



SNS COLLEGE OF ENGINEERING



Kurumbapalayam(Po), Coimbatore - 641 107

Accredited by NAAC-UGC with 'A' Grade

Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

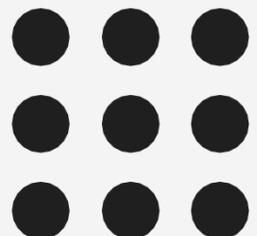
Department of Information Technology

19IT601- Data Science and Analytics

III Year / VI Semester

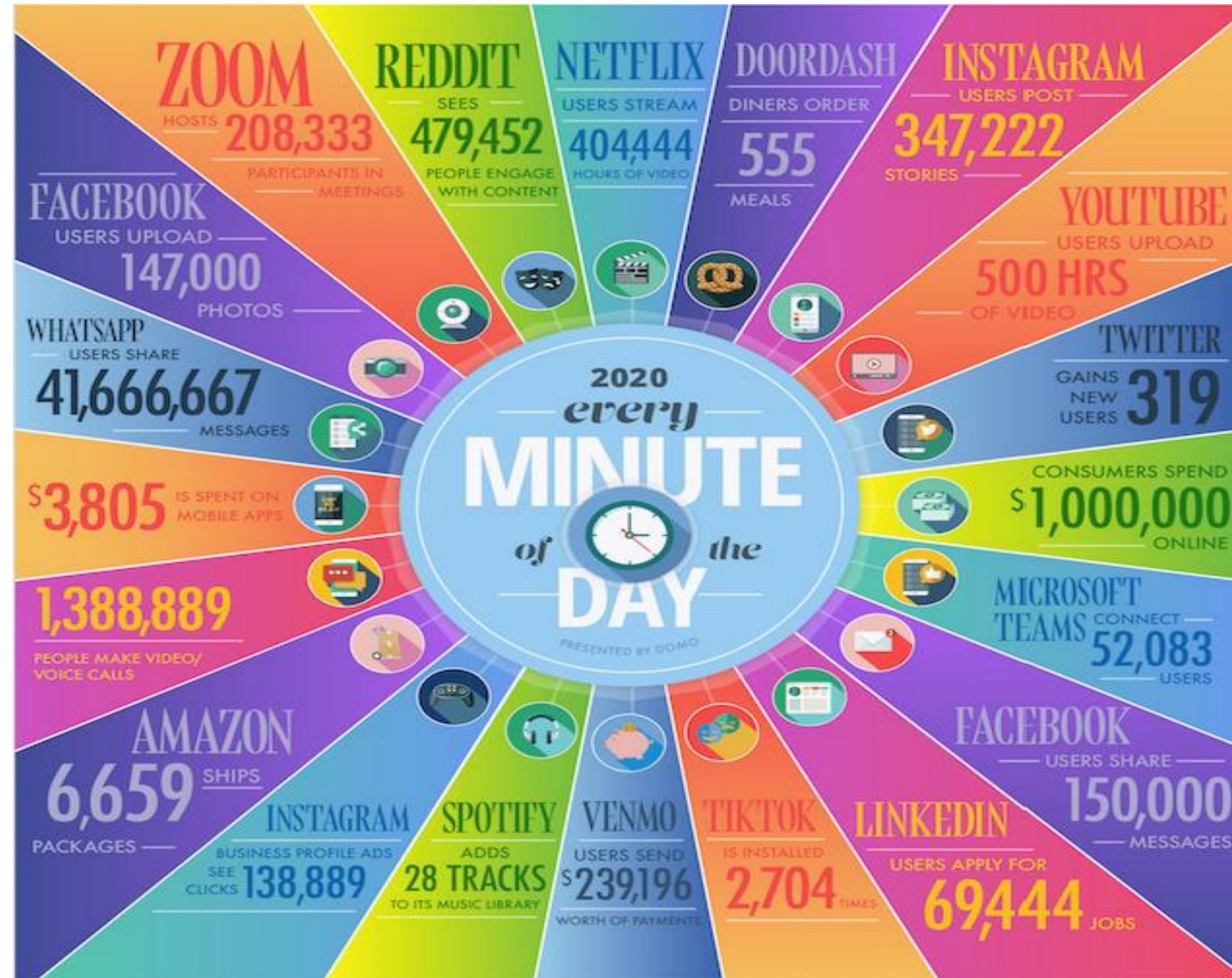
Unit 1 - Introduction

Topic 5: Big Data Characteristics



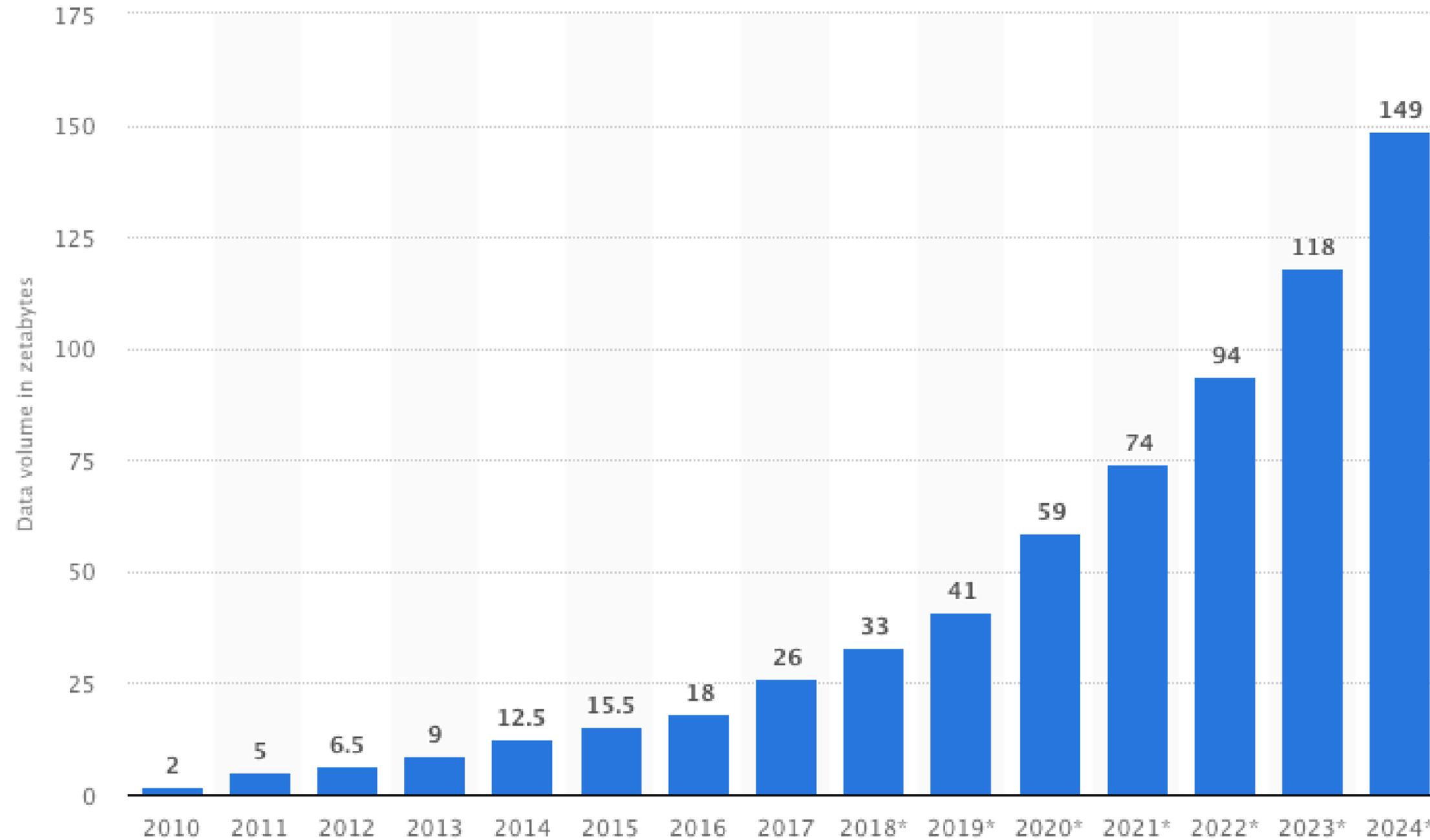
How Big is Big Data?

