research

WEBINAR

Unlock the Power of Spatial Analysis and Data Lakehouses for Better Business Decisions

tdwi

Transforming Data With Intelligence™

tdwi.org

Sponsors

DAVID STODDER

Senior Research Director
Business Intelligence
TDWI
dstodder@tdwi.org
@dbstodder

research

tdwi
Transforming Data With Intelligence™

research

WEBINAR

TDWI Research Perspectives:
Unlock the Power of Spatial Analysis and Data Lakehouses for Better Business Decisions

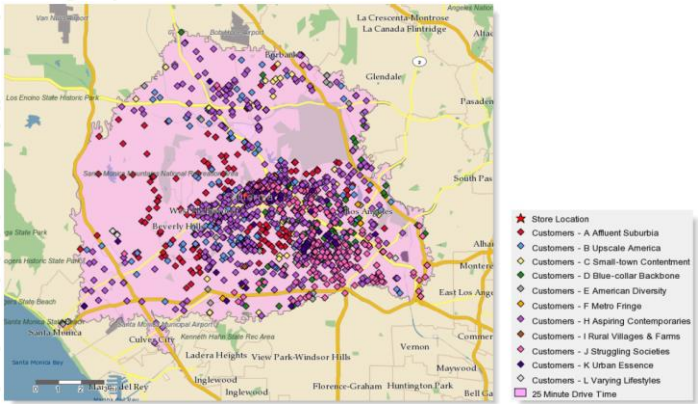
tdwi
Transforming Data With Intelligence™

tdwi.org

Spatial Analysis: Key to New Data Insights

Organizations want next-level understanding of causes, correlations, patterns, trends, and more about customers, partners, competitors, etc.

- **Digital transformation:** Reducing risks and recognizing opportunities by using geospatial data & analytics
- **Context:** Information that helps explain events, transactions, market competition, fraud and abuse
- **Avoiding surprises**: Location intelligence enables organizations to visualize situations and respond

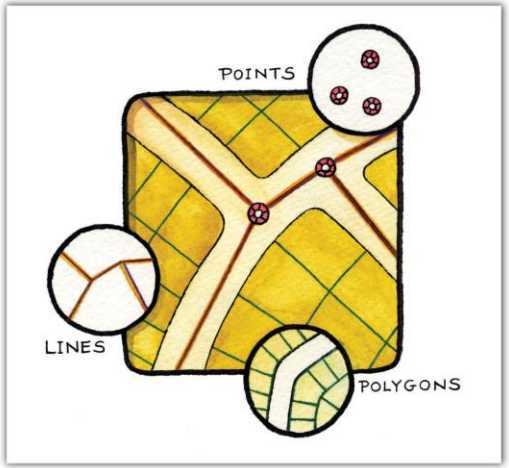


research

tdwi
Transforming Data With Intelligence™

Spatial Data: Diverse and Growing

- **Geospatial data**: Information about a specific location, often address-related; geocoding to match to specific latitude and longitude
  - Could be objects or events with a location
  - Roads, buildings, property, terrain, etc.
  - Typically combines location information (coordinates), attribute information (characteristics such as type, e.g., a house), and sometimes temporal
  - Could include raster data (e.g., image)
  - Can be static or dynamic (a moving vehicle)
- **Sources** can be varied, with volume rising



Geospatial vector data. Credit: saylordotorg.github.io

Location Intelligence for Powerful Analytics

- **Visual location intelligence**: Answering spatial analysis questions to derive new insights
  - Spatial data science: Going beyond simple inclusion of maps in business intelligence apps
  - Why did this happen at this location? Impact for geo-related entities? Efficiency and resilience
  - Location decisions for retail, real estate, inventory
  - 34%: Making it easier to discover data relationships is key to strategy (42% for future)

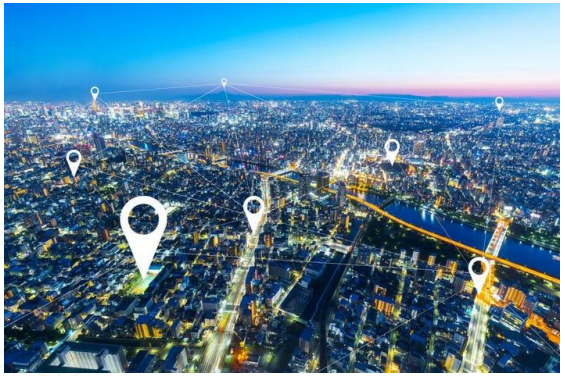


Image credit: Getty Images

- **Customer and market insights:** New dimensions for understanding marketing effectiveness, competitive situations
- **Maximizing value of new data types:** Mobile, IoT sensors, connected vehicles (logistics)

research

tdwi
Transforming Data With Intelligence™

Data Challenges for Spatial Analysis

- **Scalability and speed of access**
  - Collecting and managing new and larger data (such as spatial) is a challenge
    - 33% collect and process geospatial data
  - Use of cloud data lakes: 44% currently using and 28% plan to use
    - Cloud data warehouse: 43% current and 31% planning to use
- **Getting trusted and accurate data**
  - Diversity of sources (including global) creates challenges
- **Data integration and enrichment**
  - Reliance on legacy ETL and expert programmers slows value

research

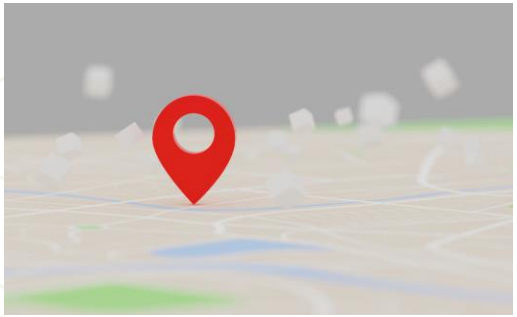
tdwi
Transforming Data With Intelligence™

