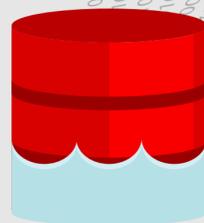


# Why are Data Lakes the Future of Big Data?



As big data gets bigger, the increasing volume of data and data sources can easily overwhelm data scientists. A data lake puts that all in one simple, cost-effective, and configurable repository.

## THE NEED FOR DATA MANAGEMENT

Data comes in all types in the digital age from countless sources



Structured or unstructured data



Sourced internally or externally

Without a smart data management strategy, organizations can quickly get overwhelmed – and fall behind the competition



**90%** of data has been generated since 2016



**95%** of businesses handle unstructured data



**4.4 GB** of data are used by Americans every minute



**50%** of businesses say that big data has changed their sales and marketing

## DATA LAKE VS DATA WAREHOUSES



Data lakes and data warehouses both act as repositories, but they are designed for very different purposes. Data warehouses work best for specific projects with set resources while data lakes are optimized for managing all incoming big data.



### Data Lake

- ▶ Structured, unstructured, and raw data
- ▶ Schema-on-read
- ▶ Large repository for data scientists
- ▶ Flexible configuration
- ▶ Low cost, high volume



### Data Warehouse

- ▶ Structured, processed data
- ▶ Schema-on-write
- ▶ Ready for reports by business users
- ▶ Fixed configuration
- ▶ Use with BI / analytics
- ▶ Cost scales with volume

VS

## DATA LAKES IN USE

Data comes from many types of sources. In fact, businesses rarely come to a data-driven decision with just one source. Instead, it takes numerous inputs, and much of that can be unstructured.

Data Lakes place these all in a single repository, saving **time, effort, and cost**



**5** Average number of data sources consulted to reach a data-driven decision



**80%** Amount of data that is unstructured and incapable of being handled by a data warehouse (original source IDC)

## DATA SOURCES



Data Lakes create a single repository for both structured and unstructured data, making it easy for data scientists to pull exactly what they need for analysis.

### Structured Data Sources

Invoices and receipts

Sensor data

Online forms

Spreadsheets

CRM profiles

### Unstructured Data Sources

Social media content

Emails

Podcasts

Security footage

Transcripts

## IS YOUR BUSINESS READY TO GET THE MOST OUT OF ITS DATA?



Go to [oracle.com/data-lake](https://oracle.com/data-lake) to learn why data lakes optimize data management in the era of big data