

© 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 1: Explore core data concepts

Agenda



Explore core data concepts



Explore roles and responsibilities in the world of data



Describe concepts of relational data

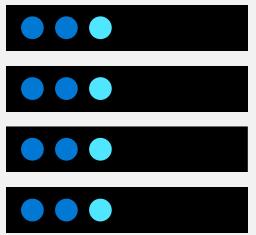


Explore concepts of non-relational data

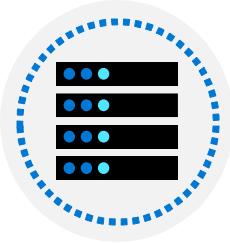


Explore concepts of data analytics

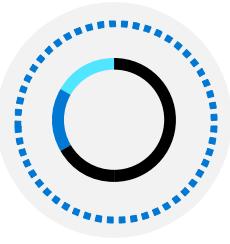
Lesson 1: Explore core data concepts



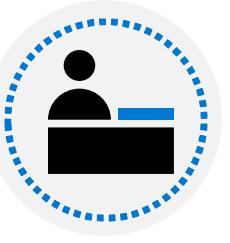
Lesson 1 objectives



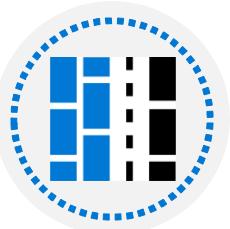
Identify how data is defined and stored



Identify characteristics of relational and non-relational data



Describe and differentiate data workloads

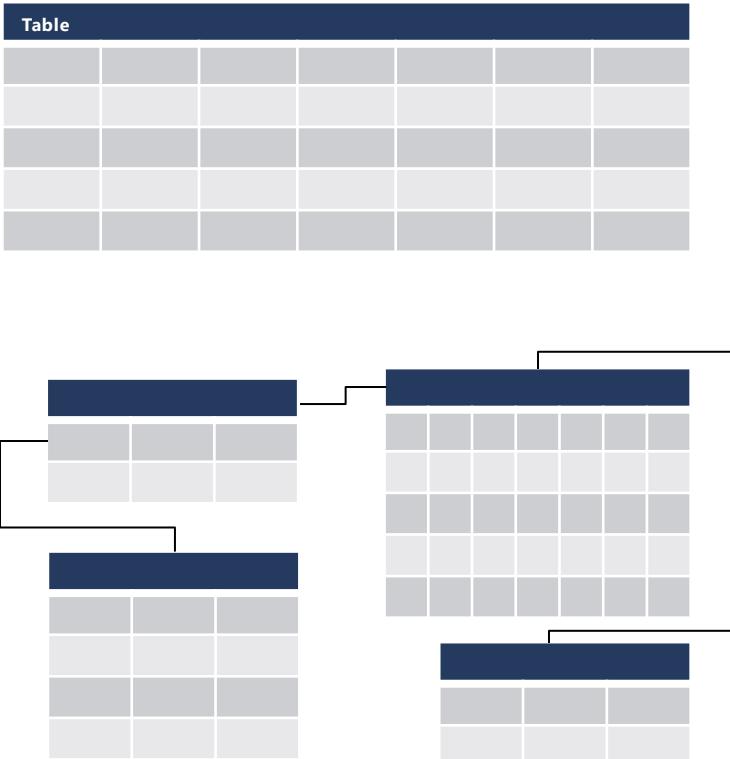


Describe and differentiate batch and streaming data

What is data?

Collection of facts, numbers, descriptions, objects , stored in a structured, semi-structured, unstructured way.

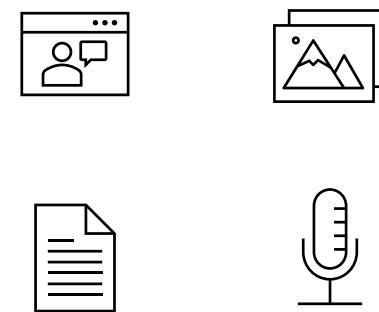
Structured



Semi-structured

```
## Document 1 ## {  
"customerID": "103248",  
"name": { "first": "AAA",  
"last": "BBB" }, "address": {  
"street": "Main Street",  
"number": "101", "city":  
"Acity", "state": "NY" },  
"ccOnFile": "yes",  
"firstOrder": "02/28/2003" }  
## Document 2 ## {  
"customerID": "103249",  
"name": { "title": "Mr",  
"forename": "AAA",  
"lastname": "BBB" },  
"address": { "street":  
"Another Street", "number":  
"202", "city": "Bcity",  
"county": "Gloucestershire",  
"country-region": "UK" },  
"ccOnFile": "yes" }
```

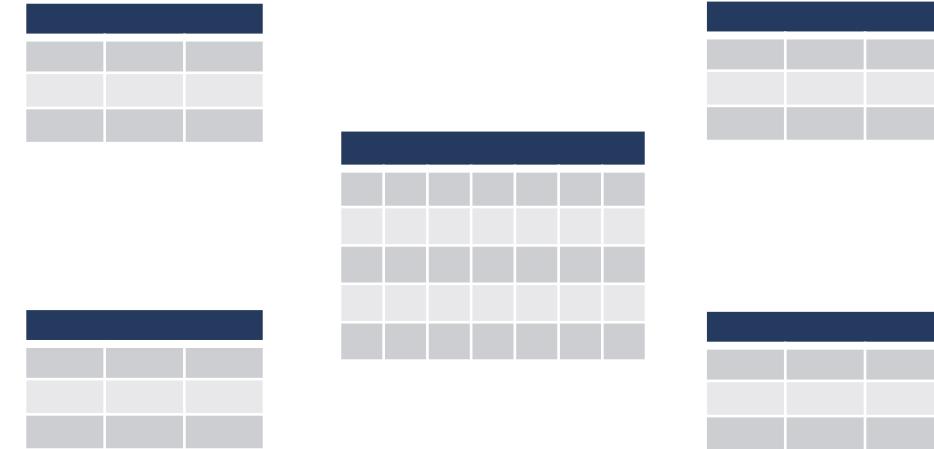
Unstructured



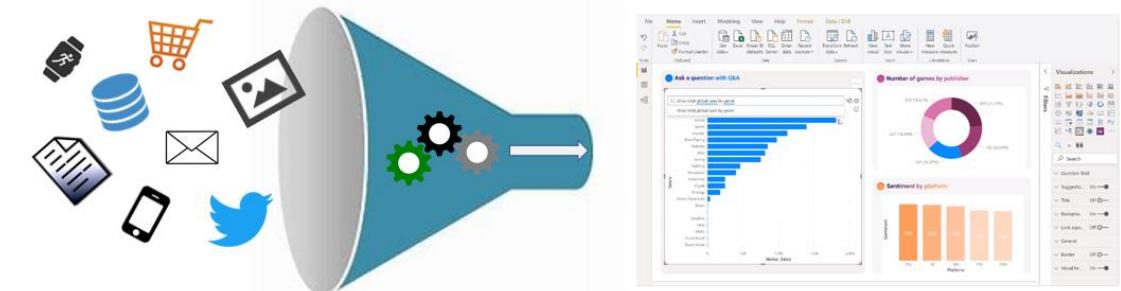
Transactional vs analytical data stores

Customer		
CustomerId	CustomerName	CustomerPhone
Orders		
OrderId	CustomerId	OrderDate

Online Transactional Processing (OLTP)



Online Analytical Processing (OLAP)



Transactional workloads

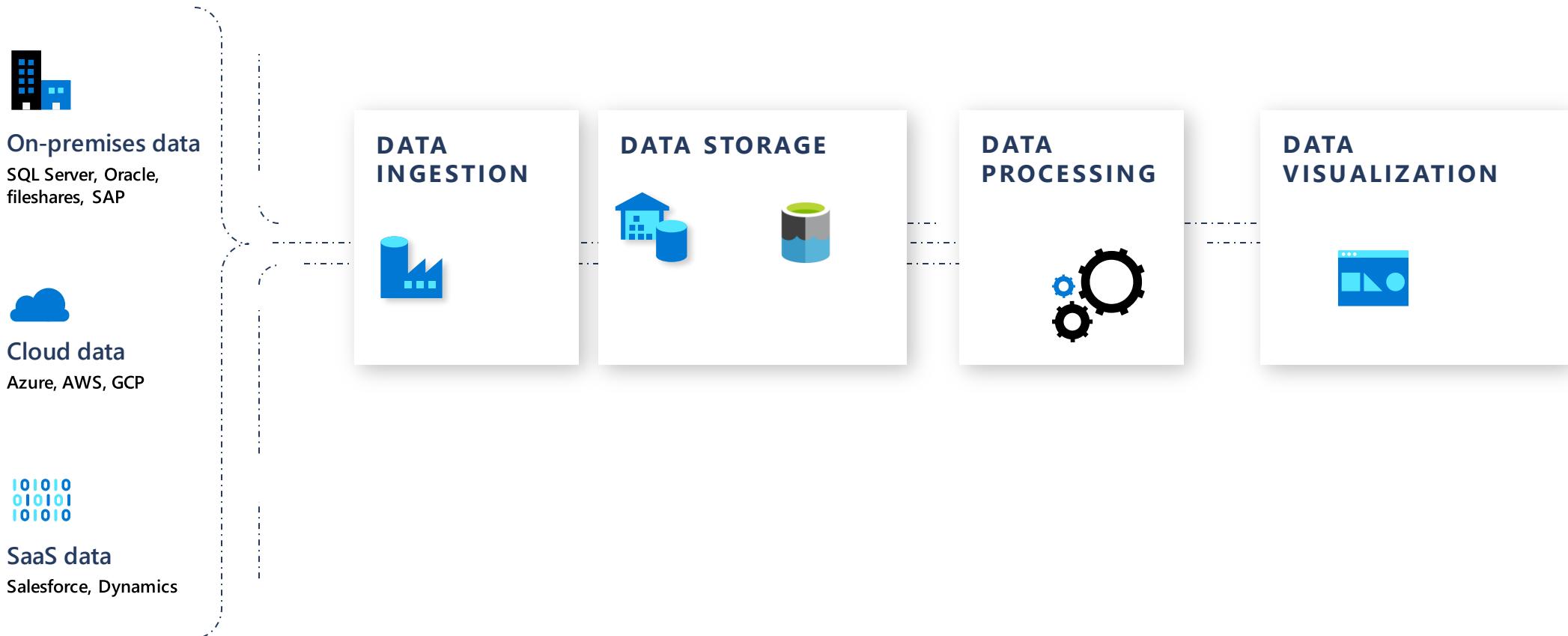
Customer		
CustomerID	CustomerName	CustomerPhone
Orders		
OrderID	CustomerID	OrderDate

Account	
CustomerID	Balance
5558	1000
6023	1500

Transfers					
TransactionID	FromAccount	ToAccount	TransactionAmount	OrderDate	TransactionDescription
982801	6023	5558	500	DD/MM/YY	Transfer 500 from account 6023 to account 5558

```
BEGIN TRANSACTION
UPDATE Account
SET Balance = Balance -500
WHERE CustomerID=6023;
UPDATE Account
SET Balance = Balance +500
WHERE CustomerID=5558;
INSERT INTO Transfers (Fromaccount,ToAccount,TransactionAmount,TransactionDescription)
VALUES (6023,5558,500,'Transfer 500 from account 6023 to account 5558')
COMMIT TRANSACTION
```

Analytical System



Batch Data / Streaming Data

BATCH

🕒 101010 101010 101010
010101 010101 010101
101010 101010 101010

🕒 101010 101010 101010
010101 010101 010101
101010 101010 101010

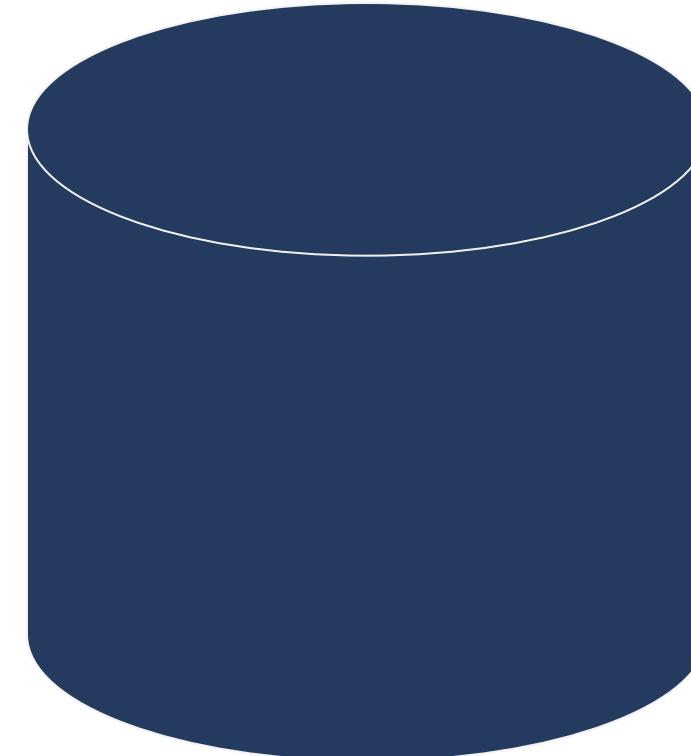
🕒 101010 101010 101010
010101 010101 010101
101010 101010 101010

STREAMING

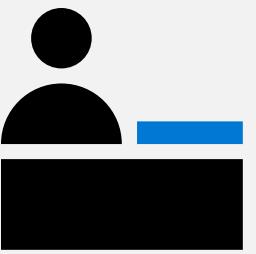
101010 101010 101010
010101 010101 010101
101010 101010 101010

