



Sponsored by
AMD

Microsoft GSI Partner Enablement

Empowering Business Growth: Modernizing Data and Analytics with Microsoft Fabric in the AI Era

<Presenter>

<Date>



Your trainers today



Presenter 1



Presenter 2

Course Plan and Learning Objectives

Target Audience
Technical

Course Prerequisites
Prior knowledge of Azure Data platform

Suggested Supplementary Certification
DP-203

Morning session

Time	Topic
0830-0900	Check-in & registration
0900-0915	Introduction and housekeeping
0915-0945	Explore end to end analytics with Microsoft Fabric
0945-1030	Real-time Analytics Overview Lab: Ingesting Data using RTA
1030-1045	Break
1045-1145	From KQL to power BI Reports: Data Visualization Lab: Building reports from KQL
1145-1230	Lunch Break

Afternoon session

Time	Topic
1230-1330	Data Science in Fabric Lab: Getting started with Building a ML Model in Fabric
1330-1400	Synapse Data Warehouse Overview
1400-1430	Lakehouse Architecture in Fabric
1430-1500	Break, Q&A
1500-1530	Lab: Building a DW using Pipelines
1530-1630	Lab: Building a Data Lakehouse
1630-1700	Continuing the Learning Journey Q&A

01

What we're hearing from our customers



Today's data and analytics challenges



Scaling data and analytics across the organization while reducing costs and optimizing existing data and management



Making data more accessible to technical and non-technical users in a governed and secure way



Growing BI adoption to drive a data-driven organization across LOBs and teams



Breaking down data siloes and eliminating unnecessary data sprawl

Modernizing your data and analytics platform is critical for resilient business transformation

Siloed, incomplete data



Integrated, unified data and governance

Technical platform that requires advanced analytics skills



Accessible insights that scale with ease

Costly integration and ongoing maintenance



Transparent, cost-optimized SaaS solution



Sponsored by
AMD

02

Introducing Microsoft Fabric



Introducing Microsoft Fabric

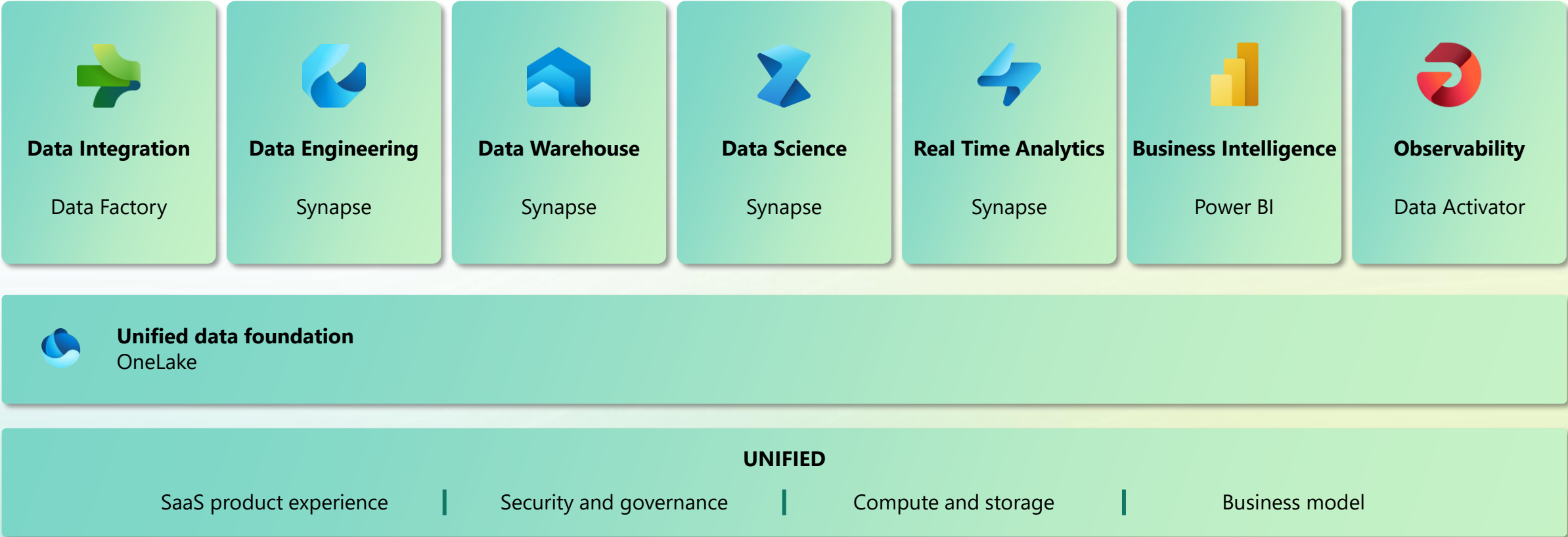
Microsoft Fabric enables you to manage your data in one place with a suite of analytics experiences that work together seamlessly, including:

- Data Factory
- Synapse Data Engineering
- Synapse Data Warehouse
- Synapse Data Science
- Synapse Real-Time Analytics
- Power BI
- Data Activator



Microsoft Fabric does it all—in a unified solution

An end-to-end analytics platform that brings together all the data and analytics tools that organizations need to go from the data lake to the business user



03

Microsoft Fabric capabilities



Microsoft Fabric Capabilities



Lake-centric and open

Align teams within a governed source of truth

Center of enablement for every user

Enable secure, democratized insights across your organization



Accelerated time to value

Leverage a powerful solution with flexibility in cost and usage

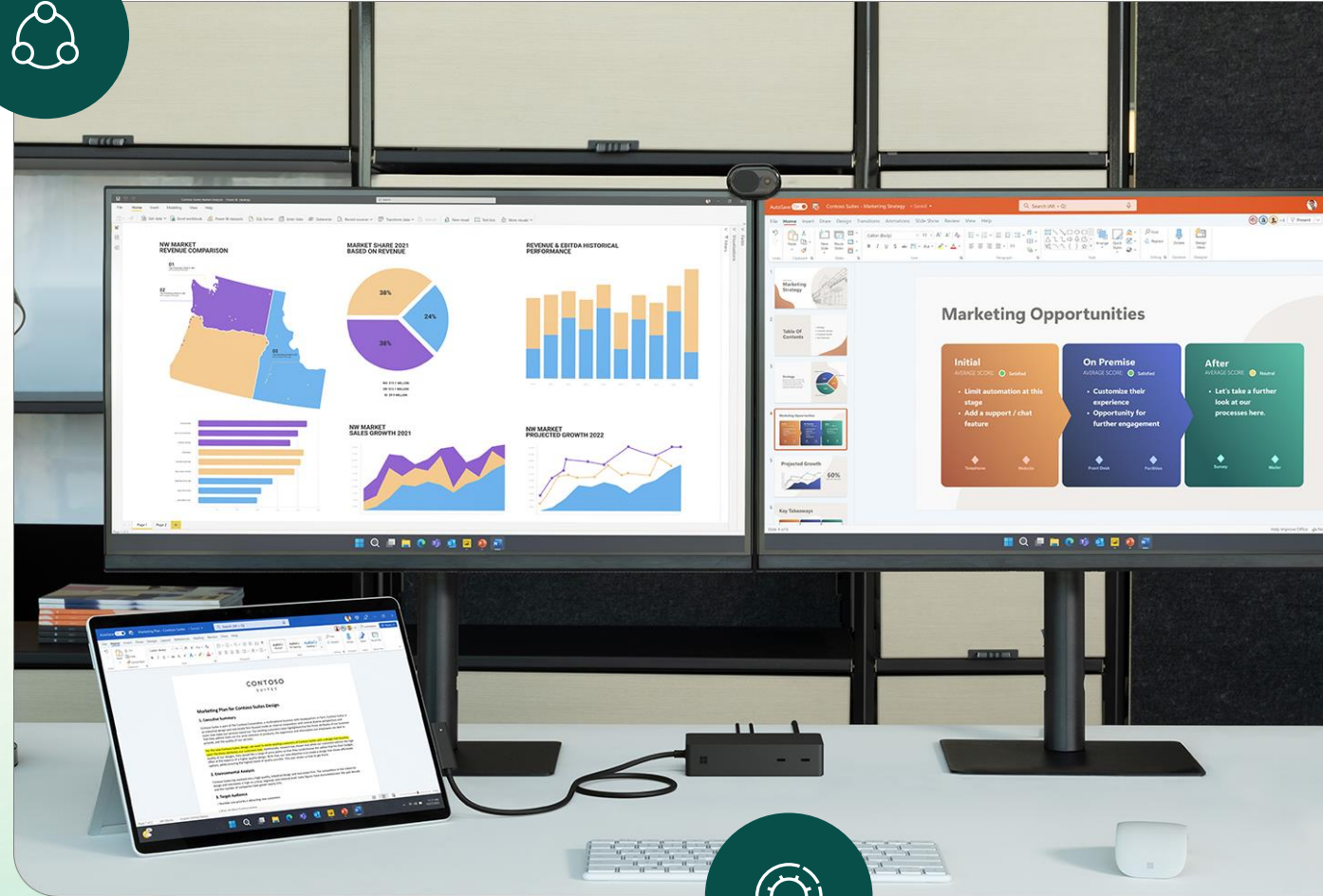


Align teams within a governed source of truth

Lake-Centric and Open



- Eliminate the integration tax with an open data format at every layer
- Ensure your teams are using the most accurate, trusted, and up-to-date data even more real-time scenarios
- Stay in control of your department's data
- Keep your data protected with out-of-the-box security, compliance, and governance
- Use native integration with Microsoft Graph Data Connect and Dataverse to transform and analyze all your Microsoft data



