New Delhi, Government of Maharashtra, DTE, MSBTE, Mumbai and  
Affiliated to Sant Gadge Baba Amravati University)

Website - [www.coemalkapur.ac.in](http://www.coemalkapur.ac.in) Email:- [coemalkapur@rediffmail.com](mailto:coemalkapur@rediffmail.com)

**Summary Report**

**National Level One Week Short Term Training Program**

On

**“Big data with Hadoop”**

In association with

**IQAC Sant Gadge Baba Amravati University,**

**ISTE Local Chapter, VBKCOE, Malkapur**

**The Institution of Engineers (India) (IEI) Aurangabad Local Center,**

**Under the Association of Computer Science and Engineering (ACSE)**

**Date: 21th to 25th Feb 2023**

Department of Computer Science and Engineering of Padm. Dr. V. B. Kolte College of Engineering, Malkapur (VBKCOE) conducted **National Level One Week Short Term Training Program On “Big data with Hadoop”** In association with IQAC Sant Gadge Baba Amravati University, ISTE Local Chapter, VBKCOE, Malkapur, The Institution of Engineers (India) (IEI) Aurangabad Local Center, Under the Association of Computer Science and Engineering (ACSE) from 21th to 25th Feb 2023. We conducted this FDP programme which enriches the knowledge of faculties, research scholar, PG and UG students.

The number of faculties, research scholar, and PG students were participated for this FDP from various disciplines, and states from India. Total 238 participants joining the session.

Programme	Department of Computer Science and Engineering
Title of course	<b>Short Term Training Program On “Big data with Hadoop”</b>
Scheduled Date	21th to 25th Feb 2023
Programme Convener	Dr. A. W. Kharche Padm. Dr. V. B. Kolte College of Engineering Malkapur Dist-Buldhana (Maharashtra)
Head of Department	Prof. S.L. Farpat
Programme coordinator	Prof. H. S. Deshpande Prof. G. V. Patil



**LSSBM's**  
**PADMASHRI DR. V.B. KOLTE COLLEGE OF ENGINEERING, MALKAPUR,**  
**DIST-BULDANA (M.S) - 443101**

(NAAC Accredited, Approved by AICTE-New Delhi, Government of Maharashtra, DTE, MSBTE, Mumbai and  
Affiliated to Sant Gadge Baba Amravati University)

Website - [www.coemalkapur.ac.in](http://www.coemalkapur.ac.in) Email:- [coemalkapur@rediffmail.com](mailto:coemalkapur@rediffmail.com)

Technical Committee	Dr. P. A. Kharat (Associate Professor) Prof. A. D. Bhople (Assistant Professor) Prof. M. U. Karande (Assistant Professor) Prof. Y. B. Jadhao (Assistant Professor) Prof. S. G. Bonde (Assistant Professor)
---------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Objectives of STTP:

- To introduce participants to the basic concepts and techniques of Big Data.
- To develop skills of using recent technologies for solving practical problems.
- To gain experience of doing independent study and research.

### Content of STTP

- Overview of Big data
- Big Data Overview & its applications
- Data Visualization in Big data
- IoT and Big data
- Hadoop , HDFS , Types of Big data
- Hands on running word count on single node cluster of Hadoop

The STTP helps create awareness and deeper understanding of recent technology such as Big Data, Big Data with Hadoop among the faculties and researchers working in various disciplines of engineering and sciences. It also helps the participants to understand and apply these technologies through several case studies and real-life examples. It provide opportunities to faculty members, research scholars and post graduate students to enhance their knowledge in research.

