

© Copyright Microsoft Corporation. All rights reserved.

FOR USE ONLY AS PART OF VIRTUAL TRAINING DAYS PROGRAM. THESE MATERIALS ARE NOT AUTHORIZED FOR DISTRIBUTION, REPRODUCTION OR OTHER USE BY NON-MICROSOFT PARTIES.

# Module 3: Explore non-relational data in Azure-

## Segment 1

# Agenda



Explore non-relational data offerings in Azure

---



Explore provisioning and deploying non-relational data services in Azure

---



Manage non-relational data stores in Azure

# Lesson 1: Explore non-relational data offerings in Azure



## Lesson 1 objectives



Explore use-cases and management benefits of using Azure Table storage



Explore use-cases and management benefits of using Azure Blob storage



Explore use-cases and management benefits of using Azure File storage



Explore use-cases and management benefits of using Azure Cosmos DB

# Explore Azure Table storage

Key (Customer ID)	Value (Customer Data)					
C1	AAAAA	BBB	101	Block Street	YY	999 888
C2	MM	NN	21	A Street	5	B Avenue
C3	DDD	EEE	FFF	111	222	66 C Road

# Explore Azure Blob Storage

## Block blobs

- Has a maximum size of 4.7TB
- Best for storing large, discrete, binary objects that changes infrequently
- Each individual block can store up to 100MB of data
- A block blob can contain up to 50000 blocks

## Page blobs

- Can hold up to 8TB of data
- Is organized as a collection of fixed sized-512 byte pages
- Used to implement virtual disk storage for virtual machines

## Append blobs

- The maximum size is just over 195GB
- Is a block blob that is used to optimize append operations
- Each individual block can store up to 4MB of data

# Explore Azure File Storage

The screenshot displays the Azure portal interface for a storage account named 'jpwstorageaccount'. The left-hand navigation pane includes sections for 'Overview' (with links to Activity log, Access control (IAM), Tags, Diagnose and solve problems, Data transfer, Events, and Storage Explorer (preview)) and 'Settings' (with links to Access keys, Geo-replication, CORS, Configuration, Encryption, Shared access signature, Firewalls and virtual networks, Private endpoint connections, Advanced security, Static website, and Properties). The main content area features a top bar with a search box and action buttons (Open in Explorer, Move, Refresh, Delete, Feedback). Below this is an information banner about Azure Monitor classic alerts. A table of account properties is shown, including Resource group (learnrg), Status (Primary: Available, Secondary: Available), Location (UK West, UK South), Subscription (change), Subscription ID, Tags (change), Performance/Access tier (Standard/Hot), Replication (Read-access geo-redundant storage (RA-GRS)), and Account kind (StorageV2 (general purpose v2)). The bottom section contains four tiles: Containers (Massively scalable data lake storage), File shares (Serverless SMB file shares, highlighted with a red box), Tables (Tabular data storage), and Queues (Effectively scale apps according to traffic). Each tile has a 'Learn more' link. At the very bottom, a 'Tools and SDKs' section lists various development tools and languages: Storage Explorer (preview), PowerShell, Azure CLI, .NET, Java, Python, and Node.js.

Home > Microsoft.StorageAccount-20200520142053 | Overview > jpwstorageaccount

**jpwstorageaccount**  
Storage account

Search (Ctrl+/)

Open in Explorer → Move Refresh Delete Feedback

**Overview**

- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Data transfer
- Events
- Storage Explorer (preview)

**Settings**

- Access keys
- Geo-replication
- CORS
- Configuration
- Encryption
- Shared access signature
- Firewalls and virtual networks
- Private endpoint connections
- Advanced security
- Static website
- Properties

**File shares**  
Serverless SMB file shares  
[Learn more](#)

**Containers**  
Massively scalable data lake storage  
[Learn more](#)

**Tables**  
Tabular data storage  
[Learn more](#)

**Queues**  
Effectively scale apps according to traffic  
[Learn more](#)

**Tools and SDKs**

- Storage Explorer (preview)
- PowerShell
- Azure CLI
- .NET
- Java
- Python
- Node.js

Classic alerts in Azure Monitor is announced to retire in 2021, it is recommended that you upgrade your classic alert rules to retain alerting functionality with the new alerting platform. For more information, see [Continue alerting with ARM storage accounts](#).

Resource group (change) : [learnrg](#)

Status : Primary: Available, Secondary: Available

Location : UK West, UK South

Subscription (change) :

Subscription ID :

Tags (change) : [Click here to add tags](#)

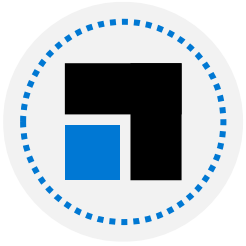
Performance/Access tier : Standard/Hot

Replication : Read-access geo-redundant storage (RA-GRS)

Account kind : StorageV2 (general purpose v2)



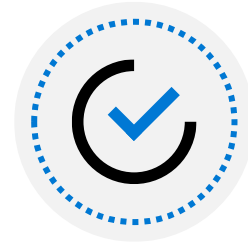
# Explore Azure Cosmos DB



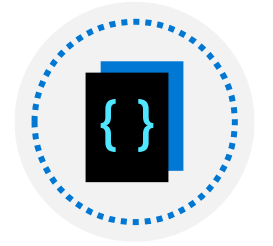
**Scalability**



**Performance**



**Availability**



**Programming  
model**

# Use cases for Azure Cosmos DB

## Web and retail

Using Azure Cosmos DB's multi-master replication model along with Microsoft's performance commitments, Data Engineers can implement a data architecture to support web and mobile applications that achieve less than a 10-ms response time anywhere in the world

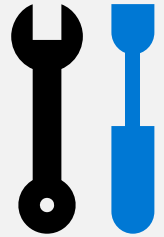
## Gaming

The database tier is a crucial component of gaming applications. Modern games perform graphical processing on mobile/console clients but rely on the cloud to deliver customized and personalized content like in-game stats, social media integration, and high-score leader boards.

## IoT scenarios

Hundreds of thousands of devices have been designed and sold to generate sensor data known as Internet of Things (IoT) devices. Using technologies like Azure IoT Hub, Data Engineers can easily design a data solution architecture that captures real-time data. Cosmos DB can accept and store this information very quickly

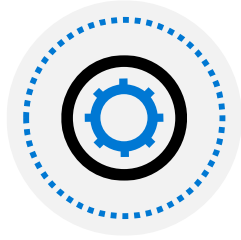
## Lesson 2: Explore provisioning and deploying non-relational data services in Azure



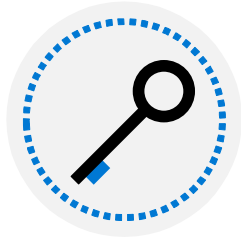
## Lesson 2 objectives



Provision non-relational data services



Configure non-relational data services



Explore basic connectivity issues



Explore data security components

# Provisioning Cosmos DB

Microsoft Azure

»

Home > Create Azure Cosmos DB Account

Create Azure Cosmos DB Account

Basics

Network

Tags

Summary

Azure Cosmos DB is a fully managed globally distributed, multi-model database service, transparently replicating your data across any number of Azure regions. You can elastically scale throughput and storage, and take advantage of fast, single-digit-millisecond data access using your favorite API among SQL, MongoDB, Apache Cassandra, Tables, or Gremlin, backed by 99.999 SLA. [learn more](#)

#### PROJECT DETAILS

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

\* Subscription

Concierge Subscription

\* Resource Group

<GUID for your resource group in your Conceirge Subscription>

[Create new](#)

#### INSTANCE DETAILS

\* Account Name

<enter a unique name>

documents.azure.com

\* API ⓘ

SQL

\* Location

<choose the location closest to you>

Geo-Redundancy ⓘ

Enable

Disable

Multi-region Writes ⓘ

Enable

Disable

Review + create

Previous

Next: Network



© Copyright Microsoft Corporation. All rights reserved.