Poll Question

- What is your biggest challenge in furthering spatial analysis and location intelligence for better business decisions?
  - Gaining buy-in from executives for investment and development
  - Having the right spatial analytics skills and training
  - Dependence on GIS experts and specialized programmers; not enough self-service for business users
  - Data quality, consistency, and accuracy concerns
  - Data management and integration of large and diverse data needed for location intelligence
  - Other (please tell us using the Ask a Question box)

Important Solution Trends for Spatial Analysis

- **Unified data architecture (e.g., a data lakehouse):** 33% say this is very important in their current data strategy and 44% for future
  - Streamlines ability to analyze new and varied data sets; perform spatial analysis directly rather than having to export to a specialized data platform
- **Ease of use and self-service**
  - Tech trends are changing what is possible for visual data interaction
  - Enabling nontechnical users to make location intelligence integral to daily decisions



research

tdwi
Transforming Data With Intelligence™

Conclusion: Enabling Better Business Decisions

- **Innovation:** Using spatial analysis for clearer and more complete insights
  - Customer intelligence: Improving targeting and segmentation; knowing where to invest for higher sales growth; being proactive in location decisions
- **Situation awareness**
  - Data relationships: Visualizing how related areas tied to locations, e.g., inventory, logistics, and supply chain processes are affected; population health impact
  - Data science for tracking trends and patterns; predictive insights
- **Future-proofing data management and self-service access**
  - Unified architecture (data lakehouse) to make spatial data discovery and analysis more efficient and complete; reducing challenge of silos
  - Data visualization for faster insights by democratized users

Thank You

David Stodder
Senior Director of Research for Business Intelligence
TDWI (www.tdwi.org)
dstodder@tdwi.org
@dbstodder

