**Type:** Interactive Exercise

**Audience:** All Levels of Learners

**Anticipated Duration:** 30 minutes

**Format:** Question and Answer

**Title:** What is Big Data?

**Author:** Shannon McWeeney, PhD

**Version:** 2.0

**Date:** January 2016

**Objective:** To have participants think critically about what defines “big data” and what the implications of this are.

**Questions [Order and use of questions will depend on responses]:**

* What does the phrase “big data” mean to you?
* Please give me a few examples of “big data”
  + Capture these on white board
* How would you know “big data” if you saw it?
* What distinguishes it from other data?
* For each characteristic, what are the challenges? Opportunities?
  + Use these characteristics to evaluate the examples on the board. Introduce the 4V’s as they recapitulate each concept (Variety, Volume, Velocity, Veracity
* What is needed to store, analyze and interpret these type of data?
* Is “Big Data” more valuable than “Little Data”?

**Alternate Approach:** If students don’t have examples, can provide examples to them and ask if it is big data

Example 1: A data set too large for Excel

Metrics on Limits for Excel: <https://support.office.com/en-za/article/Excel-specifications-and-limits-ca36e2dc-1f09-4620-b726-67c00b05040f>

Example 2: All of the Tweets from Twitter

Metrics on Usage: http://www.internetlivestats.com/twitter-statistics