FOR USE ONLY AS PART OF VIRTUAL TRAINING DAYS PROGRAM. THESE MATERIALS ARE NOT AUTHORIZED FOR DISTRIBUTION, REPRODUCTION OR OTHER USE BY NON-MICROSOFT PARTIES.

# Module 3: Explore non-relational data in Azure

## Segment 2



Demo: Create and Deploy a Cosmos DB Database

# Provisioning Data Lake Storage

## Create storage account

[Basics](#) [Advanced](#) [Tags](#) [Review + create](#)

Azure Storage is a Microsoft-managed service providing cloud storage that is highly available, secure, durable, scalable, and redundant. Azure Storage includes Azure Blobs (objects), Azure Data Lake Storage Gen2, Azure Files, Azure Queues, and Azure Tables. The cost of your storage account depends on the usage and the options you choose below. [Learn more](#)

### PROJECT DETAILS

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

\* Subscription

\* Resource group  [Create new](#)

### INSTANCE DETAILS

The default deployment model is Resource Manager, which supports the latest Azure features. You may choose to deploy using the classic deployment model instead. [Choose classic deployment model](#)

\* Storage account name ⓘ

\* Location

Performance ⓘ ☒ Standard ☐ Premium

Account kind ⓘ

Replication ⓘ

Access tier (default) ⓘ ☐ Cool ☒ Hot

[Home](#) > [New](#) > Create storage account

## Create storage account

[Basics](#) [Advanced](#) [Tags](#) [Review + create](#)

### SECURITY

Secure transfer required ⓘ ☐ Disabled ☒ Enabled

### VIRTUAL NETWORKS

Allow access from ☒ All networks ☐ Selected network  
 ⓘ All networks will be able to access this storage account. [Learn more](#)

### DATA LAKE STORAGE GEN2 (PREVIEW)

Hierarchical namespace ⓘ ☐ Disabled ☒ Enabled

# Azure authentication



# Azure authentication



# Azure authentication



# Configure Storage Accounts

Create storage account

Basics

Advanced

Tags

Review + create

Azure Storage is a Microsoft-managed service providing cloud storage that is highly available, secure, durable, scalable, and redundant. Azure Storage includes Azure Blobs (objects), Azure Data Lake Storage Gen2, Azure Files, Azure Queues, and Azure Tables. The cost of your storage account depends on the usage and the options you choose below. [Learn more](#)

PROJECT DETAILS

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

\* Subscription

Visual Studio Enterprise

\* Resource group

Select existing...

Create new

INSTANCE DETAILS

The default deployment model is Resource Manager, which supports the latest Azure features. You may choose to deploy using the classic deployment model instead. [Choose classic deployment model](#)

\* Storage account name ⓘ

\* Location

East US

Performance ⓘ

☒ Standard

☐ Premium

Account kind ⓘ

StorageV2 (general purpose v2)

Replication ⓘ

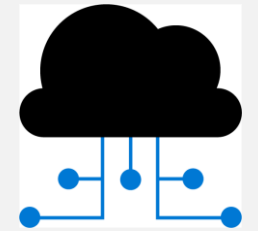
Read-access geo-redundant storage (RA-GRS)

Access tier (default) ⓘ

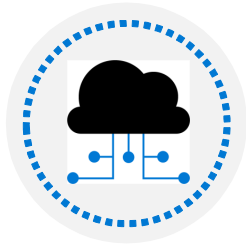
☐ Cool

☒ Hot

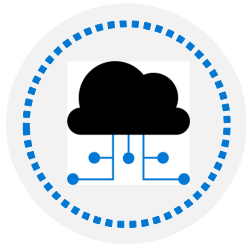
## Lesson 3: Manage non-relational data stores in Azure



## Lesson 3 objectives



Upload data to a Cosmos DB database, and learn how to query this data.



Upload and download data in an Azure Storage account.



