CHANDIGARH UNIVERSITY

Discover. Learn. Empower.

Apex Institute of Technology

Department of Computer Science & Engineering

Bachelor of Engineering (Computer Science & Engineering)

Big Data Analytics and IoT– (CST-432)

Prepared By: Ms. Bhanu Priyanka.Valluri (E8243)

DISCOVER . LEARN . EMPOWER

Big Data Analytics in IoT

Course Outcome:

Upon successful completion of this course, students will be able to:

CO Number	Title	Level
CO1	Basics of big data analytics concepts with respective to IoT along with their challenges	Understand
CO2	Implement the concepts for the development of smart systems.	Understand
CO3	Understand the use big data tools to process IoT data in various fields of communication by find a solution.	Understand

Will be covered in this lecture

Welcome to the session of

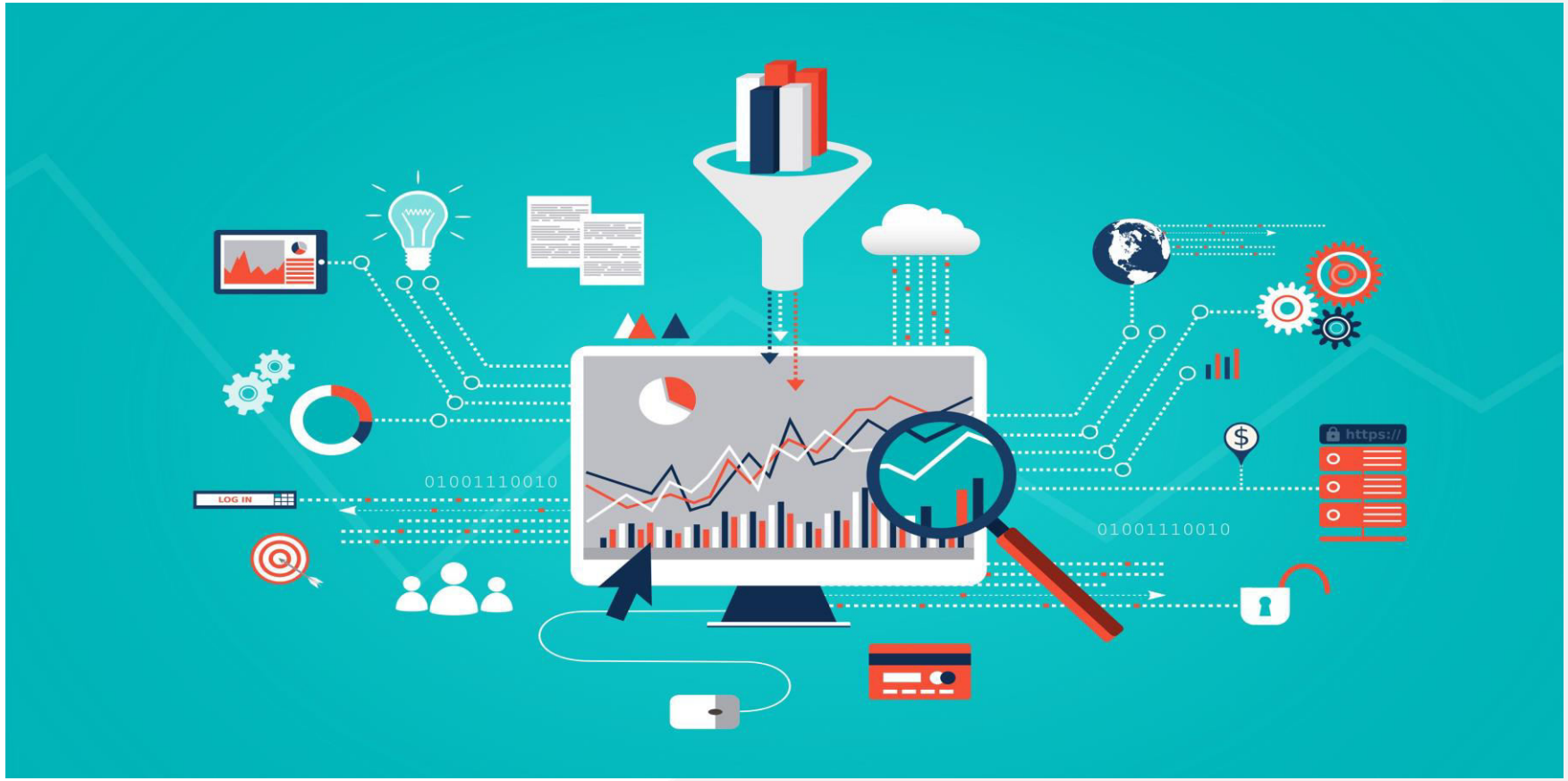
Unit-1: *Introduction to Big Data analytics in IoT Domain*

