* **Big Data**

1. **What is Big Data?**

Big data is a t+erm for collection of datasets so **large and complex** thatis difficult to process using traditional processing system.

There is limit for data size that can be processed by traditional processing systems. If the size limit exceeds we need to follow non traditional approach.

Big data is meant for analysis.

Even 1tb of data is not bigdata.

1. **Where is this big data coming from?**

Twitter generates about 12+ TB of tweet data everydata

Facebook generated 25+ TB of log data everyday

Google generates humongous data every day.

4.6 billion Camera phones worldwide.

Billions of gps enabled devices are sold every.

Etc So we need to analyze these data to get some business out of it.

1. **Why we Need RDBMS?**
2. RDBMS provides quick response. Hadoop is slow in this case. There is no replacement for this.
3. Maintains relationship between data
4. One database is utilized by all applications
5. **Why can’t we store big data in RDBMS?**
6. RDBMS can’t handle such huge data.
7. Majority of data comes in semi structured and unstructured format like video, audio etc but RDBMS is designed for structured data
8. RDBMS cannot be as scalable as hadoop.
9. RDBMS is license is costly.

What is OLTP and OLAP? Not so imp but haddop is OLAP

**Big data Definition from IBM / Types of bigdata?**

Volume: Amount Data

Velocity: Speed at which data is coming

Variety: Variety of data like text audio video images.