Big Data

Big data are those data which cannot be handled in a no sequel +sequel database.

Size of big data is in petabytes.

According to future estimation the size of big data of the entire world will be in 198 zettabyte.

Big data is such a huge ecosystem with trillions and trillions of data that if the db insertion of a transaction table is in seconds than for sure data anamolies will occur,thus company would have to migrate their database in to big data framework like apache sparks.

Big data usually includes data sets with sizes beyond the ability of commonly used software tools to [capture](https://en.wikipedia.org/wiki/Data_acquisition), [curate](https://en.wikipedia.org/wiki/Data_curation), manage, and process data within a tolerable elapsed time. Big data philosophy encompasses unstructured, semi-structured and structured data, however the main focus is on unstructured data. Big data “size” is a constantly moving target; as of 2012 ranging from a few dozen terabytes to many [zettabytes](https://en.wikipedia.org/wiki/Zettabyte) of data. Big data requires a set of techniques and technologies with new forms of [integration](https://en.wikipedia.org/wiki/Data_integration) to reveal insights from [data-sets](https://en.wikipedia.org/wiki/Data_set) that are diverse, complex, and of a massive scale.

In general terms big data needs lots of distributed and parallel computation of databases which requires high specialist computers in a bunch of 0.1K units. And than server cost of big data is additional to it thus it is important to know when to migrate to big data and than how ? will be another aspects toward bigdata framework.

Disadvantage

🡪Huge amount of data is stored.so storing exponentially growing huge dataset.

🡪Having complex of storage structure

Big data has all 3 type of data ie

Structured Data

Unstructured Data

Semi Structured Data

🡪Data processing is not faster due massive storage of data