Chapter-1: Overview of Bigdata

Link: https://www.btelligent.com/fileadmin/_processed_/e/a/csm_big-data-datenquellen_781325f164.png

...Activity...

What do you find in ?



Agenda

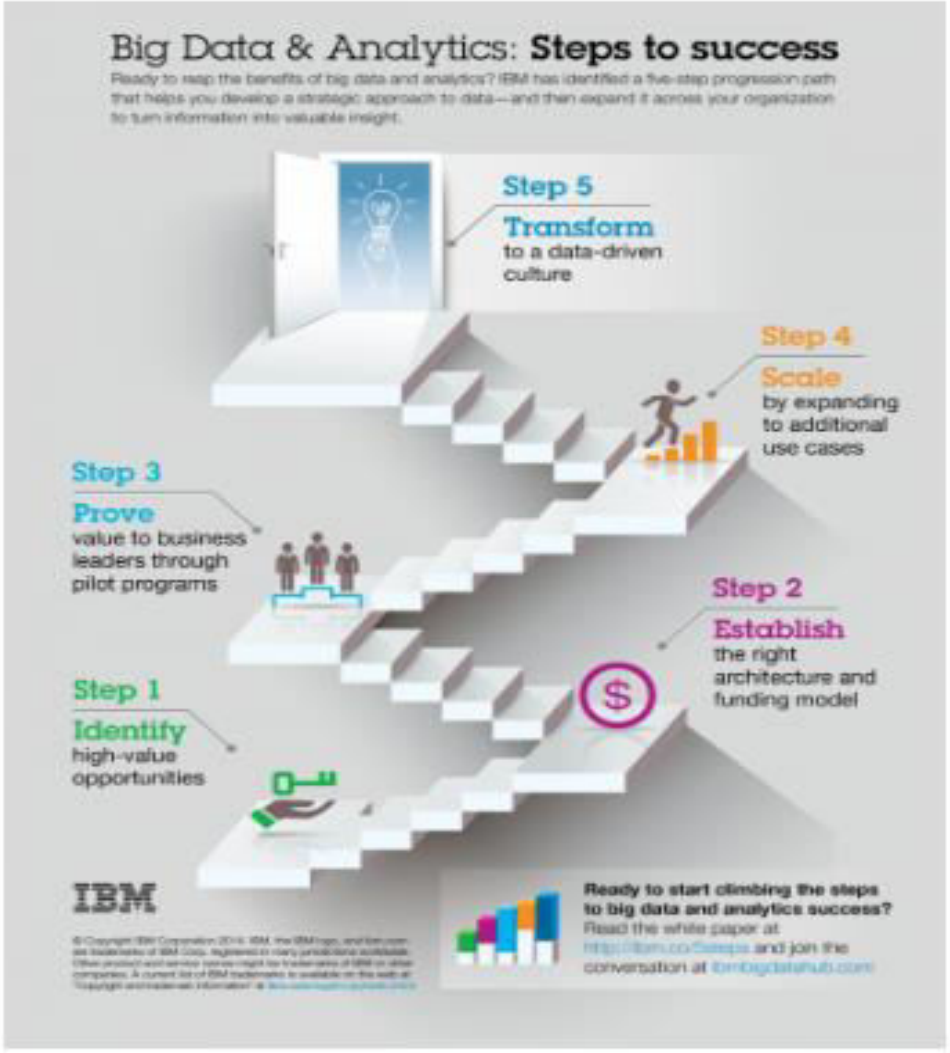
Chapter:1 __Lecture: 2

- Adoption approach,
- Data modelling,
- Implementation challenges,
- Big data user considerations,
- Big data market

Link: https://www.orangemantra.com/blog/java-plays-evolutionary-role-big-data-iot/

Adoption approach

- It provides pointers of execution plans and related challenges while adopting big data analytics.
- Adoption of big data analytics can provide firms various advantages like, it can
  - Diminish dormancy by a request of size,
  - Providing openness to data in minutes or seconds rather than hours or days,
  - It increases the ability to store data by a request of extent, moving from TBs to PBs [19–22],
  - It offers a much lower cost of acquisition and operation, the cost, is reduced by an order of magnitude as it requires less administrators.



Link :-https://cloudtweaks.com/2014/07/big-data-analytics-adoption/

Adoption approach

- There are two approaches of adopting Big Data Analytics into processes viz.
- The revolutionary approach and
- The evolutionary or hybrid approach

Revolutionary method Adoption approach

- The revolutionary method includes making a fresh out of the box new Big Data Analytics environment.
- It moves every one of the data to the new environment, and all reporting, modeling, and integration with business forms occurs in the new environment.
- This approach has been received by numerous Greenfield investigation driven organizations.

