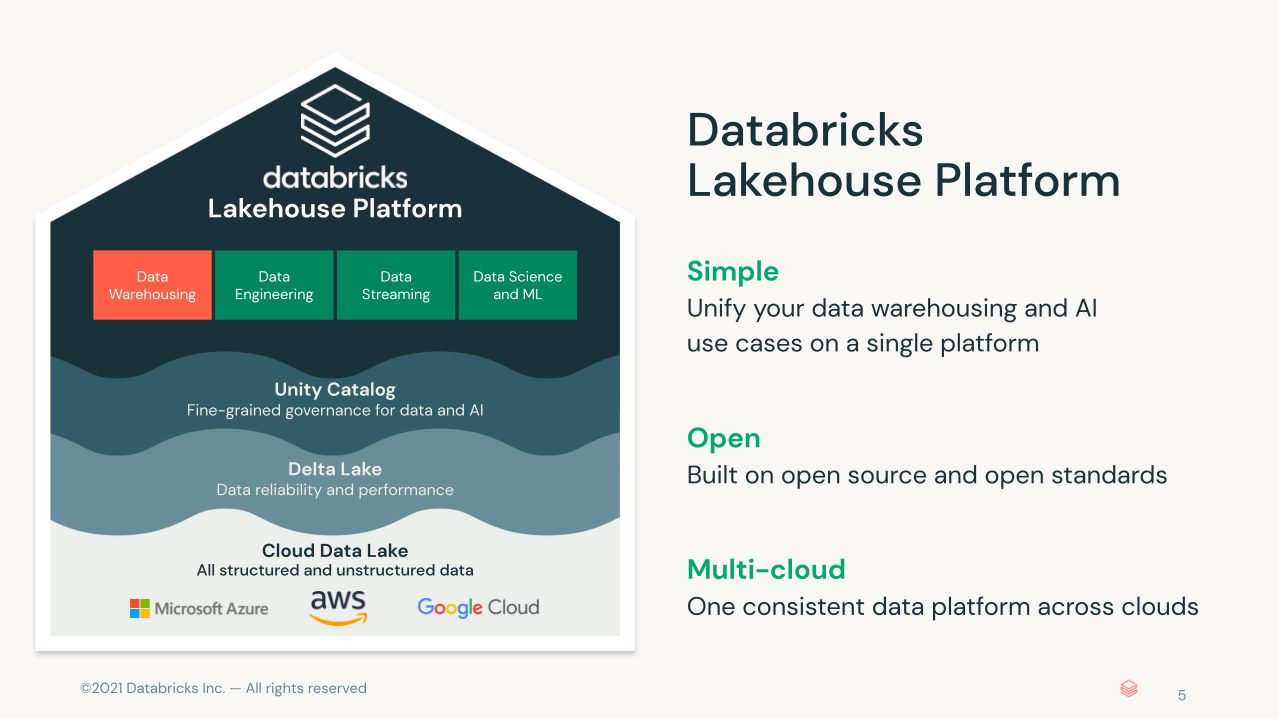
**Overview**

Azure Databricks is a data analytics platform optimized for the Microsoft Azure Cloud Services platform. Azure Databricks offers three environments for developing data-intensive applications: Databricks SQL, Databricks Data Science Engineering, and Machine Learning.

Databricks SQL provides an easy-to-use platform for analysts who want to run SQL queries in a data lake, create multiple visualization types to explore query results from different perspectives, and create and share dashboards.



**Exercise**

The objective of this lab is to know the query features (Query) of the Azure Databricks platform, using the SQL language (and the visual interfaces), exploring the Analytical potentials, and at the end, building a management panel (Dashboard).

This exercise uses a database of a Gymnastics Academy, where we will simulate an interactive exploratory analysis to evaluate the profile of the regulars, the times of use and the most frequented branches.

**Queries**

Exercise 1 – Create Database

CREATE DATABASE IF NOT EXISTS dbacademy;

GRANT ALL PRIVILEGES ON DATABASE dbacademy TO `<enter user name>`;

USE dbacademy;

CREATE TABLE intro\_to\_databricks\_sql\_gym\_logs

USING JSON

LOCATION 'wasbs://courseware@dbacademy.blob.core.windows.net/introduction-to-databricks-sql/v01/gym-logs';

Exercise 2 – Show data

SELECT

\*

FROM

dbacademy.intro\_to\_databricks\_sql\_gym\_logs;

Exercise 3 – Identify Popular Gyms

SELECT

gym,

count(gym)

FROM

dbacademy.intro\_to\_databricks\_sql\_gym\_logs

GROUP BY

gym

ORDER BY

gym;

Exercise 4 – Data Range

SELECT

from\_unixtime(min(first\_timestamp), 'd MMMM y') First\_Date,

from\_unixtime(max(last\_timestamp), 'd MMMM y') Last\_Date

FROM

dbacademy.intro\_to\_databricks\_sql\_gym\_logs;

Exercise 5 – Average Time Spent at the gym

SELECT

from\_unixtime(first\_timestamp, 'dd') as day,

avg((last\_timestamp - first\_timestamp) / 60) as avg\_time

FROM

dbacademy.intro\_to\_databricks\_sql\_gym\_logs

group by

day

ORDER BY

from\_unixtime(first\_timestamp, 'dd');