101010 101010 101010
010101 010101 010101
101010 101010 101010

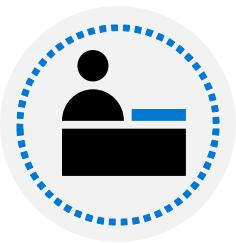
101010 101010 101010
010101 010101 010101
101010 101010 101010



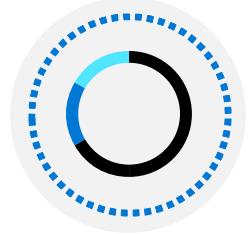
Lesson 2: Explore roles and responsibilities in the world of data



Lesson 2 objectives



Explore data job roles



Explore common tasks and tools for data job
roles

Roles in Data



Database Administrator

Database Management
Implements Data Security
Backups
User Access
Monitors performance



Data Engineer

Data Pipelines and processes
Data Ingestion storage
Prepare data for Analytics
Prepare data for analytical processing



Data Analyst

Provides insights into the data
Visual Reporting
Modeling Data for Analysis
Combines data for visualization and analysis

Common Tools – Database Administrator

Azure Data Studio

- Graphical interface for managing on-premises and cloud-based data services
- Runs on Windows, macOS, Linux

SQL Server Management Studio

- Graphical interface for managing on-premises and cloud-based data services
- Runs on Windows
- Comprehensive Database Administration tool

Azure Portal / CLI

- Tools for management and provisioning of Azure Data Services
- Manual and automation of scripts using Azure Resource Manager or Command Line Interface scripting

Common Tools – Data Engineering

Azure Synapse Studio

- Azure Portal integrated to manage Azure Synapse
- Data Ingestion (Azure Data Factory)
- Management of Azure Synapse assets (SQL Pools / Spark Pool)

SQL Server Management Studio

- Graphical interface for managing on-premises and cloud-based data services
- Runs on Windows
- Comprehensive Database Administration tool

Azure Portal / CLI

- Tools for management and provisioning of Azure resources
- Manual and automation of scripts using Azure Resource Manager or Command Line Interface scripting

Common Tools – Data Analyst

Power BI Desktop

- Data Visualization tool
- Model and Visualize Data
- Management of Azure Synapse assets (SQL Pools / Spark Pool)

Power BI Portal / Power BI Service

- Authoring and management of Power BI reports
- Authoring of Power BI dashboards
- Share Reports / Datasets

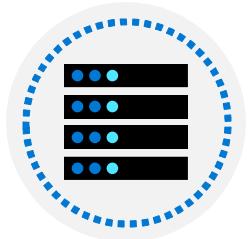
Power BI Report Builder

- Data Visualization tool for paginated reports
- Model and Visualize paginated reports

Lesson 3: Describe concepts of relational data



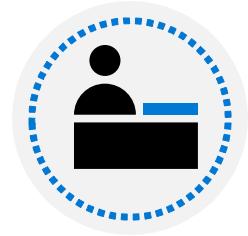
Lesson 3 objectives



Explore the characteristics of relational data



Define tables, indexes, and views



Explore relational data workload offerings in Azure

Tables

Customers		
CustomerID	CustomerName	CustomerPhone
100	Muisto Linna	XXX-XXX-XXXX
101	Noam Maoz	XXX-XXX-XXXX
102	Vanja Matkovic	XXX-XXX-XXXX
103	Qamar Mounir	XXX-XXX-XXXX
104	Zhenis Omar	XXX-XXX-XXXX
105	Claude Paulet	XXX-XXX-XXXX
106	Alex Pettersen	XXX-XXX-XXXX
107	Francis Ribeiro	XXX-XXX-XXXX

Data is stored in a table

Table consists of rows and columns

All rows have same # of columns

Each column is defined by a datatype

Normalization

Customers		
CustomerID	CustomerName	CustomerPhone
100	Muisto Linna	XXX-XXX-XXXX
101	Noam Maoz	XXX-XXX-XXXX
102	Vanja Matkovic	XXX-XXX-XXXX
103	Qamar Mounir	XXX-XXX-XXXX
104	Zhenis Omar	XXX-XXX-XXXX
105	Claude Paulet	XXX-XXX-XXXX
106	Alex Pettersen	XXX-XXX-XXXX

Data is normalized to:

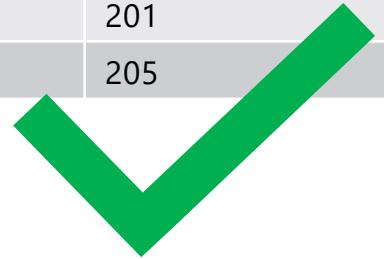
- Reduce storage
- Avoid data duplication
- Improve data quality

Orders		
OrderID	CustomerName	CustomerPhone
AD100	Noam Maoz	XXX-XXX-XXXX
AD101	Noam Maoz	XXX-XXX-XXXX
AD102	Noam Maoz	XXX-XXX-XXXX
AX103	Qamar Mounir	XXX-XXX-XXXX
AS104	Qamar Mounir	XXX-XXX-XXXX
AR105	Claude Paulet	XXX-XXX-XXXX
MK106	Muisto Linna	XXX-XXX-XXXX

Relations

Customers		
CustomerID	CustomerName	CustomerPhone
100	Muisto Linna	XXX-XXX-XXXX
101	Noam Maoz	XXX-XXX-XXXX
102	Vanja Matkovic	XXX-XXX-XXXX
103	Qamar Mounir	XXX-XXX-XXXX
104	Zhenis Omar	XXX-XXX-XXXX
105	Claude Paulet	XXX-XXX-XXXX
106	Alex Pettersen	XXX-XXX-XXXX

Orders		
OrderID	CustomerID	SalesPersonID
AD100	101	200
AD101	101	200
AD102	101	200
AX103	103	201
AS104	103	201
AR105	105	200
MK106	105	201
DB205	100	205



In a normalized database schema:

- Primary Keys and Foreign keys are used to define relationships
- No data duplication exists (other than key values in 3rd Normal Form (3NF)
- Data is retrieved by joining tables together in a query

Indexes

Customers		
CustomerID	CustomerName	CustomerPhone
100	Muisto Linna	XXX-XXX-XXXX
101	Noam Maoz	XXX-XXX-XXXX
102	Vanja Matkovic	XXX-XXX-XXXX
103	Qamar Mounir	XXX-XXX-XXXX
104	Zhenis Omar	XXX-XXX-XXXX
105	Claude Paulet	XXX-XXX-XXXX
106	Alex Pettersen	XXX-XXX-XXXX

IDX-CustomerRegion	
CustomerID	Region
100	France
101	Brazil
102	Croatia
103	Jordan
104	Spain
105	France
106	USA

An index

- Optimizes search queries for faster data retrieval
- Reduces the amount of data pages that need to be read to retrieve the data in a SQL Statement
- Data is retrieved by joining tables together in a query

View

Customers		
CustomerID	CustomerName	CustomerPhone
100	Muisto Linna	XXX-XXX-XXXX
101	Noam Maoz	XXX-XXX-XXXX
102	Vanja Matkovic	XXX-XXX-XXXX
103	Qamar Mounir	XXX-XXX-XXXX
104	Zhenis Omar	XXX-XXX-XXXX
105	Claude Paulet	XXX-XXX-XXXX
106	Alex Pettersen	XXX-XXX-XXXX

A view is a virtual table based on the result set of query

- Views are created to simplify the query
- Combine relational data into a single pane view

Orders		
OrderID	CustomerID	SalesPersonID
AD100	101	200
AD101	101	200
AD102	101	200
AX103	103	201
AS104	103	201
AR105		
MK106		
DB205		

Create the definition of a view:

```
CREATE VIEW vw_customerorders AS  
SELECT Customers.CustomerID,  
Customers.CustomerName, Orders.OrderID FROM  
Customers JOIN Orders on Customers.CustomerID  
= Orders.CustomerID
```

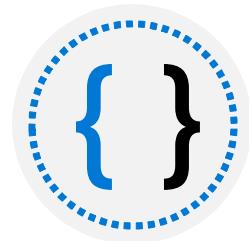
Retrieve the orders placed by customer 102 using the view:

```
SELECT CustomerName, OrderID from  
vw_customerorders WHERE CustomerID=102
```

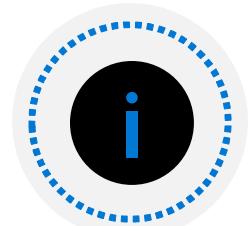
Lesson 4: Explore concepts of non-relational data

{ }

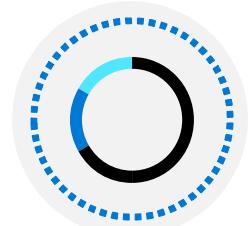
Lesson 4 objectives



Explore the characteristics of non-relational data



Define types of non-relational data



Describe NoSQL, and the types of non-relational databases

Explore characteristics of non-relational data

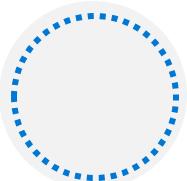
Entities

```
## Customer 1 ID: 1
Name: Mark Hanson
Telephone: [ Home: 1-999-9999999, Business: 1-888-8888888, Cell: 1-777- 7777777 ]
Address: [ Home: 121 Main Street, Some City, NY, 10110,
           Business: 87 Big Building, Some City, NY, 10111 ]
## Customer 2 ID: 2
Title: Mr
Name: Jeff Hay
Telephone: [ Home: 0044-1999-333333, Mobile: 0044-17545-444444 ]
Address: [ UK: 86 High Street, Some Town, A County, GL8888, UK,
           US: 777 7th Street, Another City, CA, 90111 ]
```

Non-relational collections can have

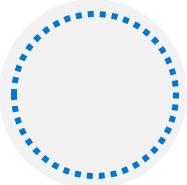
- Multiple entities in the same collection or container with different fields
- Have a different, non-tabular schema
- are often defined by labeling each field with the name it represents

Identify non-relational database use cases



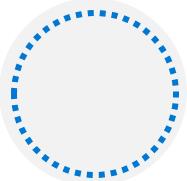
IoT and Telematics

Often require to ingest large amounts of data in frequent burst of activity, data is either semi structured or structured, often requires real time processing



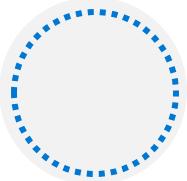
Retail and Marketing

Common scenarios for globally distributed data, document storage



Gaming

In-game stats, social media integration, leaderboards, low-latency applications



Web and Mobile

Common used with web click analytics, modern applications including bots

Types of non-relational data

What is semi-structured data?

Data structure is defined within the actual data by fields.
Format / file types include:

JSON

AVRO

ORC

Parquet

What is unstructured data?

- Does not naturally contain fields
Examples: video, audio, media streams, documents
- Often used to extract data form and categorize or identify “structures”
- Frequently used in combination with Machine Learning or Cognitive Services capabilities to “extract data” by using:
 - Text Analytics
 - Sentiment Analysis with Cognitive APIs
 - Vision API

What is NoSQL?

Loose term, to describe non-relational

Key-value
stores

Document
based

Column
family
databases

Graph
Databases

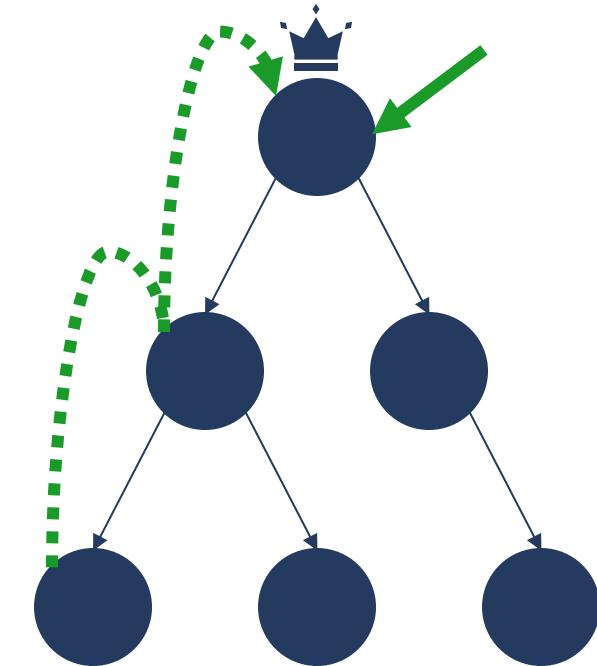
What is a graph database?

- Stores entities centric around relationships
- Enable applications to perform queries traversing a network of nodes and edges

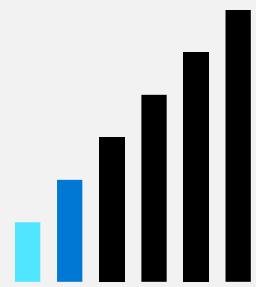
What applications require a graph database?