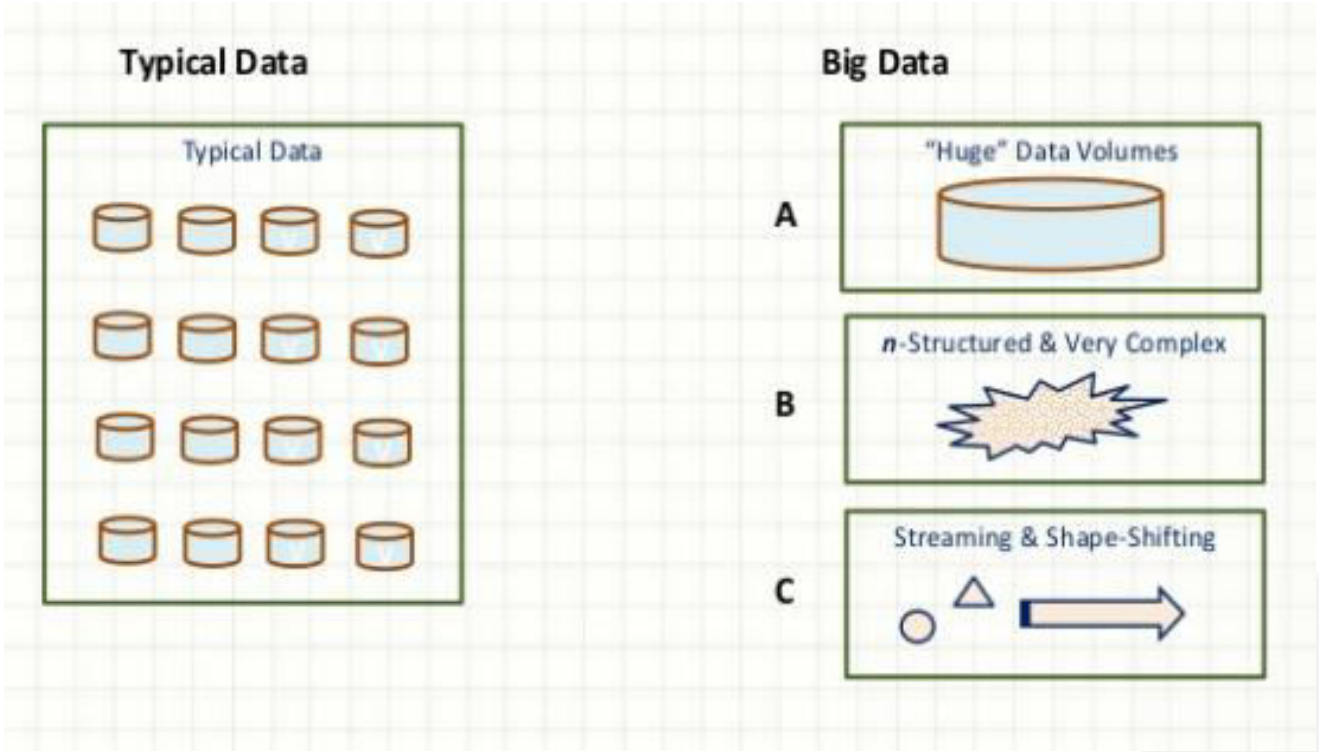
Evolutionary method Adoption approach

- Evolutionary method, Big Data turns into an input to the present BIstage.
- The data is amassed and broke down using organized and unstructured instruments, and the outcomes are sent to the data distribution center.
- Standard modeling and reporting instruments currently approach online networking estimations, use records, and other prepared Big Data thing requires sampling and processing Big Data to hold the stockroom from the gigantic volumes.

Evolutionary method Adoption approach

- The evolutionary approach has been received by develop BI organizations.
- The design has an ease section edge and in addition minimal effect on the Bi organization.
- However it can’t give the noteworthy upgrades seen by the Greenfield Administrators.
- In many cases, the kind of examination and the general end-to-end speed is constrained by the BI environment.
- This approach advanced effectively by IBM’s Information Agenda group puts the AAP design over existing BI infrastructure.
- All the Big Data courses through AAP, while conventional sources continue to give data to the data stockroom

...Activity..