Data Never Sleeps



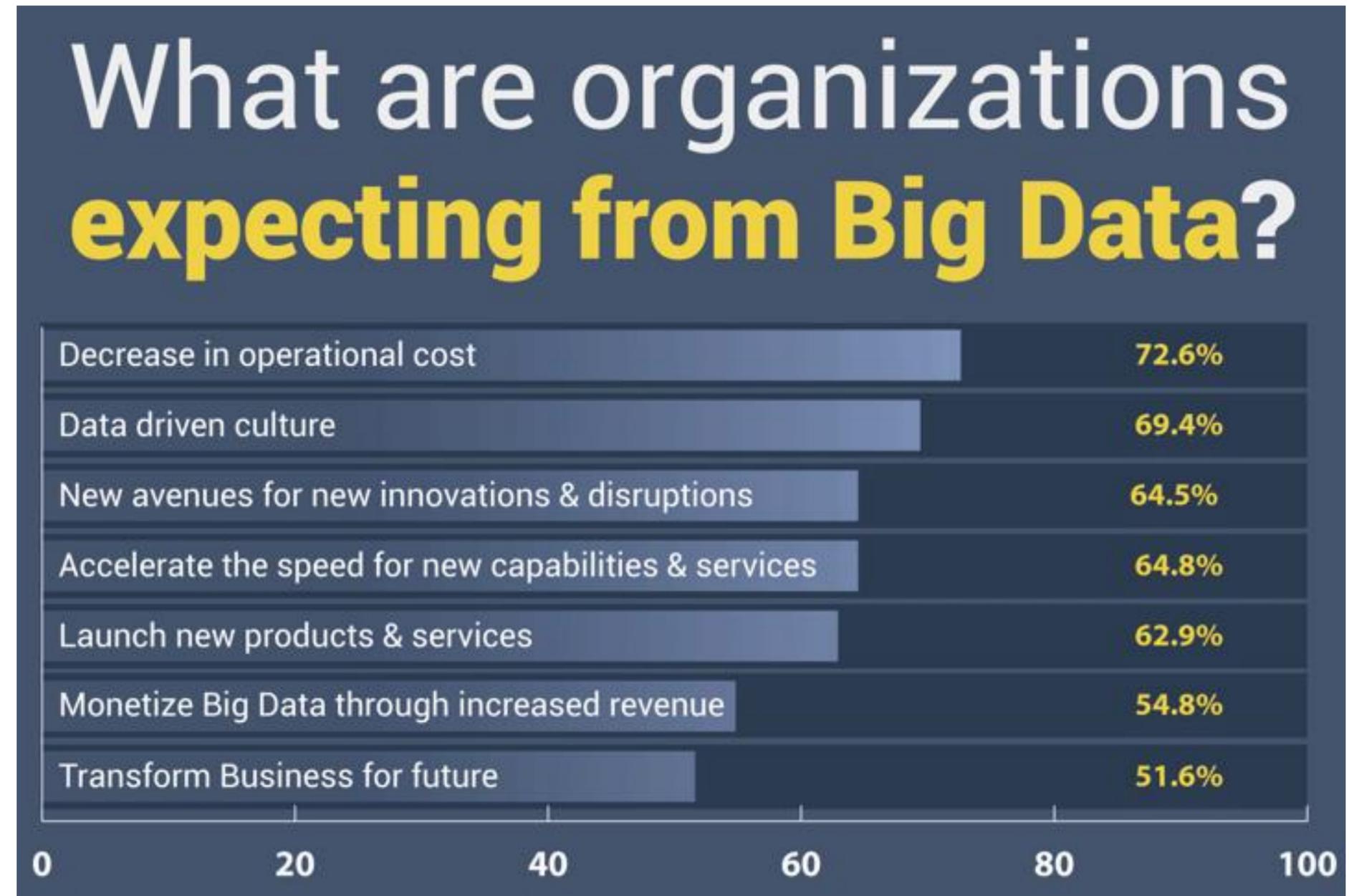


How Big is Big Data?



What to do with Big data?

