



KKR & KSR INSTITUTE OF TECHNOLOGY & SCIENCES

(Approved by AICTE, Delhi, Affiliated to JNTUK, Kakinada)

Department of Computer Science and Engineering

A ONE Day Workshop On BIG DATA ANALYTICS

Date: 20th July, 2016

Event: A One Day Workshop

Time: 10:00 AM to 3:30 PM

Venue: KKR & KSR INSTITUTE OF TECHNOLOGY & SCIENCES, Vinjanampadu.

Organized By: CSE Department

Event Description:

Big data is a collection of data sets so large and complex that it becomes different to process using on-hand database management tools. The challenges include capture, creation, storage search, sharing, analyze and visualization. Big data is the relation of greater business intelligence by storage, processing and analyzing data that was previously ignored due to the limitations of traditional data management technologies.

Session Activities:

The Session started with the basics of the data, Lots of data, Dimensions of the data.



Lots of data:

- 2.5 quintillion bytes of data are generated every day.
 - A quintillion is 10^{18}
- Data come from:
 - Social media sites
 - Sensors
 - Digital photos
 - Business transactions
 - Location based data



After she explains that Big data can be analyzed with the software tools commonly used as part of advanced analytics disciplines such as predictive analytics, data mining, text analytics and statistical analysis. Mainstream BI software and data visualization tools can also play a role in the analysis process.



Four Dimensions of Big Data:

- Volume: large volumes of data
- Velocity: Quickly moving data.
- Variety: Structured, unstructured, image etc.
- Veracity: trust and integrity is a challenge and a must and is important for big data just as for traditional relational DBMS
- Variety: Variety refers to the increasingly diversified sources and types of data requiring management and analysis. We used to store data from sources like spreadsheets and databases. Now data comes in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc.
- Validity: Like big data veracity, validity means the correct and accurate data for the intended use. The validity of big data sources and subsequent analysis must be accurate, if you are to use the results for decision making.