Enable secure, democratized insights across your organization

Center of Enablement for Every User

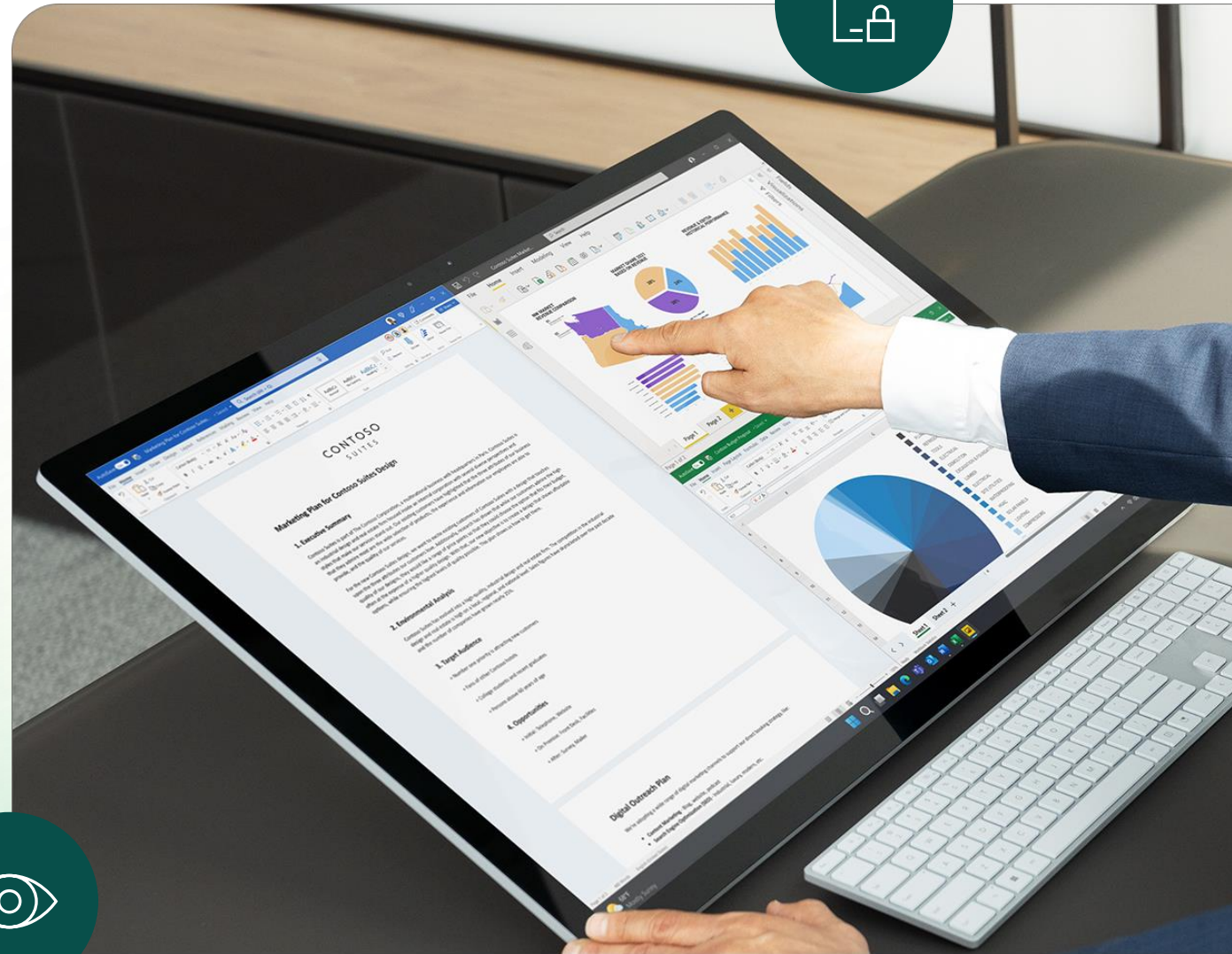
- Equip your data professionals with role-specific, self-serve tools in a unified experience to reduce time to insight
- Use conversational language to create dataflows/pipelines, generate code, build machine learning models, or visualize results at the speed of thought with Copilot in Microsoft Fabric
- Securely infuse intelligence where business users work everyday with unmatched integration with Microsoft 365 to drive better decisions and actions



Leverage a powerful solution with flexibility in cost and usage

Accelerated Time to Value

- Optimize spend through a flexible, SaaS-based pricing model
- Achieve simplified, transparent cost management
- Increase adoption through seamless sign-up, onboarding, and access to data
- Maintain and seamlessly integrate partner and/or third-party solutions



04

The business value you can expect



Get more done and accelerate time to insights with Copilot in Microsoft Fabric

- Turn your words into dataflows and data pipelines so you can intelligently integrate data from anywhere
- Get suggestions for code and entire functions in real-time
- Get a guided machine learning model creation experience to unlock more insights in your data.
- Create and tailor Power BI reports in seconds, generate DAX calculations, create narrative summaries, and ask questions about your data
- Even create your own conversational experiences that combine Azure Open AI models and your organization's data and publish as plug-ins.
- Most importantly, Microsoft Cloud runs on trust which means your data always remains your data



Future-proof your business by doing more with less



Flexible, transparent pricing model

Optimize cost management and pricing



Ease of deployment and management

Spend less time with maintenance, and more time uncovering insights



Secure, self-service data and analytics

Enable teams to self-serve data and analytics while maintaining security and compliance



05

Getting started



Empower your data professionals to move faster and unlock more value from your data



Data Engineers

- **Execute faster** with the ability to spin up a Spark VM cluster in seconds, or configure with familiar experiences like Git DevOps pipelines for data engineering artifacts
- **Streamline your work** with a single platform to build and operate real-time analytics pipelines, data lakes, lake houses, warehouses, marts, and cubes using your preferred IDE, plug-ins, and tools.
- **Reduce costly data replication** and movement with the ability to produce base datasets that can serve data analysts and data scientists without needing to build pipelines

Serve data via
warehouse or
lakehouse

Supporting experiences:



Data Factory



Data Warehouse



Data Engineering



Real-time analytics



Data Scientists

- **Quickly tune a custom model** by integrating a model built and trained in Azure ML in a Spark notebook
- **Work faster** with the ability to use your preferred data science frameworks, languages, and tools
- **Bypass engineering dependencies** with the ability to use your preferred no-code ML Ops to deploy and operate models in production
- **Tap into proven-at-scale models and services** to accelerate your AI differentiation (AOAI, Cognitive Services, ONNX integration, etc).

Serve
transformed
data

Supporting experiences



Data Science



Azure ML



Data Analysts

- **Avoid slow, progress-stagnating data wrangling** by seamlessly triggering a workflow that can unlock data engineering tools and capabilities quickly.
- **Accelerate your work** with visual and SQL based tools for self-serve data transformations and modeling as well as self-serve tools for reporting, dashboards, and data visualizations
- **Turn data into impact** with industry-leading BI tools and integration with the apps your people use everyday like Microsoft 365

Serve insights
via
embedding

Supporting experiences



Data
Warehouse



Real-time
analytics



Power BI



Data Citizens

- **Make more data-driven decisions** with actionable insights and intelligence in your preferred applications
- **Maintain access to all the data you need**, without being overwhelmed by data ancillary to your role thanks to fine grain data access management controls

Supporting experiences



Power BI



Microsoft 365

Serve data via warehouse or lakehouse



Data Stewards

- **Maintain visibility and control of costs** with a unified consumption and cost model that provides evergreen spend optics on your end-to-end data estate
- **Gain full visibility and governance** over your entire analytics estate from data sources and connections to your data lake, to users and their insights

Paradigm shift in Data Consumption

- Big Data Platforms
- Terabytes of data
- Low latency
- High freshness
- Streaming technologies
- Index everything by default



**Enterprises are still
behind in handling
real-time data
consumption**



Time for a Digital Transformation



High complexity



High costs of the existing solutions



High Skill



Integration between different teams

Time for a Digital Transformation

- Few experts to generate reports
- Long waiting lists
- Outdated data



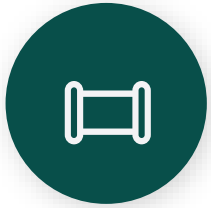
Time for a Digital Transformation

As enterprises rely more on digital technologies for operations and customer interactions, monitoring and analyzing IT systems in real-time have become essential.



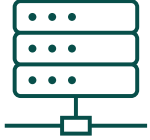


Sponsored by
AMD



Real-Time Analytics in Fabric





With Real-Time Analytics in Fabric, we help organizations scale up their analytics solution and democratize data for advanced and citizen data scientists.



Real-Time Analytics is built on top of Azure Data Explorer Engine

Real-Time Analytics in Fabric - Reducing Complexity

- One Logical Copy
- By-default streaming ingestion
- By-default indexing everything
- Time and Hash based data partitioning
- Native support for structured, semi-structured and free text data
- In-place transformation

Real-Time Analytics in Fabric - Simplifying

- Tightly integrated with Fabric experiences
- Mirroring to Lakehouse
- By-default integration with event streaming sources
- By-default integration with Power BI

Real-Time Analytics in Fabric



Auto-managed scale



Cross workload
solutions



Easy onboarding with
friendly UX

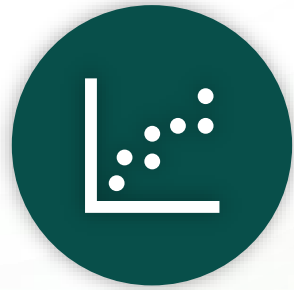


Managed experiences

o		x
	x	
o		o

Solving the Single Source of Truth problem

Highlighted industry scenarios



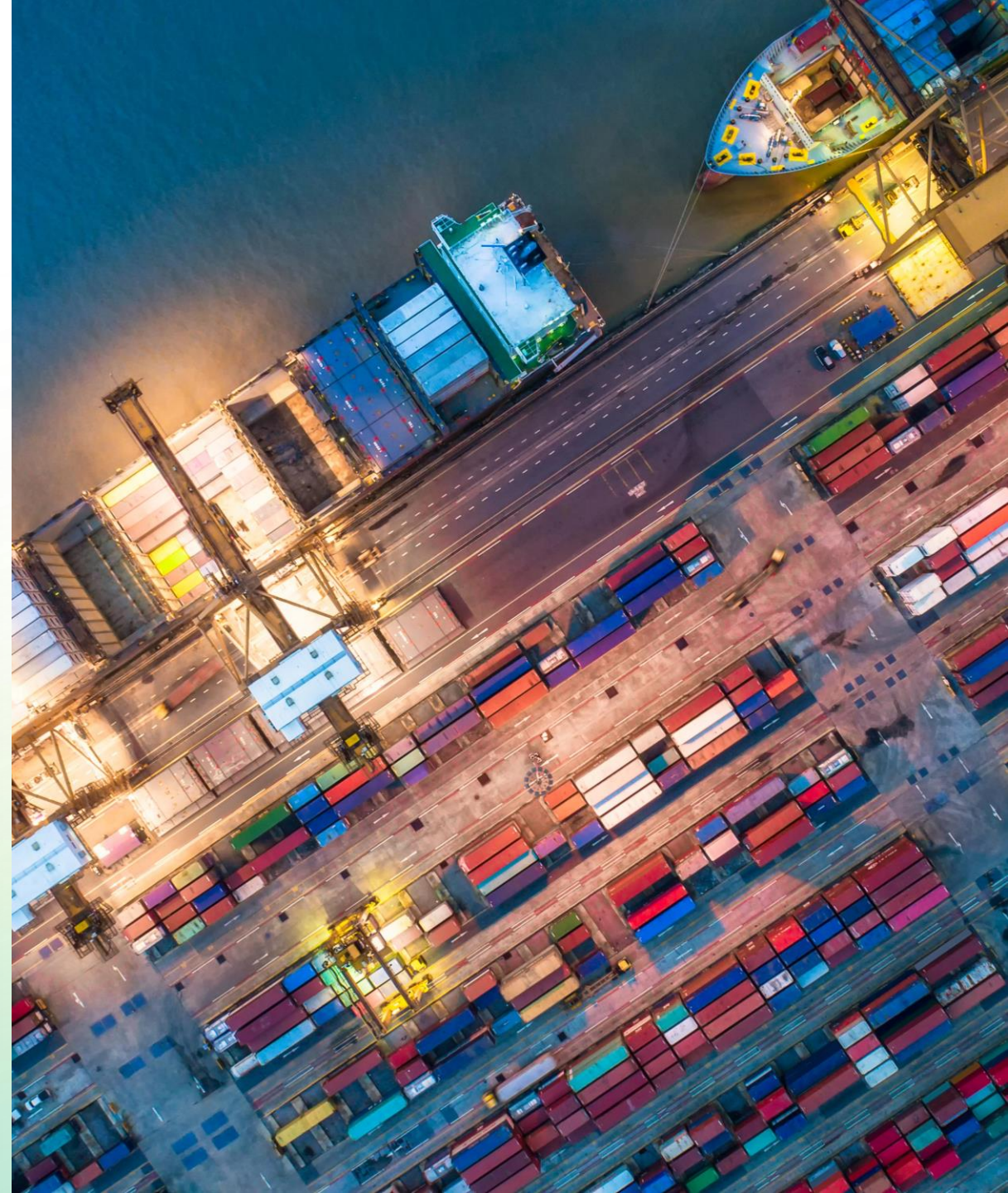
IOT Analytics



Log analytics (security)

Highlighted industry scenarios

- Cybersecurity
- Asset tracking and management
- Predictive maintenance
- Supply chain optimization
- Customer experience
- Energy management



Highlighted industry scenarios

- Inventory management
- Quality control
- Environmental monitoring
- Fleet management
- Health and safety