### Details of Speaker

Sr. No.	Name of Speaker	Designation	Organization
1.	Dr. Pradeep M. Jawandhiya	Principal	Pankaj Laddhad Institute of Technology & Management Studies, Buldhana
2.	Dr. Satish Jadhao	Assistant Professor	B. N. College of Engineering , Pusad
3.	Prof. Ms. Mayuri Kulkarni	Assistant Professor	SVKM's Institute of Technology, Dhule
4.	Mr. Yogesh P. Murumkar	CEO & Cooperate Trainer	Bharat Soft Solutions , Pune
5.	Prof. S. S. Agrawal	Assistant Professor	College of Engineering & Technology , Akola



**LSSBM's**  
**PADMASHRI DR. V.B. KOLTE COLLEGE OF ENGINEERING, MALKAPUR,**  
**DIST-BULDANA (M.S) - 443101**

(NAAC Accredited, Approved by AICTE-New Delhi, Government of Maharashtra, DTE, MSBTE, Mumbai and  
Affiliated to Sant Gadge Baba Amravati University)

Website - [www.coemalkapur.ac.in](http://www.coemalkapur.ac.in) Email:- [coemalkapur@rediffmail.com](mailto:coemalkapur@rediffmail.com)

6.	Dr. Nitin S. More	Associate Professor	MIT School of Engineering , MIT ADT University , Pune
7.	Dr. P. A. Vikhar	Associate Professor	KCE College of Engineering , Jalgaon

**Inauguration:**

Prof. G. V. Patil, Asst. Professor, Computer Science and Engineering Department welcome to all participants, explaining objectives of STTP and given brief introduction of speaker about their professional experience. Dr. Y. A. Kharche, Dean Administration, shared views regarding this STTP and congratulated to organizing team.

**Contents**

1. Review & Introduction of Big Data
2. Big Data Analysis, Challenges & Opportunities
3. Big Data Research Areas
4. Big Data Open Problems
5. Big Data Applications
6. 10 reasons why Big Data Analysis is best carrier move
7. Conclusion

**Major Sources of Big Data**

Source	Type	Volume
Log Files	Text	100 TB
Web Logs	Text	100 TB
Database	Text	100 TB
Video	Text	100 TB
Audio	Text	100 TB
Image	Text	100 TB
Text	Text	100 TB
XML	Text	100 TB
JSON	Text	100 TB
CSV	Text	100 TB
Excel	Text	100 TB
PDF	Text	100 TB
PowerPoint	Text	100 TB



**LSSBM's**  
**PADMASHRI DR. V.B. KOLTE COLLEGE OF ENGINEERING, MALKAPUR,**  
**DIST-BULDANA (M.S) - 443101**

(NAAC Accredited, Approved by AICTE-New Delhi, Government of Maharashtra, DTE, MSBTE, Mumbai and  
 Affiliated to Sant Gadge Baba Amravati University)

Website - [www.coemalkapur.ac.in](http://www.coemalkapur.ac.in) Email:- [coemalkapur@rediffmail.com](mailto:coemalkapur@rediffmail.com)

**Detail Schedule of STTP**

Date	Code	Time	Topics	Name of Speaker	Designation
<b>Day 1</b>					
<b>Inauguration Session</b>					
21 Feb 2023	1A	11.00 to 12.30 pm	Overview of Big data	Dr. Pradeep M. Jawandhiya	Principal, Pankaj Laddhad Institute of Technology & Management Studies, Buldhana
	1B	01:00 to 2:30 pm	Big data Processing tools an overview	Dr. P. A. Vikhar	Assistant Professor , KCE College of Engineering , Jalgaon
	1C	02:30 to 03:30 pm	Assignment and Feedback	-	-
<b>Day 2</b>					
22 Feb 2023	2A	11.00 to 12.30 pm	Data Visualization in Big data	Dr. Satish Jadhao	Assistant Professor B. N. College of Engineering , Pusad
	2B	01:00 to 2:30 pm	IoT and Big data	Prof. Ms. Mayuri Kulkarni	Assistant Professor SVKM's Institute of Technology, Dhule
	2C	02:30 to 03:30 pm	Assignment and Feedback	-	-
<b>Day 3</b>					
23 Feb 2023	3A	11.00 to 12.30 pm	Hadoop , HDFS , Types of Big data	Mr. Yogesh P. Murumkar	CEO & Cooperate Trainer Bharat Soft Solutions , Pune
	3B	01:00 to 2:30 pm	Hands on running word count on single node cluster of Hadoop	Mr. Yogesh P. Murumkar	CEO & Cooperate Trainer Bharat Soft Solutions , Pune
	3C	02:30 to 03:30 pm	Assignment and Feedback	-	-
<b>Day 4</b>					
24 Feb 2023	4A	11.00 to 12.30 pm	Big data and its Applications	Prof. S. S. Agrawal	Assistant Professor , College of Engineering & Technology , Akola
	4B	01:00 to 2:30 pm	Big data and its Applications	Prof. S. S. Agrawal	Assistant Professor , College of Engineering & Technology , Akola
	4C	02:30 to 03:30 pm	Assignment and Feedback	-	-



