Microsoft

Microsoft Azure Training Day: Data and Analytics

Cloud Scale Analytics with the Modern Data Warehouse
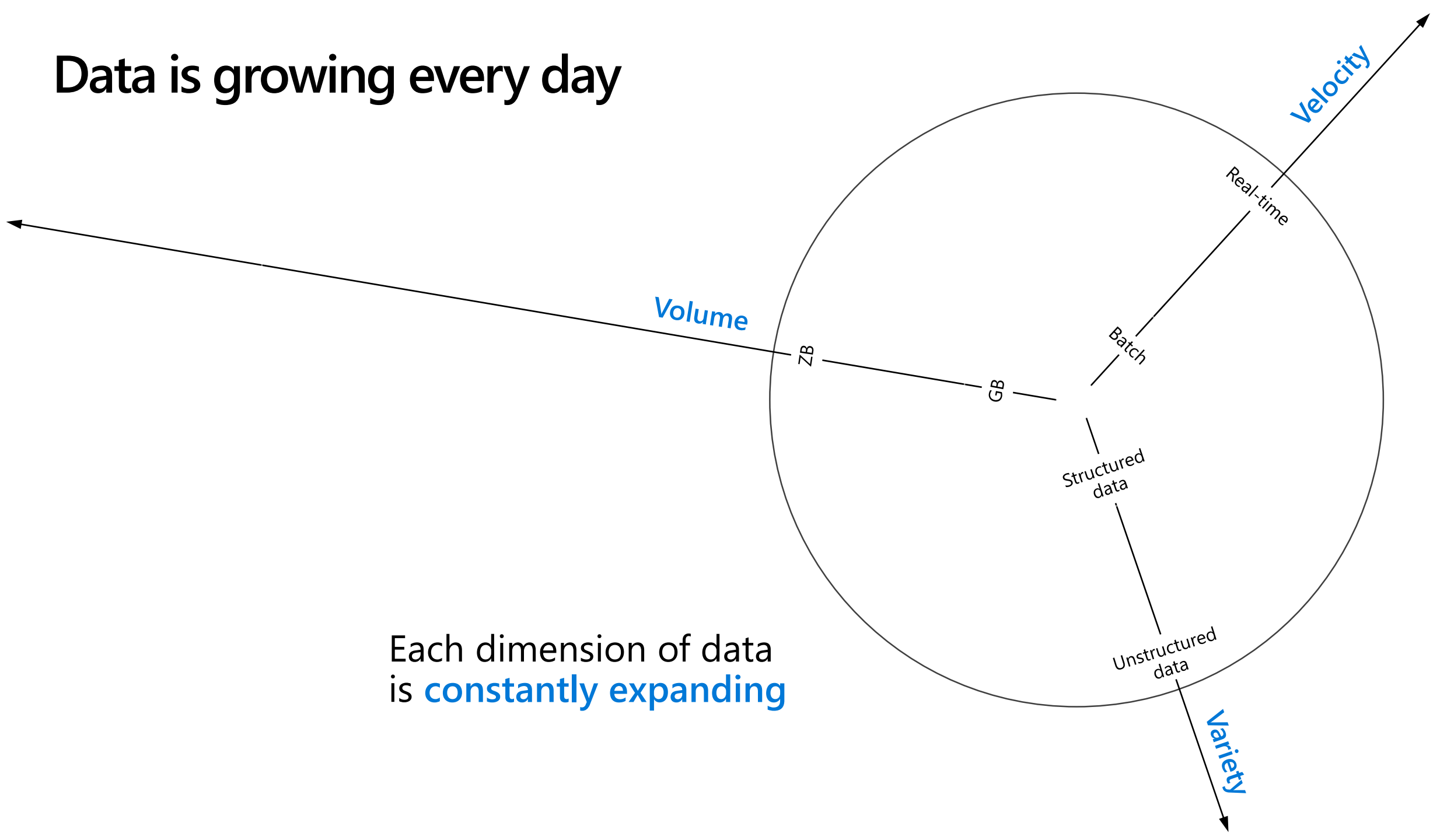
Francesco Viganò
Cloud Solution Architect – Data & AI

frvigano@microsoft.com
https://www.linkedin.com/in/francesco-viganò/

Data is growing every day

Velocity

Volume

Real-time

Batch

ZB

GB

Structured data

Unstructured data

Variety

Each dimension of data is **constantly expanding**

Get the benefits of a modern cloud data warehouse with Azure

Unlimited scale and integration

Adaptable to flexible needs
and budgets

Advanced and real-time
analytics for deeper insights

Companies that take control of their data outperform

$40k
more revenue per employee

50%+
higher average net income on revenue

$100M
in additional operating income each year

Organizations that fully harness their data outperform

Data consolidation using
Azure SQL Data Warehouse

"

This architecture performs
significantly better than the legacy
on-premises solutions it replaced.