Hands on labs

AMD 
together we advance_data analytics



The Workshop



Ingesting Data using
RTA



Building reports from
KQL



Getting started with
Building a ML Model in
Fabric



Building a DW using
Pipelines



Building a Data
Lakehouse



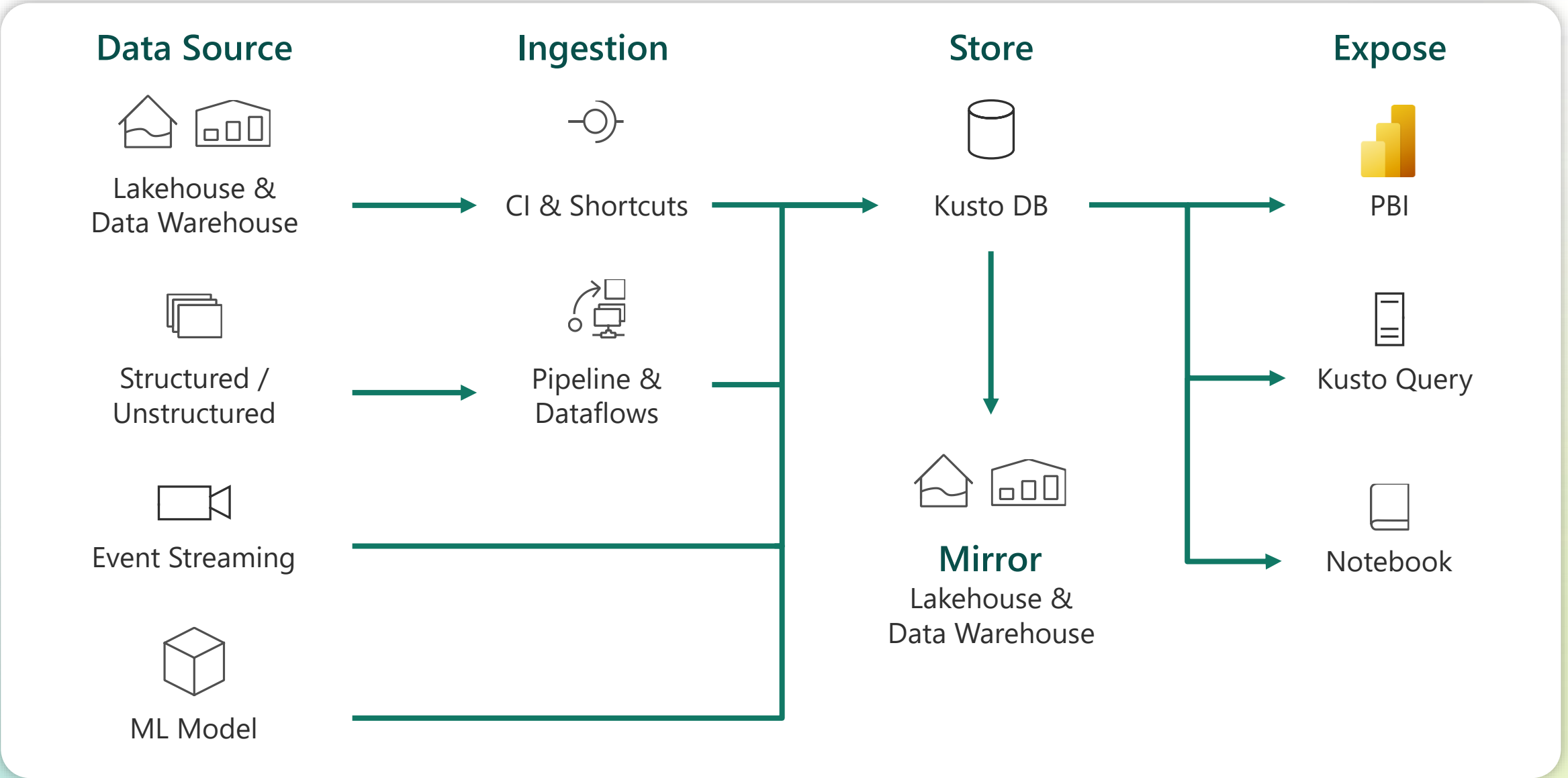
Sponsored by
AMD

Module 0

Introduction to Real-time Analytics and Setup



Common Real-time Analytics Patterns



Eventstreams

Key Capabilities

Centralized place for event data

Capturing, transforming, and routing event data.

Various source connectors

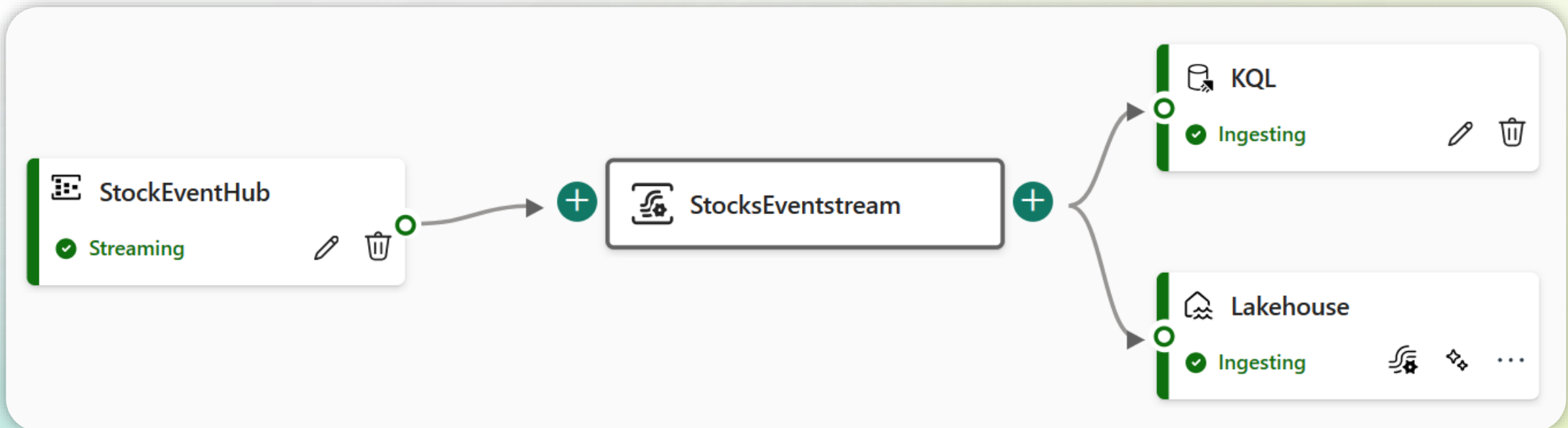
Fetch event data from diverse source, such as Custom Apps to push into Eventstreams, Azure Event Hub or Samples.

No-code experience

Drag and drop makes it intuitive and easy to use. Build an end-to-end data flow diagram.

Multiple destinations

Lakehouse, KQL Database, Custom App, Data Activator (Reflex)



Scenario Overview

A fictitious financial company named "**AbboCost**", would like to set up a stock monitoring platform to monitor price fluctuations and report on historical data.

Throughout the workshop, we'll look at how every aspect of Microsoft Fabric can be incorporated as part of a larger solution -- by having everything in an integrated solution, we'll be able to quickly and securely integrate data, build reports, create data warehouses and lakehouses, forecast using ML models, and more.



Scenario Overview

AbboCost Financial is a small firm that monitors several volatile stocks. Their current data feed provides stock prices in real time. This feed is all you have to work with.

Your goal: support the development of this fledgling company by developing a modern data analytics solution.



Microsoft Fabric End-to-End Real-Time Analytics Workshop

We'll start by rapidly generating real-time data and understanding how that data can be processed and visualized in Microsoft Fabric

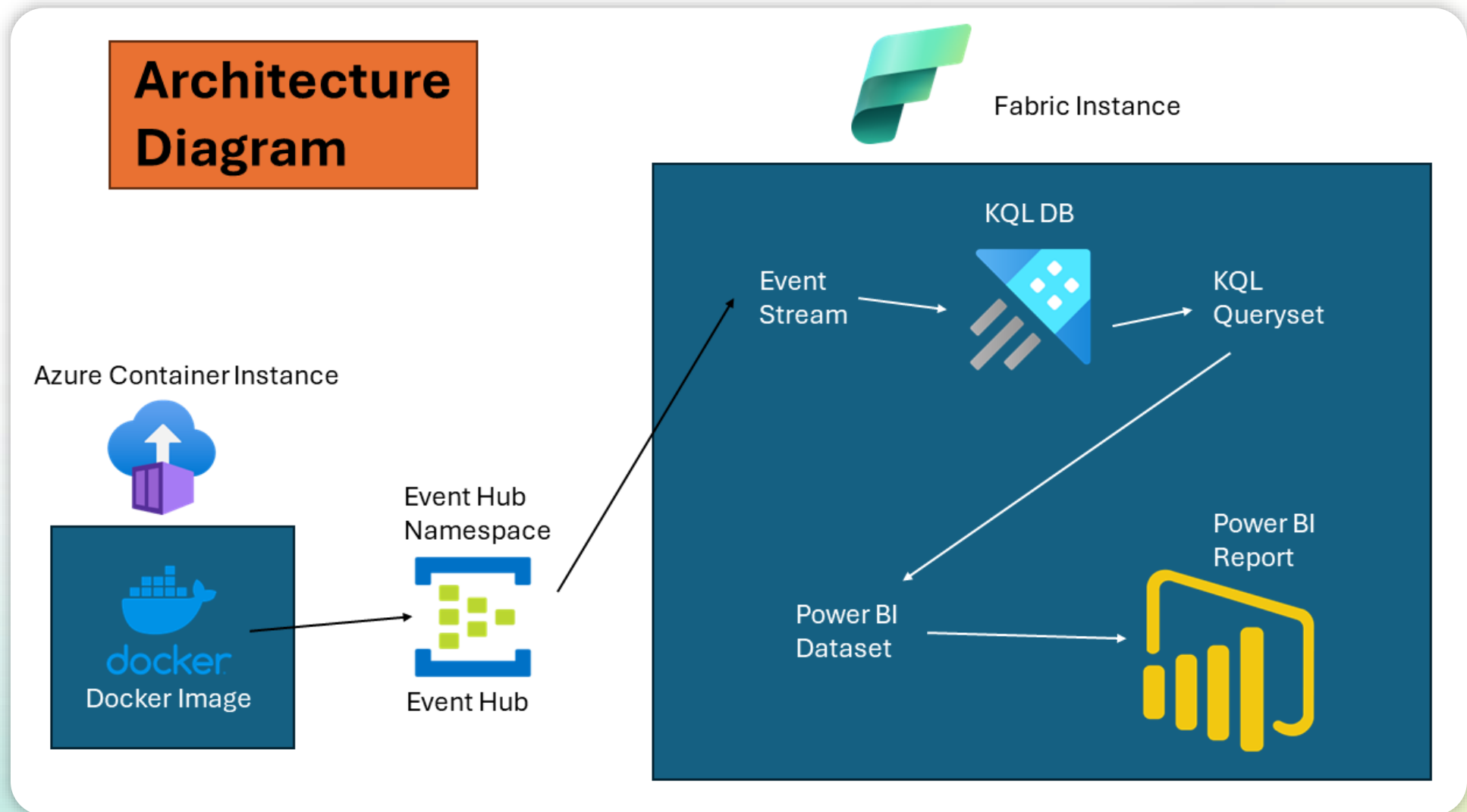
In different modules we'll explore –

- Real-time analytics
- Data warehousing
- Data lakehouse architecture
- Data science

Designed to be cohesive but flexible - all modules involve the same core scenario