research

tdwi
Transforming Data With Intelligence™

Michael Johns

Senior Specialist Solutions
Architect
Databricks

research

tdwi
Transforming Data
With Intelligence™

Matthew Forrest

VP of Spatial Data Science
CARTO

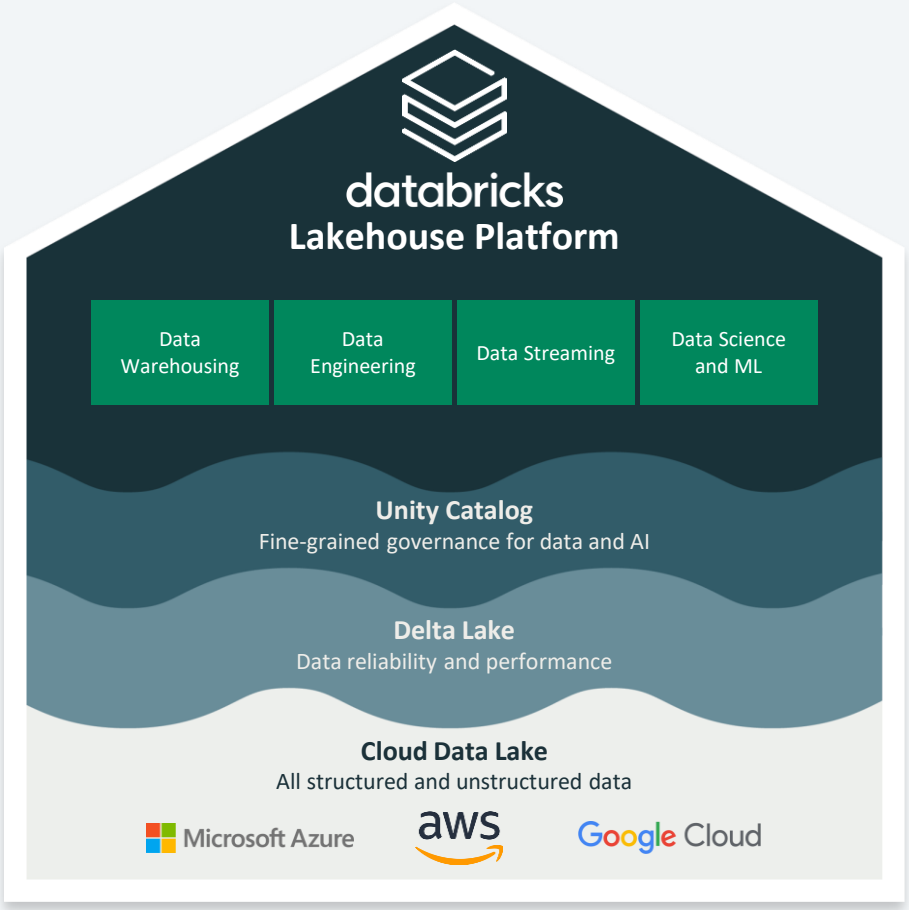
research

tdwi
Transforming Data With Intelligence™

databricks

Databricks Lakehouse Platform & Geospatial Ecosystem

Michael Johns | Geospatial Product Specialist | Databricks



Databricks Lakehouse Platform

Simple

Unify your data warehousing and AI use cases on a single platform

Scalability

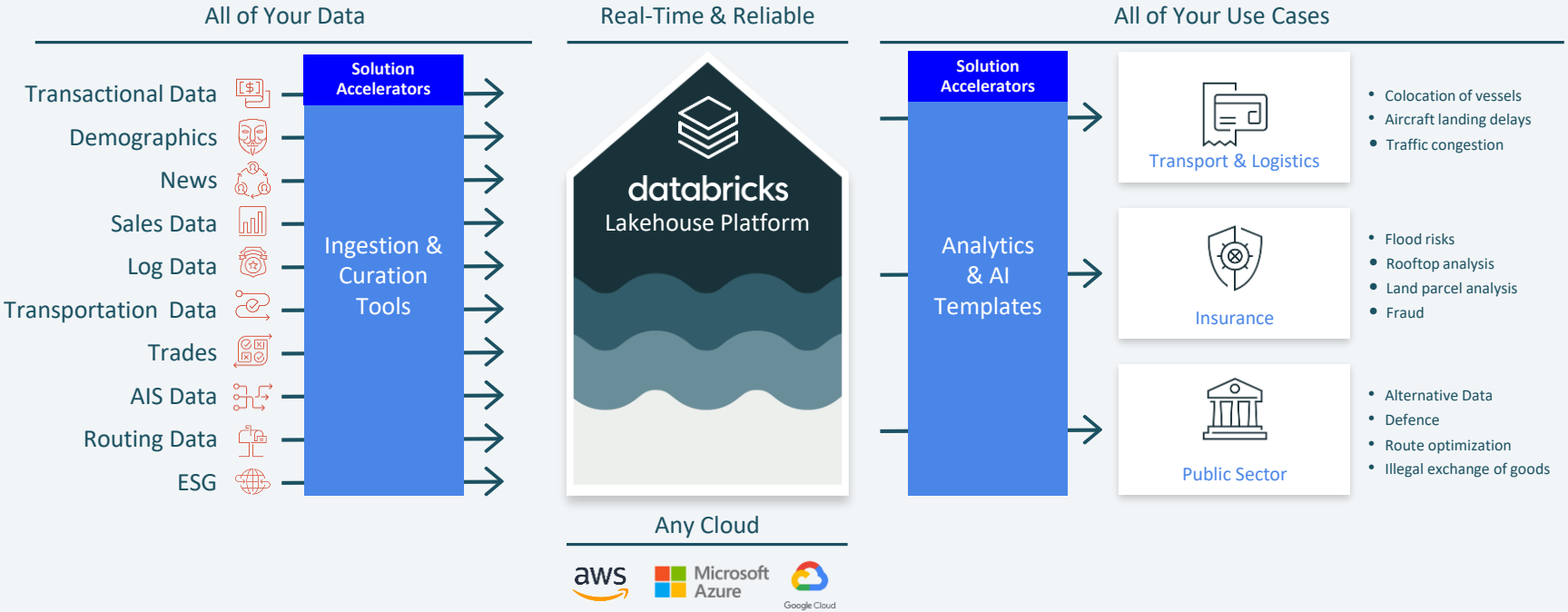
World-class performance with Data Lake economics

Open + Multicloud

Built on open source and open standards

Geospatial Analytics on Lakehouse

Helping organizations build a data asset strategy to enable multiple use cases



Scalable Geospatial Analytics with H3

Supported natively in Databricks

Grid indexing systems are ideally suited for scale

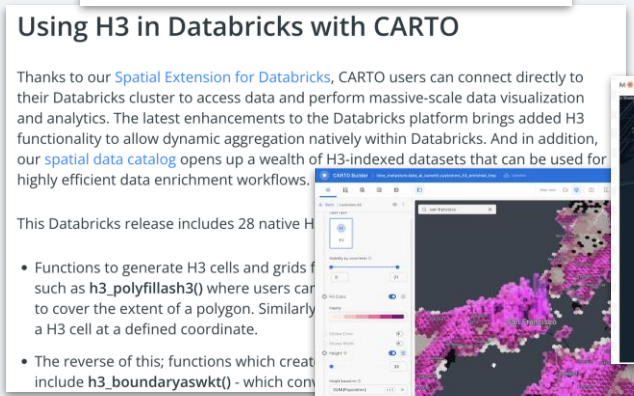
Hierarchical system offers flexibility

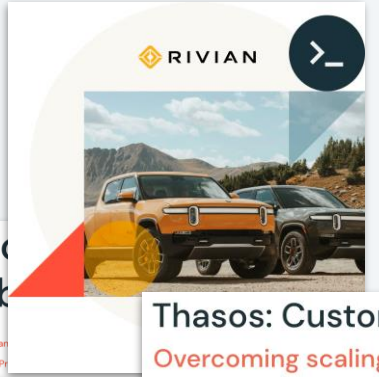
Easy and effective visualization

Precision Agriculture	Autonomous Vehicles	Retail Operations	Climate Risk/Modeling	Human Mobility
				

Databricks Solutions and more

Get your Geospatial Lakehouse journey started





Broad Platform Ecosystem

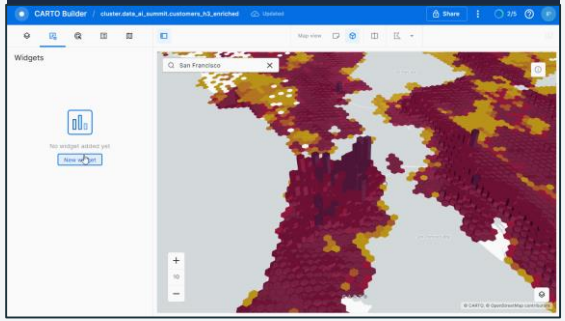
"Flexibility" to choose your own adventure for geospatial processing

Geo data
Non-geo data

What's best for my use case?

pdal
Apache Sedona
GeoPandas
RasterFrames

CARTO + databricks



- H3 integration
- Analytics Toolbox
- pydeck-carto visualization
- Spatial Features for enrichment
- Databricks Data Marketplace

+

CARTO helps Databricks enable large-scale geospatial analytics for the non-traditional user by providing access to powerful geospatial data, analytics & visualization capabilities natively on Databricks.