- Advance analytics
- Business Intelligence
- Better Informed Decisions
- BI to increase sales, profit, customers
- Identify Business Risks
- Predict New Business Opportunities



Big Data Characteristics

- V's of Big Data
- In early 3 V's
 - Volume
 - Velocity
 - Variety
- Commonly 4 V's
 - Veracity
- And now 5 V's
 - Value



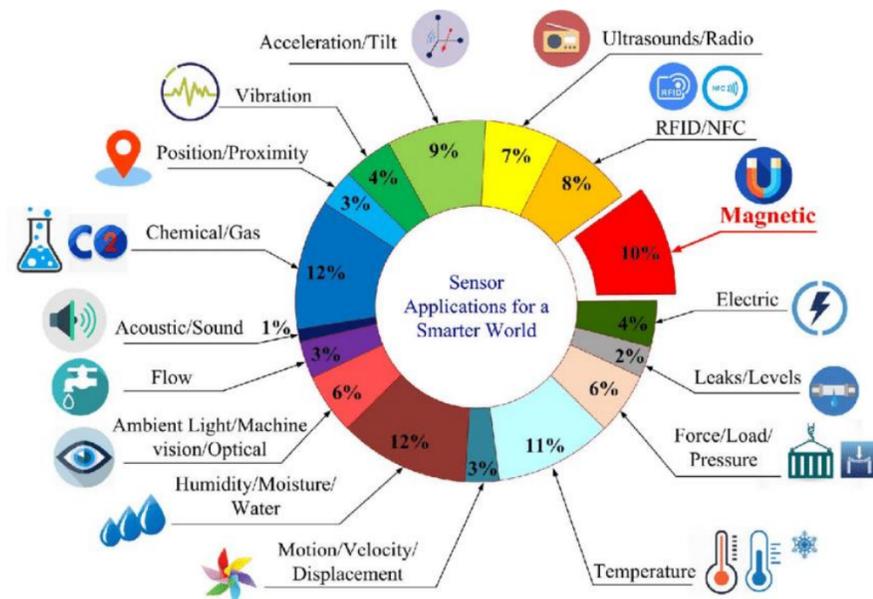
Big Data Characteristics

Volume

- Amount of Data Generated

Sources

- Social Media
- IoT Sensor Data
- Mobile Devices (Smart Devices)
- Financial Records
- Stock Market
- Weather data