Basic Architecture of sample workshop solution



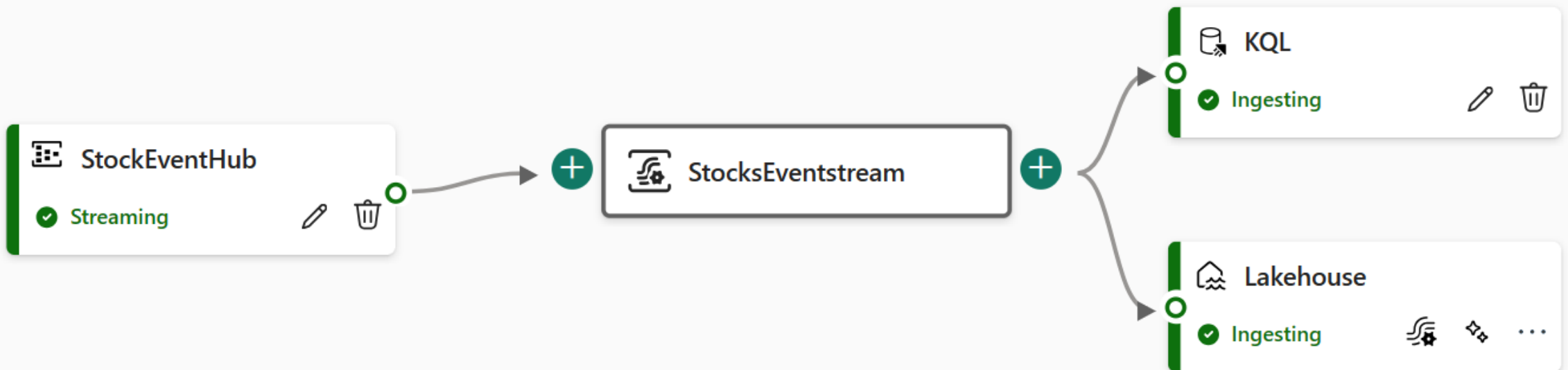
Module 0 – Introduction & Setup

Goals: Understand the solution and deploy the stock price feed generator.

Choice: Decide to deploy the generator as an Azure container or run in a Jupyter notebook.

In general, we recommend deploying as an Azure container if possible, but this requires an Azure subscription.

The data feed will produce 3 pieces of information: stock symbol, current price, and timestamp.



Module 0 – Introduction & Setup

Here are two ways to deploy this application:

- Option 1 - Deploy the app via Jupyter notebook
- Option 2 - Deploy the app via Azure Container Instance

When possible, we recommend deploying the stock price generator via Azure Container Instance (option 2).

- This is because the default Spark cluster will consume a large number of resources.

Module 0 – Introduction & Setup

Instructor-led Shared Environment:

Obtain the following information from the instructor:

- The Event Hub namespace
- The Event Hub name
- The name of the SAS key
- The primary key of the SAS key
- The consumer group name to use, which is unique to each participant

The screenshot displays the Azure Event Hubs configuration interface. A central dialog box titled "Connection settings" is highlighted with a green border. It contains the following fields:

- Event Hub namespace ***: ehns-a8b41e-fabricworkshop
- Event Hub ***: stockeventhub
- Connection**: Create new connection (with a refresh icon)
- Connection name**: {"endpoint":"ehns-a8b41e-fabricworkshop","entityPath":"st...
- Authentication kind**: Shared Access Key (with a dropdown arrow)
- Shared Access Key Name**: stockeventhub_sas
- Shared Access Key**: A masked field with dots and a toggle icon.

At the bottom of the dialog are "Create" and "Cancel" buttons. To the right, a sidebar titled "Azure Event Hubs" is also highlighted with a green border. It contains the following fields:

- Source name ***: StockEventHub
- Cloud connection ***: Type to select a cloud connection (with a dropdown arrow and a refresh icon)
- Consumer group ***: Type to select a consumer group (with a dropdown arrow)
- Data format ⓘ**: Json (with a dropdown arrow)

Module 0 – Introduction & Setup

Instructor-led Shared Environment:

- The Event Hub namespace:
- The Event Hub name:
- The name of the SAS key:
- The primary key of the SAS key:
- The consumer group name:



Sponsored by
AMD

Module 1

KQL Database



KQL Database

Key Capabilities

Unlimited Scale (query,
ingestion, storage)

Any data source

Any data format

Structured
Semi-Structured
Free-text

Real-time data
Transformation of
complicated data structure

Streaming analytics in NRT

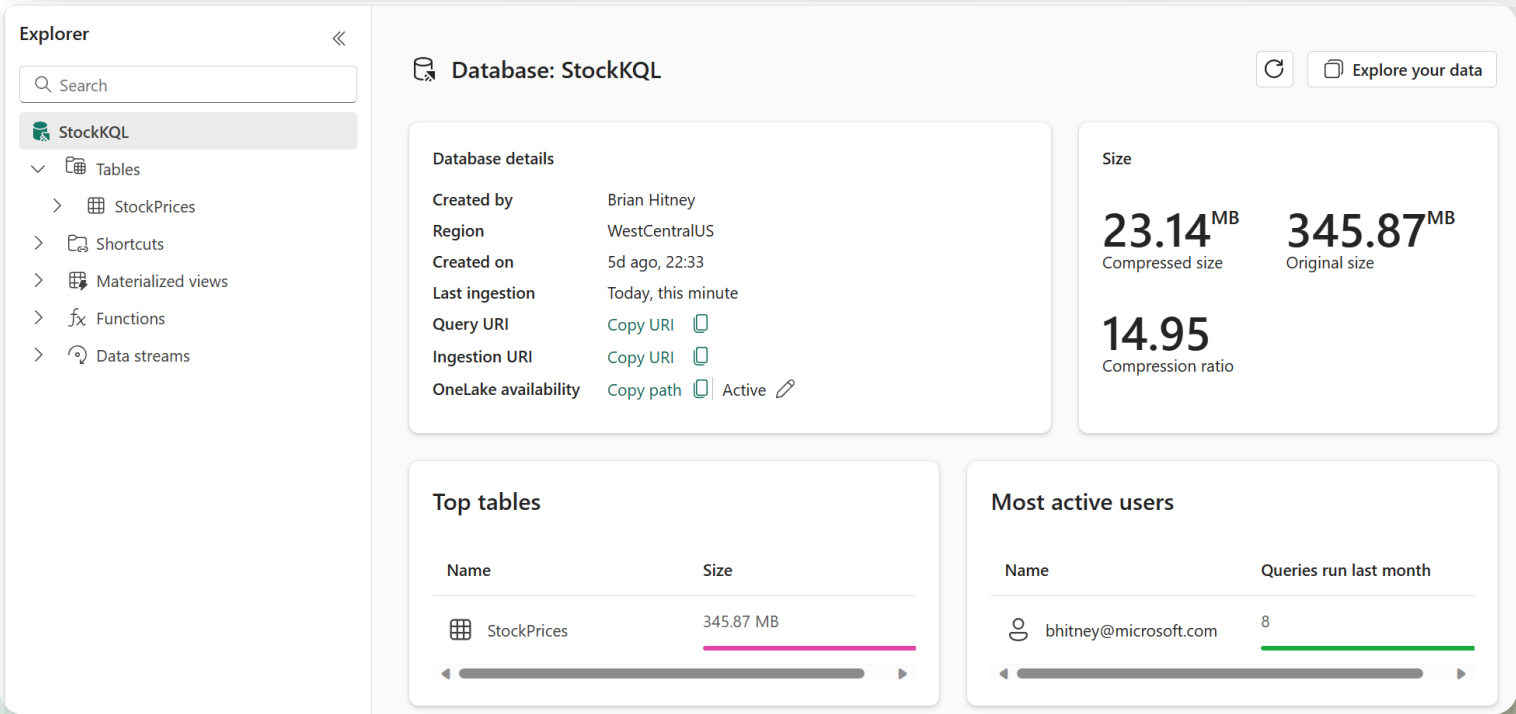
High performance
Low latency
High freshness

Timeseries database

Everything is indexed and
partitioned

Module 1 – KQL Database Setup

- **Goals:** Create and configure a KQL database to ingest data
- With data being fed to the Eventstream, this module focuses on creating and KQL database as a data sink.





Sponsored by
AMD

Module 2

KQL Queries



KQL Queryset

Overview

Uses the Kusto Query Language (KQL) for **creating queries**, views, functions, control commands, customize results and also supports many SQL functions.

Tabs in the KQL queryset can be associated with a different KQL database.

Save queries for later use or share with others to collaborate on data exploration.

You can also change the KQL database associated with any tab, allowing you to run the same query on data in different states.



What is a Kusto query?

A **Kusto query** is a read-only* request to process data and return results.

Has one or more query statements and returns data in a tabular or graph format.

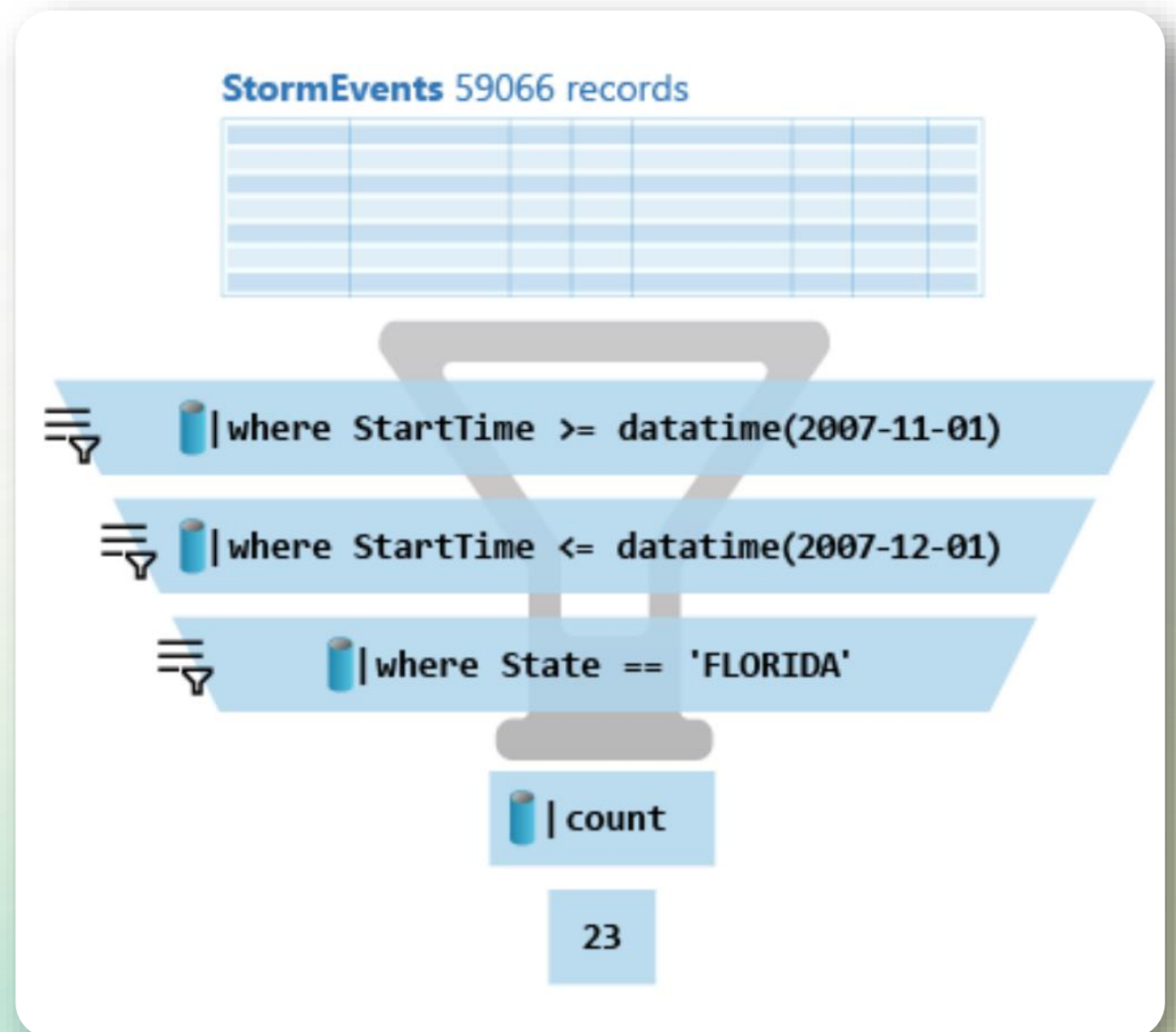
Statements are sequenced by a pipe (|). **Data flows, or is piped, from one operator to the next.**

It's like a funnel, where you start out with an entire data table.

