

Myths and Facts of Big Data Graphic Text Description

Myth #1 - Big Data is new

Fact #1 - Huge cross-references of every single word used in the Bible, called “concordances,” were in use by scholar monks for centuries well before the first databases.

Myth #2 - Big Data is made for Big Business

Fact #2 -Enterprises of all sizes are able to now leverage big data analytics thanks to recent improvements in cloud and data technology.

Myth #3 - Bigger Data is better

Fact #3 - Quality of data wins over quantity of data. What to use is often more relevant than how much to use.

Myth #4 - Our data is so messed up we can't possibly master big data.

Fact #4 - Advanced data quality, master data management, and data governance tools have made it easier to clean up the enterprise data mess.

Myth #5 - Every problem is a big data problem

Fact #5 - If you are matching a couple fields against a couple of conditions across a couple gigabytes, it isn't really a big data problem. Don't treat every analytics need as a big data effort.

Myth #6 - Big Data applications require little or no performance tuning.

Fact #6 - Big Data applications require regular tuning of the analytical and statistical models as more and more data and variables are added.

Myth #7 - Big Data is a magic 8-ball

Fact #7 - Big Data may not tell you everything. A Lot depends on the right questions and the right data for it to work.

Myth #8 - Big Data is only unstructured data

Fact #8 - Big Data does not have to be unstructured. Even voluminous structured data may be classified as Big Data because of its sheer volume.

Myth #9 - You need unstructured data to make predictions

Fact #9 - Predictive models use a combination of unstructured and structured data for training the models and make inferences.

Myth #10 - Machine learning is a concept related to Big Data

Fact #10 - The idea underlying machine learning is “using data to model an underlying process.” Machine learning algorithms can, however, provide valuable insights when used in conjunction with Big Data.

Myth #11 - Big data analytics will not require supervision by humans

Fact #11 - The adjective “unsupervised” does not mean that these algorithms run by themselves without human supervision. An analyst (or a data scientist) who is training an unsupervised learning model has to exercise a similar kind of modeling discipline as the one who is training a supervised model.