Where CARTO Adds Value

Connectivity

CARTO connects to the data warehouse, tools, and services you already use to simplify your workflow.

Premium Data

We spend time vetting providers so you don't have to, providing you with the highest-quality, most accurate and relevant data globally.

Accessibility

We make Location Intelligence accessible to all types of end user. No training or GIS experience needed.

End-to-End Offering

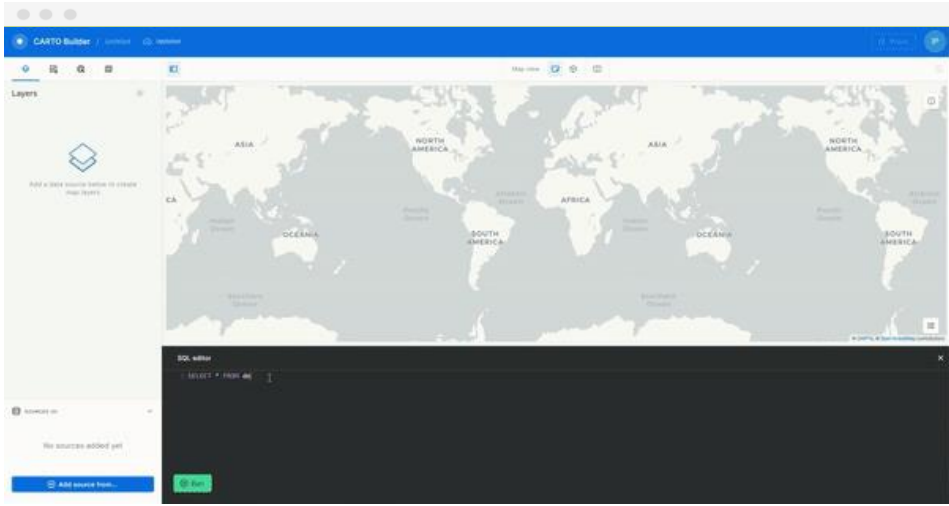
Get access to our whole tech stack with an Enterprise license: APIs, developer libraries, Python integrations, Location Data Services.

CARTO Spatial Extension for Databricks

A platform that extends Databricks to unlock Spatial Analytics

CARTO helps companies solve spatial problems using location data and analysis to understand where and why things happen, optimize processes, and predict future outcomes through Spatial Data Science.

We want to democratize access to Spatial Analytics, and we believe Databricks is where those new Spatial users are.



CARTO and Databricks

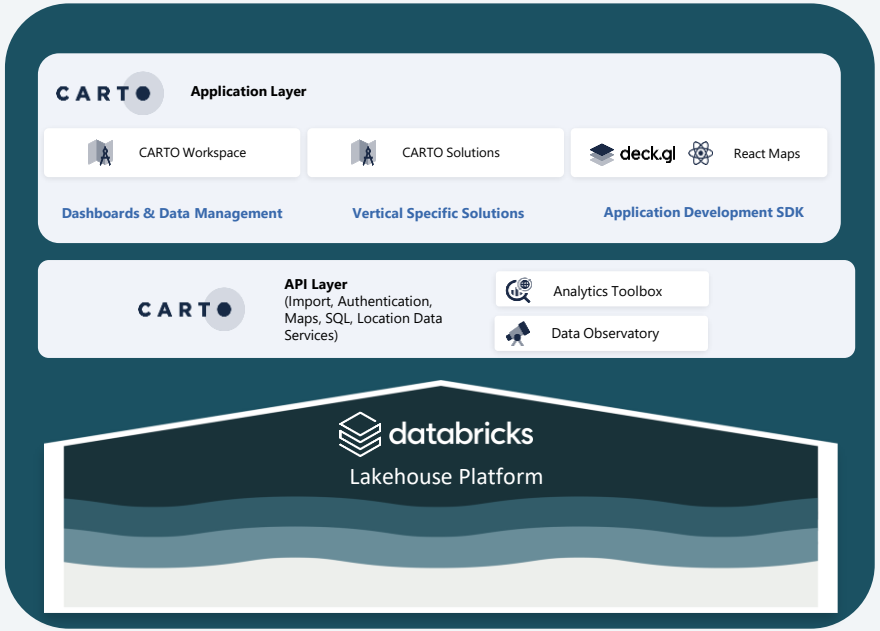
Complete geospatial platform with your Databricks Lakehouse

CARTO

Complete Spatial Analytics Toolkit
Everything you need for developing spatial dashboards and visualizations, modeling, and complete geospatial solutions.

Lightweight API Layer
CARTO has a simple API layer to query and visualize data, **pushing queries directly to Databricks**

Databricks x CARTO
Databricks provides the power of the Lakehouse Platform and CARTO provides geospatial capabilities with H3 indexing and the analytics toolbox, performing all operations within Databricks.