## Cosmos DB APIs



SQL API

MongoDB API

Cassandra API

Gremlin API

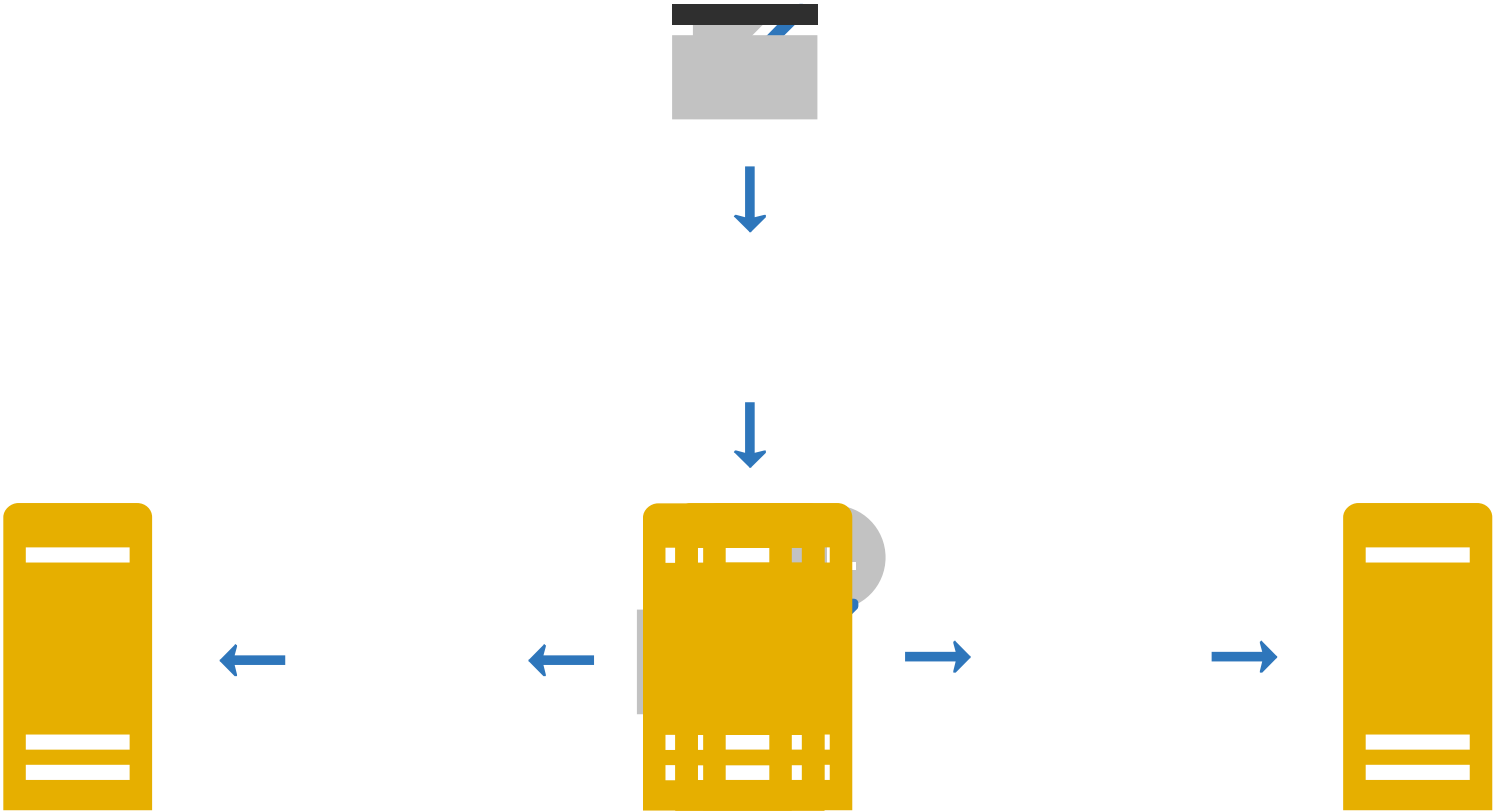
Table API

# Load data using the Cosmos DB Migration tool

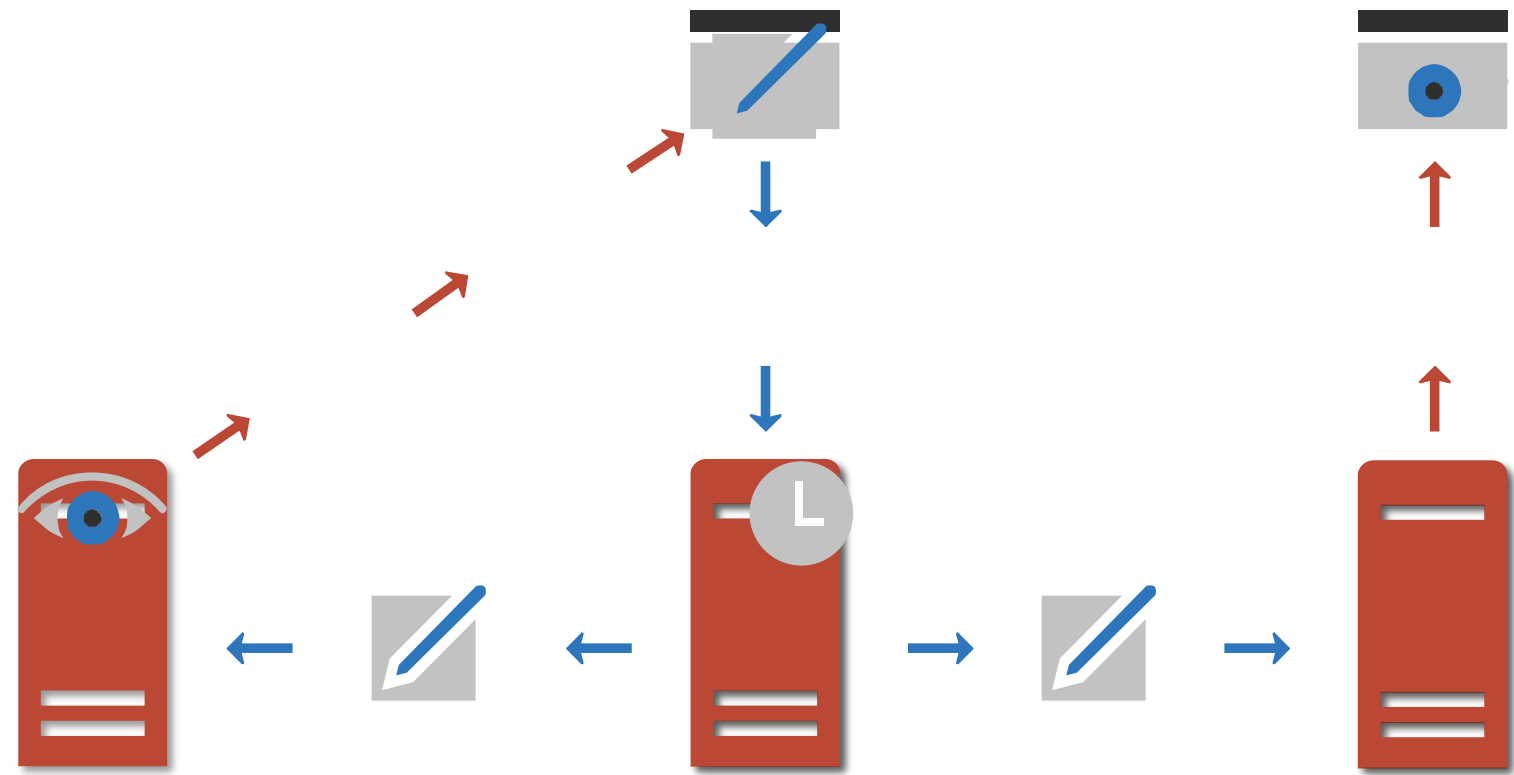
You can use the Data Migration tool to import data to Azure Cosmos DB from a variety of sources, including:

- JSON files
- MongoDB
- SQL Server
- CSV files
- Azure Table storage
- Amazon DynamoDB
- HBase
- Azure Cosmos containers

# Configure consistency



# Configure consistency



# Query Azure Cosmos DB

## SELECT Query Basics

```
SELECT <select_list>
[FROM <optional_from_specification>]
[WHERE <optional_filter_condition>]
[ORDER BY <optional_sort_specification>]
[JOIN <optional_join_specification>]
```

## Examples

```
SELECT *
FROM Products p WHERE p.id ="1"
```

```
SELECT p.id, p.manufacturer, p.description
FROM Products p WHERE p.id ="1"
```

```
SELECT p.price, p.description, p.productId
FROM Products p ORDER BY p.price ASC
```

```
SELECT p.productId
FROM Products p JOIN p.shipping
```

# Manage Azure Blob Storage

The screenshot displays the Azure portal interface for a storage account named 'contosodata'. The left-hand navigation pane lists various management options, including Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Data transfer, Events, Storage Explorer (preview), and a Settings section with options like Access keys, Geo-replication, CORS, Configuration, Encryption, Shared access signature, Firewalls and virtual networks, Private endpoint connections, Advanced security, Static website, Properties, and Locks. The main content area shows the 'Overview' tab selected, displaying account details such as Resource group (learnrg), Status (Primary: Available, Secondary: Available), Location (UK South, UK West), Subscription (change) (Freebie), Subscription ID (b581336d-00a0-41f6-bacd-a4f6b8779001), Tags (change) (Click here to add tags), Performance/Access tier (Standard/Hot), Replication (Read-access geo-redundant storage (RA-GRS)), and Account kind (StorageV2 (general purpose v2)). A blue banner at the top of the main area contains a message about classic alerts in Azure Monitor retiring in 2021. Below the account details, four storage service options are presented in a grid: Containers (highlighted with a red box), File shares, Tables, and Queues. Each option includes an icon, a title, a brief description, and a 'Learn more' link.

**contosodata**  
Storage account

Search (Ctrl+/) << Open in Explorer → Move Refresh Delete Feedback

**Overview**

- Activity log
- Access control (IAM)
- Tags
- Diagnose and solve problems
- Data transfer
- Events
- Storage Explorer (preview)

**Settings**

- Access keys
- Geo-replication
- CORS
- Configuration
- Encryption
- Shared access signature
- Firewalls and virtual networks
- Private endpoint connections
- Advanced security
- Static website
- Properties
- Locks

**Containers**  
Scalable, cost-effective storage for unstructured data  
[Learn more](#)

**File shares**  
Serverless SMB file shares  
[Learn more](#)

**Tables**  
Tabular data storage  
[Learn more](#)

**Queues**  
Effectively scale apps according to traffic  
[Learn more](#)

Classic alerts in Azure Monitor is announced to retire in 2021, it is recommended that you upgrade your classic alert rules to retain alerting functionality with the new alerting platform. For more information, [see Continue alerting with ARM storage accounts](#)

Resource group ([change](#)) [learnrg](#)

Status  
Primary: Available, Secondary: Available

Location  
UK South, UK West

Subscription ([change](#))  
[Freebie](#)

Subscription ID  
b581336d-00a0-41f6-bacd-a4f6b8779001

Tags ([change](#))  
[Click here to add tags](#)

Performance/Access tier  
Standard/Hot

Replication  
Read-access geo-redundant storage (RA-GRS)

Account kind  
StorageV2 (general purpose v2)

# Manage Azure File storage

The screenshot displays the Azure File Explorer interface for a file share named 'reports'. The left sidebar contains navigation options: Overview (selected), Access Control (IAM), Settings, Properties, Operations, and Snapshots. The main pane shows a list of directories: February Reports, January Reports, and March Reports. A red box highlights the 'Connect' button in the top toolbar. The right pane, titled 'Connect', provides instructions for connecting to the file share from Windows, Linux, or macOS. It includes a warning about 'Secure transfer required' and a PowerShell script to test the connection and mount the drive.

