**20/12/2023**

Big Data

Big data refers to extremely large and diverse collections of structured, unstructured, and semi-structured data that continues to grow exponentially over time. These datasets are so huge and complex in volume, velocity, and variety, that traditional data management systems cannot store, process, and analyse them

Big data describes large and diverse datasets that are huge in volume and also rapidly grow in size over time. Big data is used in machine learning, predictive modelling, and other advanced analytics to solve business problems and make informed decisions.

**The three Vs of big data**

VOLUME: -

the amount of data matters. With big data, you’ll have to process high volumes of low-density, unstructured data. This can be data of unknown value, such as Twitter data feeds, clickstreams on a web page or a mobile app, or sensor-enabled equipment. For some organizations, this might be tens of terabytes of data. For others, it may be hundreds of petabytes.

VELOCITY: -

Velocity is the fast rate at which data is received and (perhaps) acted on. Normally, the highest velocity of data streams directly into memory versus being written to disk. Some internet-enabled smart products operate in real time or near real time and will require real-time evaluation and action.

VARIETY: -

Variety refers to the many types of data that are available. Traditional data types were structured and fit neatly in a [relational database](https://www.oracle.com/in/database/what-is-a-relational-database/). With the rise of big data, data comes in new unstructured data types. Unstructured and semi structured data types, such as text, audio, and video, require additional preprocessing to derive meaning and support metadata.

Big Data uses: -

* Product development
* Predictive maintenance
* Fraud and compliance
* Machine learning
* Drive innovation

Customer experience