Chetan Kundavaram, Global Director, AB InBev

Migration to the cloud for
efficient business operations

"

By migrating off legacy data
infrastructure and running reports on
a more cost-efficient Azure platform,
we've reduced costs.

Roberto Pasquier, Senior Manager, Newell Brands

Using Azure SQL Data Warehouse
for **predictive analytics**

"

Because the Personalization Cloud is
in Azure, we're confident we can
handle any volume of data coming
from our clients.

Brian Muenks, IT Manager, Maritz Motivation Solutions
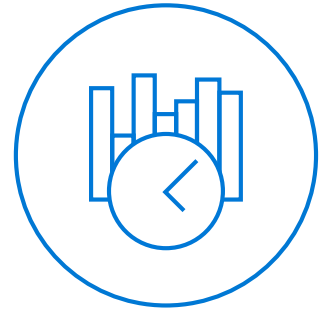
Cloud-scale Analytics

Modern data warehousing

" We want to analyze data coming from multiple sources and in varied formats...

Advanced analytics

" We want to leverage the analytics platform for advanced fraud detection...
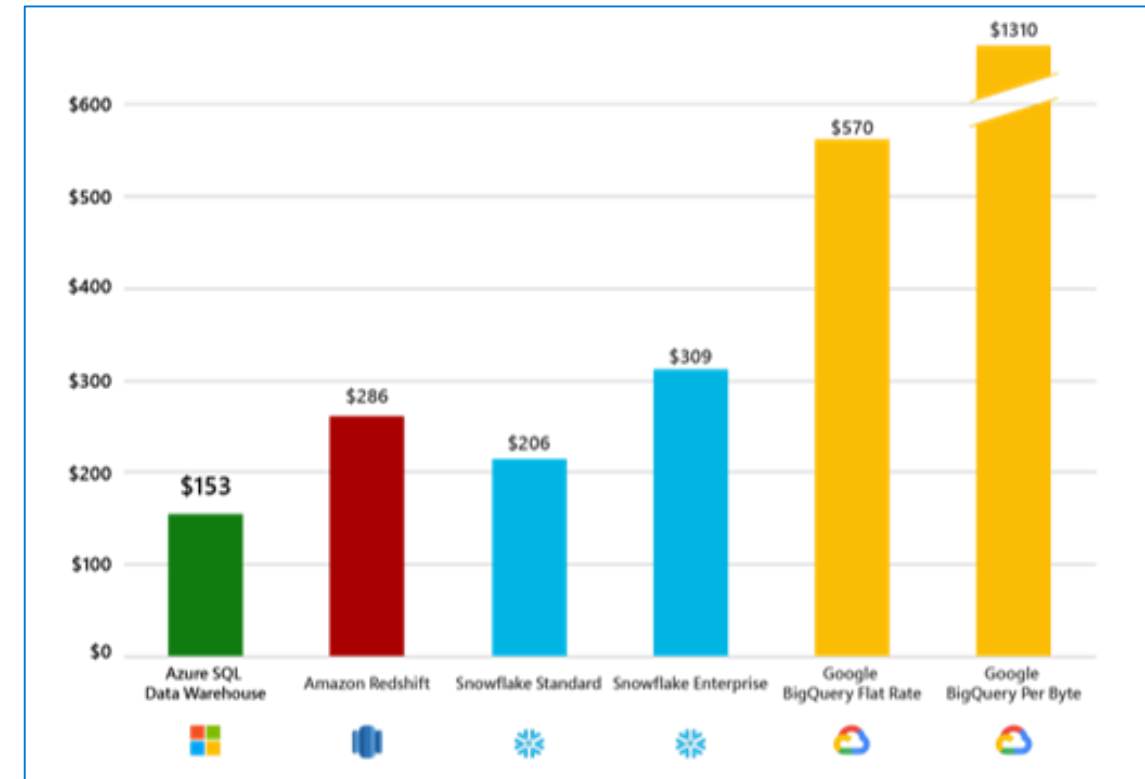
Real-time analytics

" We're trying to get insights from our devices in real-time...

