

Trino/Starburst Training: Apache Iceberg table format

Starburst Academy
v1.0.0



**Session
2 of 5**



Connection Before Content

Lester Martin - <https://about.me/lestermartin>

- Educational Engineer @ Starburst
 - Build the content
 - Teach the class
 - Repeat
- 30 years of technology experience
 - Started my journey on a TRS-80 Model III
 - Played most every role, but consider myself a programmer at my core
 - Half of career in transactional systems and the second half in analytical processing
 - A DECADE of “big data” experience to include
 - Trino/Starburst, Hadoop, Hive, Spark
 - NiFi, Kafka, Storm, Flink
 - HBase, MongoDB

Session #2 objectives

- Move beyond Hive to understand the benefits of modern table formats such as Apache Iceberg.
- Run ACID compliant transactions to modify your data and understand how this creates new table versions.
- Explore time-travel queries and table rollbacks.
- Modifying the partitioning strategy without rebuilding your table.

Session #2 objectives

- Move beyond Hive to understand the benefits of modern table formats such as Apache Iceberg.
- Run ACID compliant transactions to modify your data and understand how this creates new table versions.
- Explore time-travel queries and table rollbacks.
- Modifying the partitioning strategy without rebuilding your table.

YES... in 1 hour!!

Session #2 objectives

- Move beyond Hive to understand the benefits of modern table formats such as Apache Iceberg.
- Run ACID compliant transactions to modify your data and understand how this creates new table versions.
- Explore time-travel queries and table rollbacks.
- Modifying the partitioning strategy without rebuilding your table.

YES... in 1 hour!!

SO... we'd better get moving!!

POIs

- The lab guide has it ALL!
- The S3 credentials will expire after the weekend
- A NEW SESSION EVERY 2 WEEKS
 - ~~January 31st – Creating & querying data lake tables~~
 - February 14th – *Modern table formats & Apache Iceberg (Happy Valentine's Day!!)*
 - February 28th – *Materialized views & data pipelines*
 - March 13th – *Experience Warp Speed in action*
 - March 27th – *Transformation processing with PyStarburst*
- Copy/n/paste “issues” << if SQL errors out, verify your editor has it right

POIs

- The lab guide has it ALL!
- The S3 credentials will expire after the weekend
- A NEW SESSION EVERY 2 WEEKS
 - ~~January 31st – Creating & querying data lake tables~~
 - February 14th – Modern table formats & Apache Iceberg (Happy Valentine's Day!!)
 - February 28th – Materialized views & data pipelines
 - March 13th – Experience Warp Speed in action
 - March 27th – Transformation processing with PyStarburst
- Copy/n/paste “issues” << if SQL errors out, verify your editor has it right

Session 1 recording available

Today's approach

- I'm going to perform the labs myself & you can...
 - do them along with me
 - or just watch (and maybe do later)
- We ALL come from DIFFERENT experiences & backgrounds, so...
 - **ASK QUESTIONS**
 - **FILL UP THE CHAT!**
 - **I'LL STOP DOING THE EXERCISES TO ADDRESS QUESTIONS!!**

Today's approach

- I'm going to perform the labs myself & you can...
 - do them along with me
 - or just watch (and maybe do later)
- We ALL come from DIFFERENT experiences & backgrounds, so...
 - **ASK QUESTIONS**
 - **FILL UP THE CHAT!**
 - **I'LL STOP DOING THE EXERCISES TO ADDRESS QUESTIONS!!**





HANDS-ON LABS!!

Starburst Academy