1. **Big Data**

* Mostly unstructured.
* Only 10% data is structured the rest is unstructured data.
* Unstructured data constitutes of images, videos, website etc.
* Stored in PB, EB.
* The size of big data increases exponentially. (Daily 350 million photos are uploaded on Facebook).
* The 90% data of all that exists was created in the last 8-9 years.
* Normally stored using distributed techniques and its globally available.
* Hadoop, MapReduce, Spark etc technologies are used in Big Data.

1. **Data Lake**

* A data lake is where lots of unstructured data is kept
* Normally in the data lake the data is in a very raw form
* The raw data in a data lake is not usable and needs to be cleaned before it can be used an any meaningful information can be extracted from it.

1. **Database**

* A database is the oldest method of saving data digitally
* It contains all the detailed records of the data that’s being fed into it via a source (For Eg: a shopping website or a google form etc)
* It is designed to capture and record data (also known as OLTP [Online transaction processing])
* SQL databases store data in the form of rows and columns.

1. **Data Warehouse**

* A data warehouse is also a database but it doesn’t contain the most fresh copy of records compared to a database
* Instead it contains data that is loaded into it via ETL pipelines and then analytical processes are applied to it.
* In a way we can say that a data warehouse contains a summarized data (since the ETL process has cleaned it and only kept the useful information).