Best-in-class price per performance

Price-performance is calculated by GigaOm as the TPC-DS metric of cost of ownership divided by composite query.

Price-Performance @ 30TB
Lower is Better

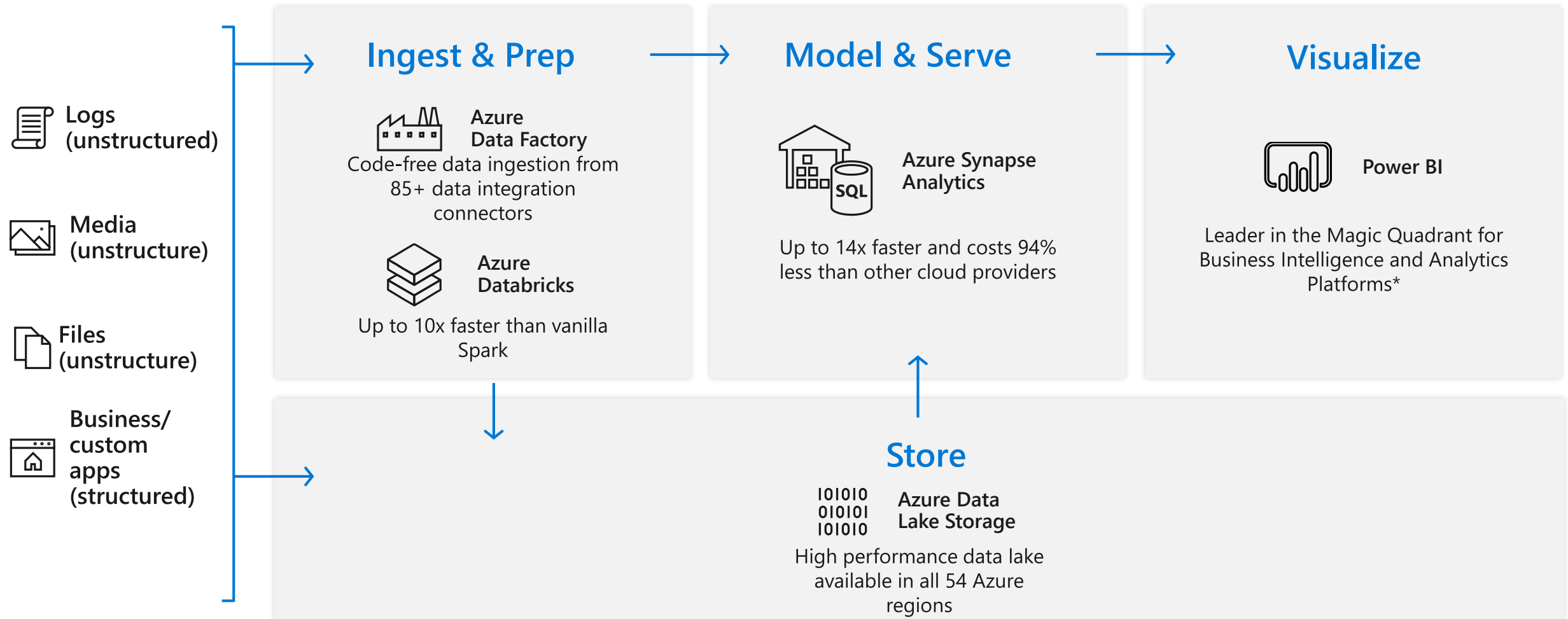|					
$153	$286	$206	$309	$570	$1310
Azure SQL Data Warehouse	Amazon Redshift	Snowflake Standard	Snowflake Enterprise	Google BigQuery Flat Rate	Google BigQuery Per Byte