LSSBM's  
**PADMASHRI DR. V.B. KOLTE COLLEGE OF ENGINEERING, MALKAPUR,  
DIST-BULDANA (M.S) - 443101**

(NAAC Accredited, Approved by AICTE-New Delhi, Government of Maharashtra, DTE, MSBTE, Mumbai and  
Affiliated to Sant Gadge Baba Amravati University)

Website - [www.coemalkapur.ac.in](http://www.coemalkapur.ac.in) Email:- [coemalkapur@rediffmail.com](mailto:coemalkapur@rediffmail.com)

Day 5					
25 Feb 2023	5A	11.00 to 12.30 pm	Inside of Big Data and Hadoop as per Industry Prospective	Dr. Nitin S. More	Associate Professor , MIT School of Engineering , MIT ADT University , Pune
	5B	01:00 to 2:30 pm	Hadoop Hands-on	Dr. Nitin S. More	Associate Professor , MIT School of Engineering , MIT ADT University , Pune
	5C	12:30 to 01:00 pm	Assignment and Feedback	-	-
		01:00 to 01 :30 pm	<b>Valedictory Session</b>		

**Number of participants Registered: 238**

**Outcomes of STTP:**

- Participants familiar with new technology like Big Data ,Hadoop
- Researchers get more clear idea about innovation.
- Research scholars and post graduate student's knowledge is upgraded with new area for research and development.
- Enhancement in academic excellence

**Valedictory:**

Prof. G. V. Patil, Asst. Professor, Compute Science and Engineering Department, delivered vote of thanks to all participants, Resource Persons for sharing their professional experience, Knowledge and skills. The vote of thanks also delivered towards Management and Principal Dr. A. W. Kharche sir, our associated bodies-ISTE, IEI, organizing team members who directly/indirectly involving program activities for successful completion of STTP Programme.

**Link to join all the sessions –**

Join Zoom Meeting

<https://us06web.zoom.us/j/82609473457?pwd=amZVRzZSZ2pTeDJrN2loZTBnandLdz09>

Meeting ID: 826 0947 3457

Passcode: 12345





LSSBM's  
**PADMASHRI DR. V.B. KOLTE COLLEGE OF ENGINEERING, MALKAPUR,  
DIST-BULDANA (M.S) - 443101**

(NAAC Accredited, Approved by AICTE-New Delhi, Government of Maharashtra, DTE, MSBTE, Mumbai and  
Affiliated to Sant Gadge Baba Amravati University)

Website - [www.coemalkapur.ac.in](http://www.coemalkapur.ac.in) Email:- [coemalkapur@rediffmail.com](mailto:coemalkapur@rediffmail.com)

## Glimpses of STTP



LSSBM's  
**PADMASHRI DR. V.B. KOLTE COLLEGE OF ENGINEERING,  
MALKAPUR, DIST-BULDANA- 443101**

(Approved by AICTE-New Delhi, Government of Maharashtra, DTE, MSBTE, Mumbai and Affiliated to S.G.B. Amravati University)

Department of Computer Science and Engineering

Organize

**One Week National Level  
Short Term Training Program (STTP)  
on  
Big Data with Hadoop**

In association with IQAC Sant Gadge Baba Amravati University,  
ISTE Local Chapter, VBKCOE, Malkapur  
The Institution of Engineers (India) (IEI) Aurangabad Local Center  
Under the Association of Computer Science and Engineering (ACSE)

21/02/2023 to 25/02/2023

**Tools for Data Visualization.**

- 1 – Tableau
- **Tableau** is one of the leading software of the data visualization market, enabling various types of graphs creation, charts, dashboards, stories, maps and other elements without programming.
- It includes various functions of descriptive and inferential statistics with analytical charts generation. Allows other tools integration such as Excel, SQL, SAP, Amazon, and others.
- Other alternatives to Tableau, are **Qlik** and **Chartio**, which we will not describe here, but that allow visual data analysis without using programming processes.