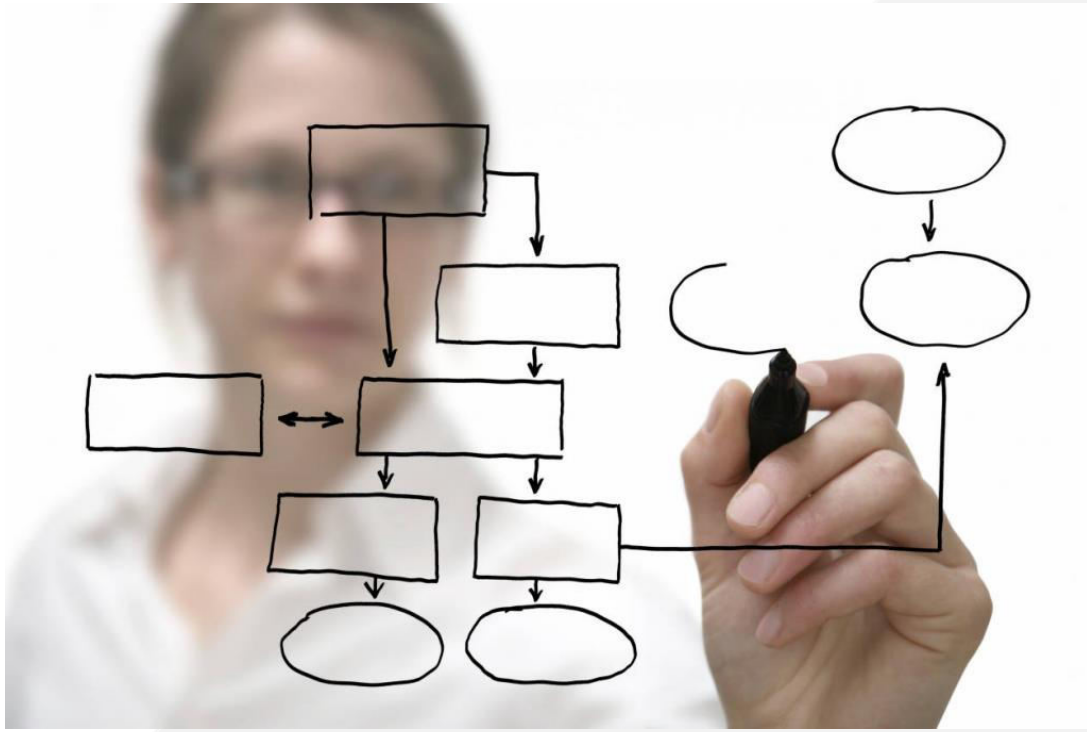


What do you understand from the fig...

Data modelling

----We always search for meanings , is in it???

- Process of creating a data model for an information system by applying formal data modeling techniques.
- Process used to define and analyze data requirements needed to support the business processes.
- Therefore, the process of data modeling involves professional data modelers working closely with business stakeholders, as well as potential users of the information system.



Link: https://www.rt-consulting.com/data-modeling

Types of Data Modeling

Conceptual data: Describes WHAT the system contains

Logical Data: describes HOW the system will be implemented, regardless of the DBMS

Physical Data: describes HOW the system will be implemented using a specific DBMS.

Importance of Data Modelling

- Are a communication tool
- Give an overall view of the database
- Organize data for various users
- Are an abstraction for the creation of good database

What is Data Modelling ???

- Data Model is a collection of conceptual tools for describing data, data relationships, data semantics and consistency constraint.
- A data model is a conceptual representation of data structures required for data base and is very powerful in expressing and communicating the business requirements.
- A data model visually represents the nature of data, business rules governing the data, and how it will be organized in the database

Implementation Challenges

There are three broad categories

- Merging internal and external data
- Big Data veracity
- Information lifecycle management

Link;-https://www.hiddenbrains.com/blog/big-data-analytics-challenges-implementation.html

Merging internal and external data

- Due to the access of more and more data, its now possible to firms to tap on consumers likes and dislikes, their demographics and tastes and commerce needs.
- It becomes necessary to understand that this information should be fully utilized by organizations to grow their businesses and for that they should merge this data with already present information with them.
- While merging this newly acquired data, they should be careful and closely monitor how the data is being used and how it is being aggregated.
- All this occurs as we radically change the rules on data privacy, redefine MDM, and encounter new concerns relating to data quality.

Big Data veracity

- Customer data comes from a variety of “biased” samples with different levels of data quality.
- It is vital to homogenize this data for the optimum usage.
- As it is homogenized, we must establish confidence levels on raw data, as well as aggregations and inferences, in order to understand and remind users of the“biases “built into the sourced data

Information lifecycle management

- This is a lot more data than we have ever encountered before.
- Current analytics systems are not capable of ingesting,storing, and analyzing these volumes at the required velocities.
- How to store, analyze, and use this data in real-time or near real-time, this is a lot more data than we have ever encountered before.
- Current analytics systems are not capable of ingesting, storing, and analyzing these volumes at the required velocities.
- To decide to store only samples of the data or use Hadoop for the storage and retrieval of large volumes of unstructured data.