Business requirements:

- OLTP apps with highly **correlated** data.
- Easy **updates** to single or many objects.
- **Flexible data modelling.**
- Data requirements that **evolve**.
- **Hierarchical data structures.**



Lesson 5: Explore concepts of data analytics



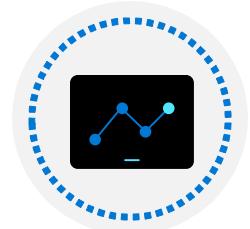
Lesson 5 objectives



Learn about data ingestion and processing

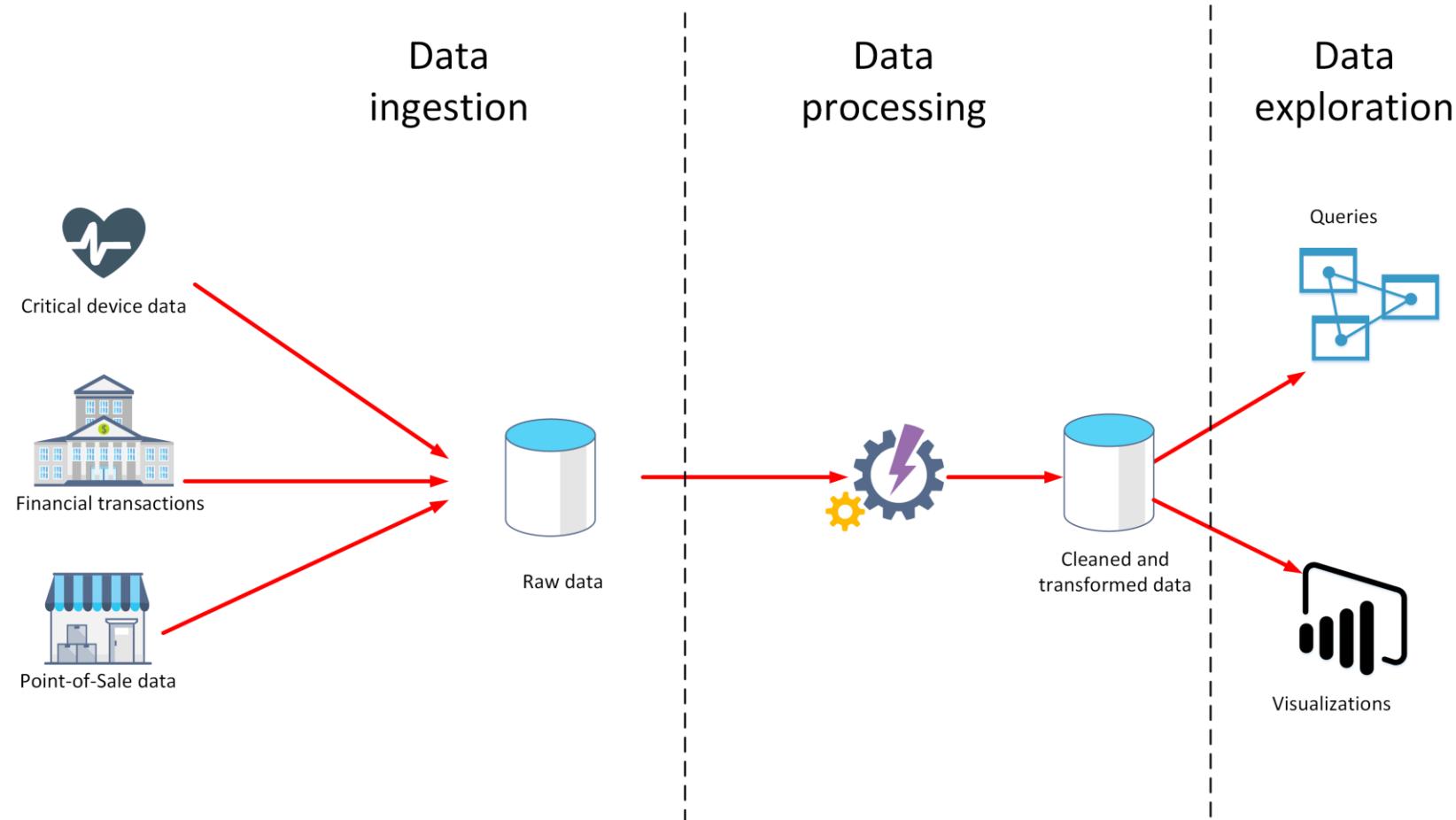


Explore data visualization

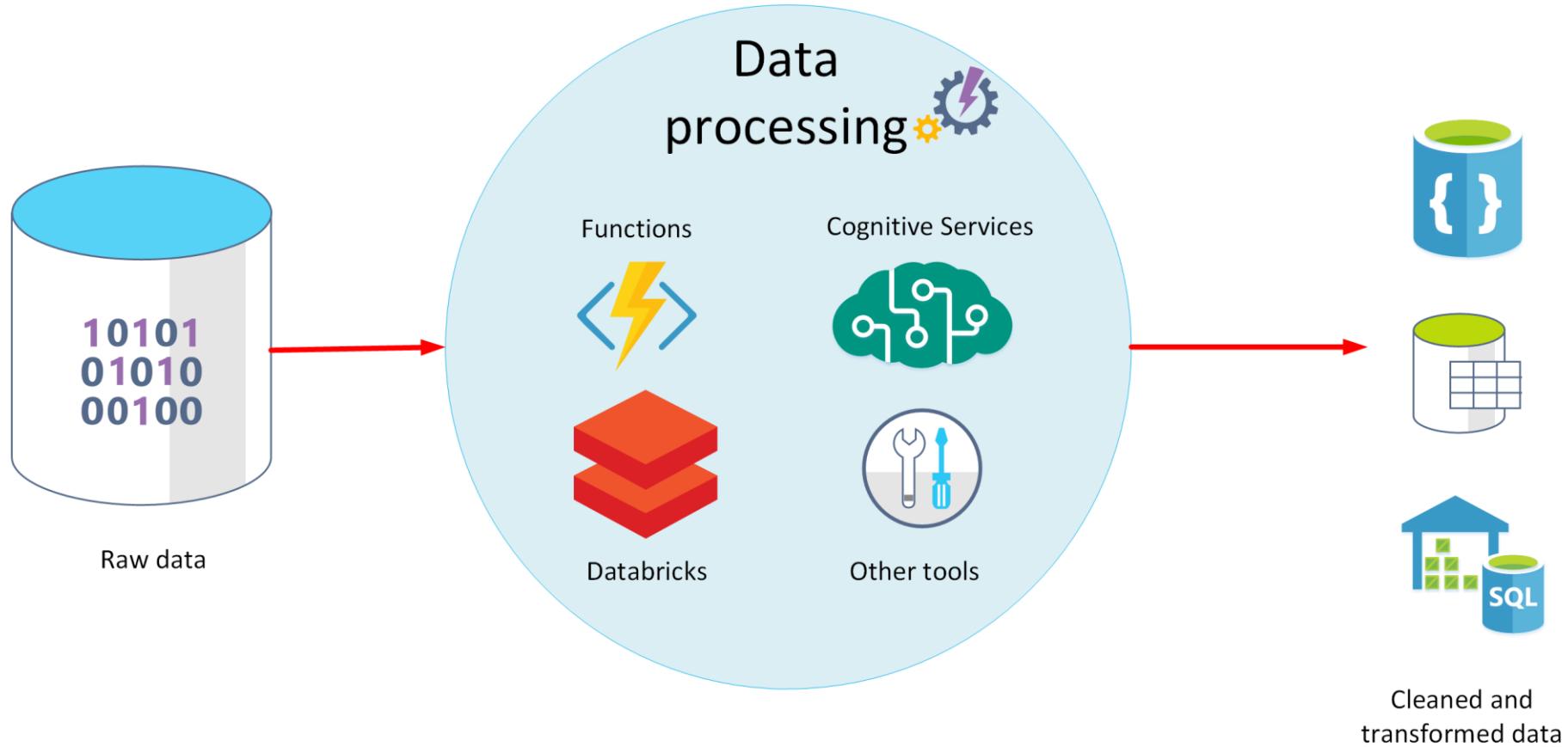


Explore data analytics

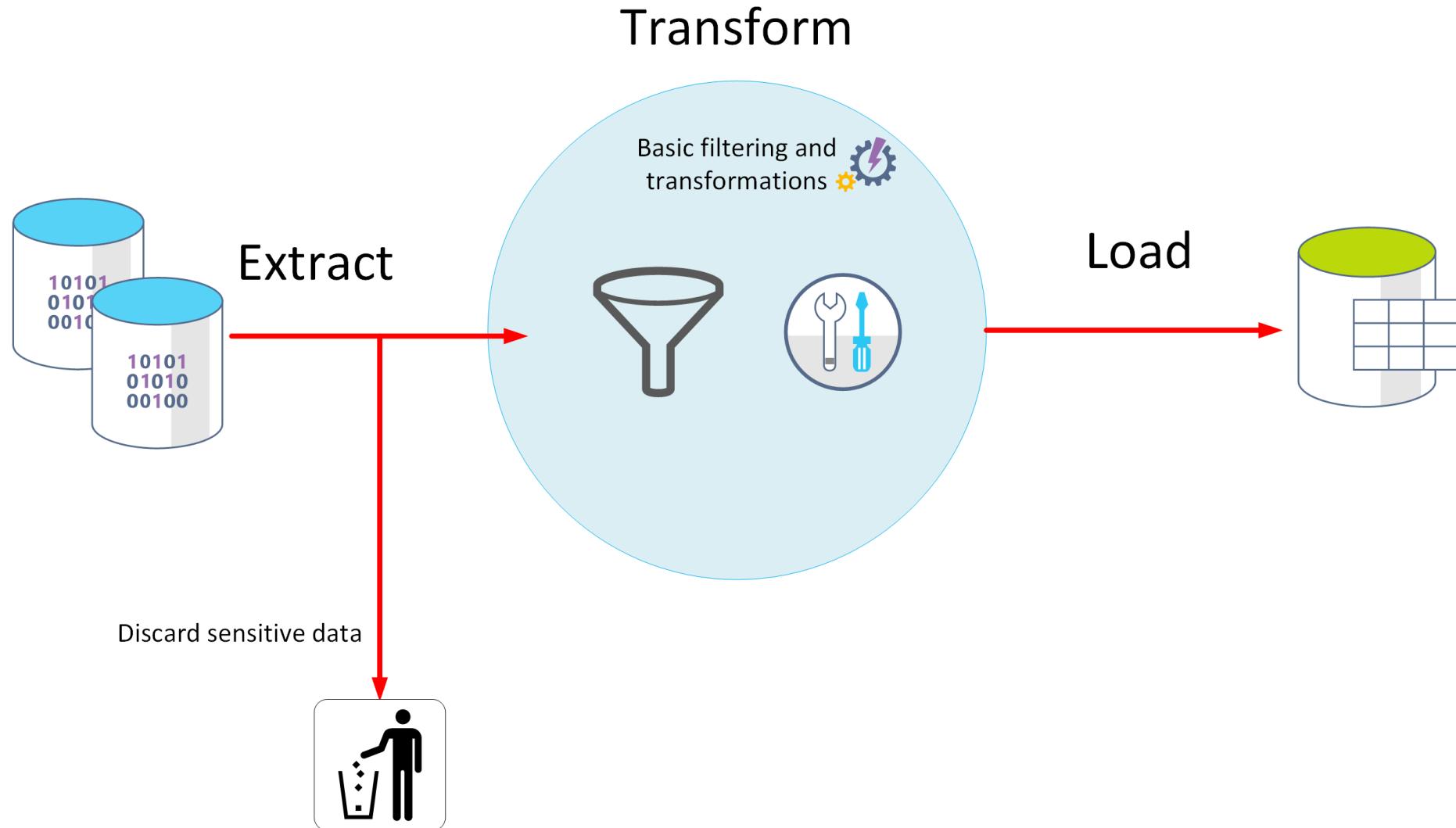
What is data ingestion?



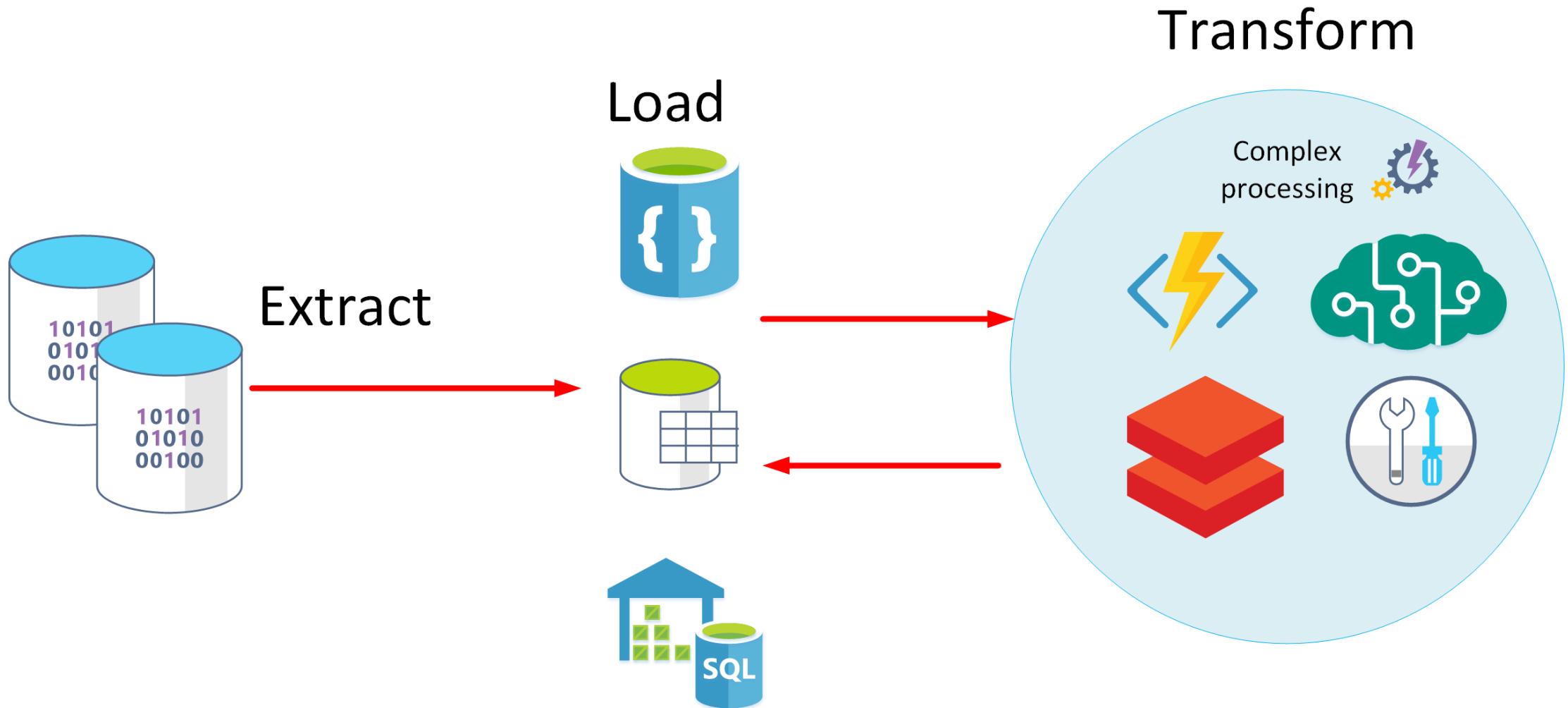
What is data processing?