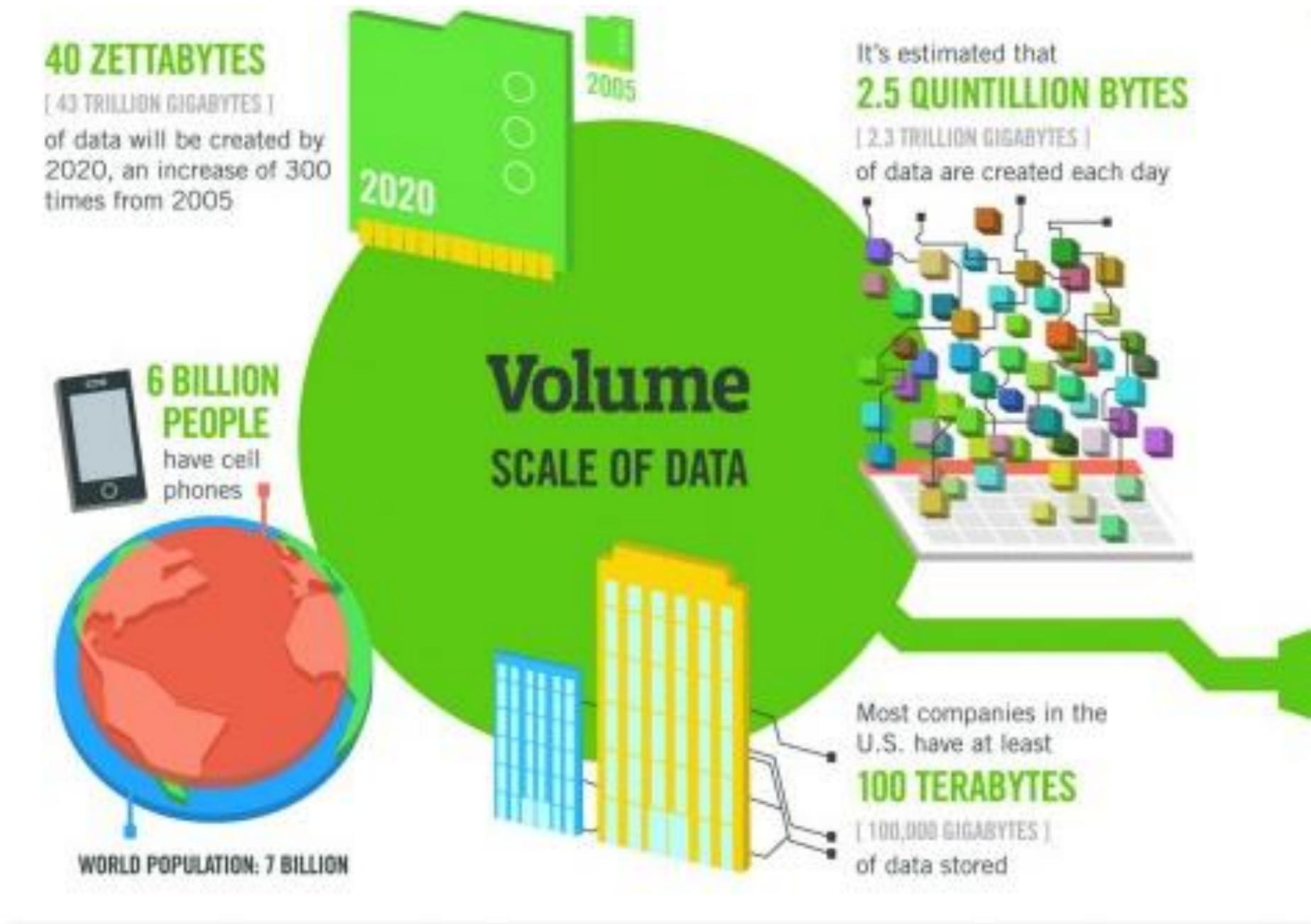
BYTE OF STORAGE

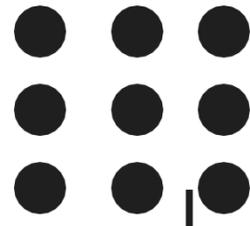
- Megabyte**: 1MB = 20 pages = 1 file
- Gigabyte**: 1GB = 256 Mp3 songs
- Terabyte**: 1TB = 250 HD movies
- Petabyte**: 20PB = Data processed by Google in 1 day
- Exabyte**: 5EB = All the words ever spoken by mankind
- Zettabyte**: 1ZB = Entire planet's digital content
- Yottabyte**: 1YB file would take 11 trillion years to download

Something bigger than Yottabyte?
Yes! The proposed standards are named to be Hellabyte and Brontobyte.

