

**GNARLY**  
Data\_Waves

PRESENTED BY  **dremio**

Episode 55

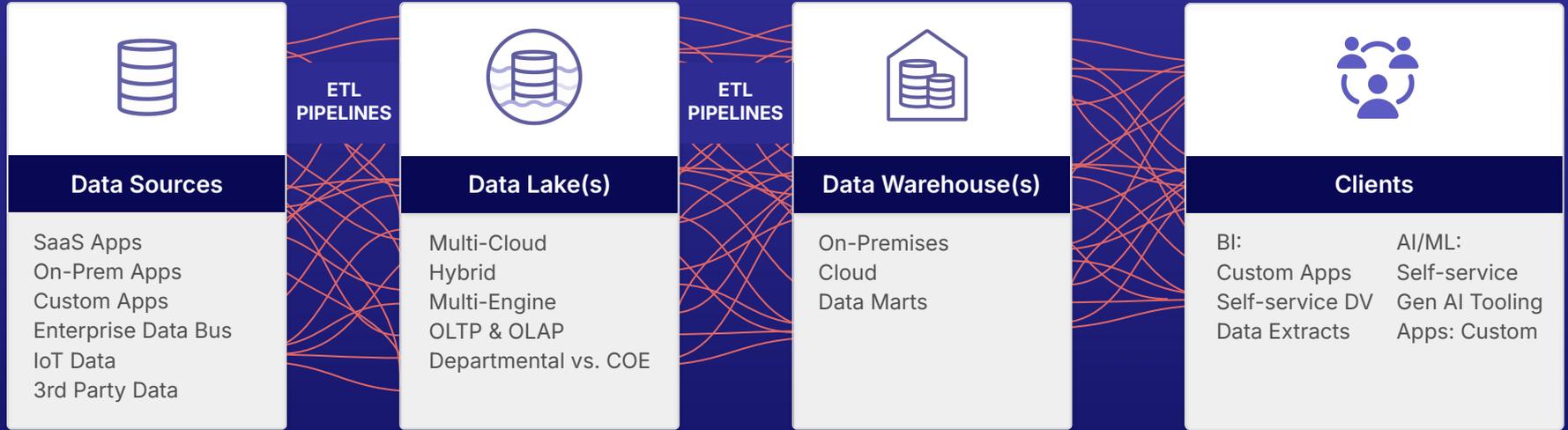
# Unite Data Across Dremio, Snowflake, Iceberg, and Beyond



**Maeve Donovan**

Senior Product Marketing Manager @ Dremio

# Introduction



**Data lifecycle and management remains complex, especially for large organizations**  
Duplicate copies, 'expert' ETL, "dark data", governance complexity, not self service

