



LAKEHOUSE WITH DELTA LAKE

PART 3: GET STARTED WITH DELTA LAKE



Install Delta in Three Steps

- Set-Up Spark: <https://youtu.be/Gqw5rrfyfV4>
- Check the required Delta version based on your Spark Version.
E.g. with Spark 3.5.x you can use Delta 3.0.0 and 3.1.0
 - [Releases — Delta Lake Documentation](#)
- `pip install delta-spark==3.1.0` or chosen version

Delta Lake version	Apache Spark version
3.1.x	3.5.x
3.0.x	3.5.x
2.4.x	3.4.x
2.3.x	3.3.x
2.2.x	3.3.x
2.1.x	3.3.x

Summary

- Write or overwrite a delta table as follows:
 - `sdf.write.format("delta").mode(„append").save(path)`
 - `sdf.write.format("delta").mode("overwrite").save(path)`
- Check the Details:
 - `deltaTable = DeltaTable.forPath(spark, path_delta)`
 - `deltaTable.detail().show(20, False)`
- Check the history: `deltaTable.history().show(20, False)`
 - Remember all parquet tables are kept
- Transform Parquet to Delta:
 - `DeltaTable.convertToDelta(spark, f"parquet.`{path_parquet}`")`
- Read:
 - `spark.read.format("delta").option("versionAsOf", 2).load(path_delta).show()`