Big Data Under Considerations

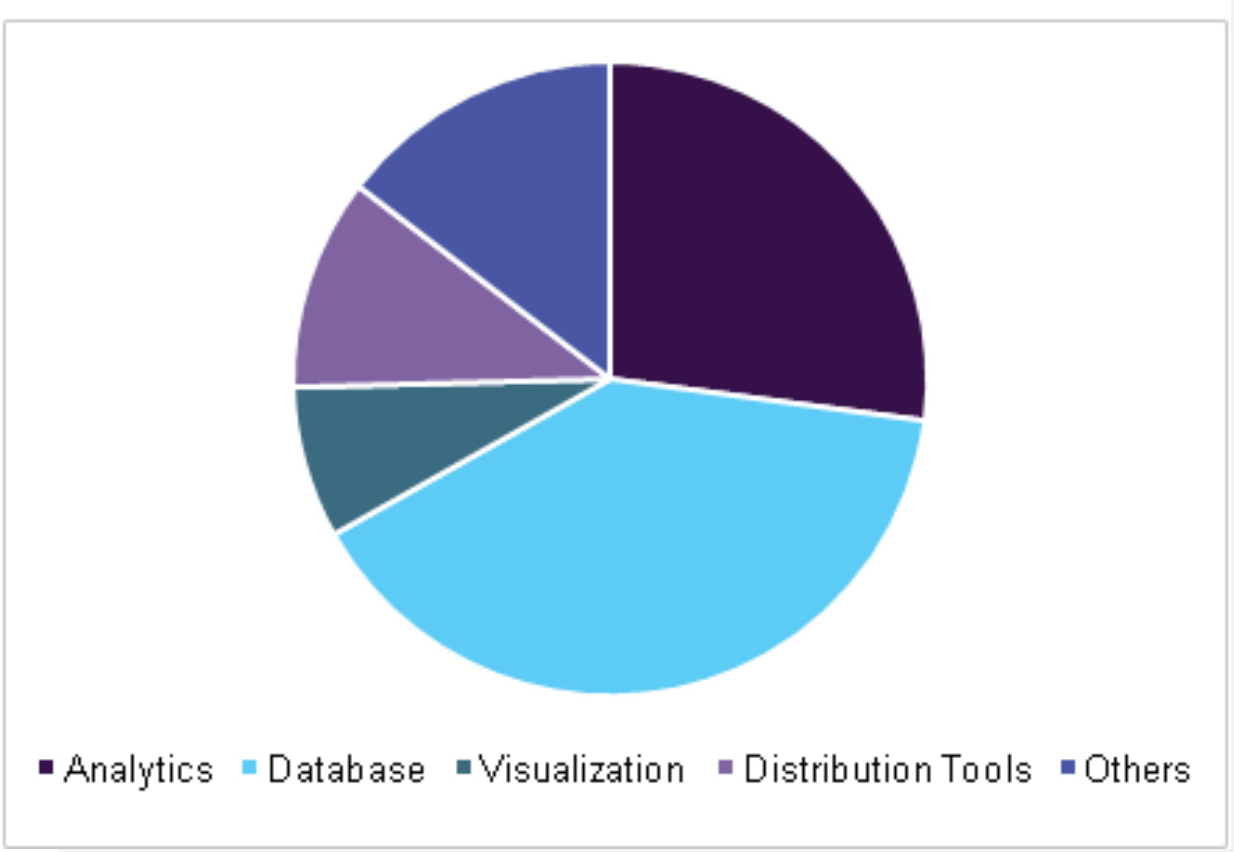
- **Data protection**: The Big Data currently created having a considerable measure of data of our private life's, most of them having a privilege to kept personal Expanding, that were requested to keep a harmony among the measure of private information we reveal and the continuance of Big Data-controlled applications and administrations works.
- **Data safety**: Even in the event that we choose we are glad for someone to have our data for a specific reason, would we be able to confide in them to protect it.
- **Data intolerance**: When everything is known, will it turn out to be satisfactory to discriminate against individuals in view of data we have on their lives.