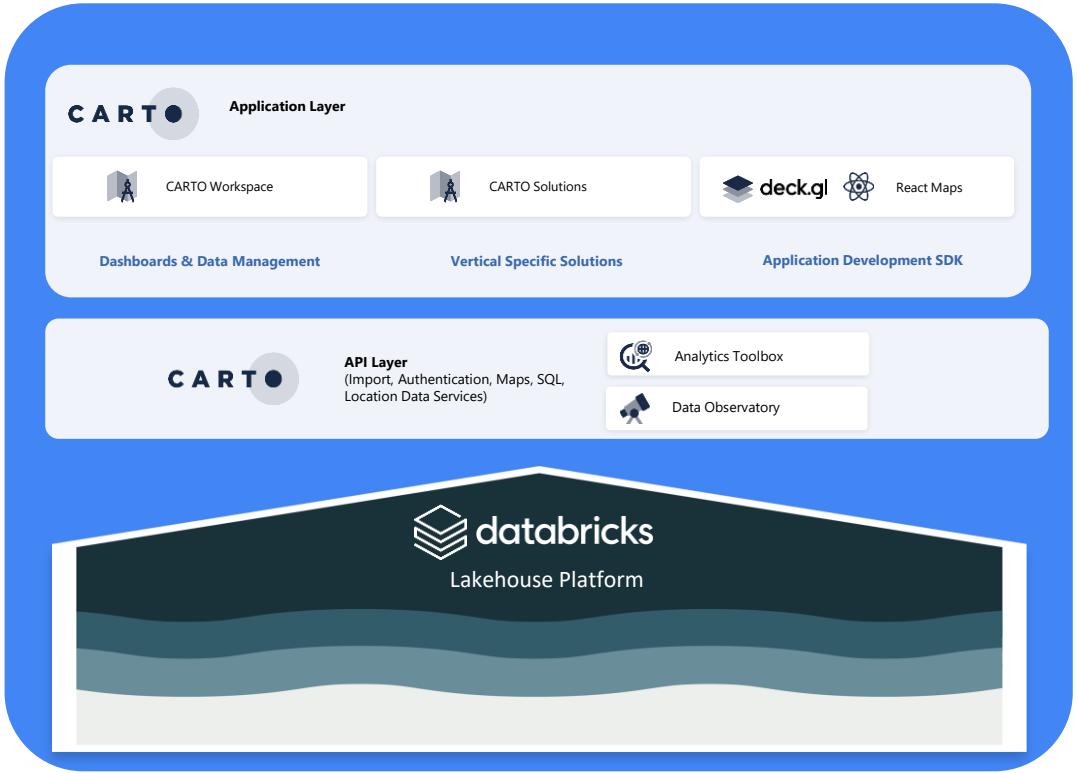
aws Microsoft Azure Google Cloud

CARTO

Complete Spatial Analytics Toolkit
Everything you need for developing spatial dashboards and visualizations, modeling, and complete geospatial solutions.

Lightweight API Layer
CARTO has a simple API layer to query and visualize data, **pushing queries directly to Databricks**

Databricks x CARTO
Databricks provides the power of the Lakehouse Platform and CARTO provides geospatial capabilities with H3 indexing and the analytics toolbox, performing all operations within Databricks.



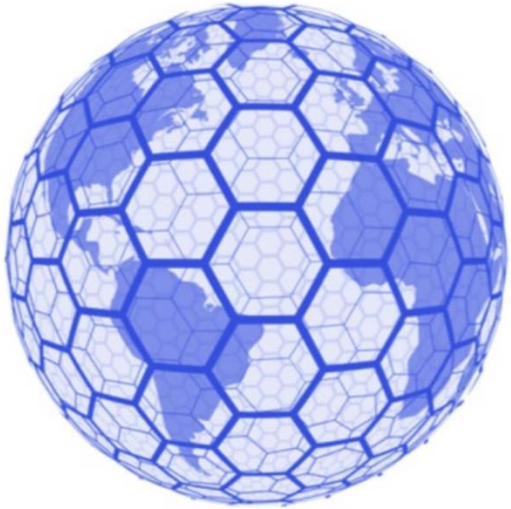
A primer on H3 and spatial indexes

Spatial analytics, without the geometry, at scale

Three major advantages:

- Visualization and approachability
- Scale and performance
- Analytics

Spatial indexes are a type of a Global Discrete Grid system, or a grid system that covers the entire globe.



A primer on H3 and spatial indexes

Geometries vs Spatial Indexes:

What do they look like?

POLYGON((-96.196141 41.125515, -96.195606 41.125514, -96.181864 41.125507, -96.177078 41.125474, -96.167733 41.125456, -96.160565 41.125456, -96.154682 41.125429, -96.151094 41.125414, -96.138848 41.125395, -96.138454 41.125394, -96.138381 41.125394, -96.137158 41.125391, -96.130043 41.125377, -96.1301...))*

*This represents about a **10th** of the geometric description of a census tract, which we have truncated for readability

