

# WHAT ARE **BIG DATA**?

- Every two days we create as much data as we did from the beginning of time until 2000.
- And the amount of data we're creating continues to increase rapidly.
- By 2020, the amount of digital information available will have grown to 50 zettabytes.
- Data are traditionally limited to spreadsheets or databases, but it can now include anything from databases to photos, videos, sound recordings, written text and sensor data.



**ATA BIG DATA BIG**

# DATA COME IN **TWO** FORMS

- **Structured** data refer to data with a high degree of organization like those found in a relational database.
- **Unstructured** data usually refer to information that doesn't reside in a traditional row-column database. Metadata, Twitter tweets, and other social media posts are good examples of unstructured data.

## *Structured Data*



|       |       |       |       |       |
|-------|-------|-------|-------|-------|
| 0.103 | 0.176 | 0.387 | 0.300 | 0.379 |
| 0.333 | 0.384 | 0.564 | 0.587 | 0.857 |
| 0.421 | 0.309 | 0.654 | 0.729 | 0.228 |
| 0.266 | 0.750 | 1.056 | 0.936 | 0.911 |
| 0.225 | 0.326 | 0.643 | 0.337 | 0.721 |
| 0.187 | 0.586 | 0.529 | 0.340 | 0.829 |
| 0.153 | 0.485 | 0.560 | 0.428 | 0.628 |

## *Unstructured Data*



