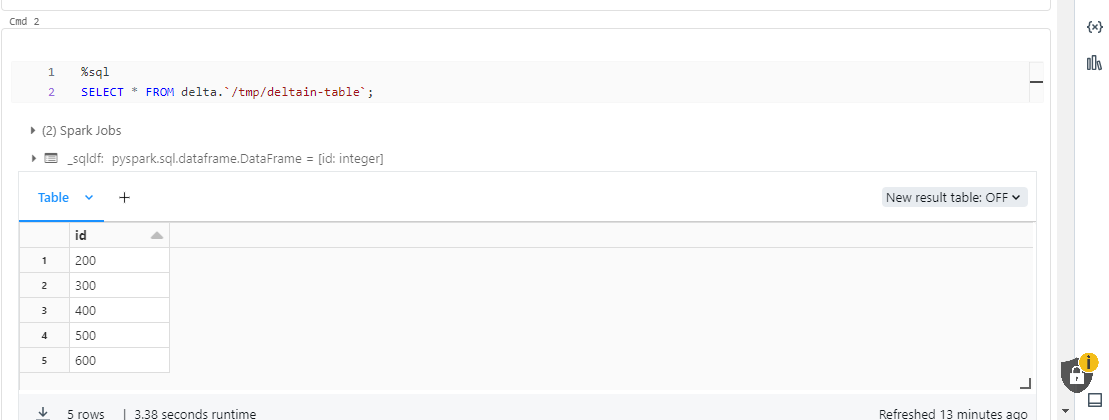
**AZURE ASSIGNMENT**

**DELTA LAKE**

Delta Lake is the optimised storage layer that provides the foundation for storing data and tables in the Databricks lakehouse. Delta Lake is open source software that extends Parquet data files with a file-based transaction log for ACID transactions and scalable metadata handling. Delta Lake is fully compatible with Apache Spark APIs, and was developed for tight integration with Structured Streaming, allowing you to easily use a single copy of data for both batch and streaming operations and providing incremental processing at scale.

**Select \* from table**

To create a Delta table, write a DataFrame out in the delta format. You can use existing Spark SQL code and change the format from parquet, csv, json, and so on, to delta.

****

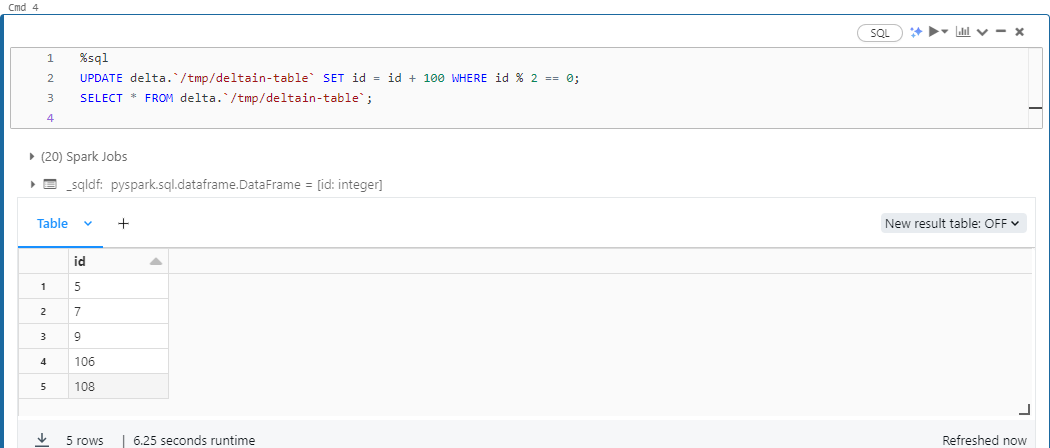
**Update table data**

Delta Lake supports several operations to modify tables using standard DataFrame APIs.