The data is filtered/manipulated at each step and then fed into the following step.

Each time the data passes through another operator, it's filtered, rearranged, or summarized.

* Feb 2024: [Update command](#) now in preview



KQL Basic Operations

... | count

Counts records in input table (e.g. T)

... | take 10

Get few records - to familiarize yourself with the data. No actual order ensured.

... | where Timestamp > ago(10m) and UserId == 'abcdef'

Filtering on a specific fields

... | project Col1, Col2, ...

Choose some columns (great if input table has dozens of columns)

... | extend NewCol1=Col1+Col2

Introduces new calculated columns

... | sort by Timestamp ASC

Sorts result (also can use Order By)

... | render timechart

Render data into a visual plot while exploring

KQL Capabilities

Schema is

Relational, Lightweight, Dynamic

Databases

Authorization boundary

Transaction boundary

But not query boundary!

Supporting cross-database and cross-cluster queries

Tables

Rectangular

Columns

Supported types: boolean, integer, real, decimal, dates, timespan, string, dynamic (JSON)

Shortcuts (external data)

Stored functions (views)

Materialized views

Module 2 – KQL Querysets

Goals: Interact with the underlying data via KQL queries

Begin writing queries to summarize data using Kusto Query Language. Similar to T-SQL, KQL is designed for data exploration and summarization.

The screenshot displays the StockKQL interface. On the left, the 'Database' pane shows a tree view with 'StockKQL' and 'StockPrices'. The 'StockPrices' table is expanded, showing columns: symbol (string), price (real), timestamp (datetime), EventProcessedUtcTime (...), PartitionId (int), and EventEnqueuedUtcTime (...). The main query editor contains the following KQL script:

```
1 StockPrices
2 order by timestamp asc, symbol asc
3 extend pricedifference = round(price - prev(price, 8), 2)
4 extend percentdifference = round(round(price - prev(price, 8), 2) / prev(price, 8), 4)
5 summarize arg_max(pricedifference, timestamp, price) by symbol
6
```

Below the query editor, the results are displayed as a table with columns: symbol, pricedifference, timestamp, and price. The table contains 8 records. The status bar indicates 'Done (1.013 s)' and '8 records'.

symbol	pricedifference	timestamp	price
WHO	93.6	2024-01-11 02:03:09.7710	906.4
TDY	340.42	2024-01-11 15:59:27.9920	931.53
IDK	422.45	2024-01-11 15:59:29.0080	1,013.37
TMRW	391.8	2024-01-11 15:59:30.0260	982.22
WHAT	198.73	2024-01-11 22:32:48.1600	857.41
IDGD	224.98	2024-01-12 12:00:20.7270	822.89
BCUZ	166.83	2024-01-12 16:00:46.1170	983.48
WHY	126.86	2024-01-12 16:00:49.1640	972.64

On the right side of the results table, there is a 'Columns' pane with a search bar and a list of columns: symbol, pricedifference, timestamp, and price. All columns are checked. Below the columns list, there are sections for 'Row Groups' and 'Values', both with a 'Drag here to set row groups' and 'Drag here to aggregate' prompt respectively.



Sponsored by
AMD

Module 3

Power BI



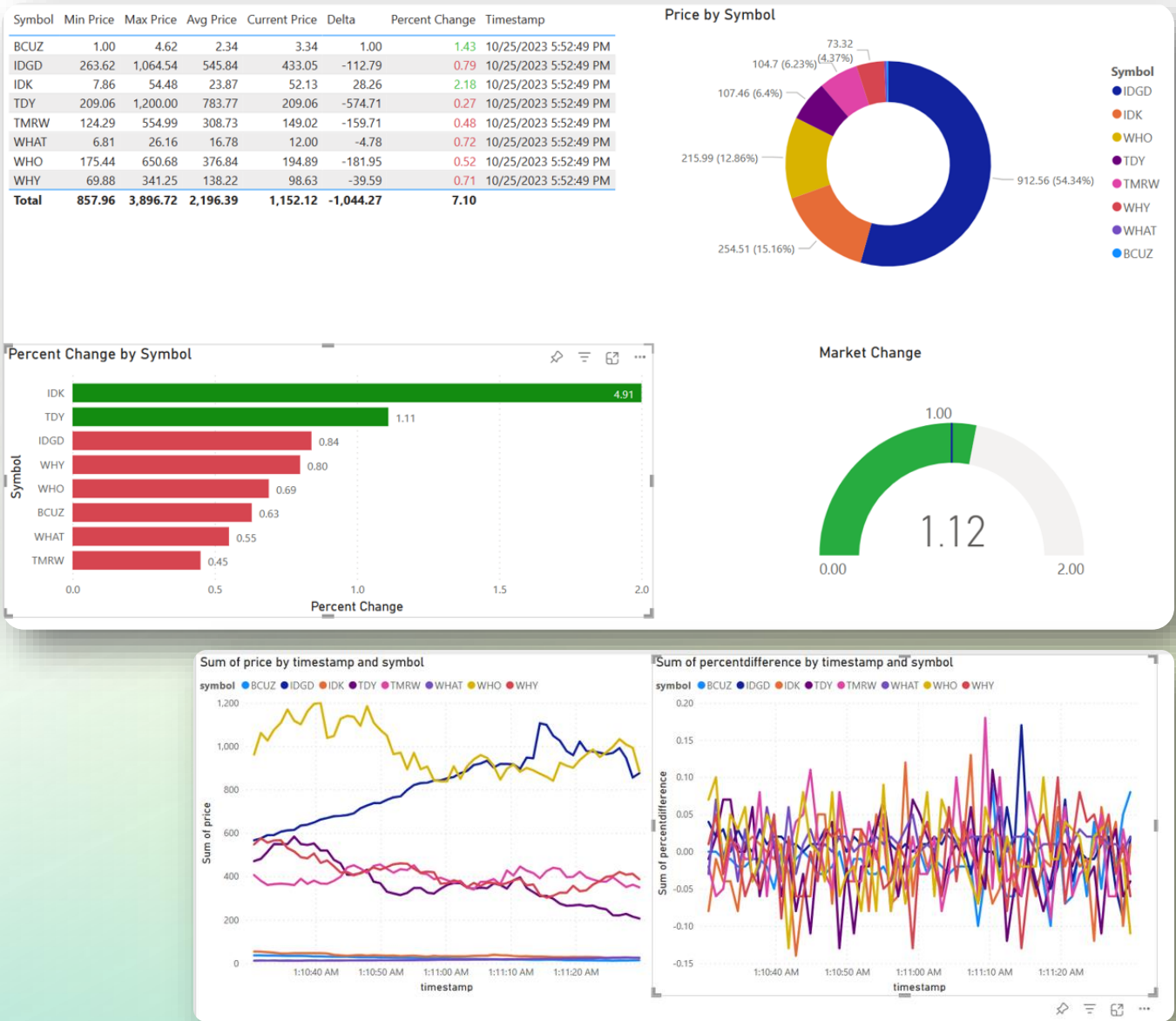


Power BI: The bridge between data and decisions



Module 3 – Reports in Power BI

- **Goals:** Build a reports in Power BI to display real-time data.
- KQL is the ideal engine to power fast, real-time dashboards. Using a combination of KQL summarizations and Power BI functions, compelling visuals can be created.