Home > contosodata | File shares >

**reports**  
File share

Search (Ctrl+/)

Connect Upload Add directory Refresh Delete share Edit qu

Overview

Access Control (IAM)

Settings

Properties

Operations

Snapshots

Search files by prefix

Name	Type
February Reports	Directory
January Reports	Directory
March Reports	Directory

### Connect

reports

⚠ 'Secure transfer required' is enabled on the storage account. SMB clients must support 3.0 encryption to connect. Click here to learn more about connecting Azure files.

Windows Linux macOS

Drive letter

Z

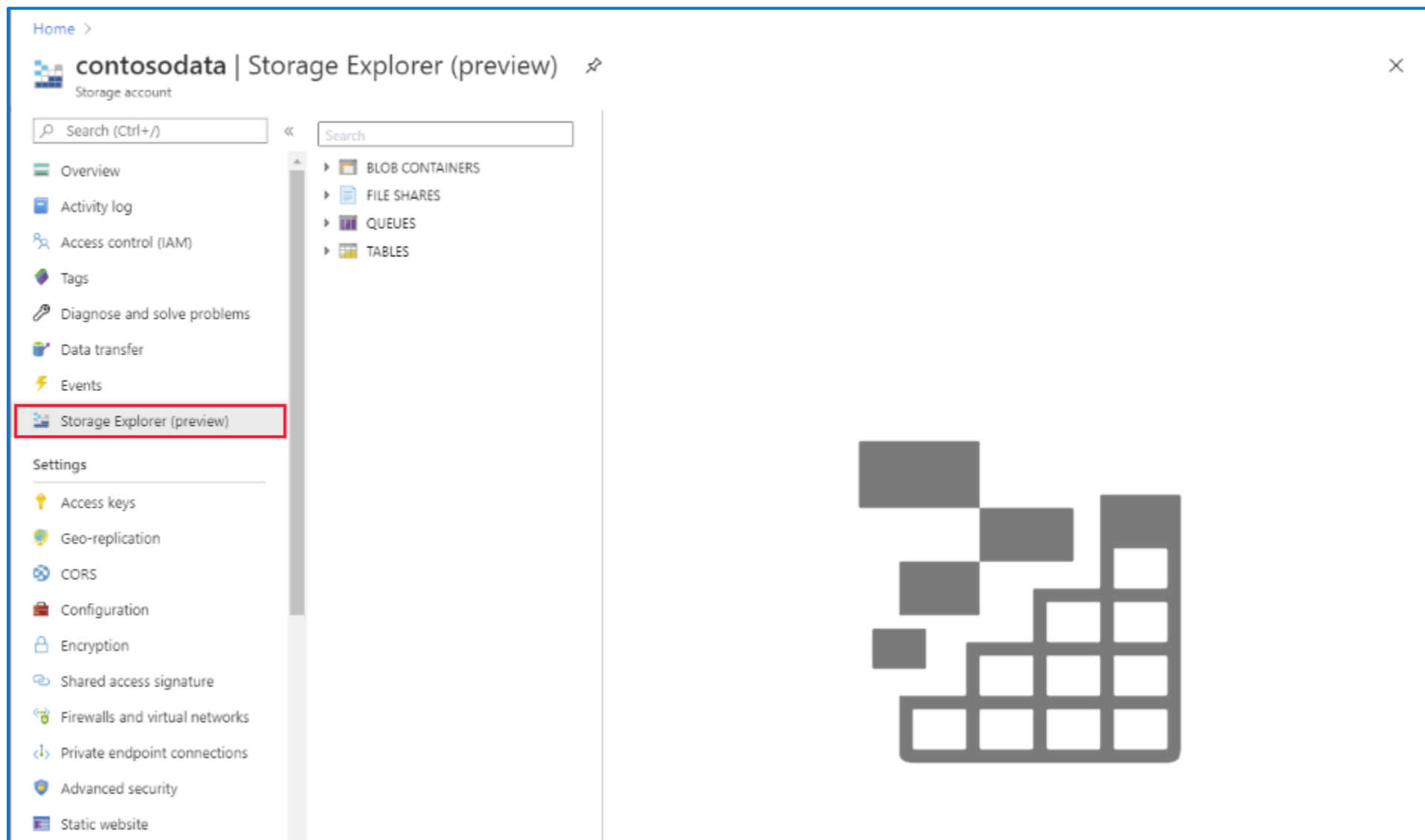
To connect to this Azure file share from Windows, run these PowerShell commands from a normal (not elevated) PowerShell terminal:

```
$connectTestResult = Test-NetConnection -ComputerName  
contosodata.file.core.windows.net -Port 445  
if ($connectTestResult.TcpTestSucceeded) {  
    # Save the password so the drive will persist on reboot  
    cmd.exe /C "cmdkey  
/add:"contosodata.file.core.windows.net"  
/user:"Azure\contosodata"
```

This script will check to see if this storage account is accessible via TCP port 445, which is the port SMB uses. If port 445 is available, your Azure file share will be persistently mounted. Your organization or internet service provider (ISP) may block port 445, however you may use Azure [Point-to-Site \(P2S\) VPN](#), Azure [Site-to-Site \(S2S\) VPN](#), or [ExpressRoute](#) to tunnel SMB traffic to your Azure file share over a different port.

[Learn how to circumvent the port 445 problem \(VPN\)](#)

# Copying to Azure Storage







© Copyright Microsoft Corporation. All rights reserved.

FOR USE ONLY AS PART OF VIRTUAL TRAINING DAYS PROGRAM. THESE MATERIALS ARE NOT AUTHORIZED FOR DISTRIBUTION, REPRODUCTION OR OTHER USE BY NON-MICROSOFT PARTIES.

# Module 4: Explore modern data warehouse analytics

# Agenda



Examine components of a modern data warehouse

---



Explore data ingestion in Azure

---



Explore data storage and processing in Azure

---



Get started building with Power BI

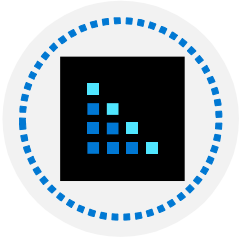
# Lesson 1: Examine components of a modern data warehouse