# The Challenges of Data Accessibility and Analytics

## Siloed Data



Inhibits discovery, analytics, and time-to-value

## Expensive Data Access



Rising costs due to increased demand, data size, and consumption models

## Limited Self-Service



Users dependent on data teams, lack easy-to-use tools

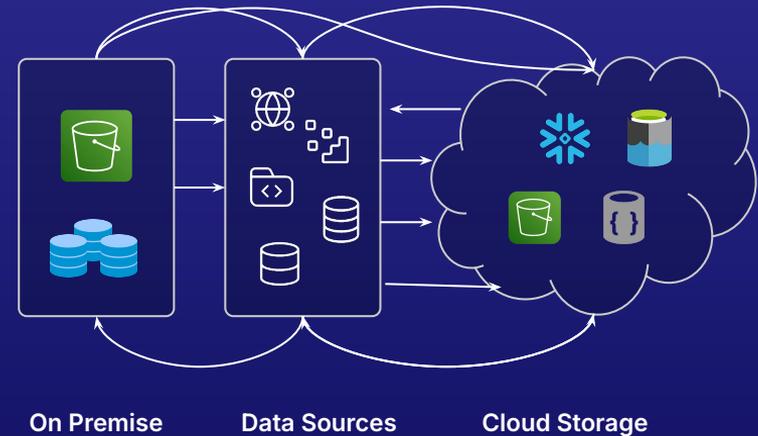
## Vendor Lock-In Risks



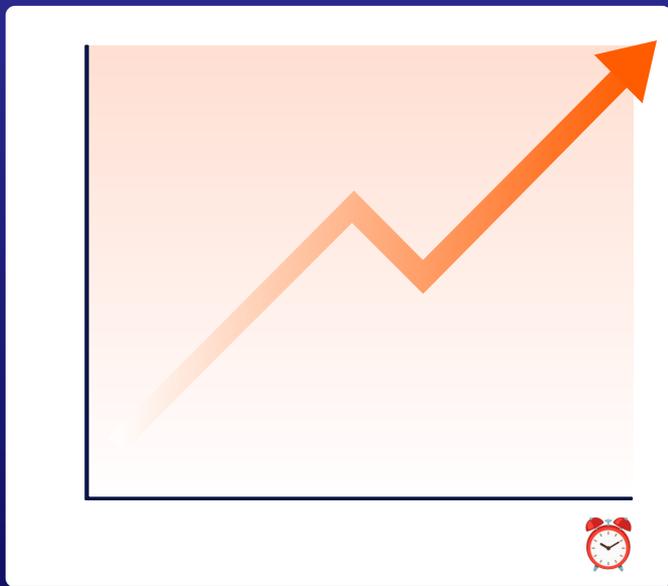
Proprietary platforms limit flexibility and tool choice

# Data Sprawl Across Multiple Platforms

- Data scattered across Snowflake and other sources
- Challenges in discovery
- Issues with cross-silo analytics



# Rising Costs of Analytics



- Increasing demand for data access inflating costs
- Cost management issues with consumption-based models

# Bottlenecks in Data Access

- Users encounter delays and limitations in accessing data
- Rely on data teams to prepare data for analysis
- Limits self-service capabilities

# Risks of Vendor Lock-in



- Inflexibility of closed data ecosystems
- Challenges in bringing the right tools to the problems