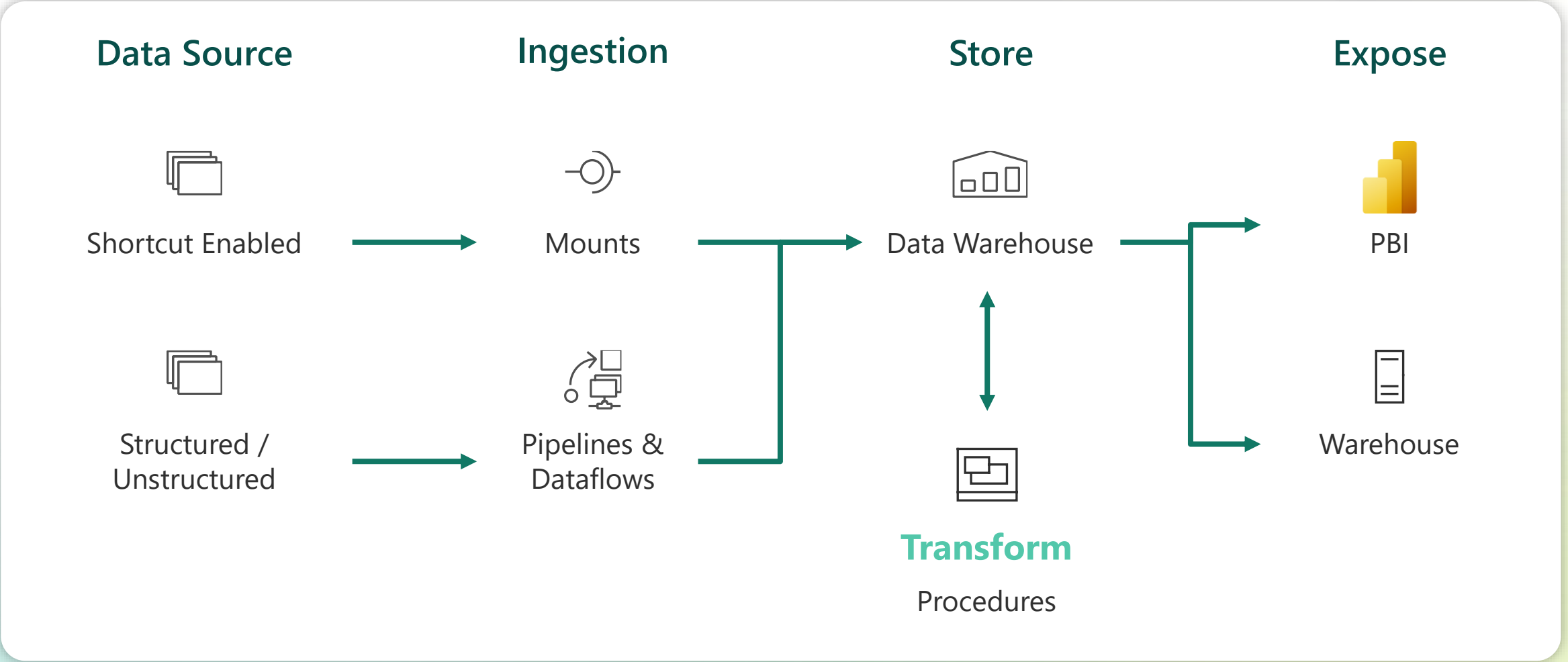


Sponsored by
AMD

Module 4 Synapse Data Warehouse



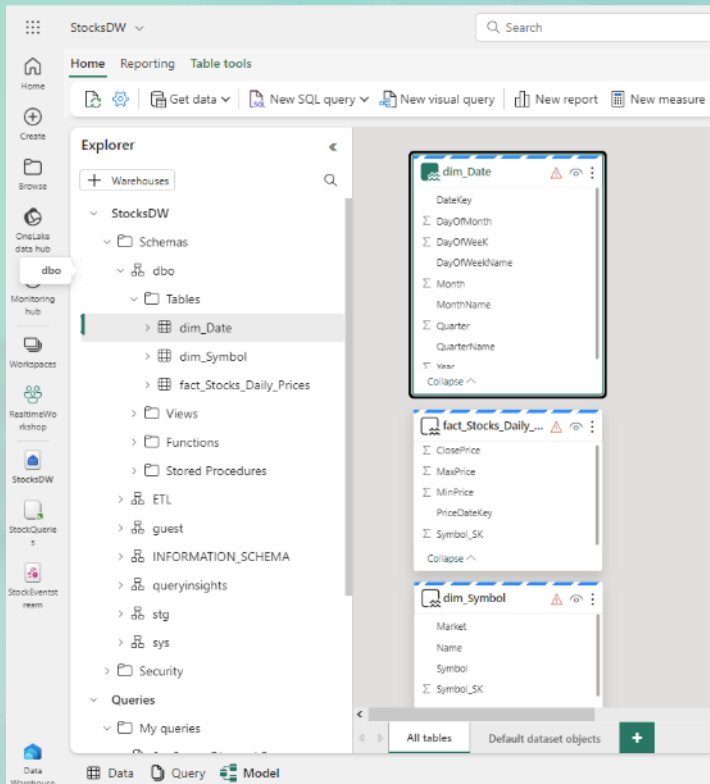
Common Data Warehouse Patterns



Data warehouse | Overview

Enterprise scale data warehouse with open standard format

No knobs performance with minimal set-up and deployment, no configuration of compute or storage needed



Key Capabilities

- Simple and intuitive warehouse experiences for both beginner and experienced data professionals
- Lake-centric warehouse **stores data in OneLake in open Delta format** with easy data recovery and management
- Fully integrated with all Fabric workloads out-of-the box
- Data loading and transforms at scale, with full multi-table transactional guarantees provided by the SQL engine
- Virtual warehouses with cross-database querying and a fully integrated semantic layer
- Enterprise-ready platform with end-to-end performance and usage visibility, with built-in governance and security
- Flexibility to build data warehouse or data mesh based on organizational needs and choice of no-code, low-code, or T-SQL for transformations

Fabric Data Warehouse VS Fabric Lakehouse

Decision Guide

<https://learn.microsoft.com/en-gb/fabric/get-started/decision-guide-data-store>

Search for: Fabric Warehouse Decision Guide

	Data warehouse	Lakehouse	Power BI Datamart	KQL Database (Eventhouse)
Data volume	Unlimited	Unlimited	Up to 100 GB	Unlimited
Type of data	Structured	Unstructured,semi-structured,structured	Structured	Unstructured, semi-structured, structured
Primary developer persona	Data warehouse developer, SQL engineer	Data engineer, data scientist	Citizen developer	Citizen Data scientist, Data engineer, Data scientist, SQL engineer
Primary developer skill set	SQL	Spark(Scala, PySpark, Spark SQL, R)	No code, SQL	No code, KQL, SQL
Data organized by	Databases, schemas, and tables	Folders and files, databases, and tables	Database, tables, queries	Databases, schemas, and tables
Read operations	T-SQL, Spark (supports reading from tables using shortcuts, doesn't yet support accessing views, stored procedures, fuctions etc.)	Spark,T-SQL	Spark,T-SQL,Power BI	KQL, T-SQL, Spark, Power BI
Write operations	T-SQL	Spark(Scala, PySpark, Spark SQL, R)	Dataflows, T-SQL	KQL, Spark, connector ecosystem
Multi-table transactions	Yes	No	No	Yes, for multi-table ingestion. See update policy .
Primary development interface	SQL scripts	Spark notebooks,Spark job definitions	Power BI	KQL Queryset, KQL Database
Security	Object level (table, view, function, stored procedure, etc.), column level, row level, DDL/DML, dynamic data masking	Row level, table level (when using T-SQL), none for Spark	Built-in RLS editor	Row-level Security



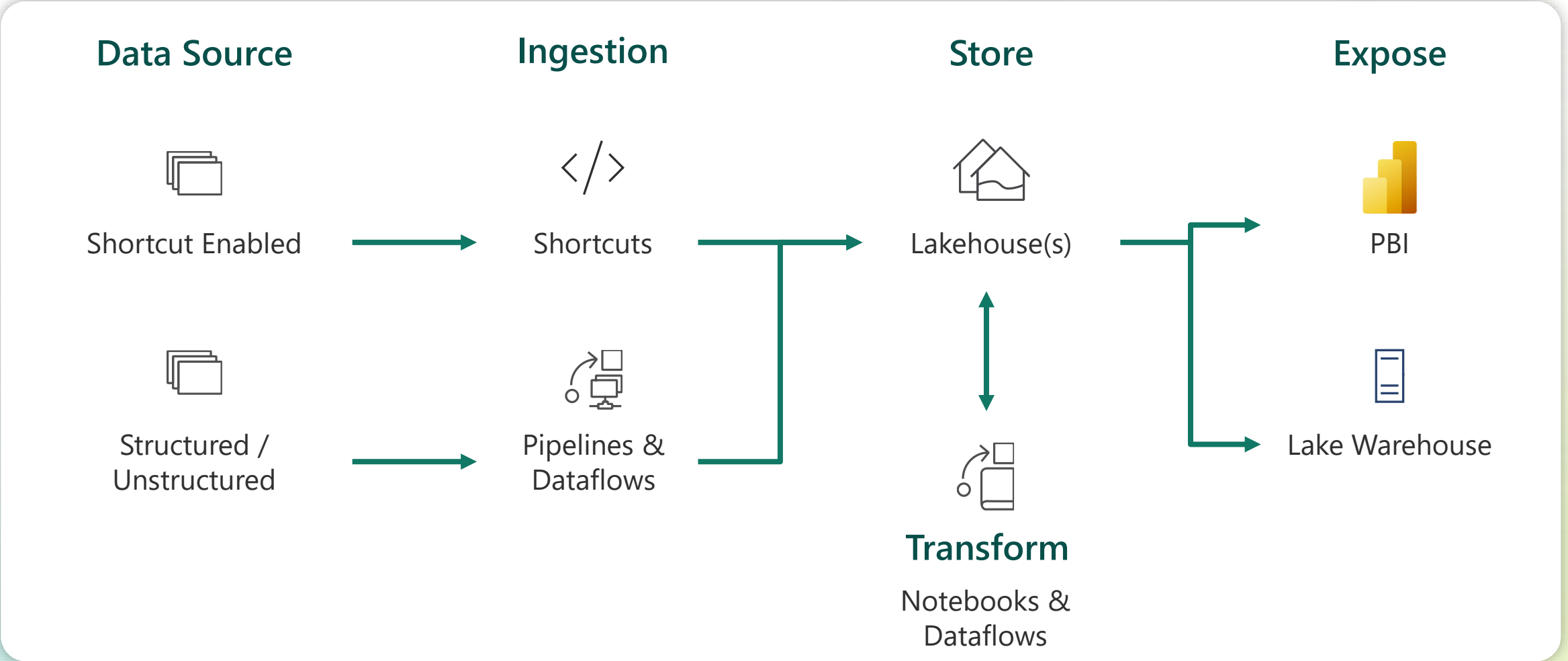
Sponsored by
AMD

Module 5

Data Lakehouse



Common Lakehouse Patterns





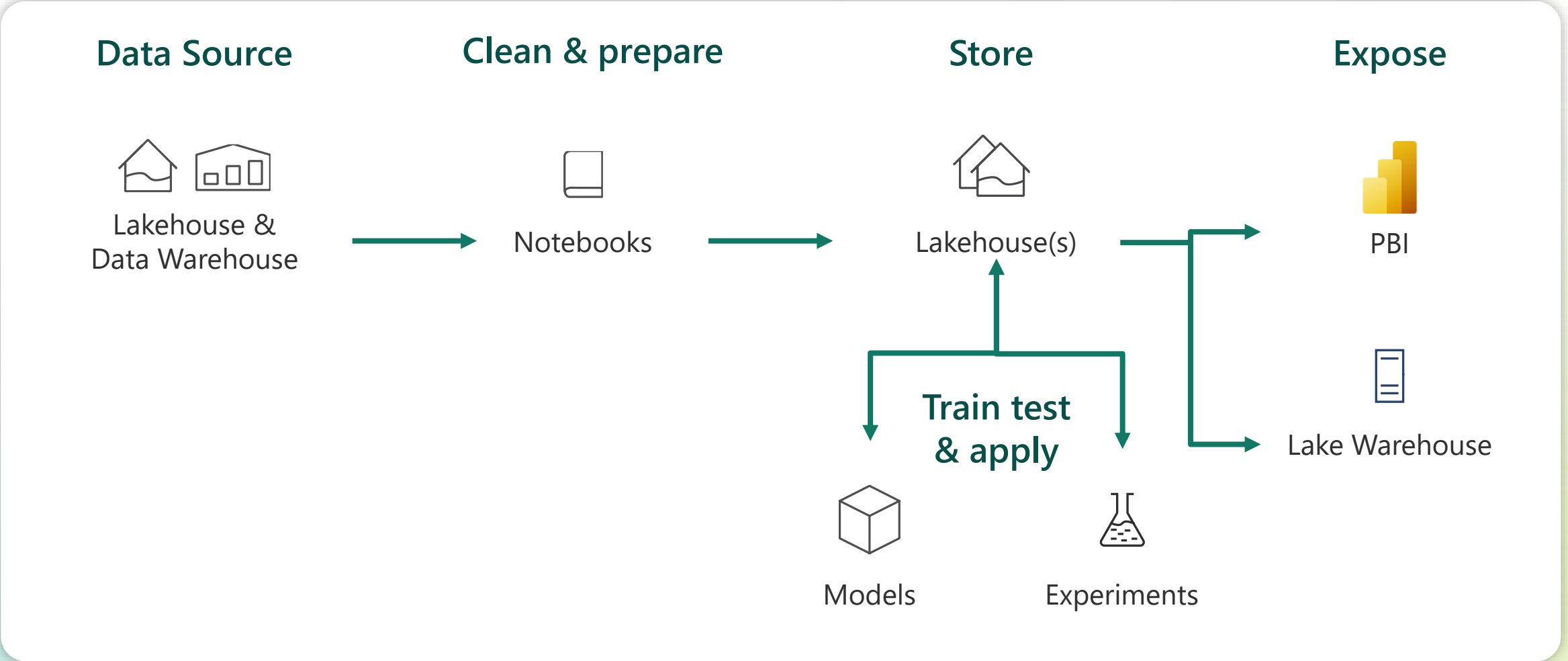
Sponsored by
AMD

Module 6

Data Science



Common Data Science Patterns



Explore, clean and prepare

Faster exportation, cleansing and transformation of data with built-in experiences as part of the data science workload

Data Science workloads supports data cleansing, transformation, exploration, and featurization by leveraging built-in experiences on Spark as well as Python based tools

Key Capabilities:

- Accelerate tedious and mundane tasks with Data Wrangler, a code tool to prepare data and generate Python code
- Simplify data analytics, capture and exploit data semantics as users perform transformations on datasets
- Leverage pre-installed open-source libraries for building visualizations like matplotlib, seaborn, Plotly, and more

The screenshot displays the Databricks workspace interface. At the top, the header shows 'DS 1 - Build Model' and 'Saved'. The left sidebar contains navigation icons for Home, Create, Browse, OneLake data hub, Monitoring hub, Workspaces, Realtime Workshop, and Data Engineering. The main area is divided into a 'Lakehouses' panel on the left and a notebook editor on the right. The 'Lakehouses' panel shows a tree view under 'StocksLakehouse' with tables like 'dim_date', 'dim_symbol', 'fact_stocks_daily_prices', 'raw_stock_data', 'raw_stock_test', 'stocks_hour_agg', 'stocks_minute_agg', 'stocks_minute_agg_dataflow', and 'stocks_prediction', as well as a 'Files' section with 'stockhistory'. The notebook editor shows a code cell with the following content:

```
1 fig2 = m.plot_components(forecast)
```

Below the code, a status bar indicates the command was executed successfully in 852 ms. The notebook displays three line plots:

- Top Plot:** A line plot showing 'trend' on the y-axis (ranging from 849 to 856) against 'ds' on the x-axis (ranging from 2024-03-07 to 2024-03-14). The trend shows a steady decline.
- Middle Plot:** A line plot showing 'weekly' on the y-axis (ranging from -150 to 150) against 'Day of Week' on the x-axis (labeled Sunday through Sunday). The plot shows a cyclical pattern with peaks on Tuesday and Thursday.
- Bottom Plot:** A line plot showing 'daily' on the y-axis (ranging from -75 to 50) against 'Hour of day' on the x-axis (labeled from 00:00:00 to 00:00:00). The plot shows a cyclical pattern with a peak around 20:34:17.