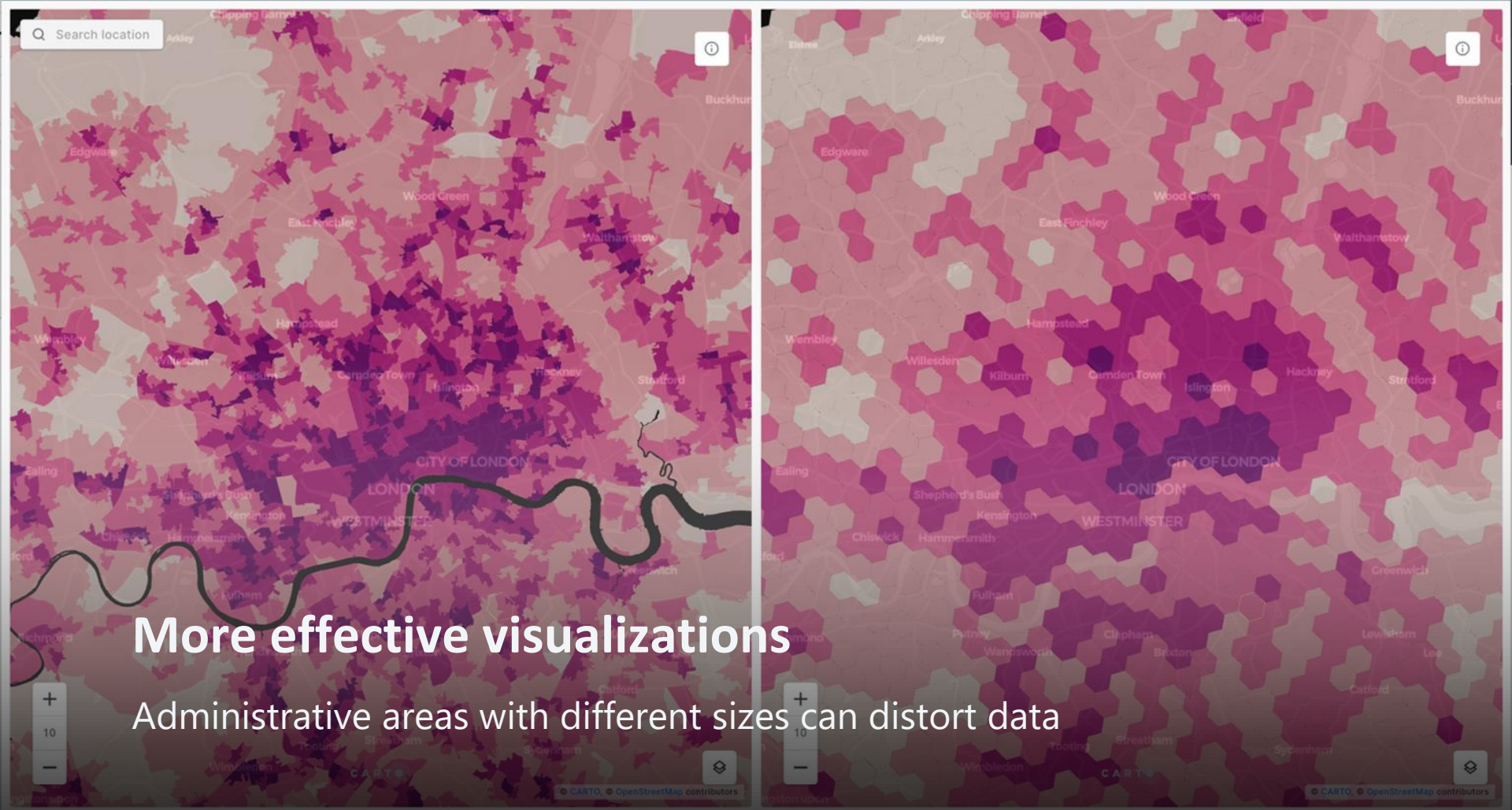
8a2aa84ec307fff

Performance comparisons

Example ETL use case	Geometries	Spatial Indexes	Gain of spatial index over geometries
Processing time	12 days	7 hours	**98% time saved**
Data transferred into the Database	4 TB	1.5 TB	**62% less data transferred**
RAM to process the largest file	256 GB	28 GB	**89% reduction in RAM**
Time to process a spatial join with population	26 minutes	3 seconds	**99% less time**
Time to generate a tileset	23 minutes	1.5 minutes	**94% time saved**
Population coverage	15.48%	15.48%	**0% coverage lost**

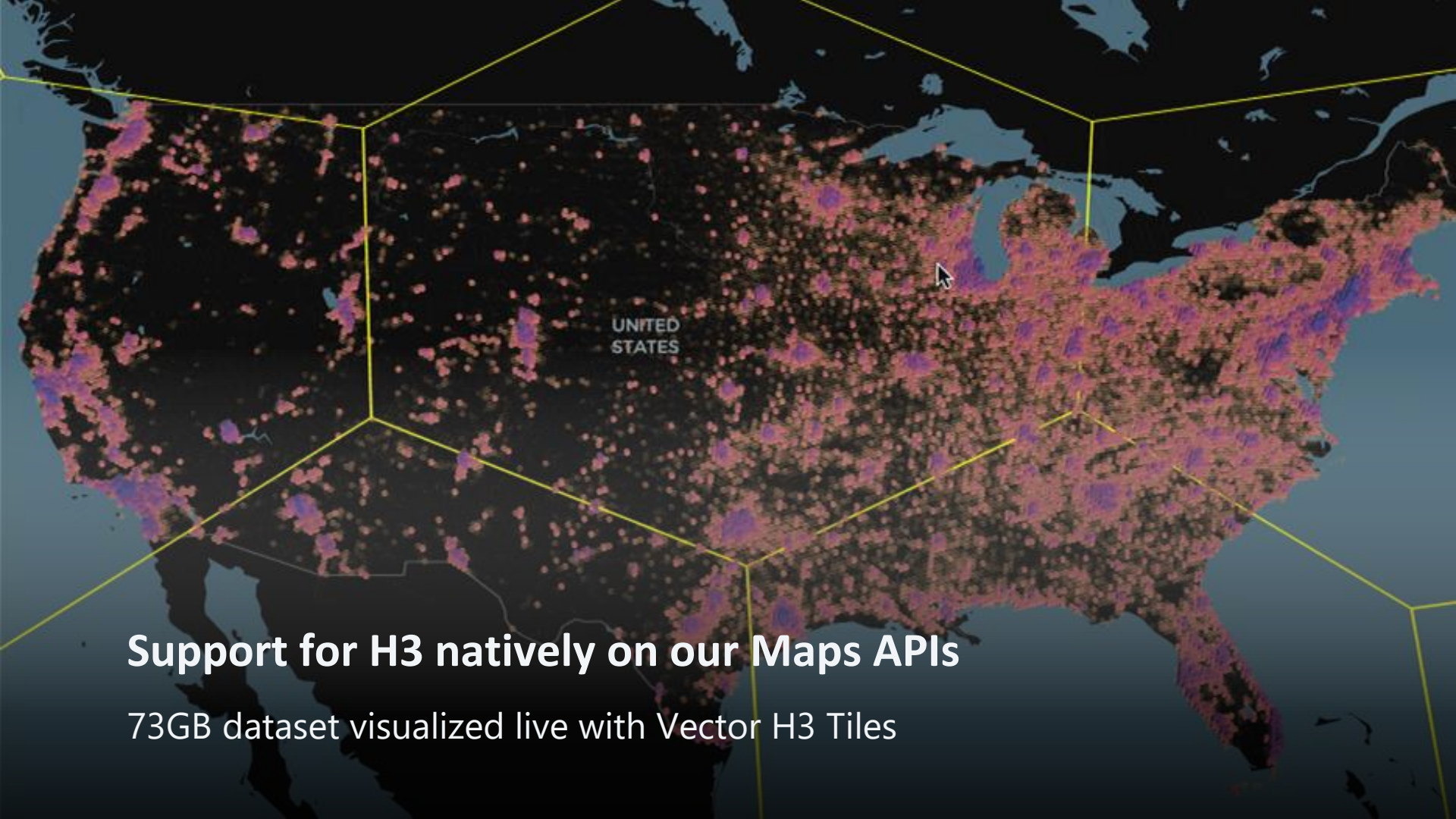
Estimated reduction in cloud data warehouse bill by 85%