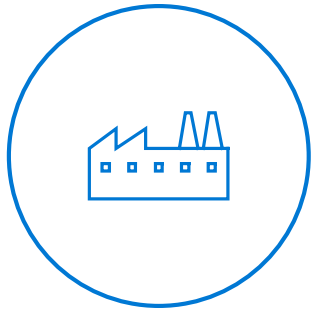Results based on GigaOm's TPC-DS results, published in February 2019

Azure Modern Data Warehouse Processes

Best end-to-end ecosystem to turn your data into actionable insights

Unparalleled performance

Logs (unstructured)

Media (unstructure)

Files (unstructure)

Business/ custom apps (structured)

Ingest & Prep

Azure Data Factory

Code-free data ingestion from 85+ data integration connectors

Azure Databricks

Up to 10x faster than vanilla Spark

Model & Serve

Azure Synapse Analytics

Up to 14x faster and costs 94% less than other cloud providers

Visualize

Power BI

Leader in the Magic Quadrant for Business Intelligence and Analytics Platforms*

Store

Azure Data Lake Storage

High performance data lake available in all 54 Azure regions
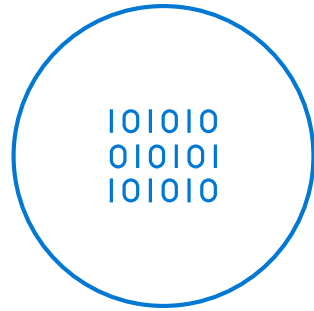
Modern Data Warehouse

Azure Data Factory

Connect to hybrid, other cloud, and SaaS applications out-of-the-box

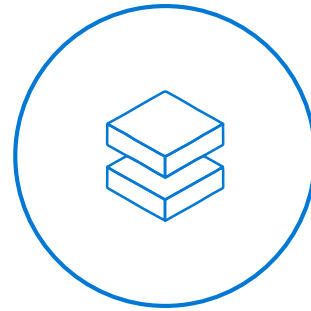No code/little code Cloud ETL

85+ Connectors at no extra charge

Azure Data Lake Storage

Analytics optimized cloud storage

Combined hierarchical file and object stores

Enterprise-grade security with Azure Active Directory

Azure Databricks

Native integration with all MDW services

Collaborative workspaces for data scientists, data engineers, and business analysts

Enterprise-grade security with Azure Active Directory

Azure SQL Data Warehouse

Separation of compute and storage

Intelligent workload management and enhanced concurrency

Industry leading price-performance and security features

Deep integration with DataOps model (CI/CD and dev tools)

