



**LAKIREDDY BALI REDDY COLLEGE OF ENGINEERING
(AUTONOMOUS)**

L.B. REDDY NAGAR, MYLAVARAM – 521230. A.P. INDIA
Affiliated to JNTUK Kakinada & Approved by AICTE, New Delhi
Accredited By NAAC, Accredited By NBA Tier-I & Certified by ISO 9001:2015
<http://www.lbrce.ac.in>, Phone: 08659 – 222933, Fax: 08659 – 222931 Extn:109

DEPARTMENT OF INFORMATION TECHNOLOGY

Report on “Industry Expectation on Data Engineers”

Event Type: Workshop (Online)

Date / Duration: 4th – 6th June,2021, Three days

Resource Person: Mr. Battu Mahesh Reddy

Lead Software Engineer,

Epam System India.

Name of Coordinator(s): Dr. Satuluri Naganjaneyulu, Professor., Dept. Of IT, LBRCE-Mylavaram.

Co-Cordinators: Dr.B. Rama Devi, Associate Professor., Dept. of IT, LBRCE-Mylavaram.

Mr. Annapareddy V N Reddy, Assistant Professor, Dept.of.IT, LBRCE-Mylavaram.

Target Audience: B.Tech. (Second year& Third year)

Students Total no of Participants: Internal – 200 External-55

Objective of the event: To extend practical knowledge in Industry Expectation

Outcome of event:

At the end of the event students can able to Understand

- Develop, construct, test and maintain architectures
- Align architecture with business requirements
- Data acquisition
- Develop data set processes
- Use programming language and tools
- Identify ways to improve data reliability, efficiency and quality
- Conduct research for industry and business questions
- Use large data sets to address business issues
- Deploy sophisticated analytics programs, machine learning and statistical methods
- Prepare data for predictive and prescriptive modelling
- Find hidden patterns using data
- Use data to discover tasks that can be automated
- Deliver updates to stakeholders based on analytics

Description / Report on Event:

The three-day workshop conducted through Google Meet on the first day and through GoToMeeting on remaining 2 days from 10.00A.M to 12.00.P.M

Day 1 (04/06/2021):

Data engineers are responsible for finding trends in data sets and developing algorithms to help make raw data more useful to the enterprise. This IT role requires a significant set of technical skills, including a deep knowledge of SQL database design and multiple programming languages. But data engineers also need communication skills to work across departments to understand what business leaders want to gain from the company's large datasets.

Day 2 (05/06/2021): Next, in day-2 sessions covered the following topics:

According to Dataquest, there are **three main roles** that data engineers can fall into. These include:

- **Generalist:** Generalists are typically found on small teams or in small companies. In this setting, data engineers wear many hats as one of the few “data-focused” people in the company. Generalists are often responsible for every step of the data process, from managing data to analysing it. Dataquest says this is a good role for anyone looking to transition from data science to data engineering, since smaller businesses won't need to worry as much about engineering “for scale.”
- **Pipeline-centric:** Often found in midsize companies, pipeline-centric data engineers work alongside data scientists to help make use of the data they collect. Pipeline-centric data engineers need “in-depth knowledge of distributed systems and computer science,” according to Dataquest.
- **Database-centric:** In larger organizations, where managing the flow of data is a full-time job, data engineers focus on analytics databases. Database-centric data engineers work with data warehouses across multiple databases and are responsible for developing table schemas.

Day 3 (06/06/2021): Last day, regular sessions, quiz, feedback and valedictory completed the

topics covered in the last sessions include:

The skills on your resume might impact your salary negotiations — in some cases by more than 10 or 15 percent, depending on the skill. According to data from , the following data engineering skills are associated with a significant boost in reported salaries:

- Scala:
- Apache Spark:
- Data warehouse:
- Java:
- Data modelling:
- Apache Hadoop:
- Linux:
- Amazon Web Services (AWS):
- ETL (extra, transform, load):
- Big data analytics:
- Software development:

Feedback / Suggestions: The feedback survey (through Google sheets) was completed by 200 attendees, a response rate of about 75 percent. The survey’s findings on outcomes, satisfaction, overall workshop feedback, and recommendations for improvement.