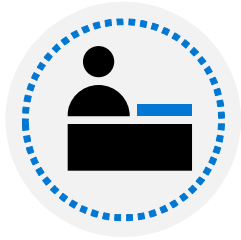
# Lesson 1 objectives



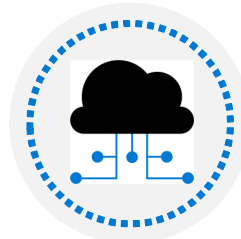
Explore data warehousing concepts



Explore Azure data services for modern data warehousing

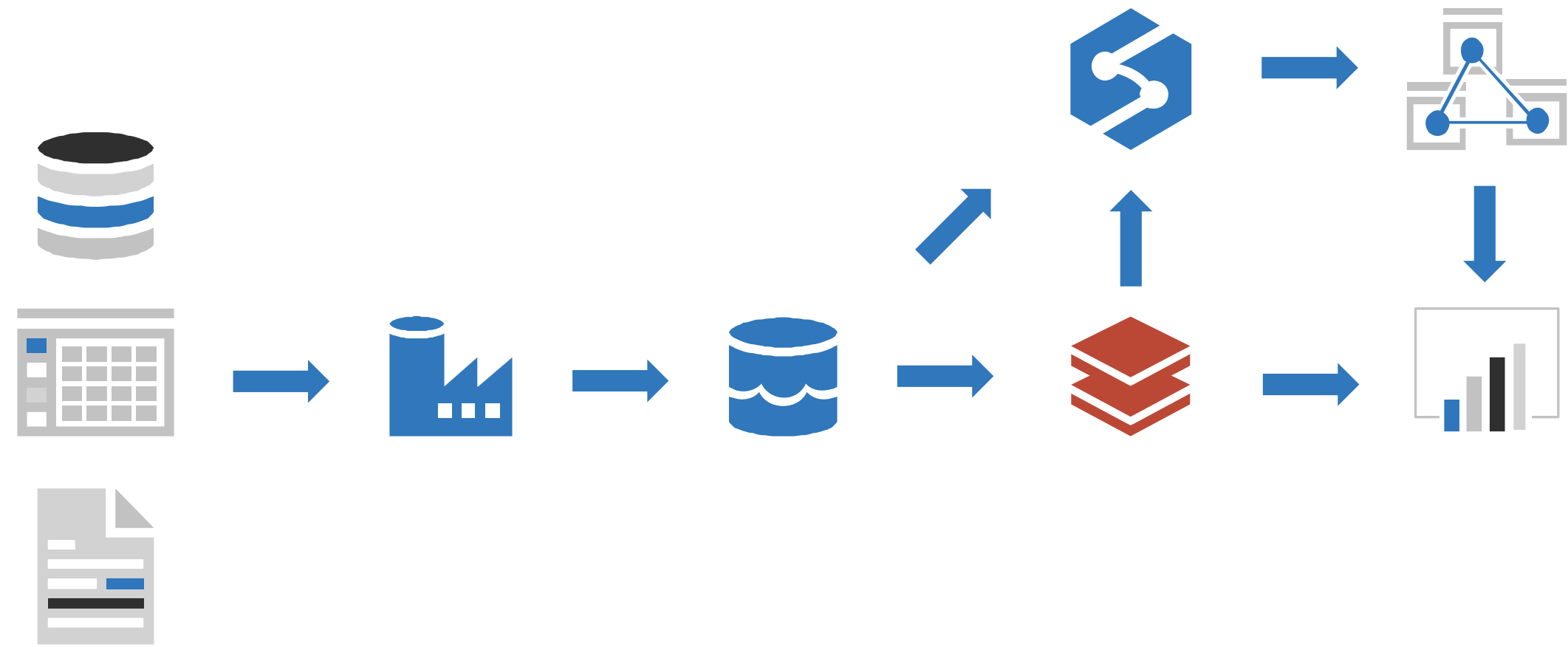


Explore modern data warehousing architecture and workload



Explore Azure data services in the Azure portal

# Modern data warehouse components



# What is modern data warehousing?

## Ingest & Prep



### Azure Data Factory

Code-free data transformation and ingestion from 90+ data integration connectors



### Azure Databricks (Data prep)

Up to 10x faster than vanilla Spark

## Model & Serve



### Azure Synapse Analytics (Data Warehouse)

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

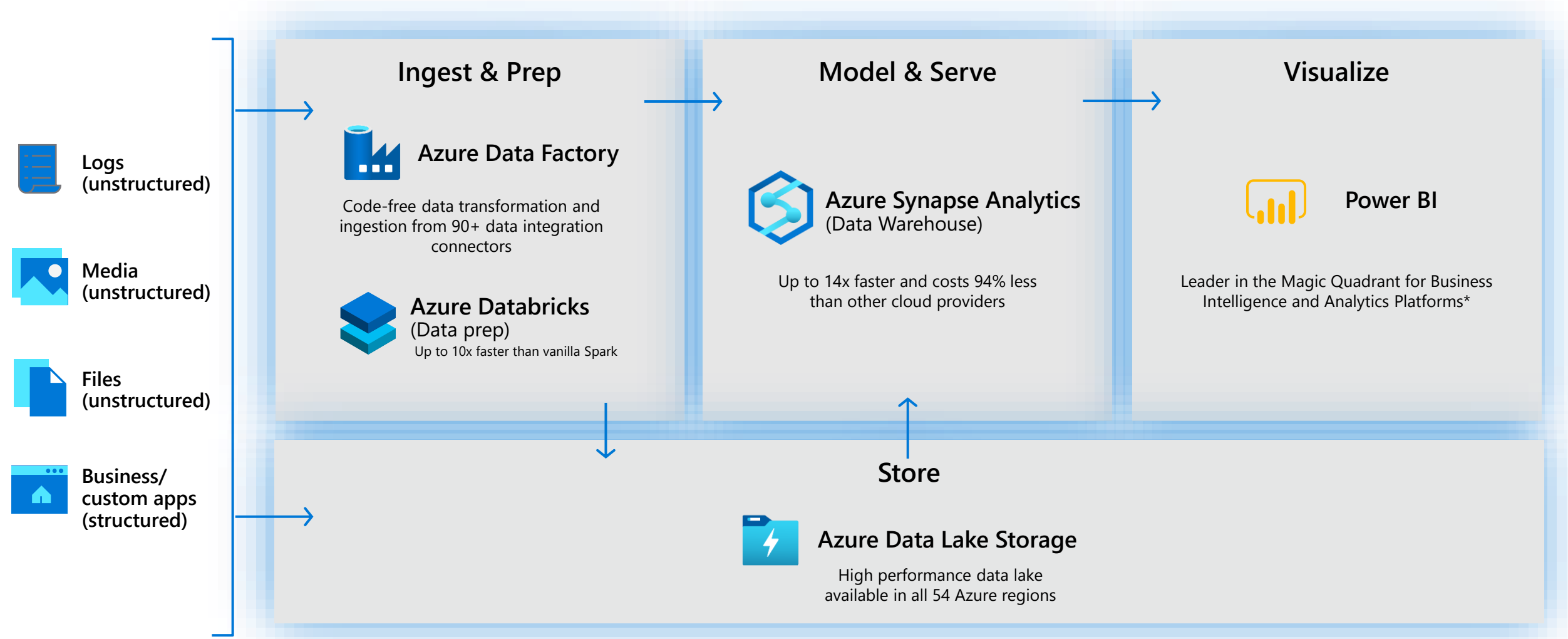


### Azure Data Lake Storage

High performance data lake available in all 54 Azure regions



# Combine batch and stream processing



# Explore Azure data services for modern data warehousing

## What is Azure Data Factory

*A cloud-based data integration service that allows you to orchestrate and automate data movement and data transformation.*

# What is Azure Data Lake Storage?

- A repository of data for your Modern Data Warehouse
- Organises data into directories for improved file access
- Supports POSIX and RBAC permissions
- It is compatible with Hadoop Distributed File System

## Store



### Azure Data Lake Storage

High performance data lake  
available in all 54 Azure regions

# What is Azure Databricks?



## Apache Spark-based platform

Simplifies the provisioning and collaboration of Apache Spark-based analytical solutions



