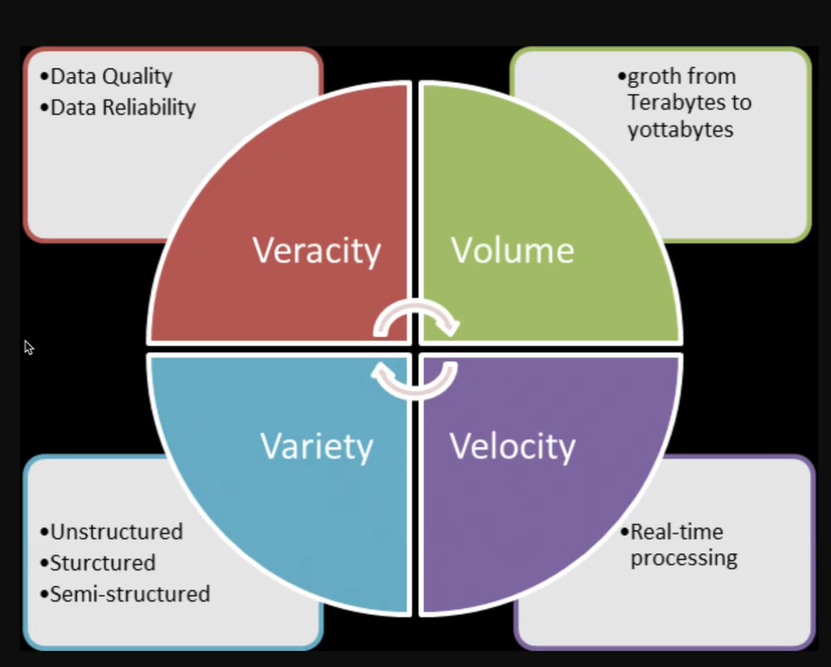
**BIG DATA**

* Big Data refers to large and complex datasets that traditional data-processing software cannot handle efficiently.
* **Sources of Big Data**
* Social media platforms (e.g., Twitter, Facebook)
* IoT devices (smart homes, wearables)
* Transaction records (e-commerce, banking)
* Sensors and logs (climate, machinery)
* Streaming services (Netflix, YouTube)



* **Datasets are characterized by the 4 Vs:**



* **Volume**:

Massive amount of data.

* **Velocity**:

Speed of data generation and processing.

* **Variety**:

Different types of data (text, audio, video, etc.).

* **Veracity**:

Trustworthiness of data.

* **BIG DATA ARCHITECTURE**

A typical architecture includes:

1. **Data Sources** (e.g., logs, APIs)
2. **Data Ingestion** (e.g., Kafka, Flume)
3. **Storage** (e.g., HDFS, NoSQL)
4. **Processing** (e.g., Spark, MapReduce)
5. **Analytics** (e.g., ML models, dashboards)

* **FUTURE OF BIG DATA**
* Integration with **AI & Machine Learning**
* Rise of **Edge Computing** and **Federated Learning**
* Increased focus on **Data Ethics** and **Regulation**