What is ETL?



What is ELT?



Explore Data Visualization

Power BI: A collection of software, services, apps, and connectors.



Explore Data Analytics

Descriptive



Diagnostic



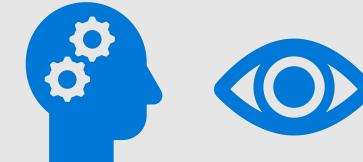
Predictive



Prescriptive



Cognitive





© 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 2: Explore relational data in Azure

Agenda



Explore relational data offerings in Azure



Explore provisioning and deploying relational database offerings in Azure



Query relational data in Azure

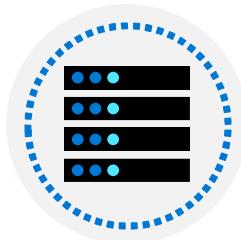
Lesson 1: Explore relational data offerings in Azure



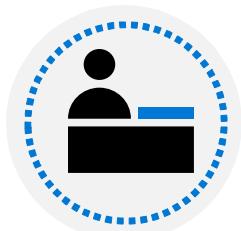
Lesson 1 objectives



Explore relational data offerings in Azure



Explore provisioning and deploying of relational database offerings in Azure



Query relational data in Azure

What are Azure Data Services?

SQL Server on Azure Virtual Machines



Best for re-hosting and apps requiring OS-level access and control
Automated manageability features and OS-level access

Infrastructure as a Service

Azure SQL Managed Instance



Best for modernizing existing apps
Offers high compatibility with SQL Server and native VNET support

Platform as a Service

Azure SQL Database



Best for building new apps in the cloud
Pre-provisioned or serverless compute and Hyperscale storage to meet demanding workload requirements

SQL Server on Azure virtual machines



Customer challenge

I want to migrate to the cloud as fast as possible but maintain operating system control and complete SQL Server functionality



Solution

Get the combined performance, security, and analytics of SQL Server, backed by the flexibility, security, and hybrid connectivity of Azure

Key features

- SQL Server and OS server access
- Expansive SQL and OS versions
- Windows, Linux, Containers
- File stream, DTC, and Simple Recovery model
- SSAS, SSRS, and SSIS

Azure differentiators

- Free Extended Security Updates for SQL Server 2008/R2
- Automated Backups and Security Updates
- Point in Time Restore with Azure Backup
- Accelerated storage performance with Azure Blob Caching
- 435 percent overall return on an Azure IaaS investment over five years¹

SQL Server on Azure VM Deployment choices

Deployment Choices

- Marketplace pre-installed SQL Server on Windows or Linux
- Install your own SQL Server
- Lift and Shift with Azure Migrate (Azure Site Recovery)

Resource Provider

- Unlock Licensing and Edition Flexibility
- Automated Backups and Security Updates
- Manage VMs through Azure SQL in portal

Sizes and Storage Performance

- Memory or Storage optimized sizes for best performance
- Data and log on Premium Storage Managed Disks
- Azure Blob Read Caching for data disks

- Tempdb on local SSD
- Ultra disks for extremely low latency needs

Networking and Security

- Virtual Networks to integrate with on-premises
- Advanced Data Security services (Preview)

HADR

- Azure VM built-in HA
- Azure Storage built-in DR
- Azure Backup and Automated backups to Azure Blob Storage
- File-Snapshot Backups

- Failover Cluster Instance with Azure Premium File Share
- Always On Availability Groups with Cloud Witness
- Hybrid Availability Group Secondary replicas
- HADR on RedHat Linux with Pacemaker and fencing

IaaS vs PaaS



Business continuity



High availability



Automated backups



Long term backup retention



Geo-replication



Scale



Advanced security



Version-less



Built-in monitoring



Built-in intelligence

Azure SQL DB



Customer challenge

I want to build modern apps, potentially multi-tenanted, with the highest uptime and predictable performance



Solution

Azure SQL Database is a highly scalable cloud database service with built-in high availability and machine learning

Key features

- Single database or elastic pool
- Hyperscale storage (100TB+)
- Serverless compute
- Fully managed service
- Private link support
- High availability with AZ isolation

Azure differentiators

Industry highest availability SLA of 99.995%

Industry only business continuity SLA with 5 second RPO and 30 second RTO

Price-performance leader for mission-critical workloads while costing up to 86 percent less than AWS RDS (GigaOm)

Azure SQL DB Service Tiers

General purpose

Most business workloads

Remote storage

IOPS

\$

Serverless*



Business critical

Workloads that require low latency,
fast recovery, and a readable
secondary

Local storage

IOPS++

\$\$\$

In-memory



Hyperscale*

Most business workloads with
highly scalable storage and read-
scale requirements

Local + remote storage

IOPS+

\$\$

Unlimited storage



*Not in managed instance

Azure SQL DB Managed Instance



Customer challenge

I want to migrate to the cloud, remove management overhead, but I need instance-scoped features (Service Broker, SQL Server Agent, CLR...)



Solution

Managed instance combines leading security features with SQL Server compatibility and business model designed for on-premises customers

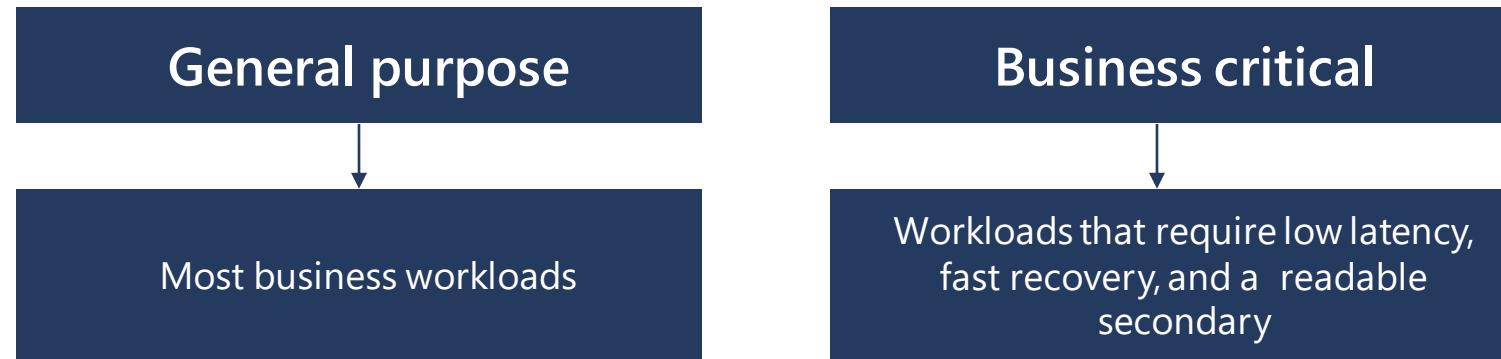
Key features

- Single instance or instance pool
- SQL Server surface area (vast majority)
- Native virtual network support
- Fully managed service
- On-premise identities enabled with Azure AD and AD Connect

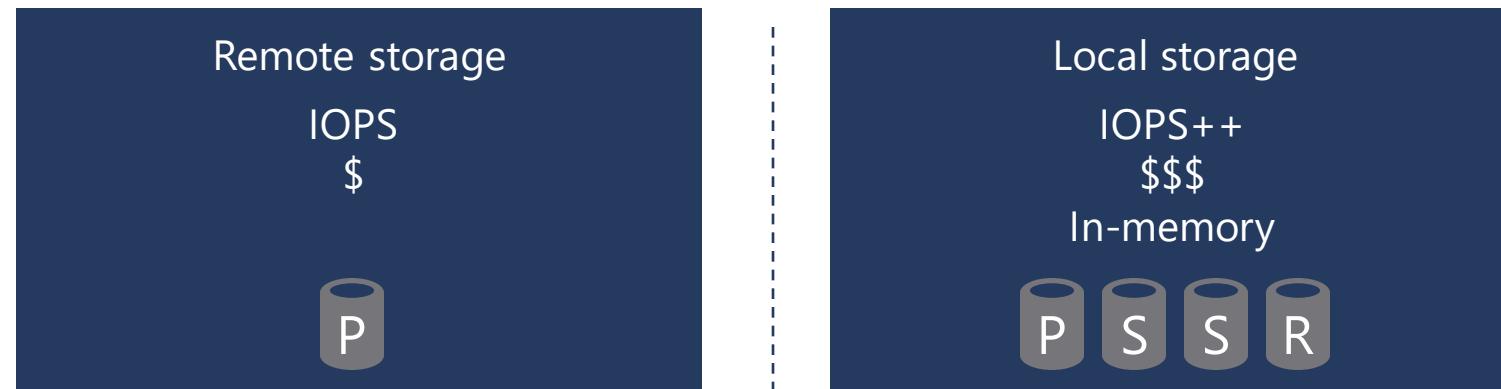
Azure differentiators

- Near zero downtime migration using log shipping
- Fully managed business continuity with failover groups
- Projected return on investment of 212 percent over three years¹
- The best of SQL Server with the benefits of a managed service

Managed Instance Service Tiers



vCore model Independent scalability



Azure SQL Managed instance or DB



Azure SQL managed instance

Single instance

SQL Server surface area (vast majority)

Native virtual network support

Fully managed service

Instance pool

Pre-provision compute resources for migration

Enables cost-efficient migration.

Ability to host smaller instances (2Vcore)

Currently in public preview



Azure SQL Database

Single database

Hyperscale storage (up to 100TB)

Serverless compute

Fully managed service

Elastic pool

Resource sharing between multiple databases to price optimize

Simplified performance management for multiple databases

Fully managed service

PostgreSQL, MariaDB, MySQL



PostgreSQL is the most popular and wanted database for modern apps



MySQL is a leading open source relational database for LAMP stack apps



MariaDB is a community-developed fork of MySQL with strong focus on the user community

Benefits of Azure Database for MySQL, PostgreSQL, MariaDB



Fully managed community database

Take advantage of a fully managed service while still using the tools and languages you're familiar with



Built-in high availability for lowest TCO

Ensure your data is always available without the need for additional costs



Intelligent performance and scale

Improve performance with built-in intelligence and up to 16TB storage and 20K IOPs



Industry-leading security and compliance

Protect your data with enhanced security features including Advanced Threat Protection



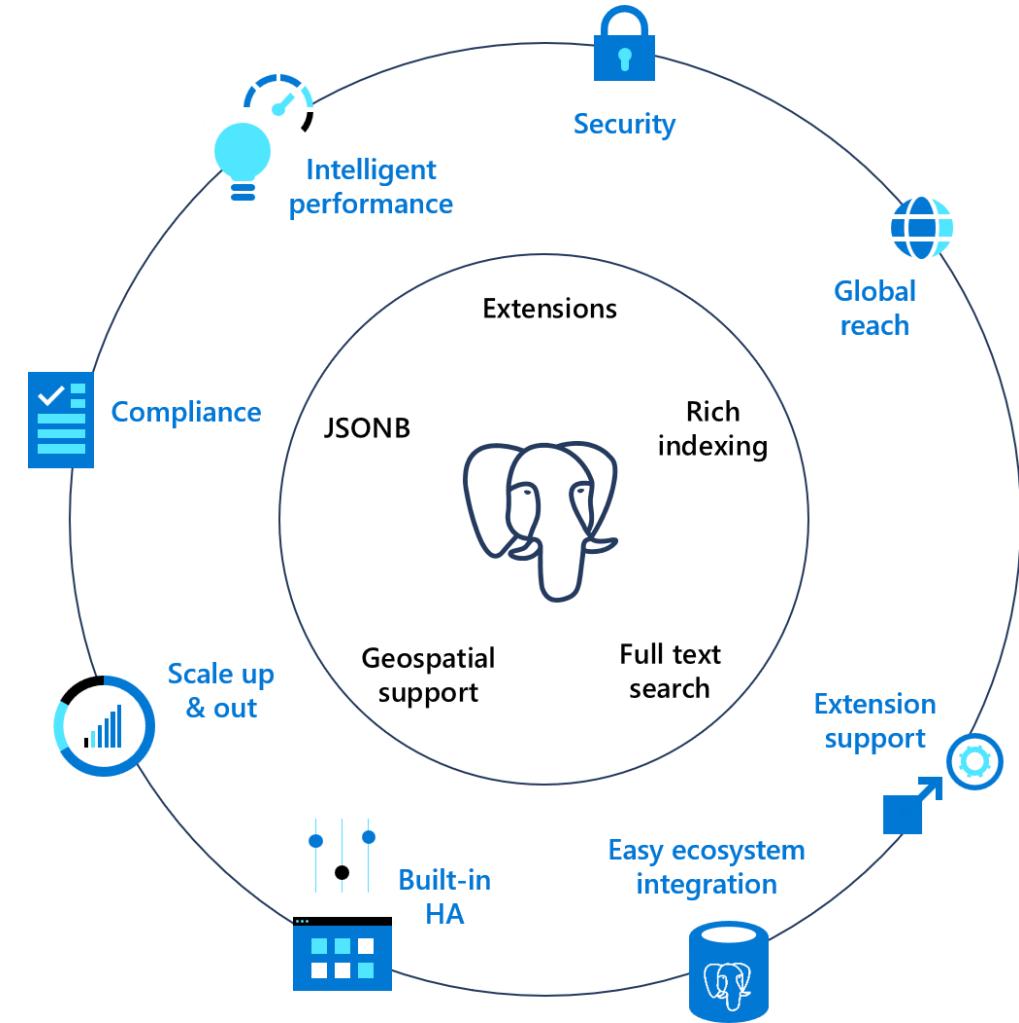
Integration with the Azure ecosystem

Build apps faster with Azure services and safeguard your innovation with Azure IP Advantage