AN INTRODUCTION TO
Spatial
Indexes
CARTO



More effective visualizations

Administrative areas with different sizes can distort data



Support for H3 natively on our Maps APIs

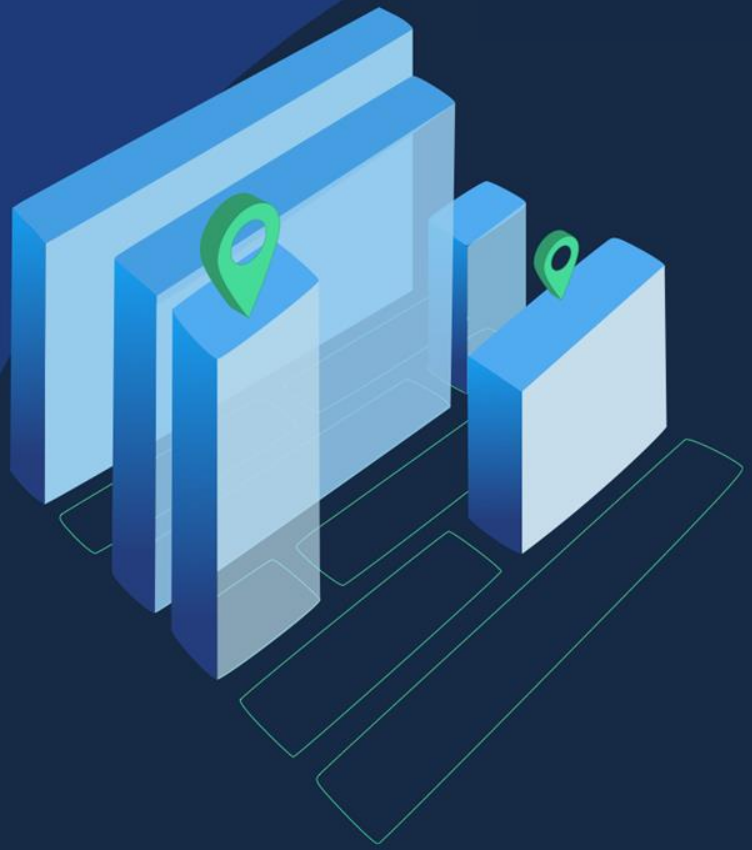
73GB dataset visualized live with Vector H3 Tiles

Spatial Indexes 101

GeoParquet

Introducing GeoParquet:

