1.

* Social Media
* webpages
* Mobile apps
* Documents
* Sensor data
* Machine log data
* data stored like photos,videos etc..

2.

**Volume:**It stores a large amount of data.

**Velocity:**It can be accessed with high speed.

**Variety:**It can store data of various fields

3.

**Vertical Scaling:**

Vertical scaling means that you scale by adding more power (CPU, RAM) to an existing machine**.** The procedure of adding multiple hardware to a single system is known as Vertical Scaling**.**

The addition of multiple hardware (like ram, hard disk) to a single system and will be very complex to maintain the hardware when they are added to a single system.

**Horizontal Scaling:**

Horizontal scaling means that you scale by adding more machines into your pool of resources. The procedure of storing and processing data in a distributed architecture is known as Horizontal Scaling

The best and suitable solution to handle Big Data is to store the data in a distributed architecture, where users can add multiple systems or nodes when there is an increase in data. Since the systems are connected in a distributed architecture, the performance of the processing data will be very high compared to those running in a single system (Vertical Scaling).

4.

Hadoop is open source software platform managed by the Apache Software .That Foundation has proven to be very helpful in storing and managing vast amounts of data cheaply and efficiently.

It is a way of storing enormous data sets across distributed clusters of servers and then running “distributed” analysis applications in each cluster.

The Hadoop platform was designed to solve problems where you have a lot of data — perhaps a mixture of complex and structured data —and it doesn’t fit nicely into tables. It’s for situations where you want to run analytics that are deep and computationally extensive, like clustering and targeting.