LSSBM's  
**PADMASHRI DR. V.B. KOLTE COLLEGE OF ENGINEERING, MALKAPUR,  
DIST-BULDANA (M.S) - 443101**

(NAAC Accredited, Approved by AICTE-New Delhi, Government of Maharashtra, DTE, MSBTE, Mumbai and  
Affiliated to Sant Gadge Baba Amravati University)

Website - [www.coemalkapur.ac.in](http://www.coemalkapur.ac.in) Email:- [coemalkapur@rediffmail.com](mailto:coemalkapur@rediffmail.com)

Zoom Meeting interface showing a grid of participants. A notification at the top states: "A participant has enabled Closed Captioning. Who can see this transcript?". The grid includes participants such as DR A KHARCHE, Kiran bonde, Dr. Najmuddin A..., Sonali chavan, Bhavneet, shabad kaur, Dr. Syed Gilani P..., Vijaya Aher, suriyapriya, Vidya Nemade, Basavaraj S Mat..., Simrandeep kaur, Ramanpreet, Prof. S. Joshi (E&C), Rohini Baingane, Diganta Baihiya, Md Shafee, Lokesh Khedekar, Bhavesh Borole, Farisa Fatema BE..., Rekha Dnyanesh..., Swapnil patil, and Goutam Kamble. A participant has enabled Closed Captioning.

Zoom Meeting interface showing a presentation slide titled "Features of Hadoop". The slide lists 8 features:

1. Cost-effective
2. Large cluster of nodes.
3. Parallel processing
4. Distributed data
5. Automatic failover management.
6. Data locality optimization
7. Heterogeneous cluster
8. Scalability

A circular diagram with an elephant icon in the center illustrates these features: Scalability, The large Cluster of nodes, Parallel processing, Distributed data, Automatic failover management, Data locality optimization, Heterogeneous cluster, and Cost-effective.



LSSBM's  
**PADMASHRI DR. V.B. KOLTE COLLEGE OF ENGINEERING, MALKAPUR,  
DIST-BULDANA (M.S) - 443101**

(NAAC Accredited, Approved by AICTE-New Delhi, Government of Maharashtra, DTE, MSBTE, Mumbai and  
Affiliated to Sant Gadge Baba Amravati University)

Website - [www.coemalkapur.ac.in](http://www.coemalkapur.ac.in) Email:- [coemalkapur@rediffmail.com](mailto:coemalkapur@rediffmail.com)

**Major Sources of Big Data**

	Video	Image	Audio	Text/numbers
Banking	High	Medium	Medium	High
Insurance	Medium	Medium	Medium	Medium
Securities and investment services	Medium	Medium	Medium	Medium
Discrete manufacturing	Medium	Medium	Medium	Medium
Process manufacturing	Medium	Medium	Medium	Medium
Retail	Medium	Medium	Medium	Medium
Wholesale	Medium	Medium	Medium	Medium
Professional services	Medium	Medium	Medium	Medium
Consumer and recreational services	Medium	Medium	Medium	Medium
Health care	Medium	Medium	Medium	Medium
Transportation	Medium	Medium	Medium	Medium
Communications and media?	Medium	Medium	Medium	Medium
Utilities	Medium	Medium	Medium	Medium
Construction	Medium	Medium	Medium	Medium
Resource industries	Medium	Medium	Medium	Medium
Government	Medium	Medium	Medium	Medium
Education	Medium	Medium	Medium	Medium

**Data ?**

- Data can be defined as a representation of facts, concepts, or instructions in a formalized manner, which should be suitable for communication, interpretation, or processing by human or electronic machine.
- In computing, data is information that has been translated into a form that is efficient for movement or processing.
- Relative to today's computers and transmission media, data is information converted into binary digital form.

**Structured Data**

**Unstructured Data**

H. O. D.  
Dept. of Computer Sci & Engg.  
Dr. V.B. Kolte College of Engineering,  
Malkapur, Dist. Buldana. (M.S.)