Big Data Characteristics

Volume

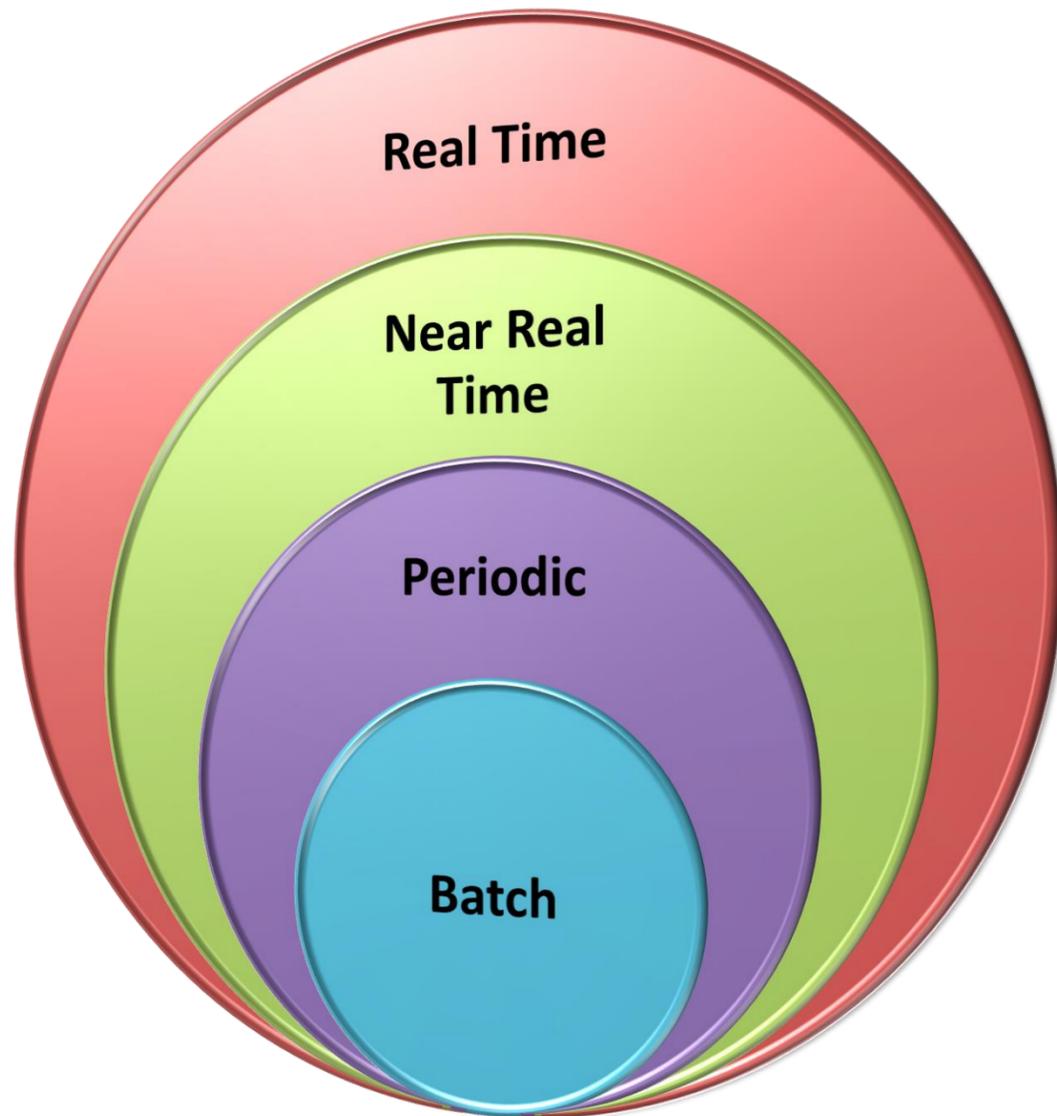




Big Data Characteristics

Velocity

- Rate at which data is received
- Speed at which data is generated



2020 This Is What Happens In An Internet Minute





Big Data Characteristics



Velocity

- Climate control,
- Social Media
- Health care,
- Stock exchange predictions,
- Banking transactions and
- IoT
- Transportation

