# Introduction

Delta lake is a framework which allows us for more efficient work with data saved in object storage (files). The main benefits are:

* **Enabling SQL** – Delta lake stores tables metadata like Hive and enables running SQL queries on them. It is also more efficient than Hive.
* **ACID transactions** – Delta lake assures ACID when performing data transformations.
* **Time travel (data versioning)** – It keeps track of what changes happen to data and we can go back to the previous version of a table, before we made some changes.
* **Provides useful functions for working with data** – functions such as:
  + Merge (combines insert, update and delete in one operation)
  + Functions helping creating Slowly Changing Dimensions
* **Protects data from corruption** – Our table is a folder consisting of multiple files. Without Delta lake we might save a new file in that folder with different columns or data types in columns and that will break the table. Delta lake makes sure problems like that don’t happen.

# Data structure

In Delta lake a table is represented as a folder consisting of multiple files.

That folder also contains logs with information about what changes happened to that table and with metadata describing that table (what enables running SQL queries on that table).

# Concepts