Azure Database for PostgreSQL

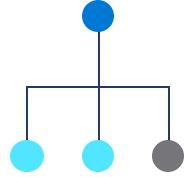
Azure builds upon the core benefits of PostgreSQL and Open Source

Azure Database for PostgreSQL is fully-managed, community PostgreSQL



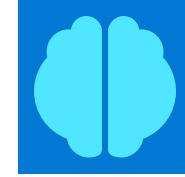
The benefits of Azure Database for PostgreSQL

Build or migrate your workloads with confidence and optimized for value



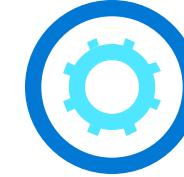
Fully managed and secure

Focus on your apps while Azure manages resource-intensive tasks, supports a large variety of Postgres versions and provides best-in-industry indemnification coverage



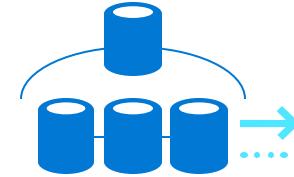
Intelligent performance optimization

Improve performance and reduce cost with customized recommendations



Flexible and open

Stay productive with your favorite Postgres extensions and leverage Microsoft's contributions to the Postgres community



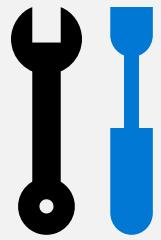
High performance scale-out with Hyperscale

Break free from the limits of single-node Postgres and scale out across 100s of nodes

Single Server

Hyperscale

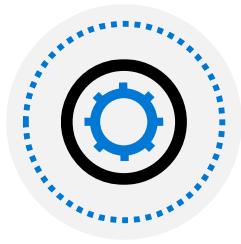
Lesson 2: Explore provisioning and deploying relational database offerings in Azure



Lesson 2 objectives



Provision relational data services



Configure relational data services

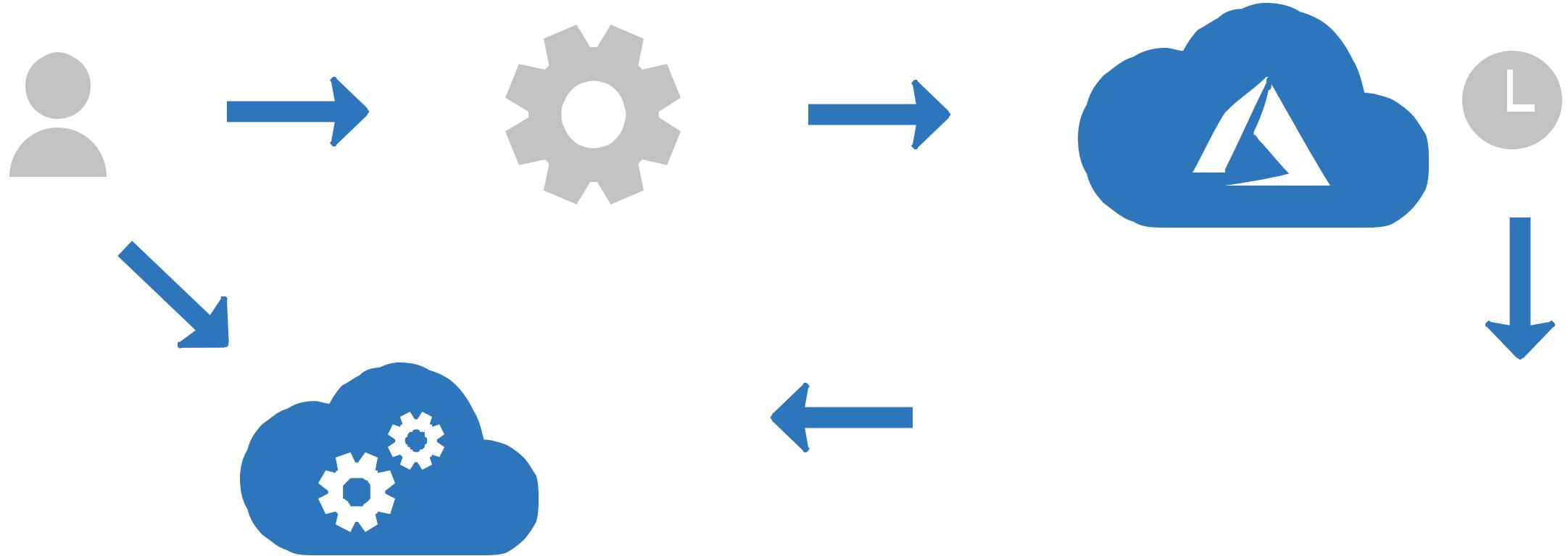


Explore basic connectivity issues



Explore data security

What is provisioning?



Configure Relational Data Services

Basics

Subscription
Resource group
Managed Instance/
Server name
Database Name (DB)
Admin Login
Password
Region
Opt-in for pools (DB)
Compute + storage

Network connectivity

Public vs Private access
VNet / Firewall rules
Connection type (MI)

Additional settings

Data source (DB)
Server Collation (MI)
Database Collation (DB)
Time zone (MI)
Opt-in for Advanced
data security (DB)

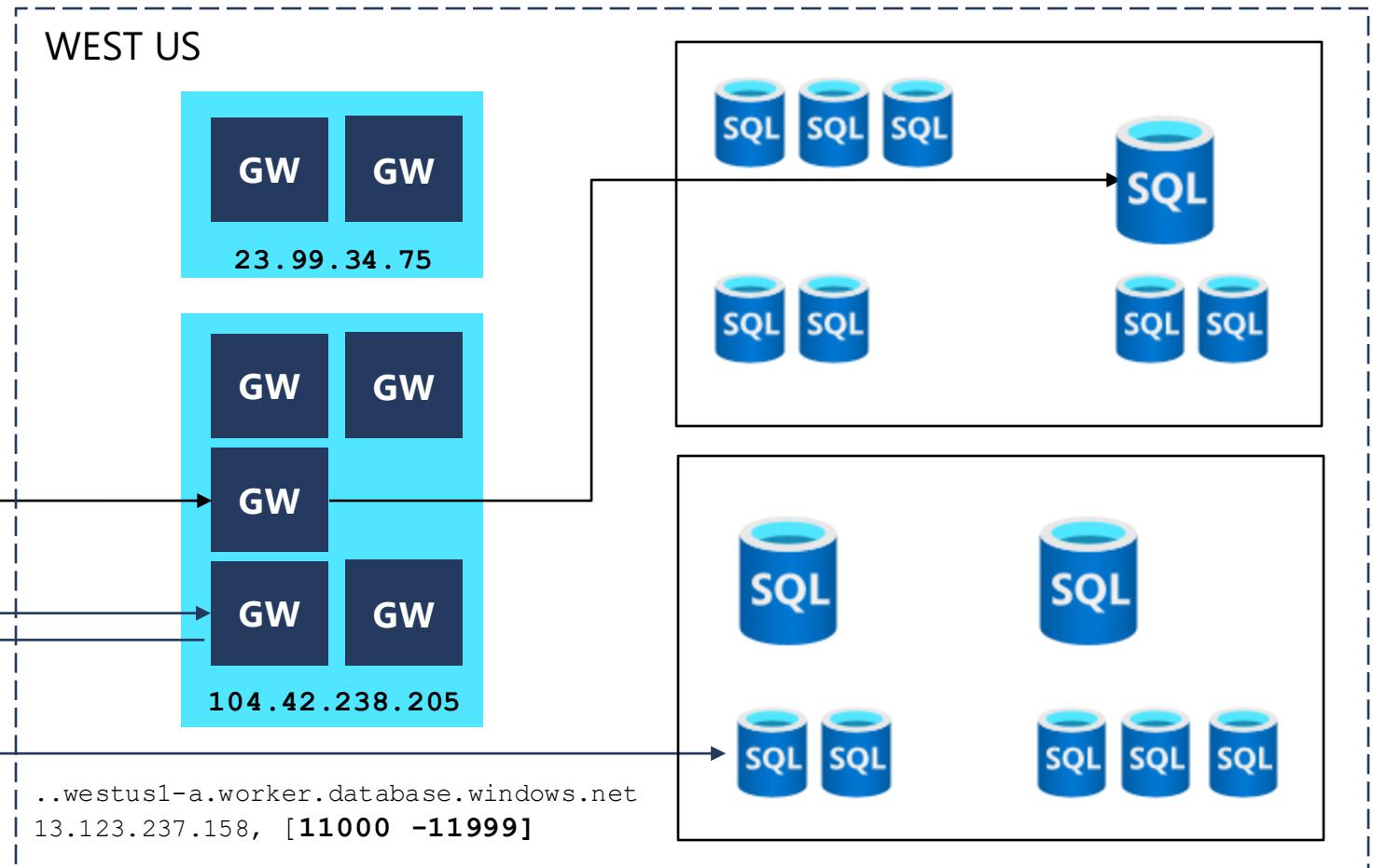
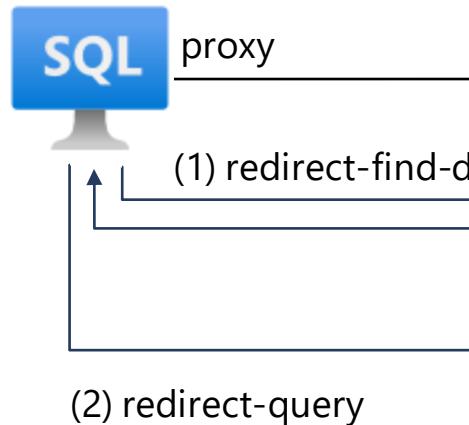
Tags (DB)

Review & create

Terms and Privacy

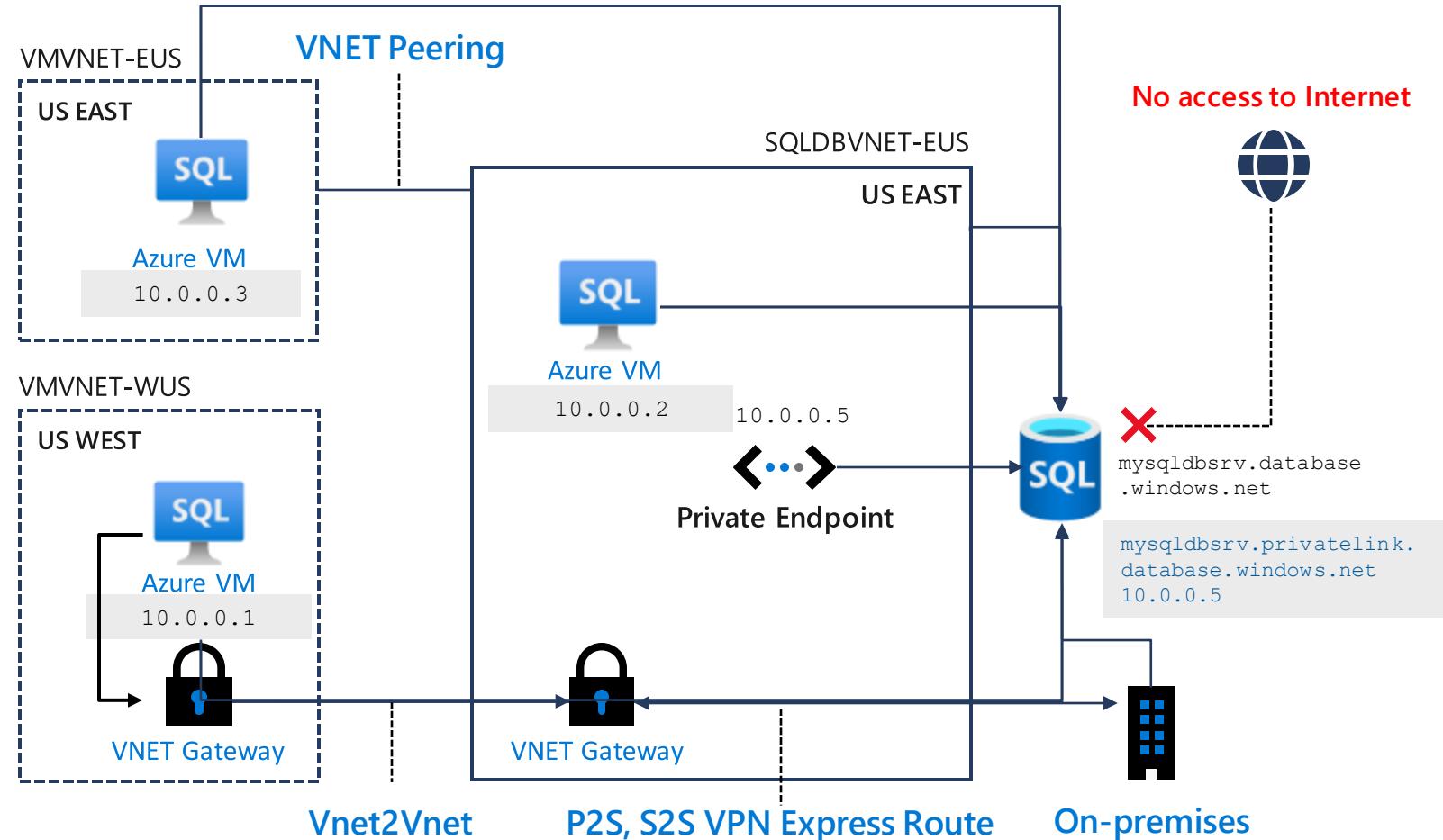
Connectivity and Firewalls

mysqlsrv.database.windows.net
westus1-a.control.database.windows.net
104.42.238.205, **1433**