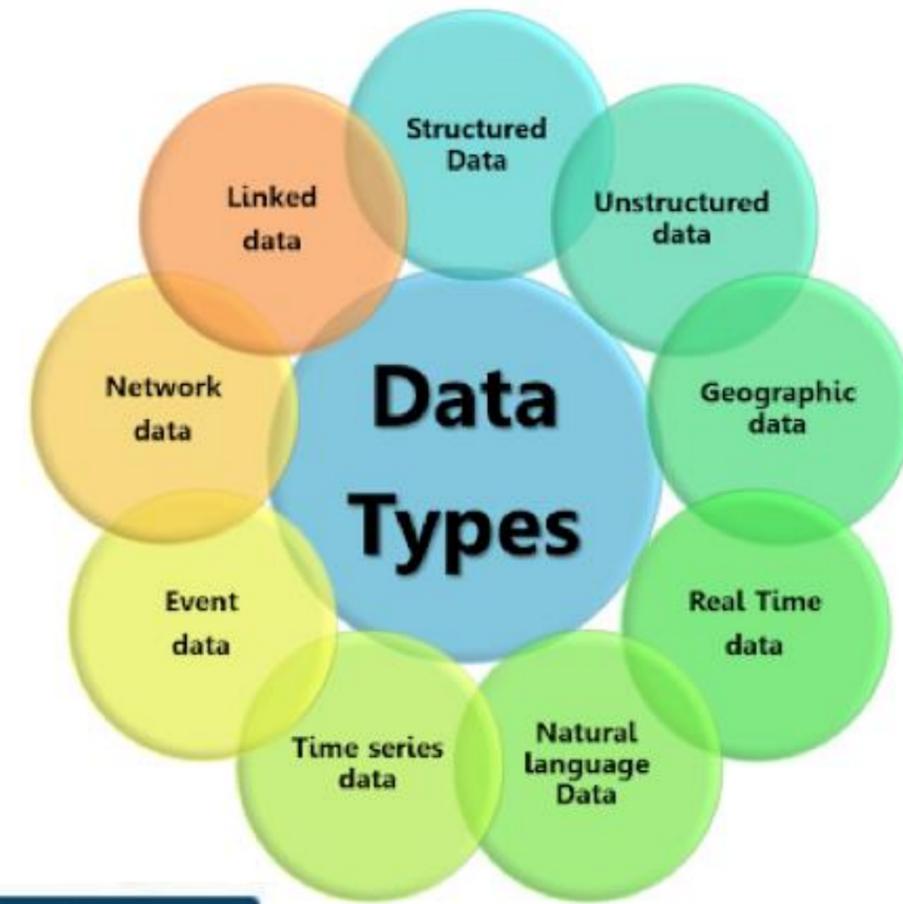
Big Data Characteristics

Variety

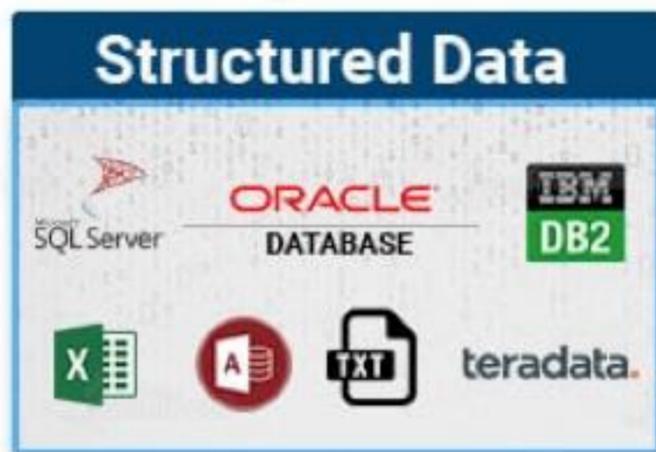
- Different types of Data
- Various forms of data

The variety of data is categorized as follows:

- Structured – RDBMS
- Semi Structured – XML, HTML, RDF, JSON
- Unstructured- Text, audio, video, logs, images



Structured Data



Unstructured Data



Semi-structured Data



Big Data Characteristics

As of 2011, the global size of data in healthcare was estimated to be

150 EXABYTES

[161 BILLION GIGABYTES]