## Enterprise Security

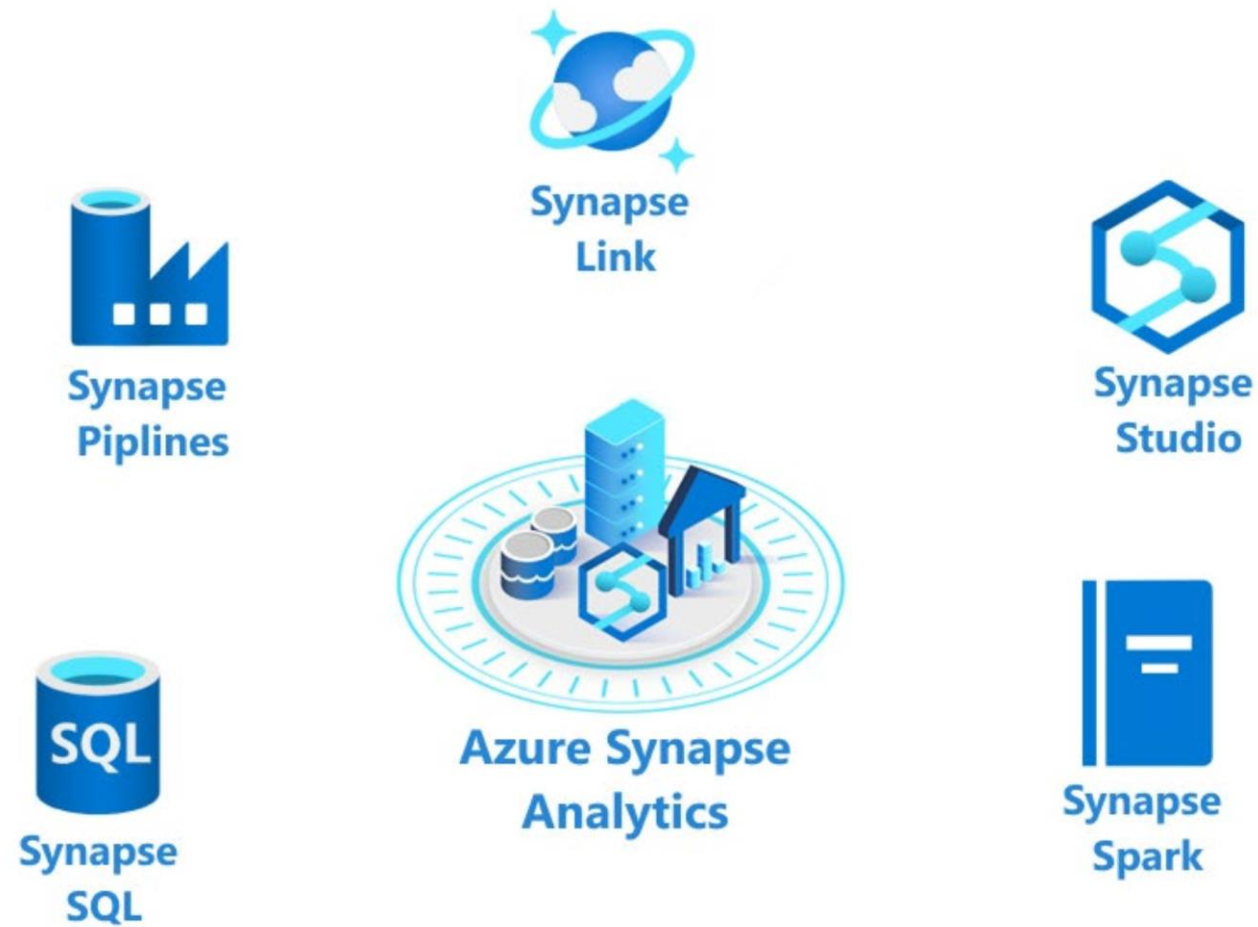
Utilizes the security capabilities of Azure.



