



PARSHVANATH CHARITABLE TRUST'S

A. P. SHAH INSTITUTE OF TECHNOLOGY

(All Branches NBA Accredited)



Characteristics of Big Data as follows:

1. Volume:

The name 'Big Data' itself is related to a size which is enormous.

Volume is a huge amount of data.

To determine the value of data, size of data plays a very crucial role. If the volume of data is very large then it is actually considered as a 'Big Data'. This means whether a particular data can actually be considered as a Big Data or not, is dependent upon the volume of data.

Hence while dealing with Big Data it is necessary to consider a characteristic 'Volume'.

Example: In the year 2016, the estimated global mobile traffic was 6.2 Exabytes(6.2 billion GB) per month. Also, by the year 2020 we will have almost 40000 ExaBytes of data.

2. Velocity:

Velocity refers to the high speed of accumulation of data.

In Big Data velocity data flows in from sources like machines, networks, social media, mobile phones etc.

There is a massive and continuous flow of data. This determines the potential of data that how fast the data is generated and processed to meet the demands.

Sampling data can help in dealing with the issue like 'velocity'.

Example: There are more than 3.5 billion searches per day are made on Google. Also, FaceBook users are increasing by 22%(Approx.) year by year.

3. Variety:

It refers to nature of data that is structured, semi-structured and unstructured data.

It also refers to heterogeneous sources.

Variety is basically the arrival of data from new sources that are both inside and outside of an enterprise. It can be structured, semi-structured and unstructured.

Structured data: This data is basically an organized data. It generally refers to data that has defined the length and format of data.

Semi- Structured data: This data is basically a semi-organised data. It is generally a form of data that do not conform to the formal structure of data. Log files are the examples of this type of data.

Unstructured data: This data basically refers to unorganized data. It generally refers to data that doesn't fit neatly into the traditional row and column structure of the relational database. Texts,



PARSHVANATH CHARITABLE TRUST'S

A. P. SHAH INSTITUTE OF TECHNOLOGY

(All Branches NBA Accredited)



pictures, videos etc. are the examples of unstructured data which can't be stored in the form of rows and columns.

4. Veracity:

It refers to inconsistencies and uncertainty in data, that is data which is available can sometimes get messy and quality and accuracy are difficult to control.

Big Data is also variable because of the multitude of data dimensions resulting from multiple disparate data types and sources.

Example: Data in bulk could create confusion whereas less amount of data could convey half or Incomplete Information.

5. Value:

After having the 4 V's into account there comes one more V which stands for Value!. The bulk of Data having no Value is of no good to the company, unless you turn it into something useful.

Data in itself is of no use or importance but it needs to be converted into something valuable to extract Information. Hence, you can state that Value! is the most important V of all the 6V's.

6. Variability:

How fast or ,available data that extent is the structure of your data is changing ?.

How often does the meaning or shape of your data!to?change?.

example : if you are eating same ice-cream daily and the taste just keep changing.