Towards geospatial compatibility between Data Clouds

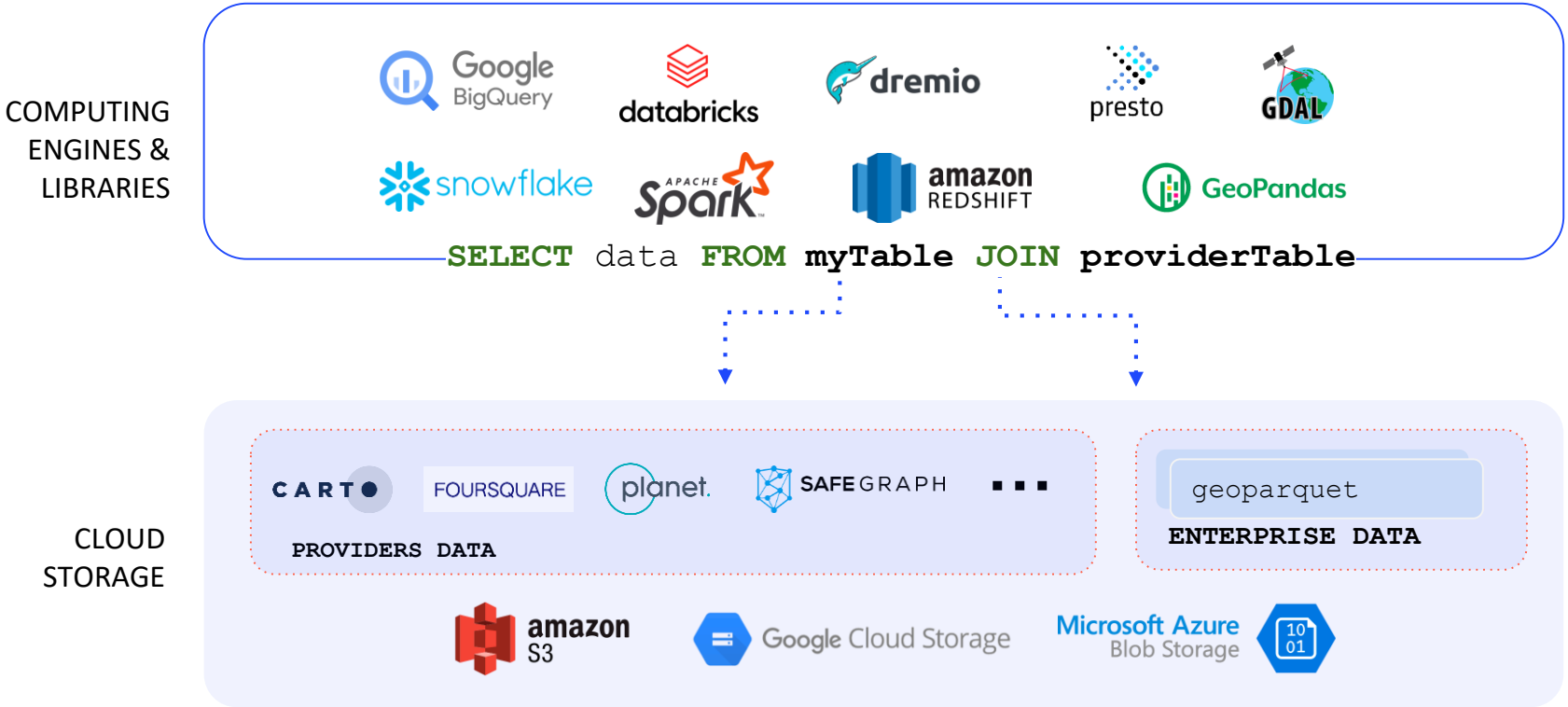


A community driven initiative

Under the umbrella of the Open Geospatial Consortium

github.com/opengeospatial/geoparquet

Geoparquet as a standard storage layer for geo



Roundtable Discussion

Michael Johns
Senior Specialist Solutions Architect
Databricks

Matthew Forrest
VP of Spatial Data Science
CARTO

David Stodder
Senior Director of Research, BI
TDWI

Discussion Topics:

- What are opportunities you see for spatial analysis to solve business questions and meet new requirements?
- How has technology for spatial analysis been evolving, both for users and data management? What are some important recent developments?
- How do Databricks and CARTO work together? Case examples?
- How can organizations move faster to take advantage of spatial analysis? What are some strategies for overcoming obstacles (and not just technology)?

research

tdwi
Transforming Data With Intelligence™

Audience Q&A with Speakers

Questions?

CONTACT INFORMATION

If you have further questions or comments:

David Stodder, TDWI
dstodder@tdwi.org

Michael Johns, Databricks
mjohns@databricks.com

Matthew Forrest, CARTO
mforrest@carto.com

Sponsors

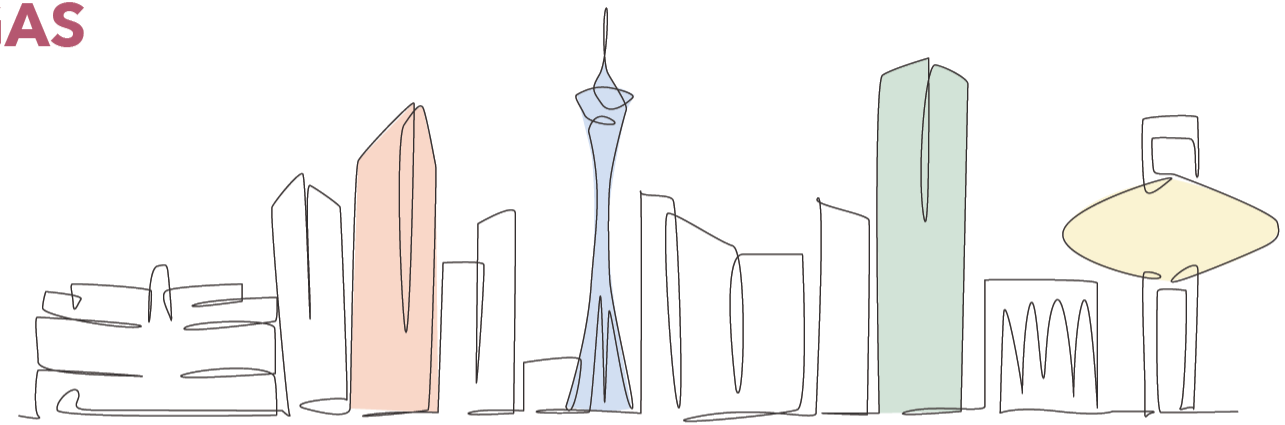
research

tdwi
Transforming Data With Intelligence™

tdwi

LAS VEGAS

Feb. 12–17



Save 30% off TDWI Las Vegas with promo code WEBINAR

Learn More and View the Agenda at tdwi.org/lasvegas