# Why Dremio?



**Unify disparate  
data sources**



**Intelligent  
semantic layer  
and powerful  
SQL platform**

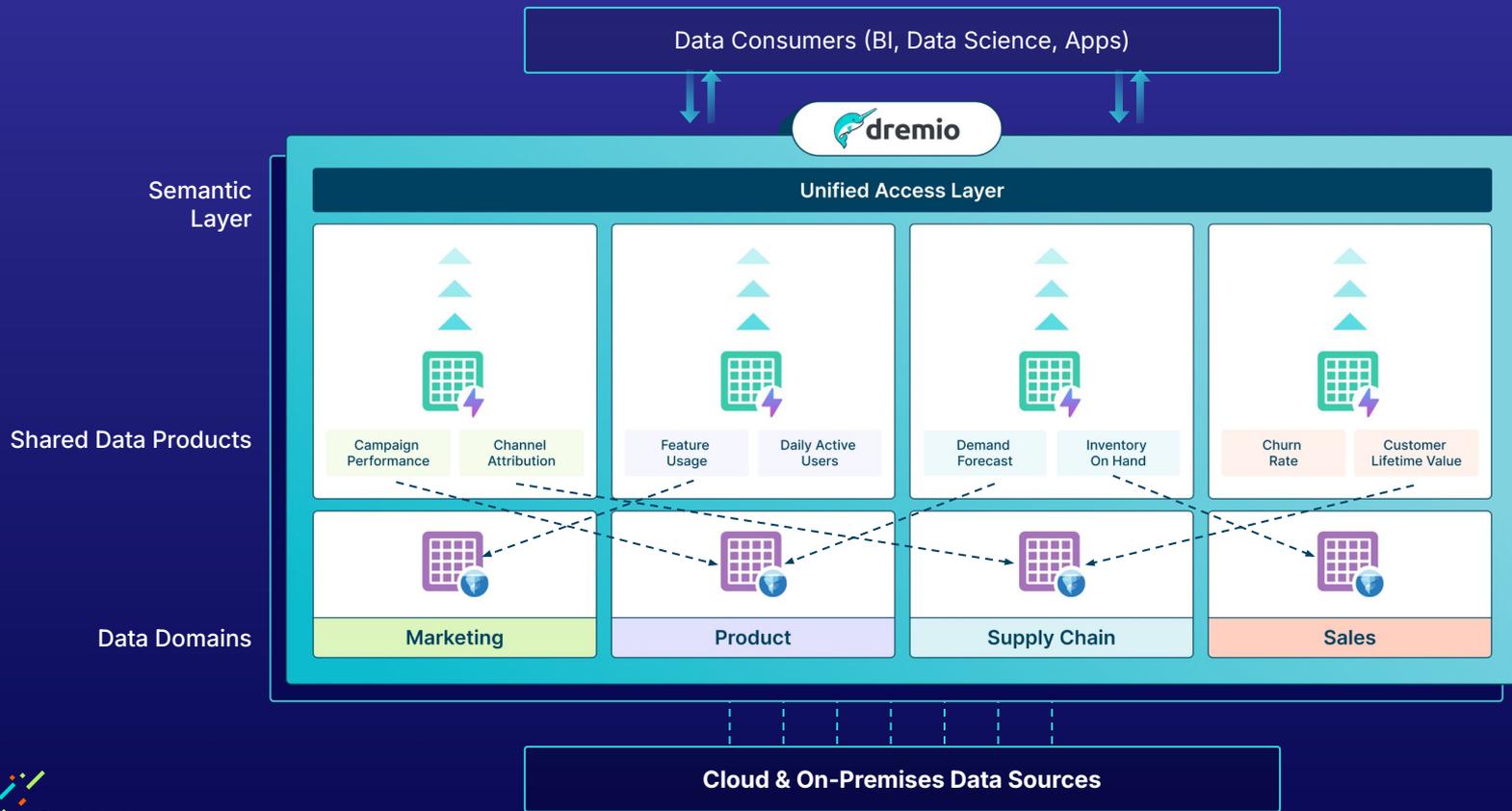


**Eliminate data  
silos and  
optimize cost**

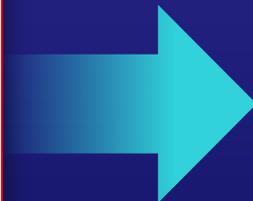


**Enable  
self-service  
analytics**

# Unified Lakehouse analytics without data copies or ETL



# From Challenges to Solutions



# Seamless Data Integration

- Unify all data without movement
- Reduce ETL
- Improved governance and security



# Maximize Cost Efficiency

- Enhance performance for BI and analytics
- Achieve significant cost savings



# Faster Time-to-Insight

- Lightning-fast, intuitive data access
- Empower users with self-service capabilities

The image displays two overlapping screenshots from a data management interface. The top screenshot shows a 'Wiki' page with a rich text editor on the left and a 'Generated wiki (Preview)' on the right. The content includes a 'Dataset Summary' for 'NYC-taxi-trips-iceberg' and a 'Column Summary' for 'pickup\_datetime'. The bottom screenshot shows a table view for 'Alliances' with columns for 'Name' and 'Alliances'. The table lists various data sources and their counts, such as 'Curated Data' (1), 'dbt Datasets' (0), 'dev' (6), 'Product' (0), 'Production Data' (0), 'Raw Data' (4), 'Sales' (1), 'Tableau Data' (2), 'Trips' (2), and 'Yelp' (4). Below the table, there are sections for 'Sources' (Arctic Catalogs) and 'Metastores'.

**Wiki**

**Dataset Summary:**

- Dataset Name: NYC-taxi-trips-iceberg
- Dataset Description: This dataset contains information about taxi trips in New York City. It includes details such as the pickup datetime, passenger count, trip distance in miles, fare amount, tip amount, and total amount.

**Column Summary:**

1. pickup\_datetime:
  - Data Type: DATETIME
  - Summary: This column represents the date and time when the taxi trip was picked up.
2. passenger\_count:
  - Data Type: INTEGER
  - Summary: This column indicates
3. trip\_distance\_mi:
  - Data Type: FLOAT
  - Summary: This column represent
4. fare\_amount:
  - Data Type: FLOAT
  - Summary: This column indicates
5. tip\_amount:
  - Data Type: FLOAT
  - Summary: This column represent

**Generated wiki (Preview)**

**Dataset Summary:**

- Dataset Name: NYC-taxi-trips-iceberg
- Dataset Description: This dataset contains information about taxi trips in New York City. It includes details such as the pickup datetime, passenger count, trip distance in miles, fare amount, tip amount, and total amount.

