What is big data ?

* Big Data is similar to small data, but bigger in size, but having data bigger it requires different approaches:
* Techniques, tools and architecture
* An aim to solve new problems or old problems in a better way . Big Data generates value from the storage and processing of very large quantities of digital information that cannot be analyzed with traditional computing techniques.

Characteristics of Big Data :

* Volume
* Velocity
* Variety

Processing Big Data :

* Integrating disparate data stores :
* Mapping data to the programming framework .
* Connecting and extracting data from storage .
* Transforming data for processing .
* Subdividing data in preparation for Hadoop MapReduce .
* Employing Hadoop MapReduce :
* Creating the components of Hadoop MapReduce jobs .
* Distributing data processing across server farms .
* Executing Hadoop MapReduce jobs .
* Monitoring the progress of job flows .

Why Big Data ?

* Growth of Big Data is needed :
  + Increase of storage capacities .
  + Increase of processing power .
  + Availability of data(different data types) .
  + Every day we create 2.5 quintillion bytes of data; 90% of the data in the world today has been created in the last two years alone .
  + FB generates 10TB daily .
  + Twitter generates 7TB of data Daily .
  + IBM claims 90% of today’s stored data was generated in just the last two years.

Big Data Analytics :

* Examining large amount of data .
* Appropriate information .
* Identification of hidden patterns, unknown correlations .
* Competitive advantage .
* Better business decisions: strategic and operational .
* Effective marketing, customer satisfaction, increased revenue .

Risks of Big Data :

* Will be so overwhelmed
* Need the right people and solve the right problems
* Costs escalate too fast
* Isn’t necessary to capture 100%
* Many sources of big data is privacy
* self-regulation
* Legal regulation

Benefits of Big Data :

* Real-time big data isn’t just a process for storing petabytes or exabytes of data in a data warehouse, It’s about the ability to make better decisions and take meaningful actions at the right time.
* Fast forward to the present and technologies like Hadoop give you the scale and flexibility to store data before you know how you are going to process it.
* Technologies such as MapReduce,Hive and Impala enable you to run queries without changing the data structures underneath.
* Big Data is already an important part of the $64 billion database and data analytics market
* It offers commercial opportunities of a comparable scale to enterprise software in the late 1980s .
* And the Internet boom of the 1990s, and the social media explosion of today.