The bottom status bar indicates 'Not connected' and 'AutoSave: On'. The bottom right corner shows 'Selected Cell 1 of 44 cells'.

Machine Learning with MLflow

Built-in model & experiment tracking enables data scientists to track and compare their different experiment runs and model versions

Train, evaluate, and score machine learning models by using built-in Experiment and Model artifacts

Key Capabilities:

- Seamless integration with MLflow for experiment tracking and model registration/deployment
- Automatically capture model metrics & parameters with built-in support for MLflow auto-logging
- SynapseML integration with Azure OpenAI service makes it easier to use Apache Spark to process millions of prompts and scale workflow
- Model registry powered by AzureML

WHO-stock-prediction

HomeView

Save as ML modelDownload run filesDelete run

5 selected

Properties				Run metrics		
Run name	Start time ↓	Duration	Status	mse	mae	
mighty_cord_8nvfvkv8	3/6/2024 9:28 ...	12s	Completed	11261.553556...	85.771074955...	
joyful_picture_5lxw7dhh	2/15/2024 10:...	11s	Completed	10790.529824...	81.405928966...	
yellow_pen_s3m2kgd0	2/15/2024 10:...	12s	Completed	9604.2773655...	77.676489984...	
careful_fennel_2n7nrb31	2/14/2024 11:...	11s	Completed	9753.3858813...	78.048121863...	
serene_root_pmp55qfk	1/27/2024 7:5...	11s	Completed	12083.421235...	87.083739908...	

Metric comparisonPerformanceTraining

mse

Run name	mse
serene...	12100
careful...	9800
yellow...	9600
joyful_p...	10800
mighty...	11300

mae

Run name	mae
serene...	87.1
careful...	78.1
yellow...	77.7
joyful_p...	81.4
mighty...	85.8

Customize columns

Filters

Properties

- Start time
- Duration
- Status
- Created by
- Source
- ML model

Metrics

- mse
- mae
- rmse
- mape

Parameters

- changepoi...
- changepoi...
- seasonality...
- weekly_sea...

Azure ML vs Fabric Data Science

Azure Machine Learning is an E2E managed ML platform for building, fine-tuning, prompt-engineering, deploying, and operationalizing ML models responsibly at scale.

Microsoft Fabric is a unified solution for analytics, encompassing data integration, data engineering, data warehouse, data science, observational analytics, and business intelligence—all built on open, lake-centric foundations. Within Fabric, there is Data Science in Microsoft Fabric, a SaaS experience to build and use ML models to enrich data as part of data and analytics workflows.

Together you get a seamless Data & AI stack for your entire Data & AI team; from your data scientists, data engineers and machine learning engineers.



Your organization is invested in Azure ML

Choose Azure Machine Learning. Attach Microsoft Fabric for your data workloads.

Optionally, Fabric DS can be used for EDA and feature engineering.



Your organization is invested in PBI and seeking impro data implementations

If you'd like to deploy models through managed endpoints, towards MLOps?

Choose Azure ML for ML workload. Attach Fabric for data workloads. Optionally, Fabric DS can be used for EDA and feature engineering.

If you'd like to deploy models through one-click flows in Fabric DW/Lakehouse or Power BI?

Choose Fabric DS for ML workload. Attach Azure ML for MLOps (CI/CD, model monitoring, etc.)



Your organization is invested in Azure Databricks

Your choice for data and ML workload to be on ADB or Fabric + AML.

Attach Azure ML if you'd like to fine-tune, prompt-engineer, and deploy pre-built LLMs (Hugging Face, AOAI, Meta).

AzureML & Fabric: Top-line product differentiation



Azure Machine Learning

Value proposition

Azure Machine Learning is a managed end-to-end machine learning platform for building, fine-tuning, deploying, and operating ML models, responsibly at scale.

Ideal personas in your organization

Machine Learning Engineers, Data Scientists

Key capabilities

- ML model building and training with any types of compute including Spark and GPUs for cloud-scale large AI workloads
- Automated ML and drag-and-drop UI
- End-to-end MLOps and repeatable ML pipelines
- Prompt experimentation and LLM fine-tuning
- Model deployment with REST API endpoints, real-time and batch inference



Microsoft Fabric

Value propositions

Microsoft Fabric is a unified solution for analytics—data integration, data engineering, data warehouse, data science, observational analytics, and business intelligence—all built on open, lake-centric foundations.

Ideal personas in your organization

Data Engineers, Data Analysts, ITPros, DBAs, Data Scientists, Cloud Architects

Key capabilities

- Data lake/data warehouse modernization
- Data transformation and unification
- Data visualization
- ML model building and training

AzureML & Fabric: Top-line product capabilities



Azure Machine Learning

Value proposition

Azure Machine Learning is a managed end-to-end machine learning platform for building, fine-tuning, deploying, and operating ML models, responsibly at scale.

Ideal personas in your organization

Machine Learning Engineers, Data Scientists

Key capabilities

- ML model building and training with any types of compute including Spark and GPUs for cloud-scale large AI workloads
- Automated ML and drag-and-drop UI
- End-to-end MLOps and repeatable ML pipelines
- Prompt experimentation and LLM fine-tuning
- Model deployment with REST API endpoints, real-time and batch inference



Data Science in Microsoft Fabric

Value propositions

Data Science in Microsoft Fabric empowers your data professionals to build and use machine learning models to enrich data as part of their data and analytics workflows, all built on top of secured and governed data in Microsoft Fabric.

Ideal personas in your organization

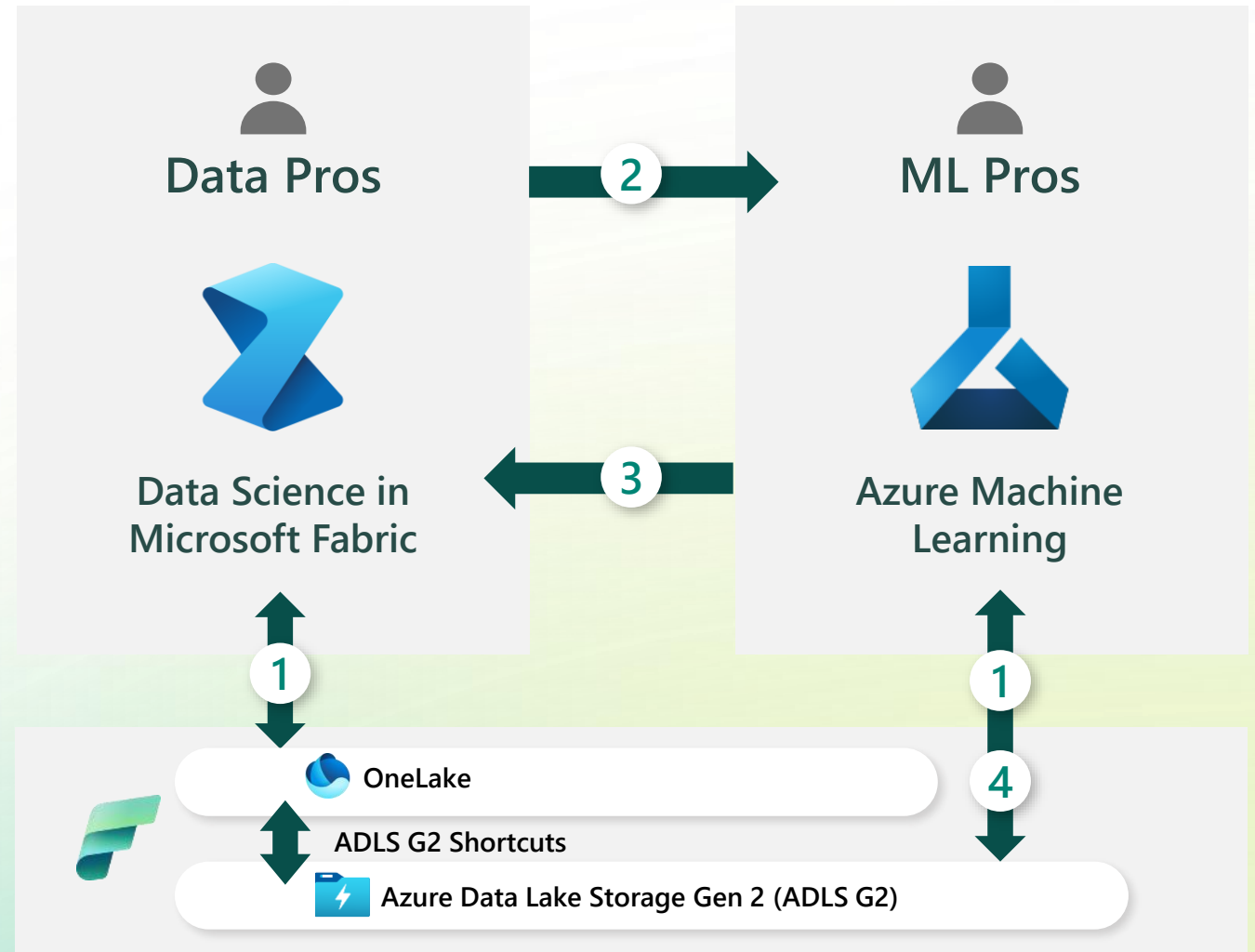
Data Scientists, Data Engineers, Data Analysts

Key capabilities

- Data prep and code generation with Data Wrangler
- Enrich data in your Lakehouse with scalable PREDICT function
- ML model building and training with SynapseML (simple and distributed ML for Spark) with open-source libraries and 1P Azure AI services
- Serving predictive insights through BI