S.No	Content	Excellent	Good	Normal
1	Content covered and Time management	98	100	2
2	Workshop fulfilled your Expectations.	115	82	3
3	Workshop was applicable to your career.	112	80	8
4	Presenters speak clearly and audibly	131	65	4
5	Knowledge gained	102	94	4
6	Command of expert over the content.	114	82	4
7	Worthiness of workshop	106	88	6
8	Overall Workshop rating.	128	66	5

- Suggested to organize this kind Workshops frequently.
- Recommendations for improving future workshops by providing more time, background information, and documentation.

Photographs :(Max 10 limited)

Day-1(Inaguration)

This screenshot shows a Zoom meeting interface. At the top, there is a grid of 25 participant icons, with 'MB' highlighted. Below the grid, a presentation slide is displayed. The slide title is 'Lakireddy Bali Reddy College of Engineering (AUTONOMOUS)'. The main content of the slide reads: 'Three days Workshop on Industry Expectation on Data Engineers 04/06/2021 to 06/06/2021'. Below this, it says 'By: Mr. Battu Mahesh Reddy, Lead Software Engineer, Epam System India' and 'Organized by Department of Information Technology'. A 'You're muted.' notification box is visible in the bottom left of the presentation area. On the right side of the Zoom window, there are controls for 'Take a screenshot', '50%' zoom, and a 'Leave' button.

This screenshot shows a Zoom meeting interface, similar to the one above. The participant grid at the top shows 'MB' highlighted. The presentation slide is on slide 6 of 8. The slide content includes a diagram comparing 'Batch Processing' (30 Min) and 'Real-Time Processing' (Less Than 1 Sec). The diagram shows 'Information' and 'Employee' leading to 'System' for both processes. A text box at the top of the slide says 'Click to add title'. A 'You're muted.' notification box is present in the bottom left. The Zoom controls on the right include 'Take a screenshot', '50%' zoom, and a 'Leave' button.

Day-2:

Contents:

- 1. why Big Data solutions?
- 2. Architecture Diagram
- 3. Batch & Real time processing
- 4. Analytics
- 5. Use cases for Batch & real time.

AVR	MB	IND	TOR	R	THO	YOR	TKR	TMA	T	TE	TKA	TSR	VNB	T	TA	T	T
TBP	BD	T	TPA	T	T	TBL	T	2P	TAL	TK	T	SPG	T	HS	PVR	TS	T
T	T	VA	MC	T	BR	TJ	T	XD	T	FB	C	T	TGA	TGD	T	TP	ORW
TPV	ISM	T	TK	TDB	TNP	T	TP	2	TKL	SL	AG	NA	TAT	MC	TC	CVK	T
NV	TK	TKY	2SL	SPN	TBV	TKK	EC	T	T	R	M	MT	VIT	T	PLS	TBB	SM
UK	B	TS	TDB	T	TGR	TY	TPOL	T	TPV	SE	TK	PV	GLS	T	TWP	T	SS

Real-time Use Cases of Hadoop

- Finance
- Security & Law
- Government
- Retail
- Sentiment Analysis
- Healthcare
- Advertisements
- Understanding Customers

Day-3(Feedback and Valedictory)

The screenshot shows a Zoom meeting interface. At the top, there is a grid of 25 participants, with 'MB' highlighted in the first row. Below the grid is a presentation window displaying a slide with a yellow background and the text: "Data Migration – Batch Uber/Ola and sensor data – Real time". The slide is part of a 9-slide presentation. The Zoom control bar at the bottom shows 'MaheshReddy Battu is presenting' and buttons for Mic, Camera, Screen, and Leave. The system tray at the bottom right shows the time as 11:18 AM on 6/9/2021.

2.

The screenshot shows a Zoom meeting interface. At the top, there is a grid of 25 participants, with 'MB' highlighted in the first row. Below the grid is a presentation window displaying a slide titled "Analytics". The slide features a flowchart for "BIG Data Analytics" that illustrates the process from "Big Data" to "Deep Analysis". The flowchart includes steps such as "Data Collection", "Data Cleaning", "Data Integration", "Data Transformation", "Data Modeling", "Data Mining", and "Data Visualization". The Zoom control bar at the bottom shows 'MaheshReddy Battu is presenting' and buttons for Mic, Camera, Screen, and Leave. The system tray at the bottom right shows the time as 11:21 AM on 6/9/2021.

Coordinator

Dr. Satuluri Naganjaneyulu

Dr.B. Rama Devi

Mr.A.V.N.Reddy

Convener

(Dr. B. Srinivasa Rao)