The Azure Big Data Landscape

Azure Data Factory

Azure Import/Export service

Azure CLI

Azure SDK

Azure IoT Hub

Azure event hubs

Kafka on Azure HDInsight

Azure SQL DB

Azure Cosmos DB

Azure Synapse

Azure Analysis Services

Power BI

Azure Blob Storage

Azure Data Lake Store

Azure HDInsight

Azure Databricks

Azure ML

ML Server

Azure Databricks

Azure Search

Azure Data Catalog

Azure Stream Analytics

Azure HDInsight

Azure Databricks

Bot service

Cognitive services

Azure ExpressRoute

Azure Active Directory

Azure network security groups

Azure key management service

Operations Management Suite

Azure Functions

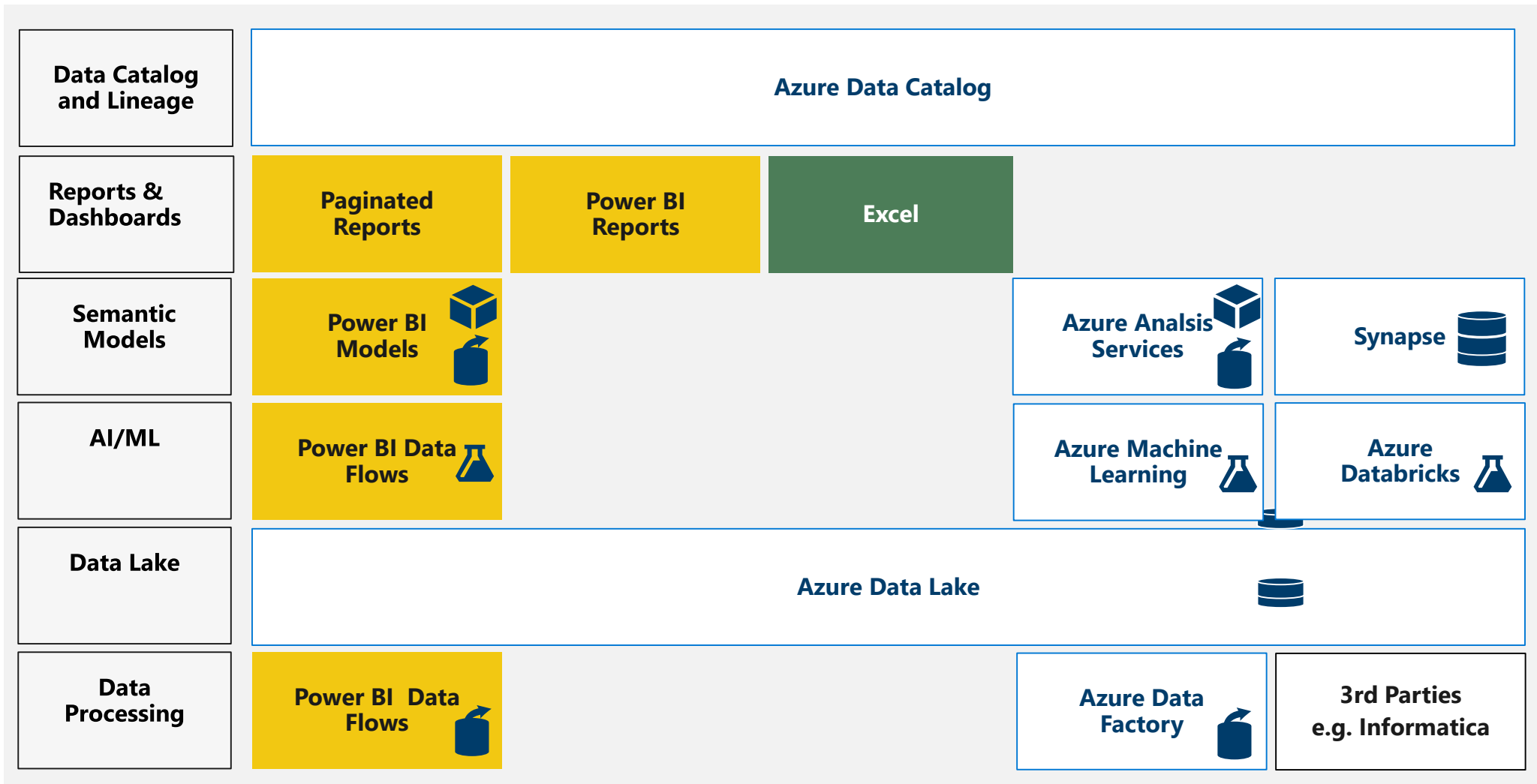Visual Studio

Data warehousing pattern in Azure
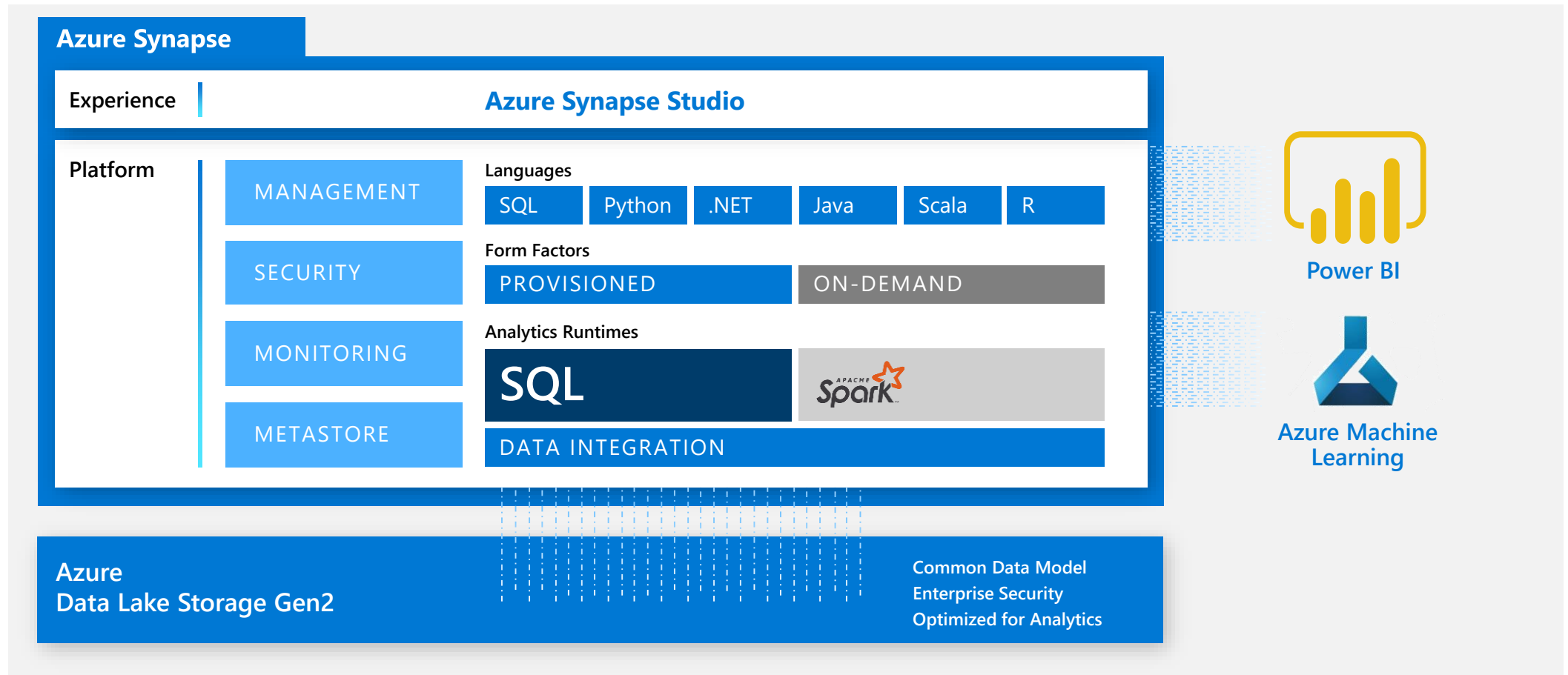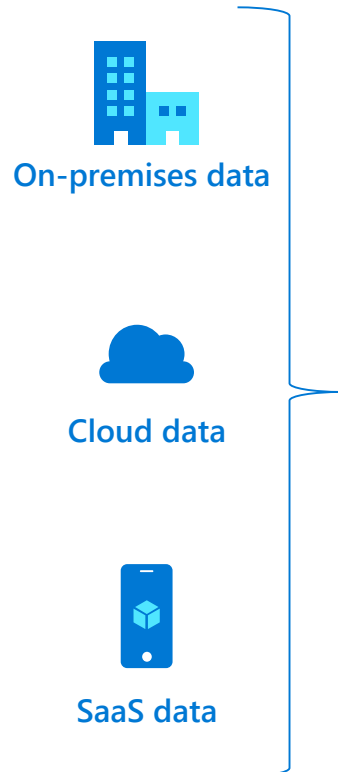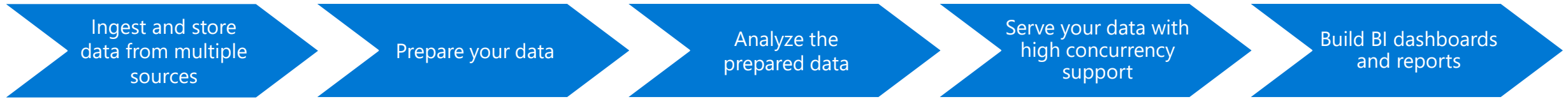## Loading and preparing data for analysis with a data warehouse
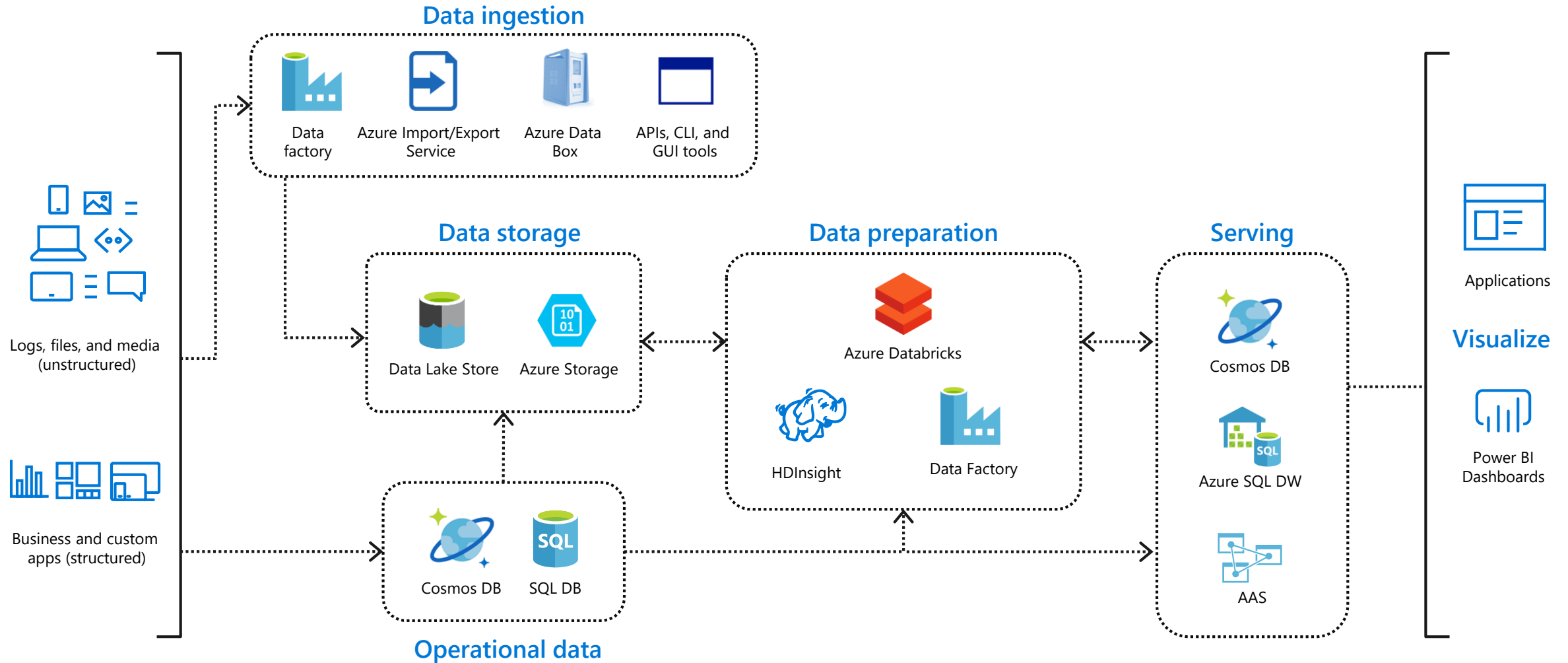
Modern Data Warehouse and Power BI