Network Security – SQL Database

- Allow access to Azure services
- Firewall Rules
- Virtual Network Rules
- Private Link



Authentication and Access Control

"Mixed Mode" authentication **forced**

SQL Auth for deployment: **server admin**

- Server-level principal for logical server for DB
- Member of sysadmin server role for MI

Need Windows Auth? Use Azure AD Authentication

Azure Managed Instance

- Azure AD Server Admin
- SQL or Azure AD Logins
- Database Users
- SQL Server Contained Database supported

Azure SQL Database

- Azure AD Server Admin
- SQL logins
- loginmanager and dbmanager roles for limited server admins
- Database Users
- Contained Database Users including Azure AD (recommended)

Azure Role Based Access Control (RBAC)

- All Azure operations for Azure SQL are controlled through RBAC
- Think of this as security rights outside the Managed Instance or Database
- Security principal and role-based system
- Scope includes subscription, resource group, and resource
- Decoupled from SQL Security (today)
- Applies to operations in Azure portal and CLI
- Allows for separation of duties for deployment, management, and usage
- Azure locks help protect resources from delete or read-only
- Built-in Azure SQL roles available to reduce need for owner

SQL DB
Contributor

SQL Managed
Instance Contributor

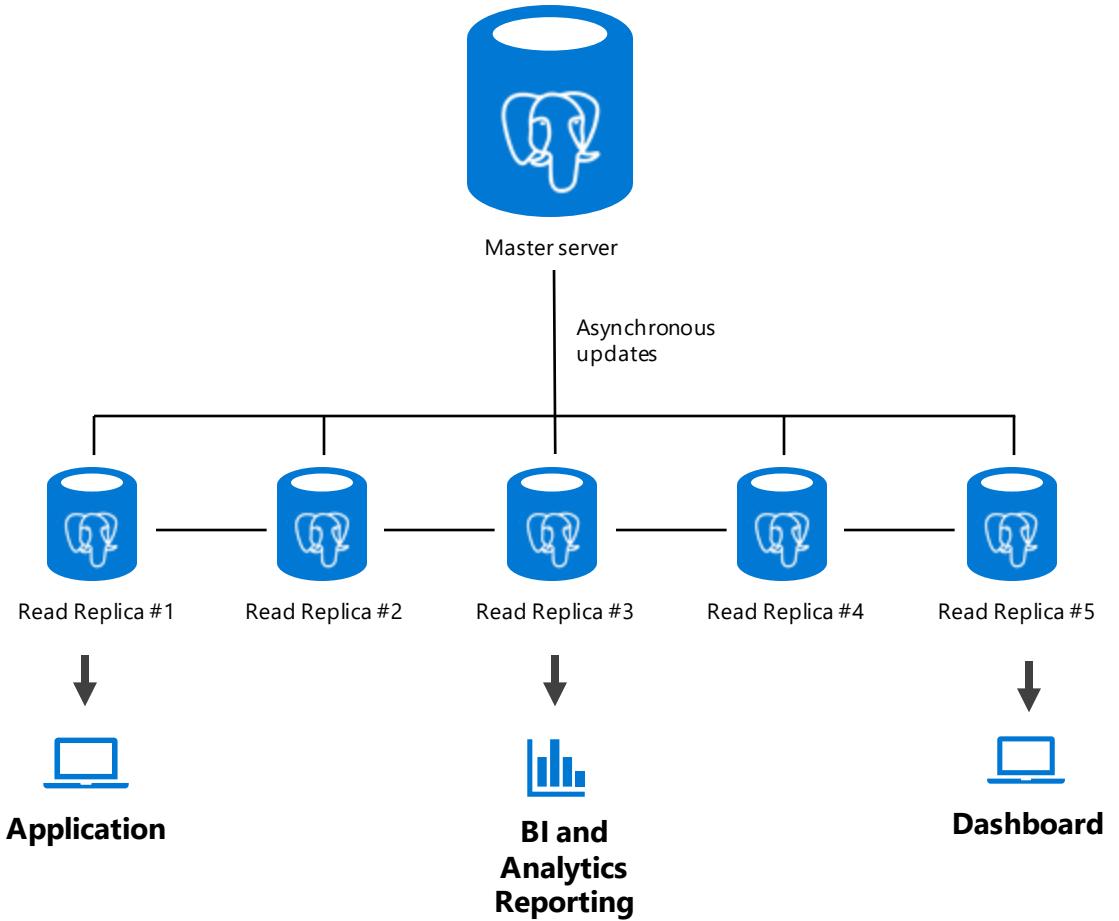
SQL Security
Manager

SQL Server
Contributor

Azure DB - read replicas

Read replicas help improve performance and scale of read-intensive workloads such as BI and analytics
Consider the read replica features in scenarios when delays in syncing data between the master and replicas are acceptable
Create a replica in a different Azure region from the master for a disaster recovery plan, where a replica replaces the master in cases of regional disasters
Data storage on replica servers grows automatically without impacting workloads

Create up to five replicas of the master server





© Copyright Microsoft Corporation. All rights reserved.

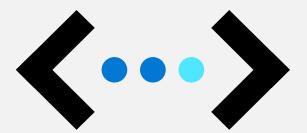
**FOR USE ONLY AS PART OF VIRTUAL TRAINING DAYS PROGRAM. THESE MATERIALS ARE NOT AUTHORIZED FOR
DISTRIBUTION, REPRODUCTION OR OTHER USE BY NON-MICROSOFT PARTIES.**



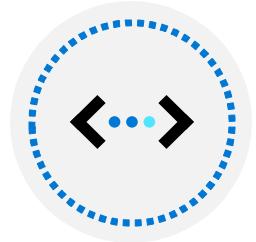
Module 2: Explore relational data in Azure- Segment 2

Demo: Provision an Azure SQL Database

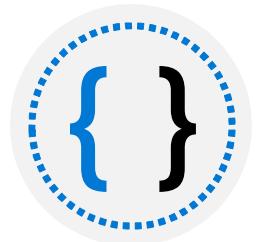
Lesson 3: Query relational data in Azure



Lesson 3 objectives



Describe query techniques for data using the SQL language



Query relational data

Introduction to SQL

- SQL is a standard language for use with relational databases
- SQL standards are maintained by ANSI and ISO
- Proprietary RDBMS systems have their own extensions of SQL such as T-SQL, PL/SQL, pgSQL

SQL Statement types

DML

- Data Manipulation Language
- Used to query and manipulate data
- SELECT, INSERT, UPDATE, DELETE

DDL

- Data Definition Language
- Used to define database objects
- CREATE, ALTER, DROP, REMOVE

DCL

- Data Control Language
- Used to manage security permissions
- GRANT, REVOKE, DENY

Use DML statements

Statement	Description
SELECT	Select/read from a table
INSERT	Insert new rows in a table
UPDATE	Edit/Update existing rows in a table
DELETE	Delete existing rows in a table

Elements of the SELECT Statement

Clause	Expression
SELECT	<select list>
FROM	<table or view>
WHERE	<search condition>
GROUP BY	<group by list>
ORDER BY	<order by list>

Example of SELECT statement

```
SELECT EmployeeId, YEAR(OrderDate) AS OrderYear  
FROM Sales.Orders  
WHERE CustomerId = 71  
GROUP BY EmployeeId, YEAR(OrderDate)  
HAVING COUNT(*) > 1  
ORDER BY EmployeeId, OrderYear;
```

Example of INSERT statement

- The INSERT ... VALUES statement inserts a new row

```
INSERT INTO Sales.OrderDetails  
    (orderid, productid, unitprice, qty, discount)  
VALUES (10255,39,18,2,0.05);
```

- Table and row constructors add multirow capability to INSERT ... VALUES

```
INSERT INTO Sales.OrderDetails  
    (orderid, productid, unitprice, qty, discount)  
  
VALUES  
    (10256,39,18,2,0.05),  
    (10258,39,18,5,0.10);
```

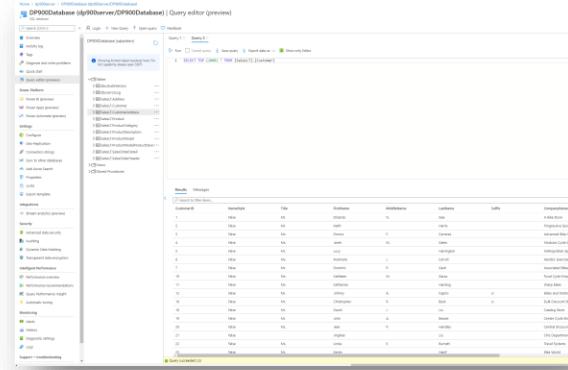
Use DDL statements

Statement	Description
CREATE	Create a new object in the database, such as a table or a view.
ALTER	Modify the structure of an object. For instance, altering a table to add a new column.
DROP	Remove an object from the database.
RENAME	Rename an existing object.

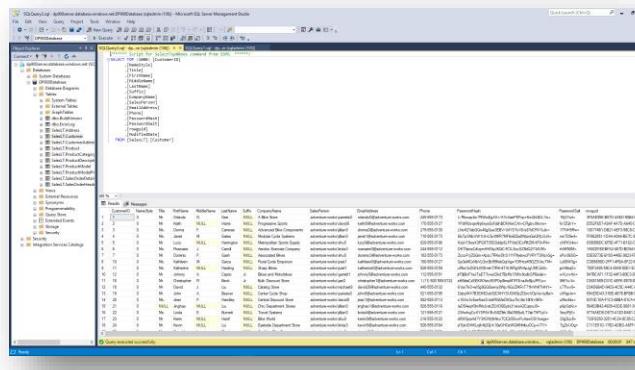
Example of CREATE statement

```
CREATE TABLE Mytable  
(Mycolumn1 int NOT NULL PRIMARY KEY, Mycolumn2  
VARCHAR(50) NOT NULL , Mycolumn2 VARCHAR(10) NOT NULL
```

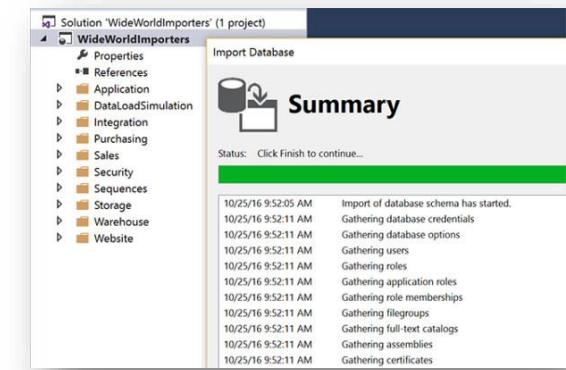
Query Tools



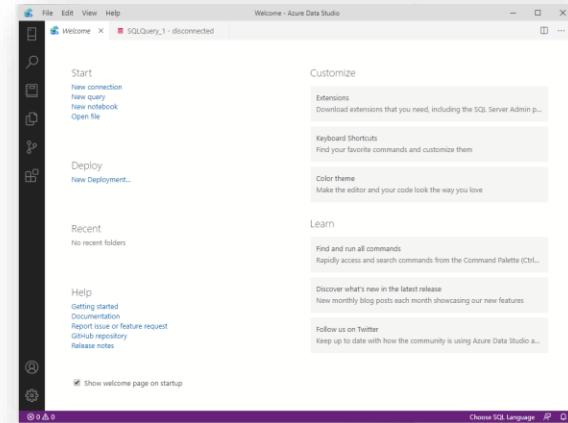
Azure Portal



SQL Management Studio



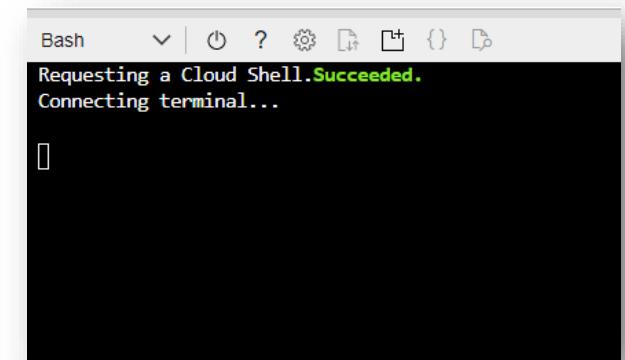
SQL Server Data Tools



Azure Data Studio

```
1> select name from sys.databases;
2> go
name
-----
master
tempdb
model
msdb
(4 rows affected)
1>
```