**Column Summary:**

**Alliances**

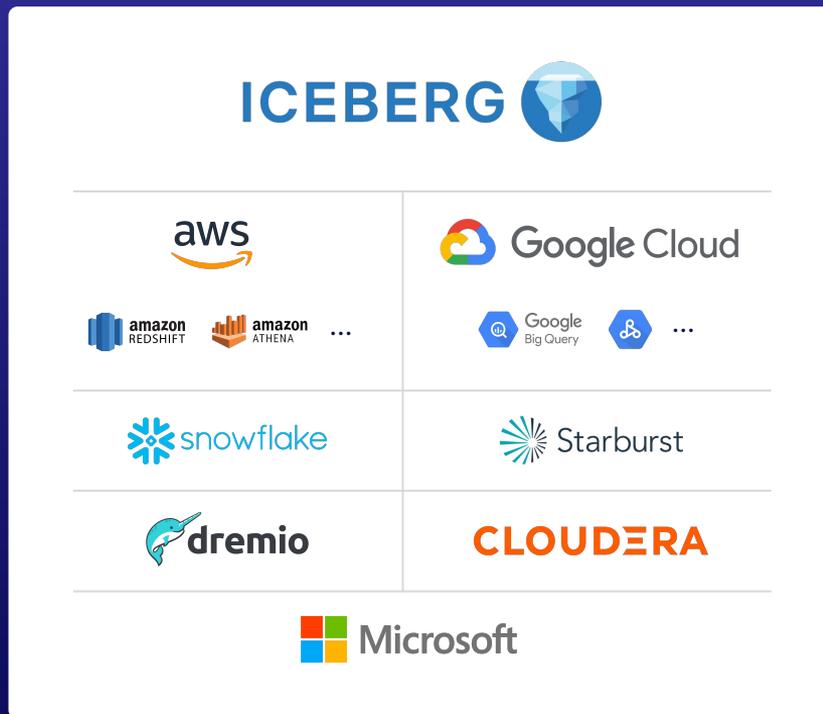
Name	Alliances
Account Mapping	
AWS Co-Sell Table	
Brett Roberts	
Henry	
MSFT Co-Sell Table	

**Sources**

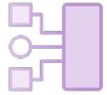
- Arctic Catalogs (1)
  - arctic
- Metastores (7)
  - Add Source

# Apache Iceberg Compatibility

- Ensure open and compatible data access
- Leverage Iceberg with Snowflake and Dremio



# Real-World Impact



Unified data access with virtualization and federation



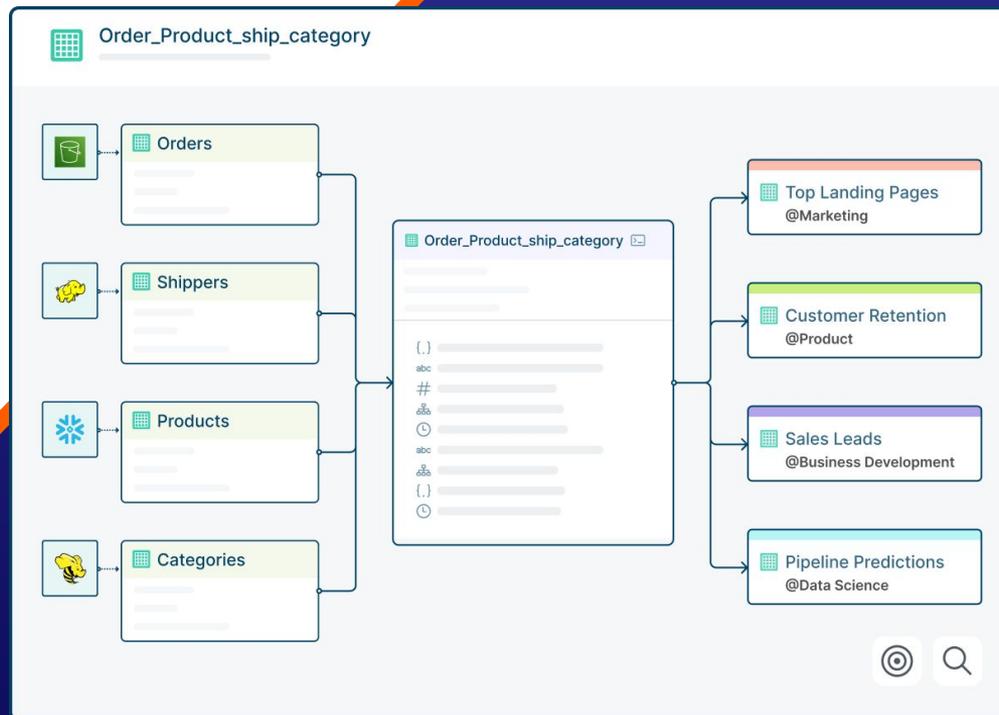
Optimize costs and performance with Reflections



