

# Data Lakes VS Data Warehouses

**Data Lakes** -> Centered repository used to *store, protect and process* a huge amount of data.

- > More *flexible* and *scalable*.
- > Stores large volumes of data in its *original format*.
- > Supports frameworks for *advanced data science*.

**Data warehouses** -> Data storage system that *centers and consolidates information from many sources*.

- > More *consistent* and *organized*.
- > Best suited for supporting *business intelligence* and *data analytics* efforts.

	Data Lake	Data Warehouse
Data Types	Any type of data (raw, unstructured and semi-structured data).	Structured, cleaned and processed data.
Use Cases	Data Science, ML and AI.	Business intelligence and reporting.
Data Structure	Schema is defined after data is stored.	Schema is defined before data is stored.
Data Access	Typically accessed by data scientists and engineers.	Typically accessed by managers and business users.