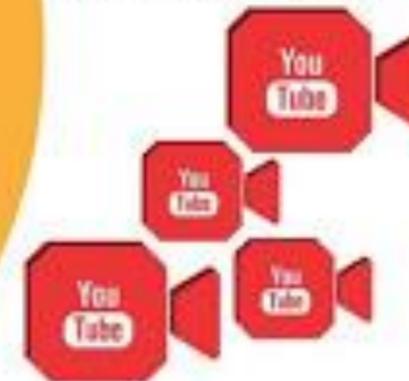
By 2014, it's anticipated there will be

420 MILLION WEARABLE, WIRELESS HEALTH MONITORS



4 BILLION+ HOURS OF VIDEO

are watched on YouTube each month



Variety
DIFFERENT FORMS OF DATA

30 BILLION PIECES OF CONTENT

are shared on Facebook every month



400 MILLION TWEETS

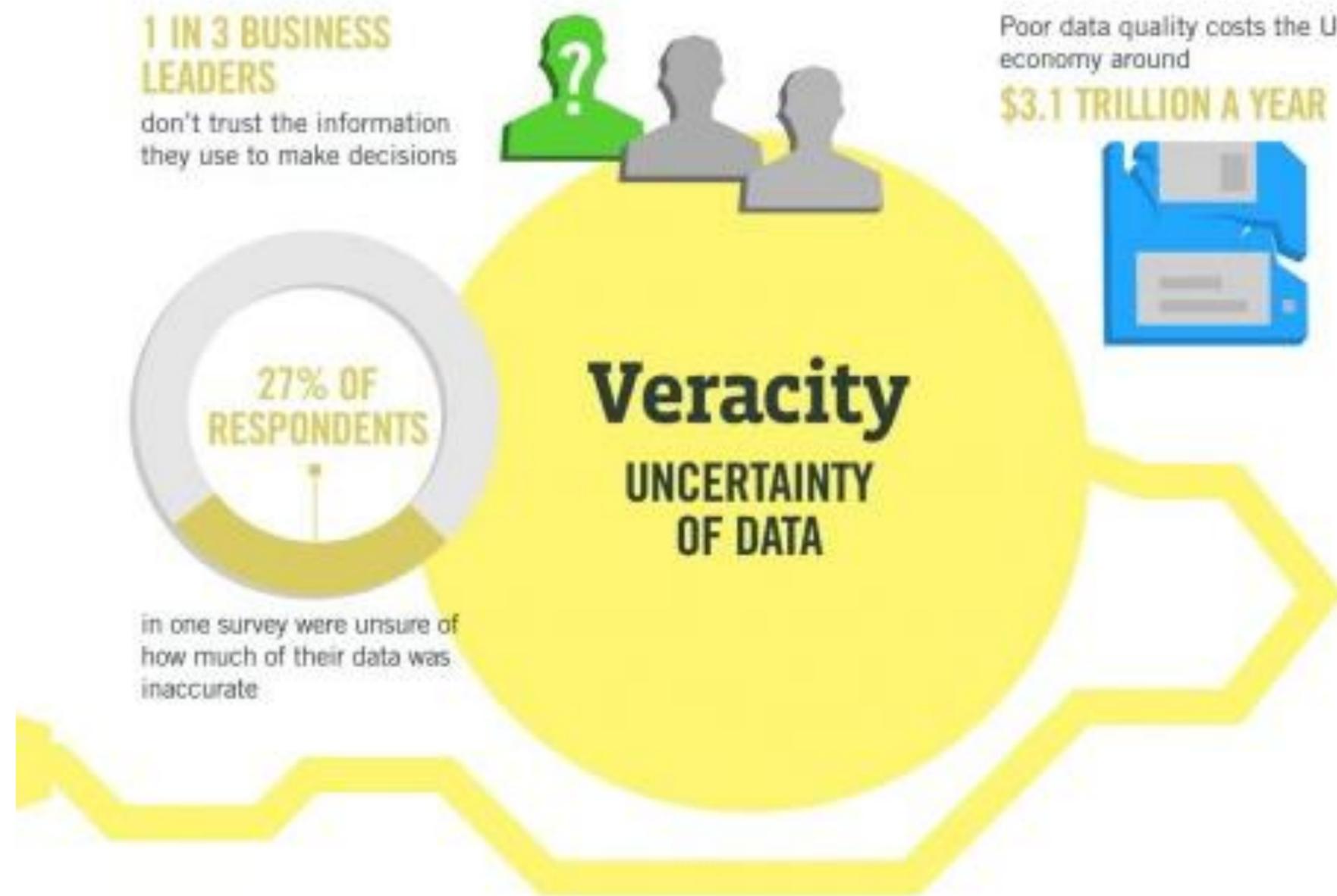
are sent per day by about 200 million monthly active users



Big Data Characteristics

Veracity

- Uncertainty
- Inconsistent or Incomplete
- Data quality is Unreliable
- How Accurate? Accuracy
- Truthfulness of data
- Data from different sources



Big Data Characteristics

Value

- Social or economic value that the data might create.
- Represents benefit of data to your business
- Business Intelligence
- Insights
- Informed decision making



The Value of Big Data

Over 80% of organizations say:

Big Data is critical to meet strategic objectives.	Sharing insights is a must-have capability for businesses.	Big Data will amplify other technology innovations.
		



THANK YOU