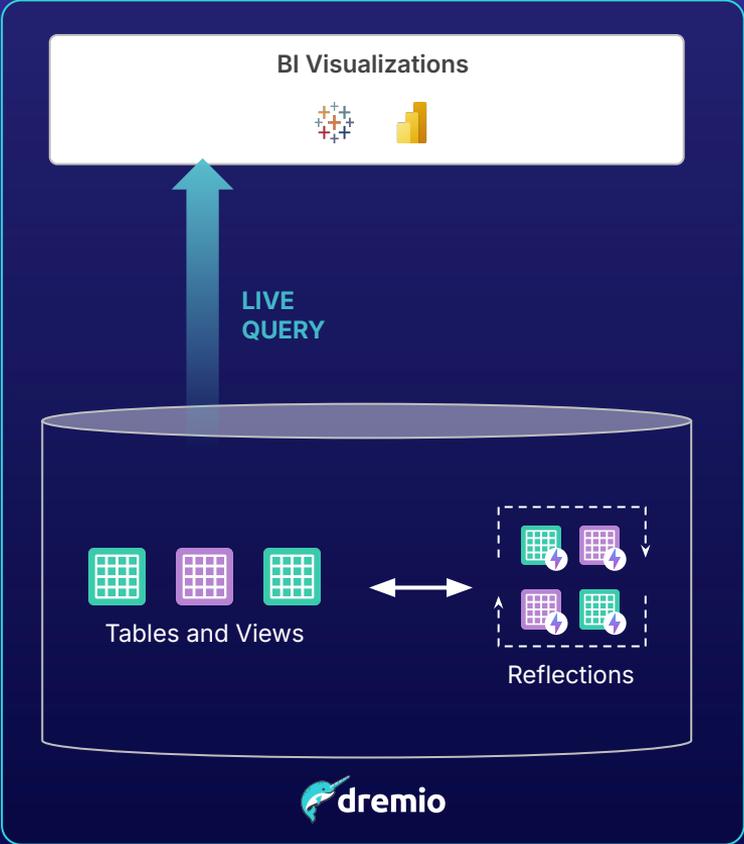
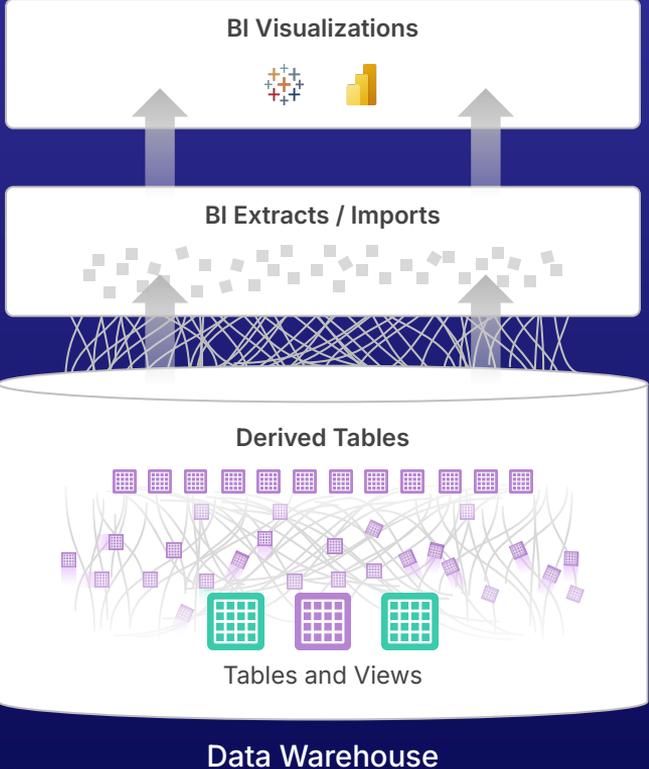
Enhanced self-service analytics capabilities