SQLCMD



Azure CLI / Cloud Shell

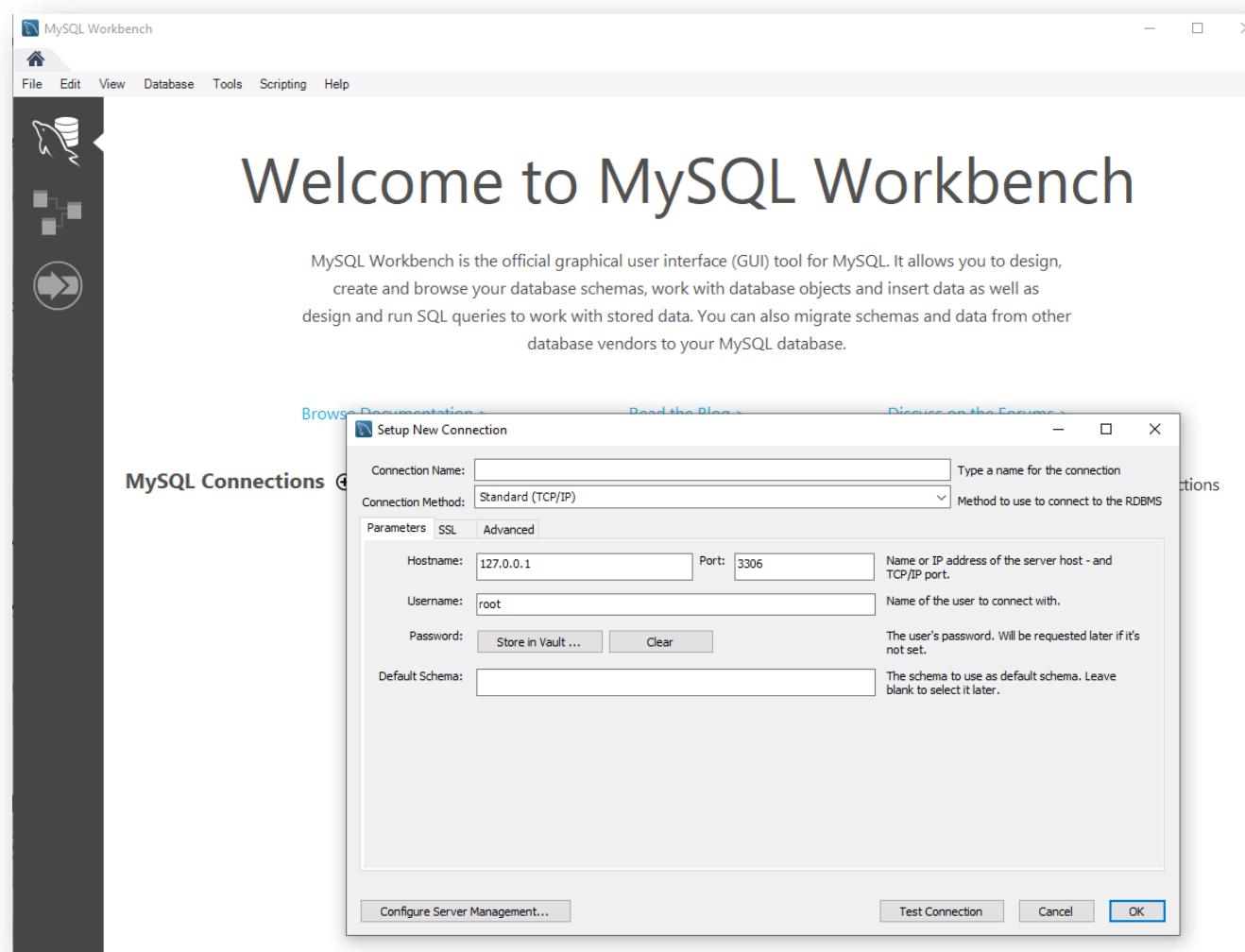
Query relational data in Azure SQL Database for PostgreSQL

[Use PSQL to query a database](#)

```
psql --host=<server-name>.postgres.database.azure.com --  
username=<admin-user>@<server-name> --dbname=postgres
```

Query relational data in Azure SQL Database for MySQL

Use MySQL Workbench to query a database



© 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

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	AAAAAA 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 shows the Azure Storage Account Overview page for the account 'jpwsstorageaccount'. The left sidebar contains navigation links for Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Data transfer, Events, and Storage Explorer (preview). The main content area displays account details: Resource group (change) : learnrg, Status : Primary: Available, Secondary: Available, Location : UK West, UK South, Subscription (change) : , Subscription ID : , and Tags (change) : Click here to add tags. A message at the top states: '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.' Below the details are four service cards: '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). At the bottom, there's a 'Tools and SDKs' section with links for Storage Explorer (preview), PowerShell, Azure CLI, .NET, Java, Python, and Node.js.

Home > Microsoft.StorageAccount-20200520142053 | Overview > jpwsstorageaccount

jpwsstorageaccount
Storage account

Search (Ctrl+ /)

Open in Explorer Move Refresh Delete Feedback

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)

Containers
Massively scalable data lake storage
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

Tools and SDKs

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

Explore Azure Cosmos DB



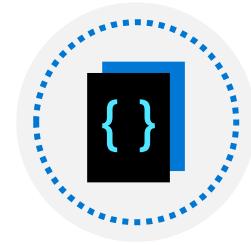
Scalability



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

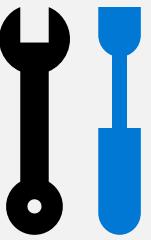
Gaming

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

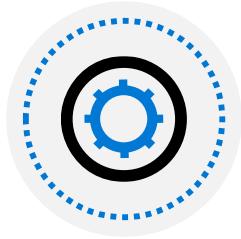
Lesson 2: Explore provisioning and deploying non-relational data services in Azure



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 99.99% 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 Concierge Subscription>

[Create new](#)

INSTANCE DETAILS

* Account Name: <enter a unique name> documents.azure.com

* API: SQL

* Location: <choose the location closest to you>

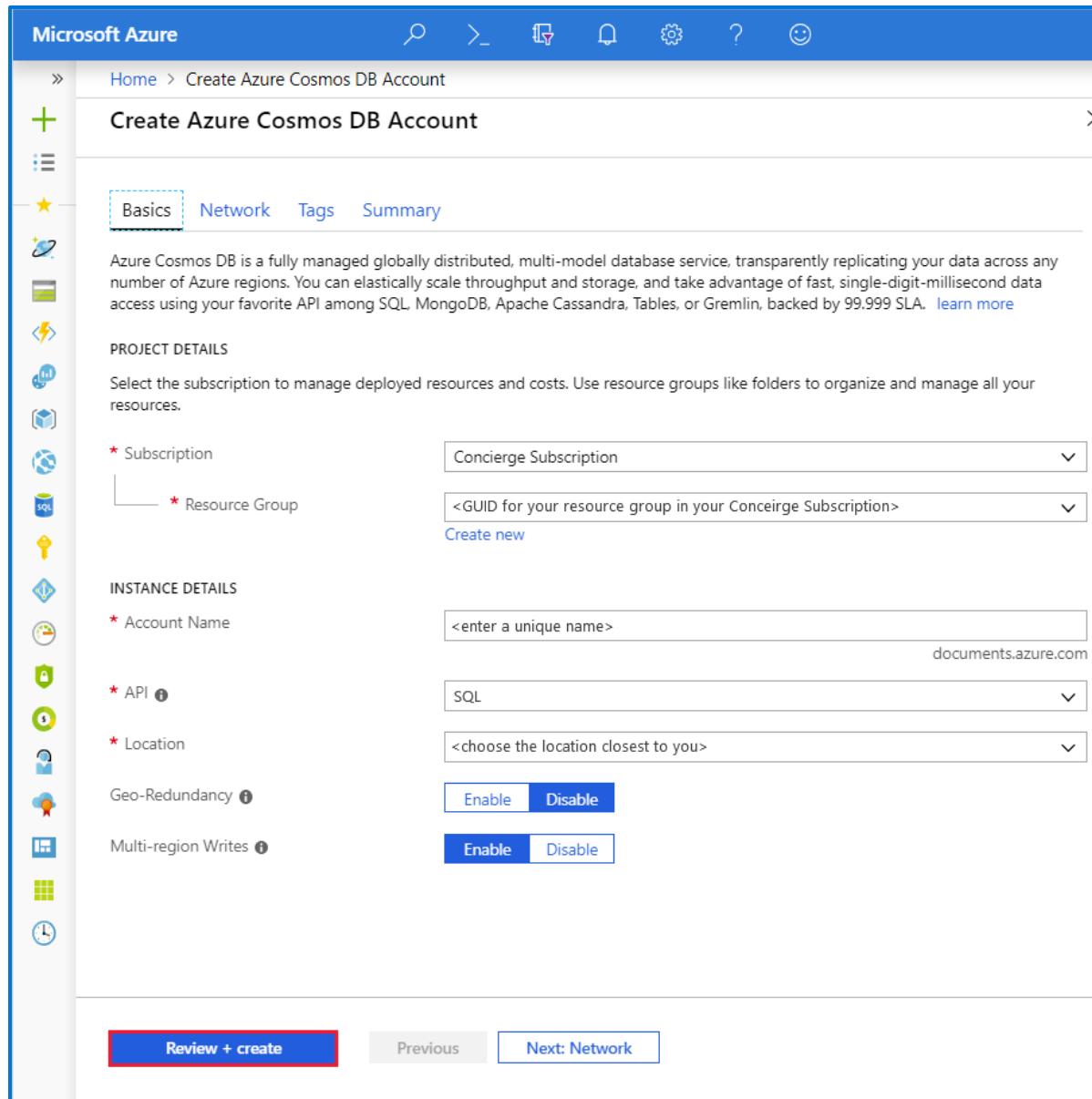
Geo-Redundancy: [Enable](#) [Disable](#)

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

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

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 i

* Location

East US

Standard Premium

Account kind i

StorageV2 (general purpose v2)

Replication i

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

Access tier (default) i

Cool Hot

Home > New > Create storage account

Create storage account

Basics Advanced Tags Review + create

SECURITY

Secure transfer required i

Disabled Enabled

VIRTUAL NETWORKS

Allow access from

All networks Selected network

i All networks will be able to access this storage account. [Learn more](#)

DATA LAKE STORAGE GEN2 (PREVIEW)

Hierarchical namespace i

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... or 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

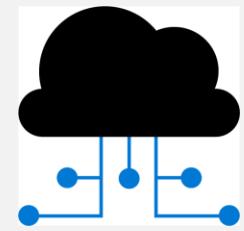
Performance: Standard Premium

Account kind: StorageV2 (general purpose v2)

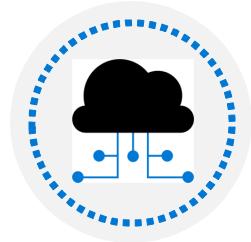
Replication: Read-access geo-redundant storage (RA-GRS)

Access tier (default): Cool Hot

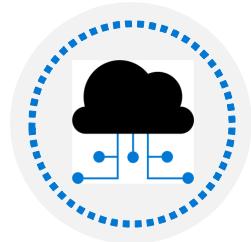
Lesson 3: Manage non-relational data stores in Azure



Lesson 3 objectives

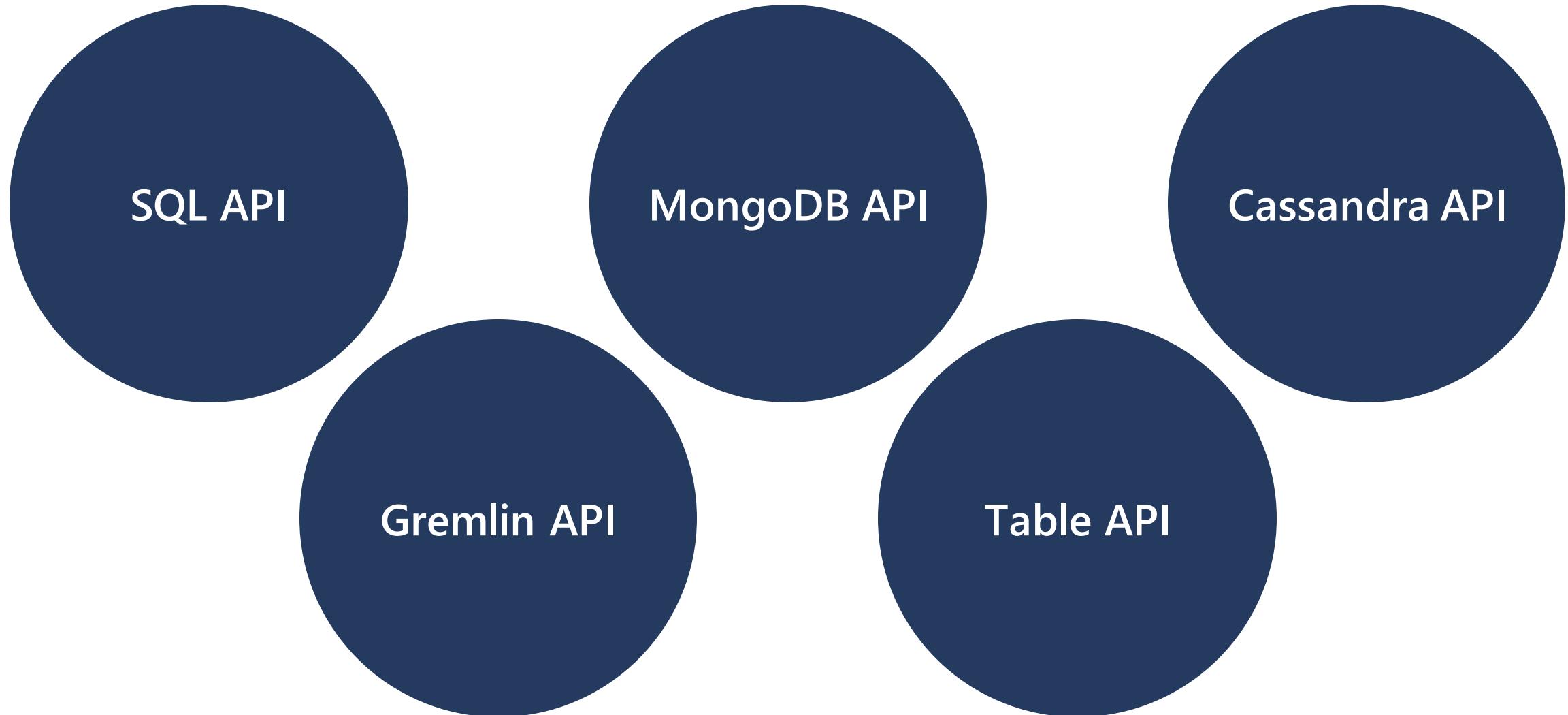


Upload data to a Cosmos DB database, and learn how to query this data.



