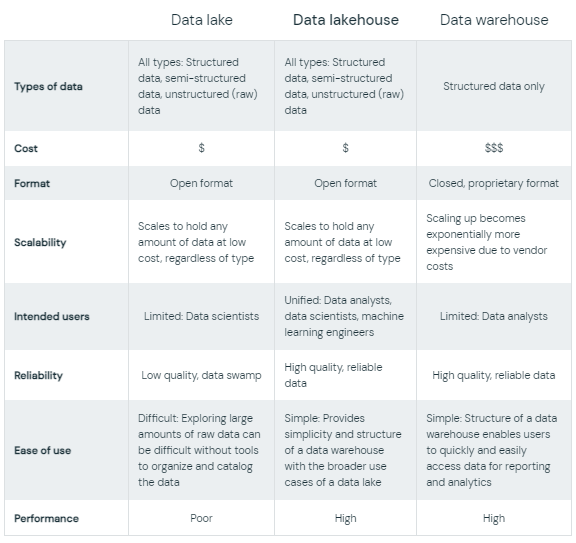
**Data Lakes Vs. Data Lakehouses Vs. Data Warehouses**

****

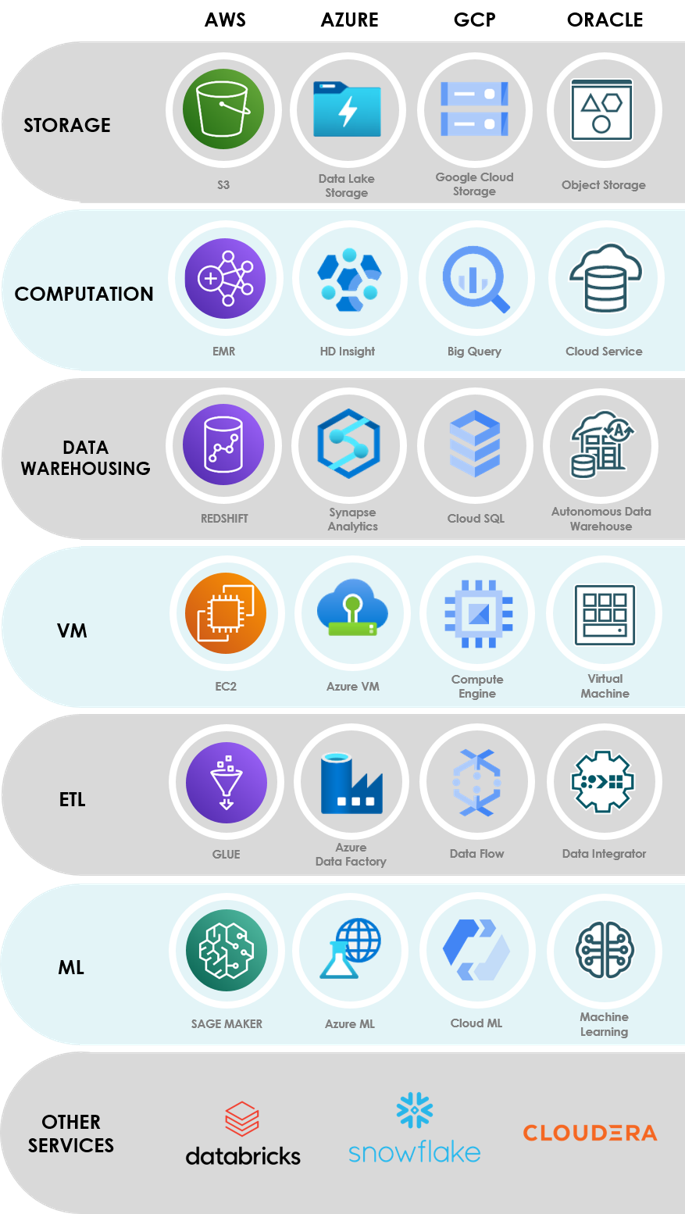
Leverage your business with Data Lake

### ***Scaling up your business with Data Lake***

In recent time, Data is become the valuable thing and every organisation Making profit or Growth along the Analytics with AI to take decisions. It is like if you got good data quality and relevant data it will give you good insights and results.

Data Lake is solution for the storing unlimited data in cloud and utilizing.