# Virtualization and Federation

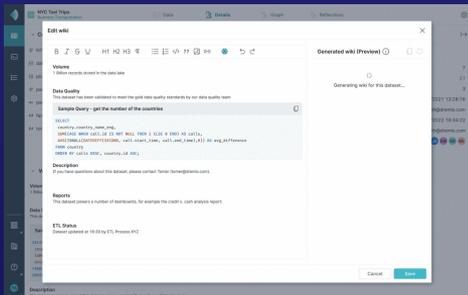
Access all data seamlessly,  
reducing ETL complexity



# Efficiency Gains

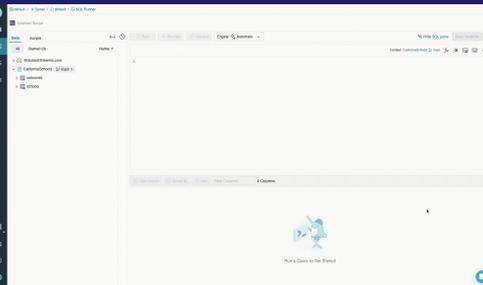


# Empowering Users



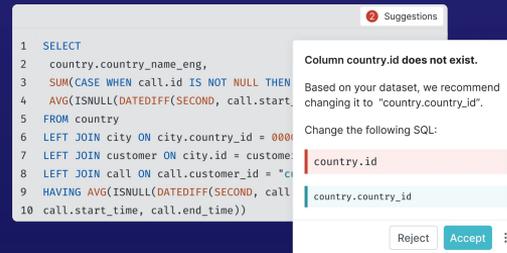
## Data Wikis & Data Labeling

Automatic generation of dataset descriptions, SQL examples, data tags for tables and views and more



## Text-to-SQL

Generate SQL queries based on natural language questions



## Debugging SQL/Data

Automatic recommend and correct SQL syntax errors

# Future-Proof Your Data

- Ensure open and compatible data formats with Iceberg
- Dremio is Apache Iceberg native
- Enterprise Hybrid Cloud Iceberg Catalog



# Conclusion



ABC Supply delivers self-service analytics using Dremio to organize domain data and accelerate development



Nomura modernized their data lake delivering unparalleled efficiency with Dremio Unified Lakehouse Platform to enable massive workload scaling.



RWE empowered analysts with quicker access to data, self-service insights, and analytics to enable collaboration and data sharing with Dremio.

The screenshot shows the Dremio website landing page. At the top, there is a navigation bar with the Dremio logo and links for Product, Solutions, Customers, Partners, Resources, Company, and Apache Iceberg. On the right side of the navigation bar, there are links for 'Login' and 'Start for Free', along with a 'Book a Meeting' button. The main content area features the heading 'Dremio with Snowflake for Easy, Cost-Effective Iceberg Analytics'. Below this heading, there is a sub-heading 'DREMIO SOLUTIONS' and a paragraph of text: 'Unify all of your data, from Snowflake to Iceberg and other sources, by combining Dremio's intelligent semantic layer with its powerful SQL platform, to eliminate data silos, optimize costs, and enable self-service analytics'. A 'Start For Free' button is located at the bottom left of the main content area. On the right side of the main content area, there are three icons: a Dremio logo, a Snowflake logo, and an Iceberg logo.

# Thank you

**GNARLY**  
Data\_Waves

PRESENTED BY  **dremio**