AzureML & Fabric: 1000ft view of product integration

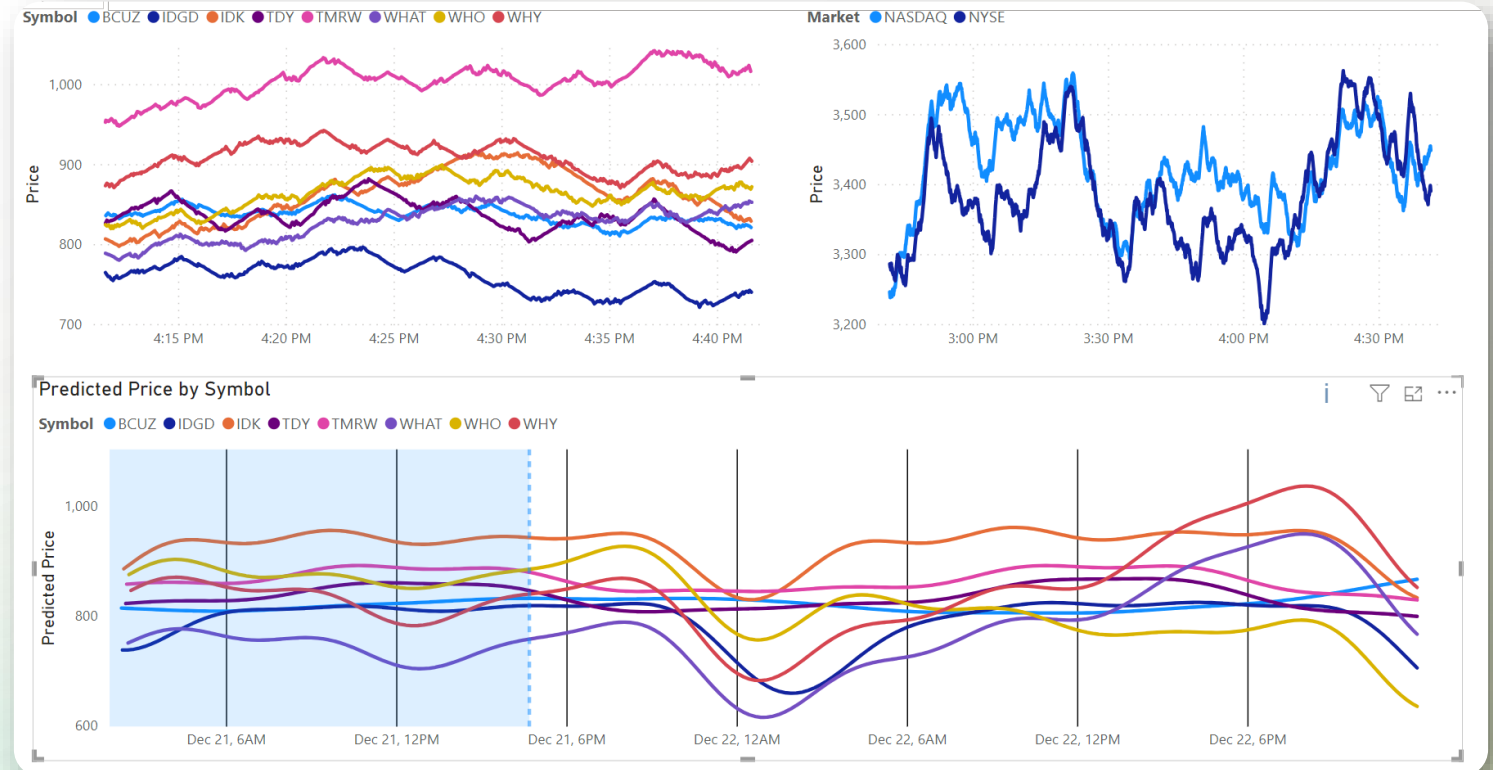
- 1 Data Pros and ML Pros access data items across all stages of the ML lifecycle by reading from and writing to OneLake. ADLS Gen 2 Shortcuts to OneLake can be used reference data from ADLS Gen 2 into OneLake.
- 2 Data Pros can train with Spark and access ML models in Fabric and deploy models using AzureML online and batch endpoints for production.
- 3 ML Pros can export models from AzureML to Fabric to enrich lakehouse or data warehouse serving data as part of analytics workflows.
- 4 Write predictions back to OneLake via Shortcuts to ADLS Gen2 and surface insights through PBI reports with Direct Lake mode.



Module 6 – Data Science

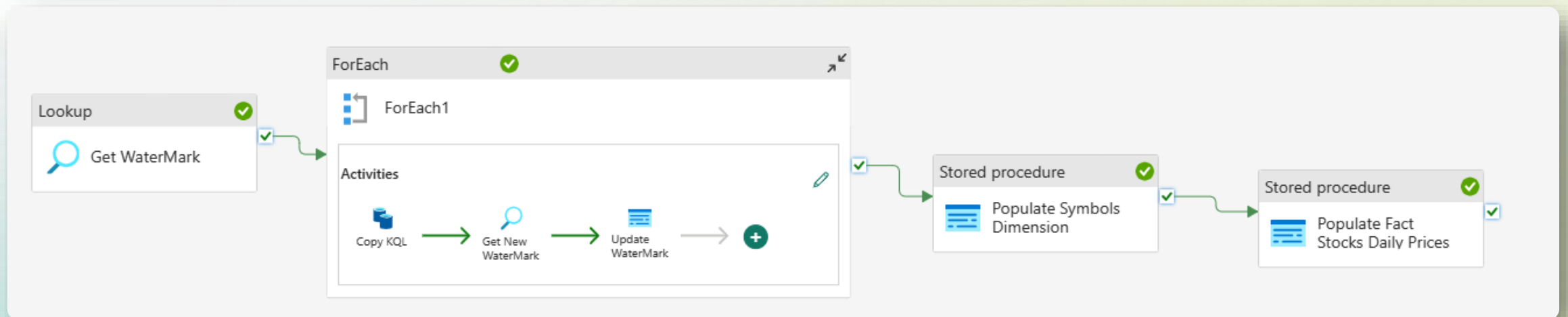
Goals: Use notebooks to build an ML model predicting the stock prices.

Using the library Prophet, data will be analyzed allowing for future predictions. These predictions will be stored in the lakehouse and integrated into the reports.



Module 4 – Data Warehouse

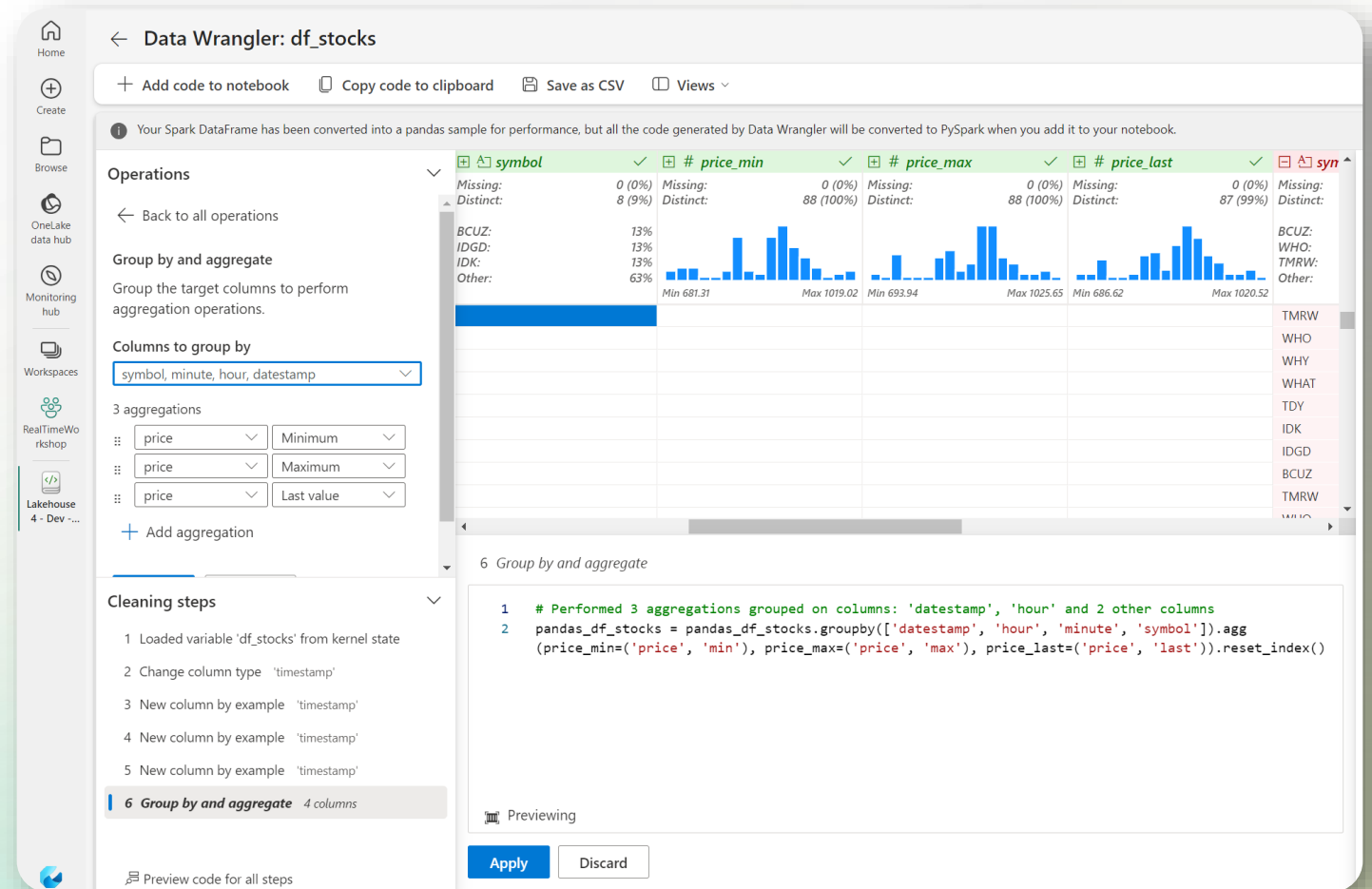
- **Goals:** Build a dimensional model and data pipeline that stores aggregated data into a data warehouse.
- Using the data warehousing capability, the data will be periodically loaded and summarized into a data warehouse for historic reporting. To accomplish this, we'll use data pipelines for orchestration of an ETL pipeline.



Module 5 – Data Lakehouse

Goals: Create a data lakehouse using data wrangler, notebooks, and Delta tables.

After modifying the Eventstream to copy data to the lakehouse, notebooks and data wrangler will be used to aggregate the data into tables that can be used for reporting and data science.





Sponsored by
AMD

Wrap-up, Q&A, & Feedback



Wrap-up

Thank you for attending! Check out the additional challenges throughout the modules, and the Extras module with unique add-ons.



Comments or suggestions? Please let us know.



Get your questions answered in the Microsoft Fabric webinar series: <https://aka.ms/fabric-webinar-series>



See the latest announcements in the Microsoft Fabric blog site: <https://aka.ms/fabric-tech-blog>



Read the implementation guide:
<https://aka.ms/Getting-Started-eBook>



Microsoft Fabric Community Resources



Community Call to Action

- ✓ Try Microsoft Fabric for free: <https://aka.ms/try-fabric>
- ✓ Join the Fabric community: <https://aka.ms/fabriccommunity>
- ✓ Share and vote for ideas to improve Fabric: <https://aka.ms/fabricideas>
- ✓ Read and comment our blog: <https://aka.ms/fabricblog>

Learn More about Microsoft Fabric

- Product announcement: <https://aka.ms/fabric>
- Digital Event at Build (videos): <https://aka.ms/build-with-analytics>
- Product website: <https://aka.ms/microsoft-fabric>
- Documentation: <https://aka.ms/fabric-docs>
- Fabric e-book: <https://aka.ms/fabric-get-started-ebook>
- Microsoft Learn: <https://aka.ms/learn-fabric>
- End-to-end scenario tutorials: <https://aka.ms/fabric-tutorials>
- Fabric Notes: <https://aka.ms/fabric-notes>



Thank you