### ***Technologies that support Data Lake***

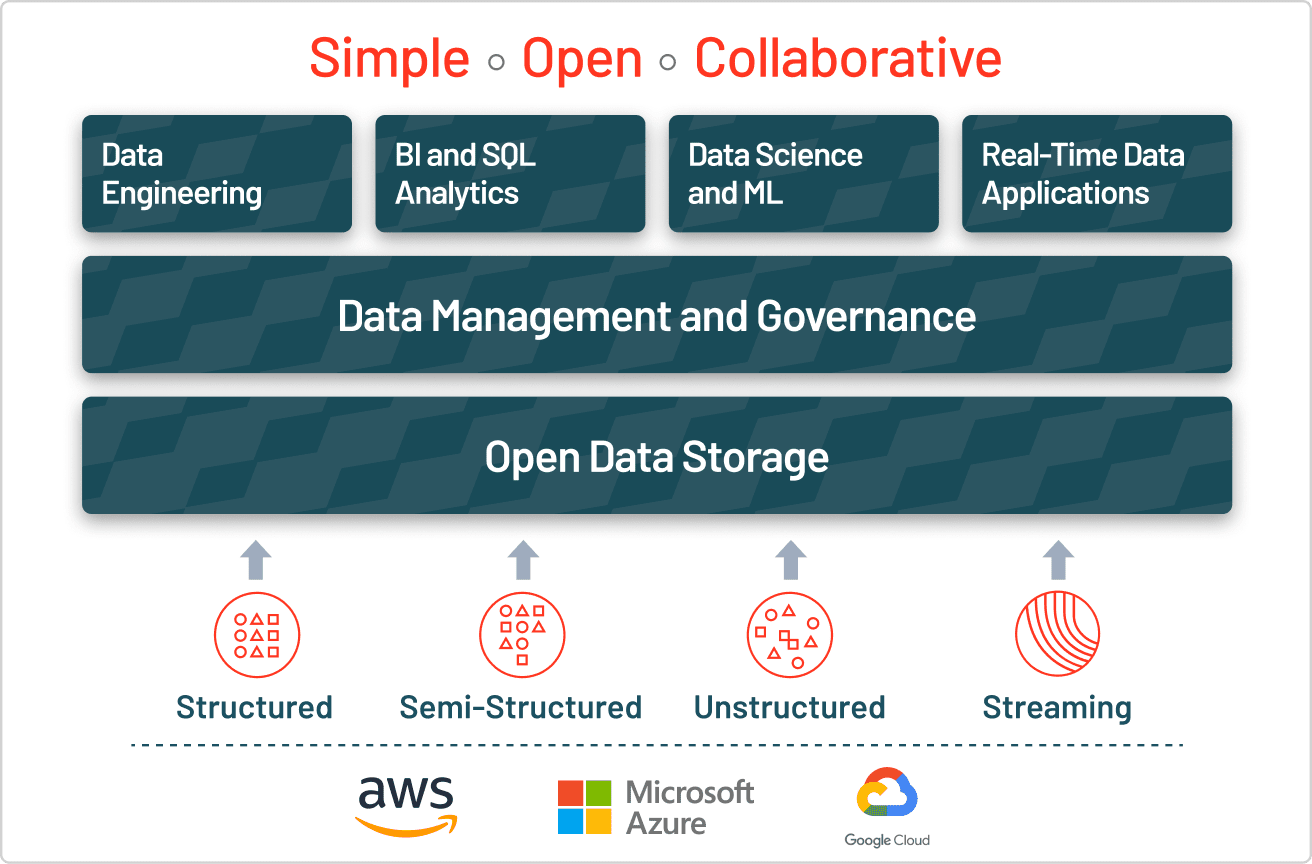


### ***What is Data Lake?***

### Data lakes provide a complete and authoritative data store that can power data analytics, business intelligence and machine learning. Data lakes can encompass hundreds of terabytes or even petabytes, storing replicated data from operational sources, including databases and SaaS platforms.

Data lakes were developed in response to the limitations of data warehouses.

i.e. a database stores the current data required to power an application. A data lake stores current and historical data for one or more systems in its raw form for the purpose of analysing the data.



### ***Journey to Data Lakes***

The rising internet usage generated large amounts of data. Traditional databases with their traditional storage are challenged for scale. Data warehouses are too expensive when it comes to storage, processing of unstructured data, and its support. Thus Data Lake comes in picture. Data Lakes allow you to import any amount of data that can come in real-time. Data is collected from multiple sources, and moved into the data lake in its original format. This process allows you to scale to data of any size, while saving time of defining data structures, schema, and transformations.   
 A distributed file system like Hadoop’s HDFS was a turning point for providing cheap storage on commodity hardware.

### ***Data Lake Market 2020-2025***

**Demand forecasts for Data Lake in the coming years vary, yet they all show a minimal double-digit compounded annual growth rate.**

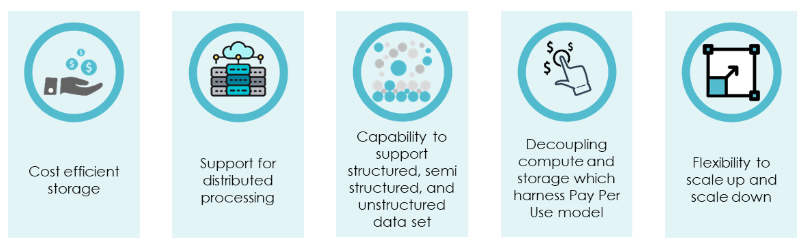
### ***Why is Data Lake widely used?***

It is an open format, so users can avoid lock-in to a proprietary system like a data warehouse, which has become increasingly important in modern data architectures. It is highly durable and low cost because of its ability to scale and leverage object storage.

It allows advanced analytics and machine learning on unstructured data, which are the most strategic priorities for enterprises today.

It has a unique ability to ingest raw data in a variety of formats (structured, unstructured, semi-structured), making it a clear choice for data storage.

### ***Data Lake Advantages***



### ***Data Lake Challenges***

The lack of some critical features:

No support for transactions,

No enforcement of data quality or governance, and

Poor performance optimizations.

Reliability issues -

Without the proper tools in place, data lakes can suffer from data reliability issues that make it difficult for data scientists and analysts to reason about the data. These issues can stem from difficulty combining batch and streaming data, data corruption and other factors.

Slow performance -

As the size of the data in a data lake increases, the performance of traditional query engines has traditionally gotten slower. Some of the bottlenecks include metadata management, improper data partitioning and others.

Lack of security features -

Data lakes are hard to properly secure and govern due to the lack of visibility and ability to delete or update data. These limitations make it very difficult to meet the requirements of regulatory bodies.