Big Data Under Considerations

- We as of now utilize acclaim keeping count to choose which person could acquire fund and coverage is intensely information oriented.
- We may hope to be investigated and evaluated in more prominent details and responsibility essentially be taken that is not done in the kind that is creating life much troublesome for the individuals who as of now has less assets and approach to data.
- Overcoming these difficulties is an essential piece of Big Data, and that should be tended to by organizations that need to exploit data.
- Inability to do as such can leave businesses helpless as far as their reputation, as well as lawfully and financially

Big Data Market

- Big data persists to pervade one's everyday life's, that have been a critical move of center from the publicity surrounding it to finding genuine incentive in its utilization.
- The estimation of big data continues to remain a test, other pragmatic difficulties including funding and rate of profitability and abilities continue to remain at the forefront for various diverse industries that are adopting big data



Big data and its application in various industries

- Banking and Securities
- Health care Providers
- Education
- Manufacturing and Natural Resources

Banking and Securities

- An investigation of 16 extends in 10 top investment and retail banks shows that the difficulties in this industry include: securities misrepresentation early warning, tick examination, card extortion detection, authentic of review trails, enterprise credit hazard reporting, exchange deceivability, client data transformation, social examination for trading, IT operations investigation,and IT strategy consistence investigation, among others.
- Big Data Technologies are heavily used by banks to “Know Your Customer” and fraud mitigation.

Health care Providers

- The healthcare area approaches gigantic measures of data yet has been tormented by disappointments in utilizing the data to check the cost of rising medicinal services and by inefficient systems that smother quicker and better human services benefits no matter how you look at it.
- This is mainly because of the way that electronic data is inaccessible, inadequate, or unusable.
- Additionally, the social insurance databases that hold wellbeing related information have made it hard to link data that can show designs helpful in the medicinal field.
- Different difficulties identified with big data include: the exclusion of patients from the decision-making process, and the utilization of data from various promptly accessible sensors.

Education

- A noteworthy test in the education industry is to incorporate big data from various sources and sellers and to use it on stages that were not intended for the varying data.
- From a pragmatic point of view, staff and institutions need to take in the new data administration and examination instruments.
- On the specialized side, there are difficulties to integrate data from various sources, on various stages and from various sellers that were not intended to work with one another.
- Politically, an issue of security and personal data protection related with big data utilized for educational designs is a test.
- Colleges everywhere throughout the world are using Big Data to assess the execution of educators and understudies.

Manufacturing and Natural Resources

- Increasing interest for regular assets including oil, farming items, minerals, gas, metals, et cetera has prompted an increase in the volume, many-sided quality, and speed of data that is a test to deal with.
- So also, extensive volumes of data from the manufacturing industry are undiscovered.
- The underutilization of this information counteracts enhanced nature of items, vitality proficiency, unwavering quality, and better net revenues.

Summary

In todays session we tried understanding the below concept :

- ☞ *What is data modelling ?*
- ☞ *What is data marketing ?*
- ☞ *Need of data modelling?*
- ☞ *What is applications of data processing and modelling?*
- ☞ *Applications of big data*

In the next session we will discuss about data modelling, implementation challenges, big data market and there features.

...Activity...

Discussion on data modelling

What are the types of data modelling and there classifications?

Guess the characters of??

References:

- ✓ https://www.simplilearn.com/how-facebook-is-using-big-data-article?source=CTAexp
- ✓ https://www.icas.com/ca-today-news/10-companies-using-big-data
- ✓ https://www.bernardmarr.com/default.asp?contentID=1076
- ✓ Bryant, R.E., Katz, R.H., Lazowska, E.D.: Big-Data Computing: Creating Revolutionary Breakthroughs in Commerce, Science and Society
- ✓ Sathi, A.: Implementation section (book 1). In: Big Data Analytics: Disruptive Technologies for Changing the Game, 1st ed. MC Press Online (2012)

Assessment Pattern

S.No.	Item	Number/semester	Marks
1	MSTs	2	20 per each
2	Quiz	2 per unit	4 per each quiz
3	Time bound surprise test	3 (one per unit)	12 per each test
4	Assignments	3 (one per unit)	10 per each Assignment
5	Engagement task (non gradable)	One per each topic	depends
6	Attendance + Engagement score	Above 90%	2
Internal (division as mentioned above points 1-6)			**40**
External			**60**
Total			**100**

THANK YOU

For queries
Email: bhanu.priyanka.e8243@cumail.in