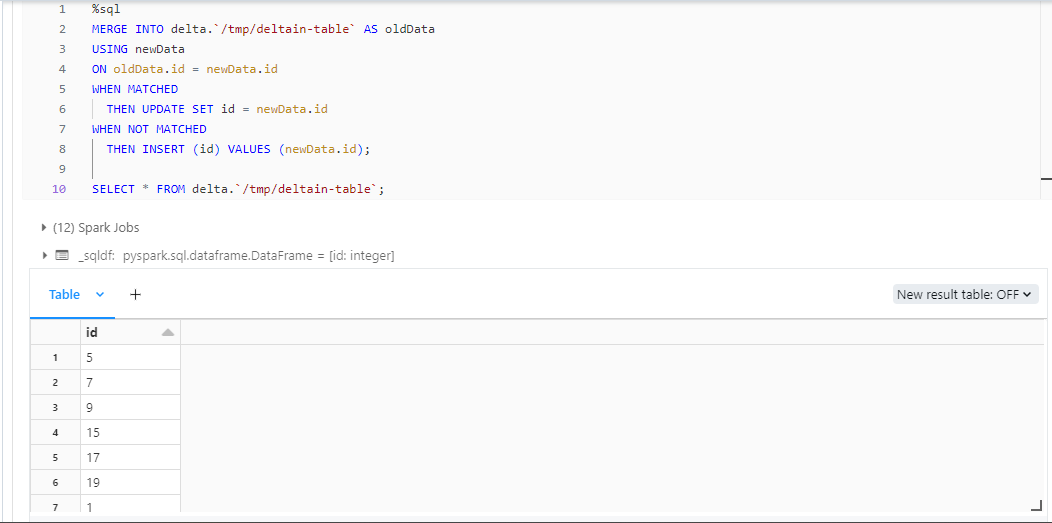
****

### **Conditional update without overwrite**

Delta Lake provides programmatic APIs to conditional update, delete, and merge (upsert) data into tables.

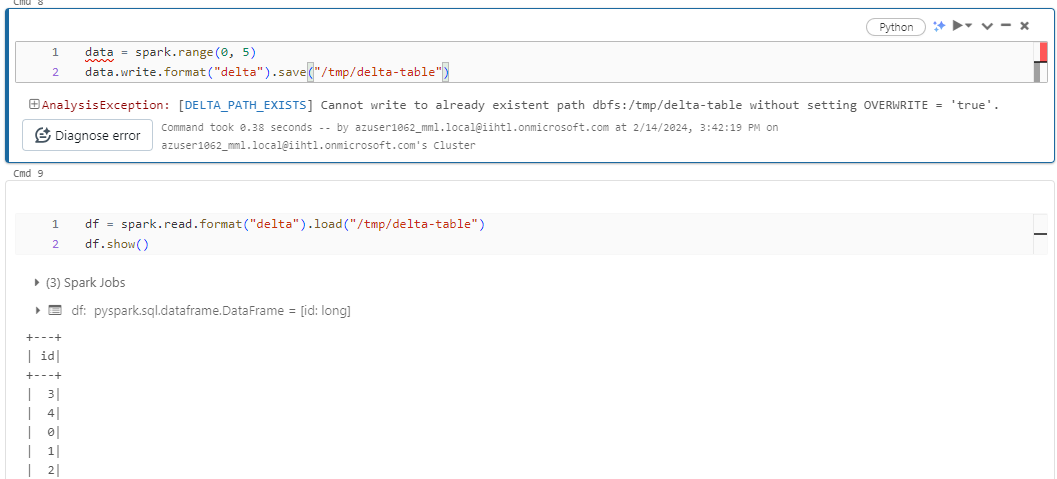
****

**Upsert (merge) new data**

****

## **Write a stream of data to a table**

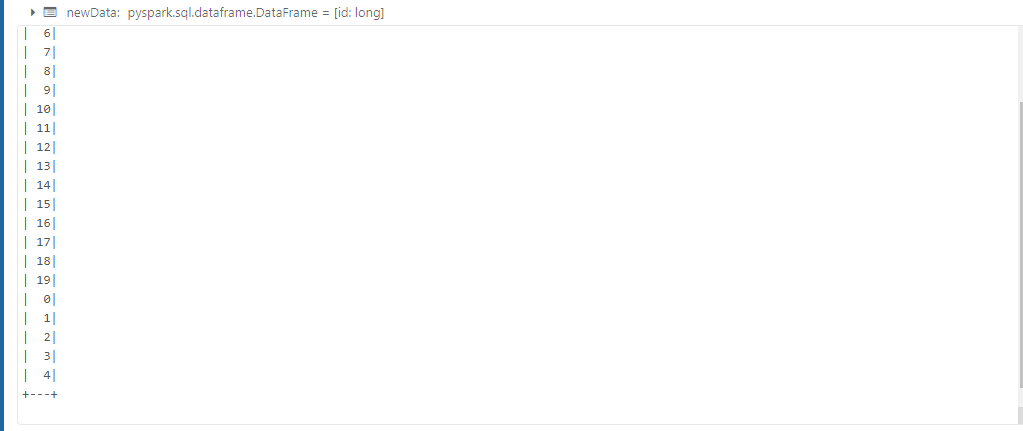
You can also write to a Delta table using Structured Streaming. The Delta Lake transaction log guarantees exactly-once processing, even when there are other streams or batch queries running concurrently against the table. By default, streams run in append mode, which adds new records to the table:



### **Conditional update without overwrite**

## Delta Lake provides programmatic APIs to conditional update, delete, and merge (upsert) data into tables.



****