**Big Data**

Big Data is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so large size and complexity that none of traditional data management tools can store it or process it efficiently. Big data is also a data but with huge size.

**Sources of Big Data**

Stock Exchange is an example of Big Data that generates about one terabyte of new trade data per day.

The statistic shows that 500+terabytes of new data get ingested into the databases of social media site Facebook, every day. This data is mainly generated in terms of photo and video uploads, message exchanges, putting comments etc.

**Types of Big Data**

1. Structured
2. Unstructured
3. Semi-structured

**Structured**

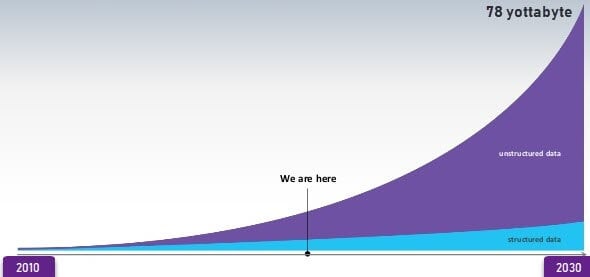
Any data that can be stored, accessed and processed in the form of fixed format is termed as a ‘structured’ data. Over the period of time, talent in computer science has achieved greater success in developing techniques for working with such kind of data (where the format is well known in advance) and also deriving value out of it. Common examples of structured data are Excel files or SQL databases. Each of these has structured rows and columns that can be sorted. Structured data depends on the existence of a data model – a model of how data can be stored, processed and accessed

**Unstructured**

Any data with unknown form or the structure is classified as unstructured data. In addition to the size being huge, un-structured data poses multiple challenges in terms of its processing for deriving value out of it. A typical example of unstructured data is a heterogeneous data source containing a combination of simple text files, images, videos etc. Now day organizations have wealth of data available with them but unfortunately, they don’t know how to derive value out of it since this data is in its raw form or unstructured format.

**Semi-structured**

Semi-structured data can contain both the forms of data. We can see semi-structured data as a structured in form but it is actually not defined with e.g. a table definition in relational [DBMS](https://www.guru99.com/what-is-dbms.html). Example of semi-structured data is a data represented in an XML file.



Data Growth over the years

**Characteristics of Big Data**

Big data can be described by the following characteristics:

* Volume - There is a massive amount of data generated every second.
* Velocity - The speed at which data is generated, collected, and analyzed
* Variety - The different types of data: structured, semi-structured, unstructured
* Value - The ability to turn data into useful insights for your business
* Veracity - Trustworthiness in terms of quality and accuracy

**Advantages of Big Data:**

* Businesses can utilize outside intelligence while taking decisions

Access to social data from search engines and sites like Facebook, twitter are enabling organizations to fine tune their business strategies.

* Improved customer service

Traditional customer feedback systems are getting replaced by new systems designed with Big Data technologies. In these new systems, Big Data and natural language processing technologies are being used to read and evaluate consumer responses.

* Early identification of risk to the product/services, if any
* Better operational efficiency

**Points to remember**

* Big Data definition: Big Data meaning a data that is huge in size. Big data is a term used to describe a collection of data that is huge in size and yet growing exponentially with time.
* Big Data analytics examples include stock exchanges, social media sites, jet engines, etc.
* Big Data could be 1) Structured, 2) Unstructured, 3) Semi-structured
* Volume, Variety, Velocity, and Variability are few Big Data characteristics
* Improved customer service, better operational efficiency, Better Decision Making are few advantages of Big data