Analytics using Azure Synapse

| Ingest and store data from multiple sources | Prepare your data | Analyze the prepared data | Serve your data with high concurrency support | Build BI dashboards and reports |

On-premises data

Cloud data

SaaS data

Azure Synapse

Experience — **Azure Synapse Studio**

Platform

| MANAGEMENT |
| SECURITY |
| MONITORING |
| METASTORE |

Languages

| SQL | Python | .NET | Java | Scala | R |

Form Factors

| PROVISIONED | ON-DEMAND |

Analytics Runtimes

| SQL | Spark |

| DATA INTEGRATION |

Azure
Data Lake Storage Gen2

Common Data Model
Enterprise Security
Optimized for Analytics

Power BI

Azure Machine Learning

Things to note

- > • There are no right or wrong solutions, only optimal solutions

- > • Lead with certain solutions and customize based on customer scenarios

- > • Customer voice and product and service maturity govern lead solutions

- > • Consider price and performance, ease of use, and ecosystem acceptance as factors

- > • Everything is fluid - a lead solution today might be non-optimal tomorrow, based on the factors above and new releases

Modern Data Warehouse considerations

Big Data and advanced analytics

Security

Enables the modern data warehouse to control access in order to protect sensitive data and maintain desired compliance

Automation

Enables all components of the modern data warehouse solution to be controlled, deployed, and monitored programmatically

Monitoring

Provides insights into the status and health of the data warehouse solution

aka.ms/DATA10   #MSIgnite