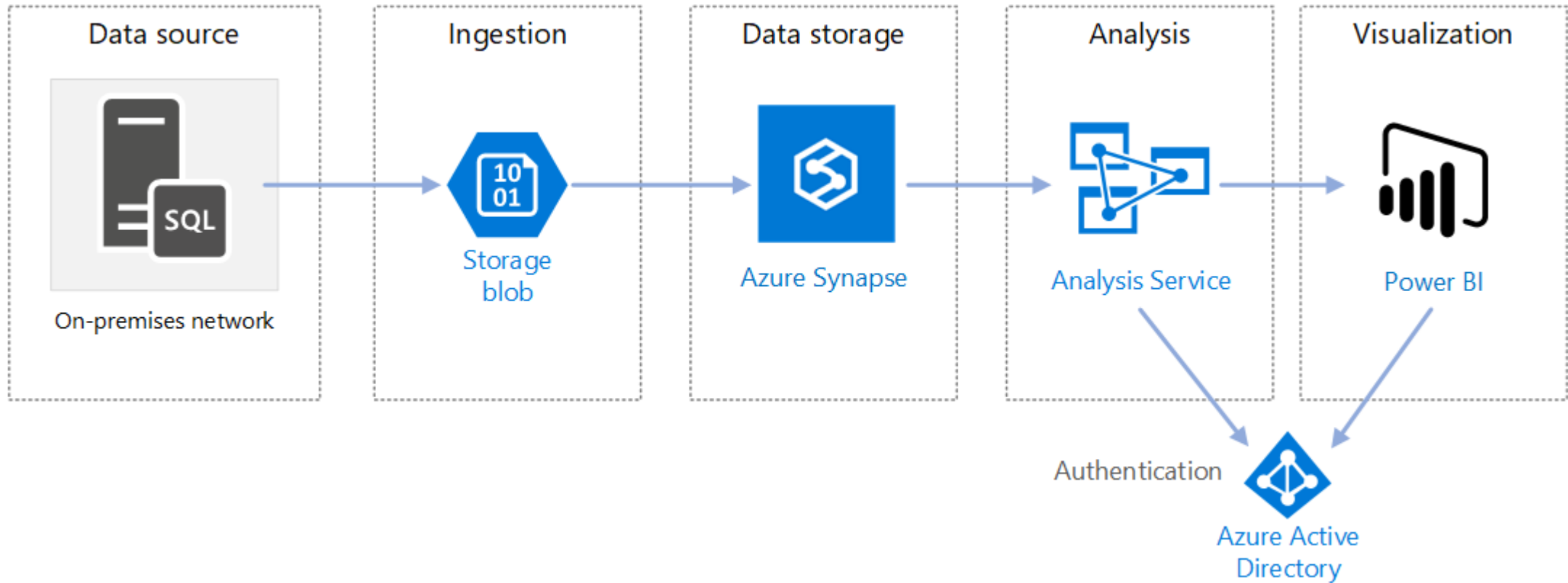
## Integration with Azure services

Can integrate with a variety of Azure data platform services and Power BI

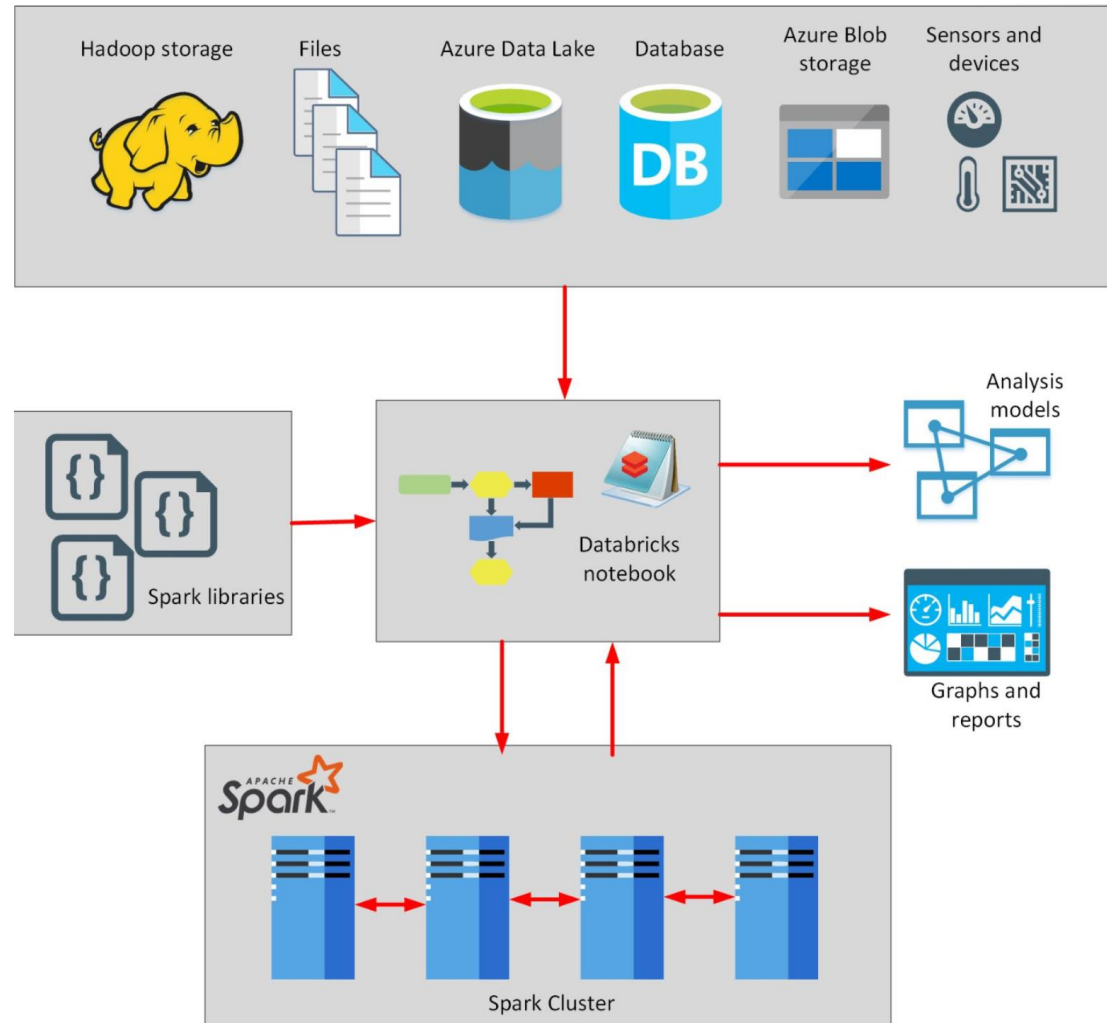
# What is Azure Synapse Analytics?



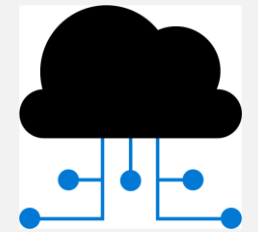
# What is Azure Analysis Services?



# What is Azure HDInsight?

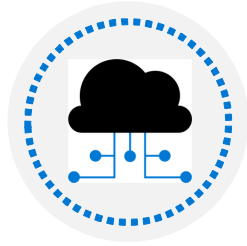


## Lesson 2: Explore data ingestion in Azure





## Lesson 2 objectives



Describe data ingestion in Azure



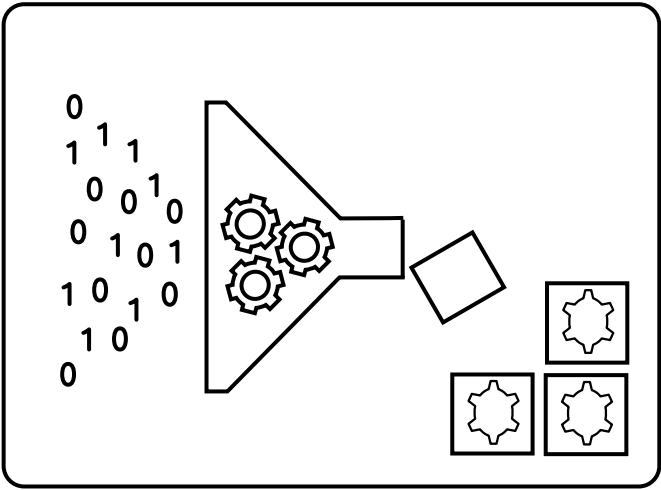
Describe components of Azure Data Factory



