GNARLY Data_Waves

PRESENTED BY dremio

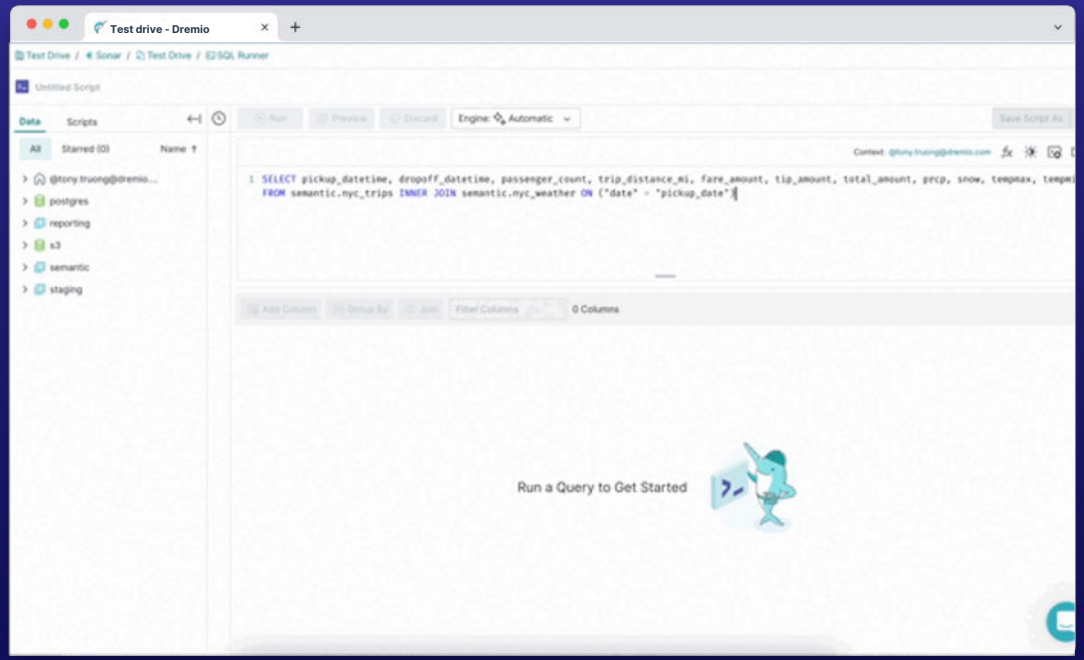
EPISODE 11

Apache Iceberg Office Hours

Experience the data lakehouse with Dremio Test Drive

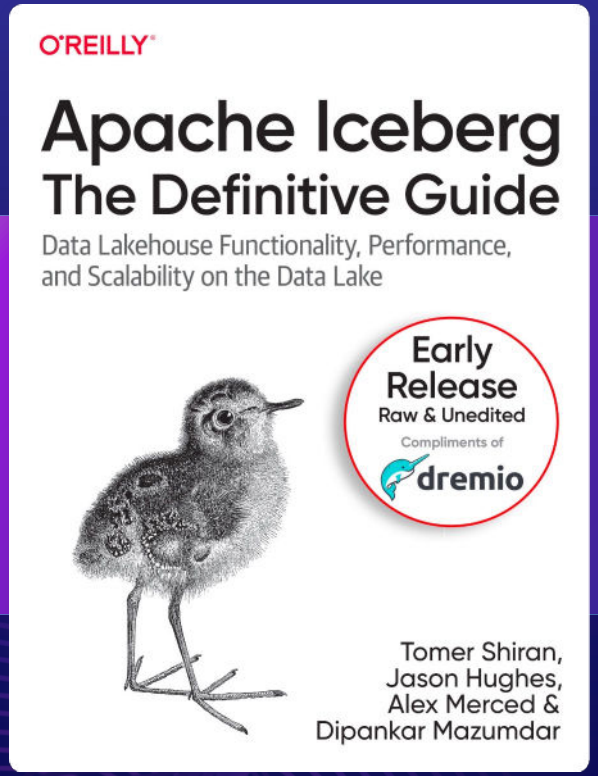
- Sub-second query on 1 billion rows of data joining Amazon S3 with a Postgres database
- Connect to Tableau or Power BI and build a dashboard with this dataset
- Everything hosted by Dremio - 100% free for you

Start Test Drive



GNARLY Data_Waves

Apache Iceberg: The Definitive Guide



Upcoming shows

Register now

EPISODE 12

How to Modernize Hive to the Data Lakehouse with Dremio and Apache Iceberg

April 11, 2023 at 8AM PST | 11AM EST | 4PM GMT

EPISODE 13

Making the Move: Five Factors to Consider When Migrating from Hadoop to the Data Lakehouse

April 18, 2023 at 8AM PST | 11AM EST | 4PM GMT

EPISODE 14

Enabling data mesh with Dremio Arctic and Data as Code

April 25th, 2023 at 8AM PST | 11AM EST | 4PM GMT

EPISODE 15

Getting Started with Dremio

May 2, 2023 at 8AM PST | 11AM EST | 4PM GMT

EPISODE 16

Automatic Data Optimization with Dremio Arctic

May 2, 2023 at 8AM PST | 11AM EST | 4PM GMT

GNARLY Data_Waves

PRESENTED BY dremio

EPISODE 11

Apache Iceberg Office Hours

Jason Hughes
Director of Product Management

Alex Merced
Developer Advocate

Dipankar Mazumdar
Developer Advocate

Ajantha Bhat
Open Source Software Engineer

Dmitri Bourlatchkov
Senior OSS Developer

Scott Cowell
Software Engineer

dremio | ICEBERG

Apache Iceberg 1.2.0

- Added AES GCM encrpytion stream spec (#5432)
- Added support for Delta Lake to Iceberg table conversion (#6449, #6880)
- Added support for `position_deletes` metadata table (#6365, #6716)
- Added support for scan and commit metrics reporter that is pluggable through catalog (#6404, #6246, #6410)
- Added support for branch commit for all operations (#4926, #5010)
- Added `FileIO` support for ORC readers and writers (#6293)
- Updated all actions to leverage bulk delete whenever possible (#6682)
- Updated snapshot ID definition in Puffin spec to support statistics file reuse (#6272)
- Added human-readable metrics information in `files` metadata table (#5376)
- Fixed incorrect Parquet row group skipping when min and max values are `NaN` (#6517)
- Fixed a bug that location provider could generate paths with double slash (`//`) which is not compatible in a Hadoop file system (#6777)
- Fixed metadata table time travel failure for tables that performed schema evolution (#6980)