

Kyle Adrian R. Mascuñana BSIT3 A

I've learned that big data is used when the data gathered in a certain location has many data files. For example hospitals and clinics make almost 2.3k exabytes annually so the data is so big that the computer or the device that is used to store that data can't handle the massive input of data so we are going to use big data. To classify any data as big data we should identify if that data has 5 V's such as Volume of data, Velocity of data, Variety of data, Veracity of data and the Value of data. For volume, the massive amount of data gathered from a certain location.

For velocity, We can use hospitals as an example. Velocity is the gathered patients records and test results and this data is generated at a very high speed. For the Variety, it has three parts such as Structured Data. Excel data, Semi-Structured Data such as Log Files of the patients, and Unstructured Data that has X ray images that are stored. Veracity is used if the data gathered has accuracy and trustworthiness. And lastly the Value of the data. Big data is used using Frameworks such as Cassandra, Hadoop and Spark. For hadoop, it can store big data and can break down data to other devices to always secure a copy if one of the devices broke or has been damaged.