See how to use Azure Data Factory to load data into a data warehouse

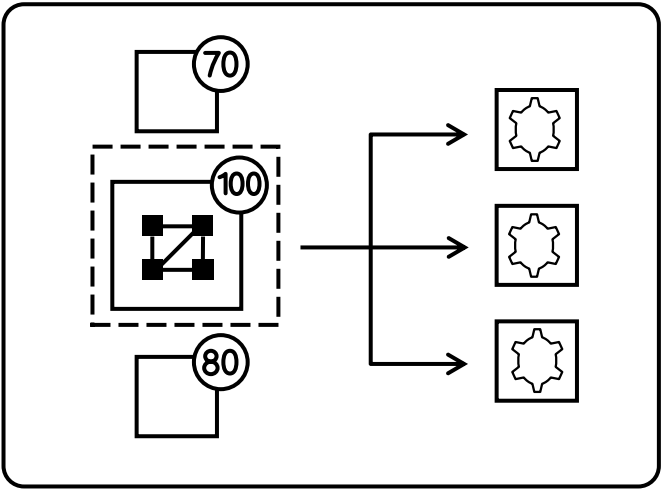
# Describe data ingestion in Azure

## ADF



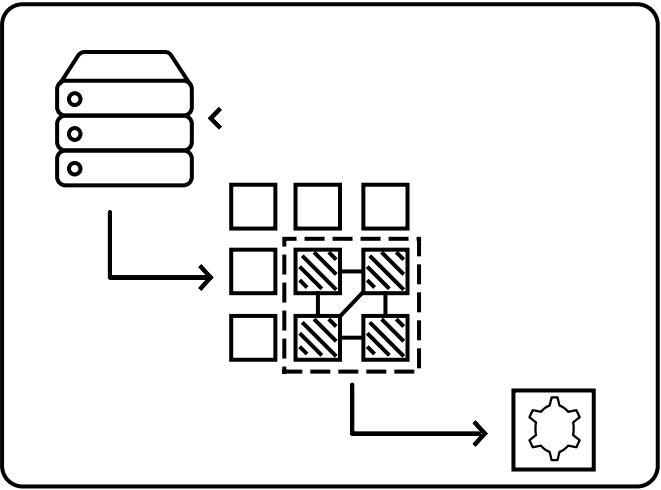
Heterogenous

## PolyBase



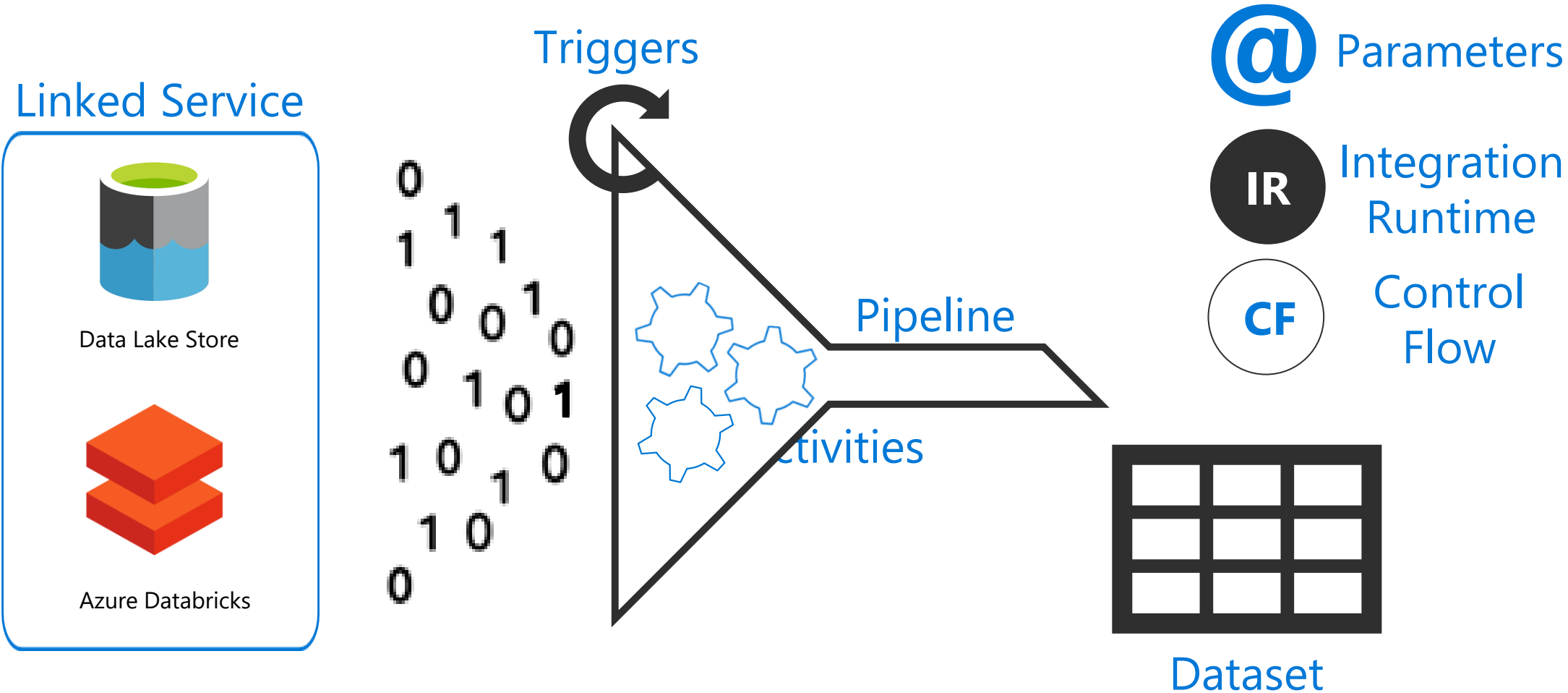
File based

## SSIS



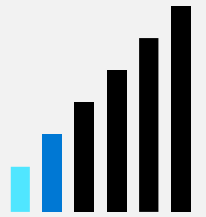
Heterogenous

# Describe components of Azure Data Factory



# Demo: Load data into Azure Synapse Analytics

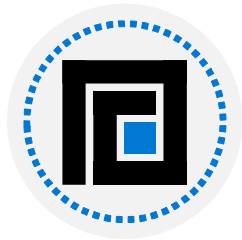
## Lesson 3: Explore data storage and processing in Azure



## Lesson 3 objectives



Describe data processing options for performing analytics in Azure

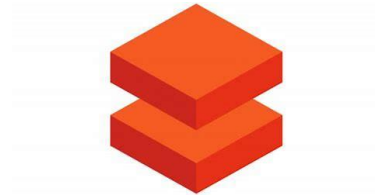


Explore Azure Synapse Analytics

# Data processing options for performing analytics in Azure



Azure Synapse  
Analytics



Azure Databricks



Azure HDInsight



Azure Data Factory



Data Lake Store

# Explore Azure Synapse Analytics



Synapse  
Pipelines



Synapse  
Link



Synapse  
Studio



Azure Synapse  
Analytics



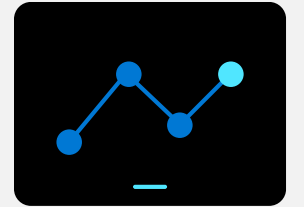
Synapse  
SQL



Synapse  
Spark



## Lesson 4: Get started building with Power BI



## Lesson 4 objectives



Learn how Power BI services and applications work together



Explore how Power BI can make your business more efficient

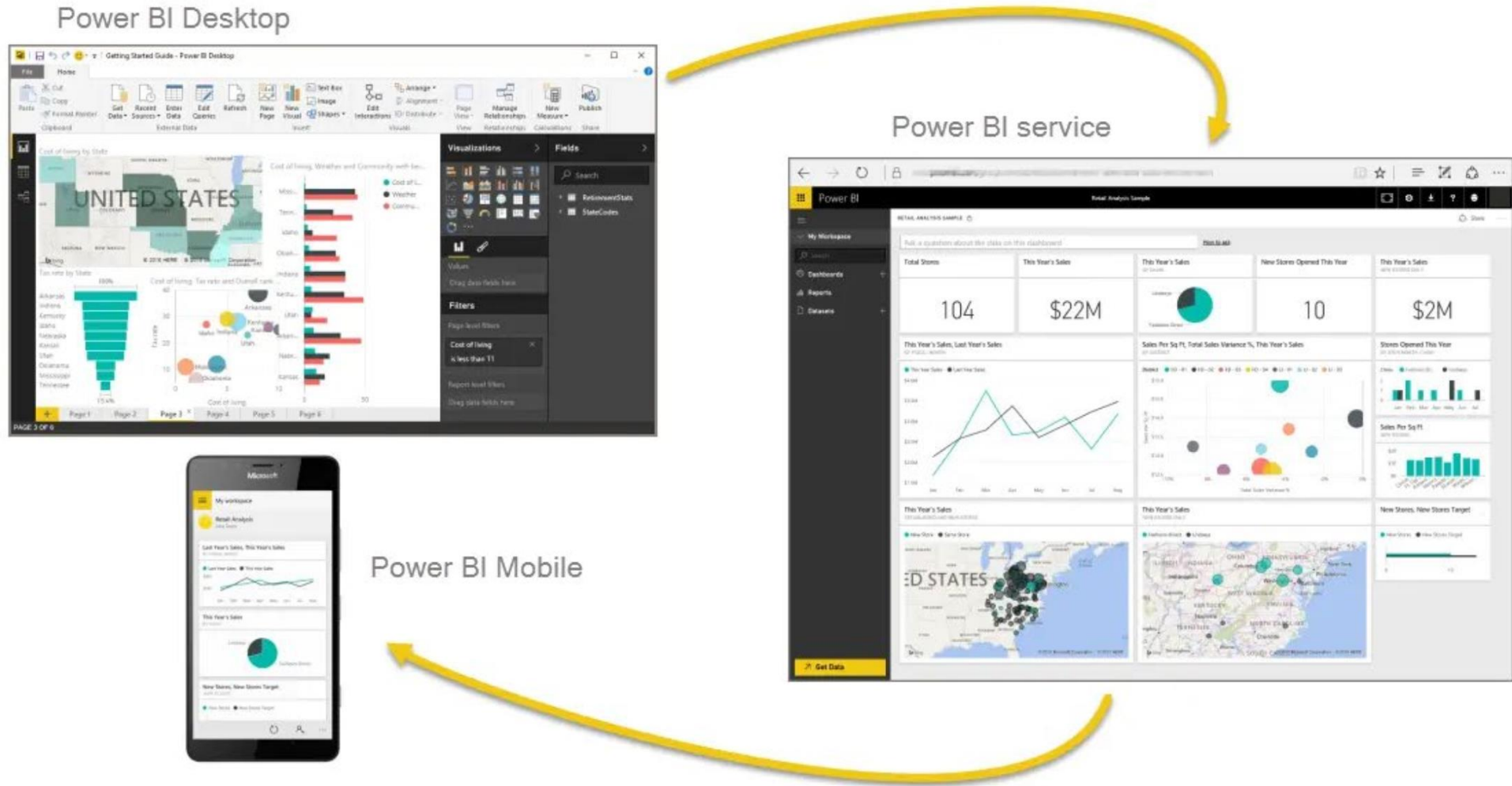


Learn how to create compelling visuals and reports.

# Learn how Power BI services and applications work together



# Explore how Power BI can make your business more efficient



# Learn how to create compelling visuals and reports.