Upload and download data in an Azure Storage account.

Cosmos DB APIs

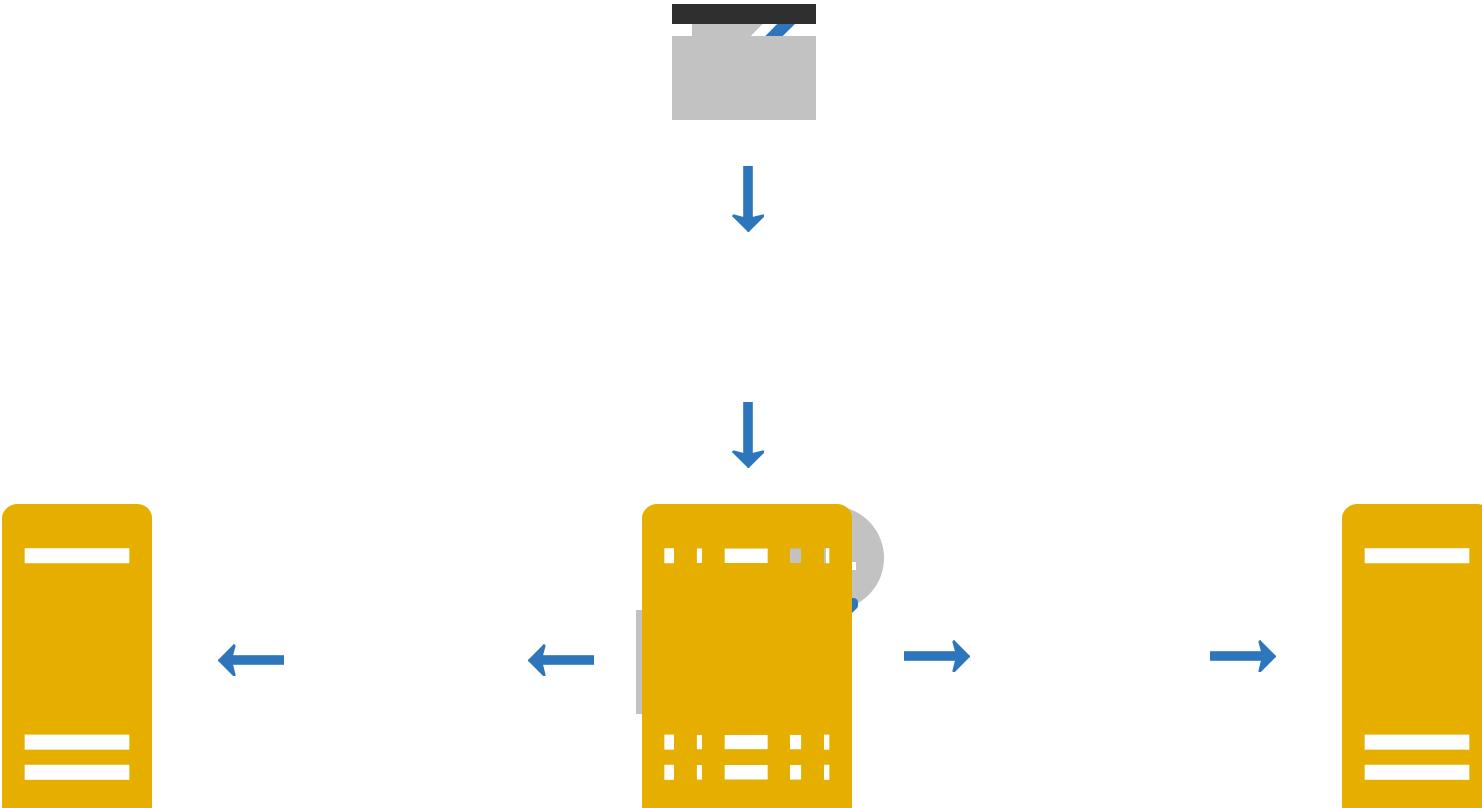


Load data using the Cosmos DB Migration tool

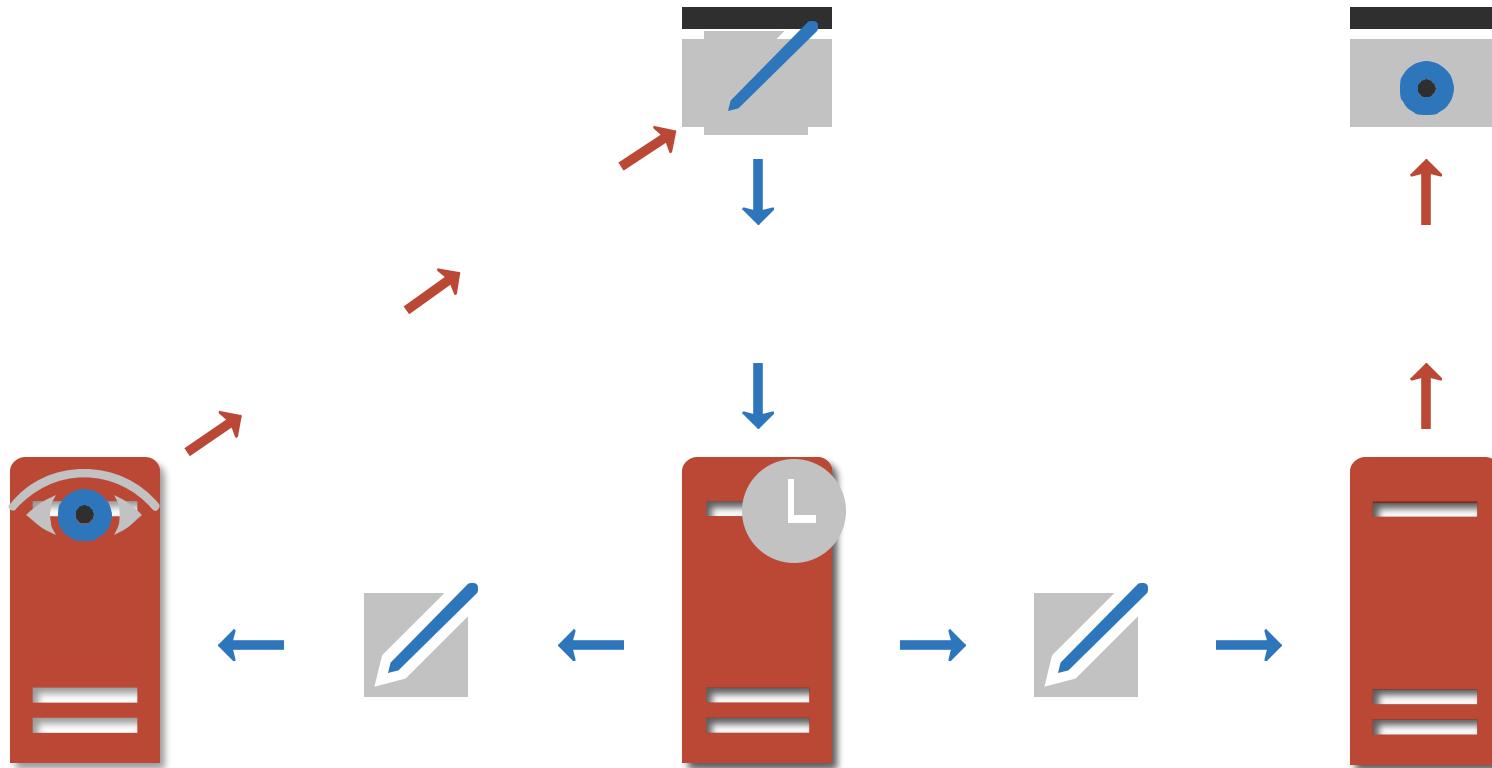
You can use the Data Migration tool to import data to Azure Cosmos DB from a variety of sources, including:

- 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 shows the Azure Storage account management interface for the account 'contosodata'. The left sidebar contains navigation links for 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), and a search bar.

The main content area displays account details:

- 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)

A message at the top right indicates: "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."

The bottom section features four service cards:

- Containers**: Scalable, cost-effective storage for unstructured data. (This card is highlighted with a red border.)
- File shares**: Serverless SMB file shares. [Learn more](#)
- Tables**: Tabular data storage. [Learn more](#)
- Queues**: Effectively scale apps according to traffic. [Learn more](#)

Manage Azure File storage

The screenshot shows the Azure Storage Explorer interface. On the left, a sidebar lists navigation items: Home, contosodata | File shares, reports (selected), Overview, Access Control (IAM), Settings, Properties, Operations, and Snapshots. The main area displays a file share named 'reports'. A red box highlights the top navigation bar, which includes 'Connect' (with a gear icon), 'Upload', 'Add directory', 'Refresh', 'Delete share', and 'Edit query'. Below this is a search bar labeled 'Search files by prefix'. A table lists three directory entries: 'February Reports', 'January Reports', and 'March Reports', under the 'Name' and 'Type' columns respectively. To the right, a 'Connect' panel is open. It shows the 'reports' file share and includes a warning message: "'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." Below this are tabs for 'Windows' (selected), 'Linux', and 'macOS'. Under the 'Windows' tab, a 'Drive letter' dropdown is set to 'Z'. Below it, PowerShell commands are provided to connect from a Windows 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"
```

A note below explains the script's purpose: "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." At the bottom, a link reads "Learn how to circumvent the port 445 problem (VPN)".

Copying to Azure Storage

Home >

contosodata | Storage Explorer (preview) Storage account

Search (Ctrl+/) Search

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

BLOB CONTAINERS FILE SHARES QUEUES TABLES





© 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

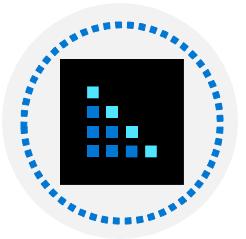
Lesson 1: Examine components of a modern data warehouse



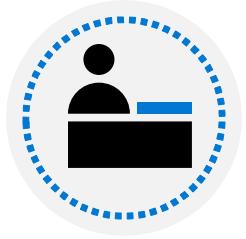
Lesson 1 objectives



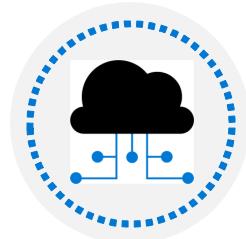
Explore data warehousing concepts



Explore Azure data services for modern data warehousing

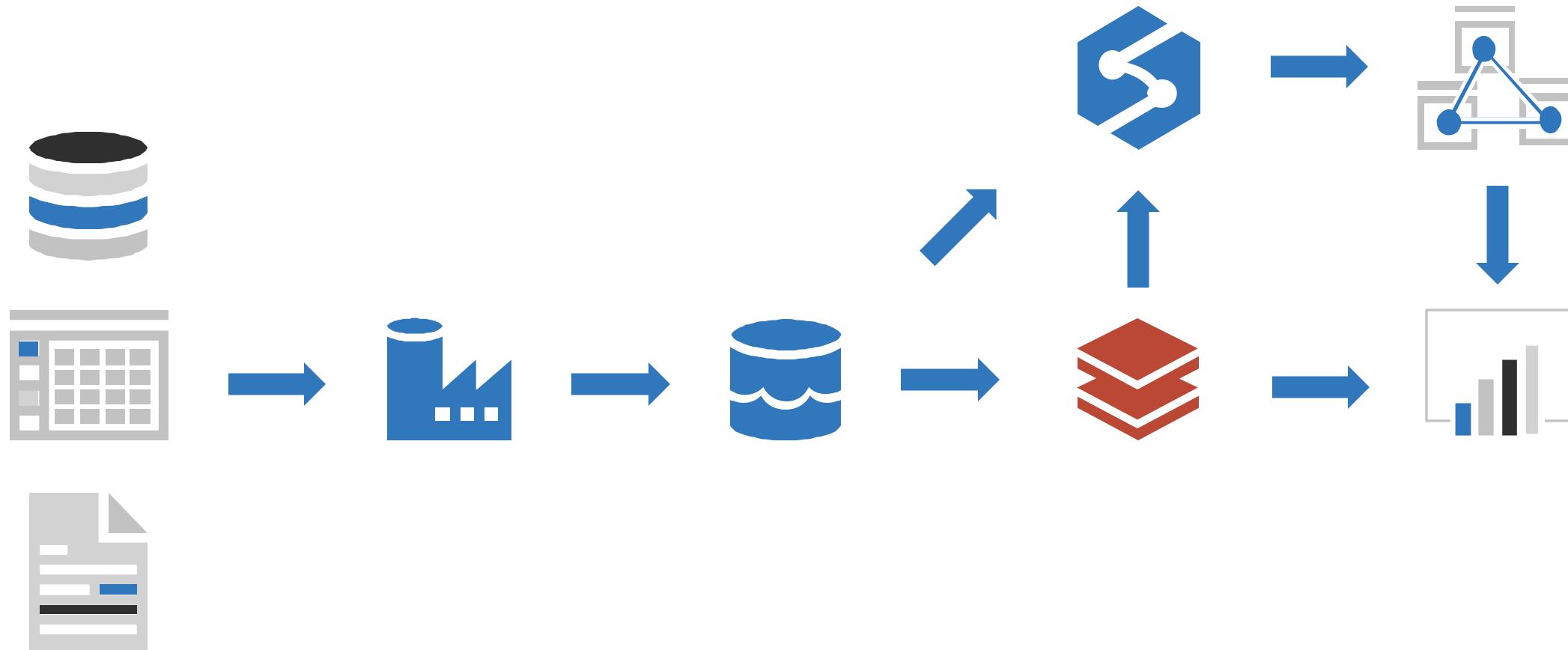


Explore modern data warehousing architecture and workload

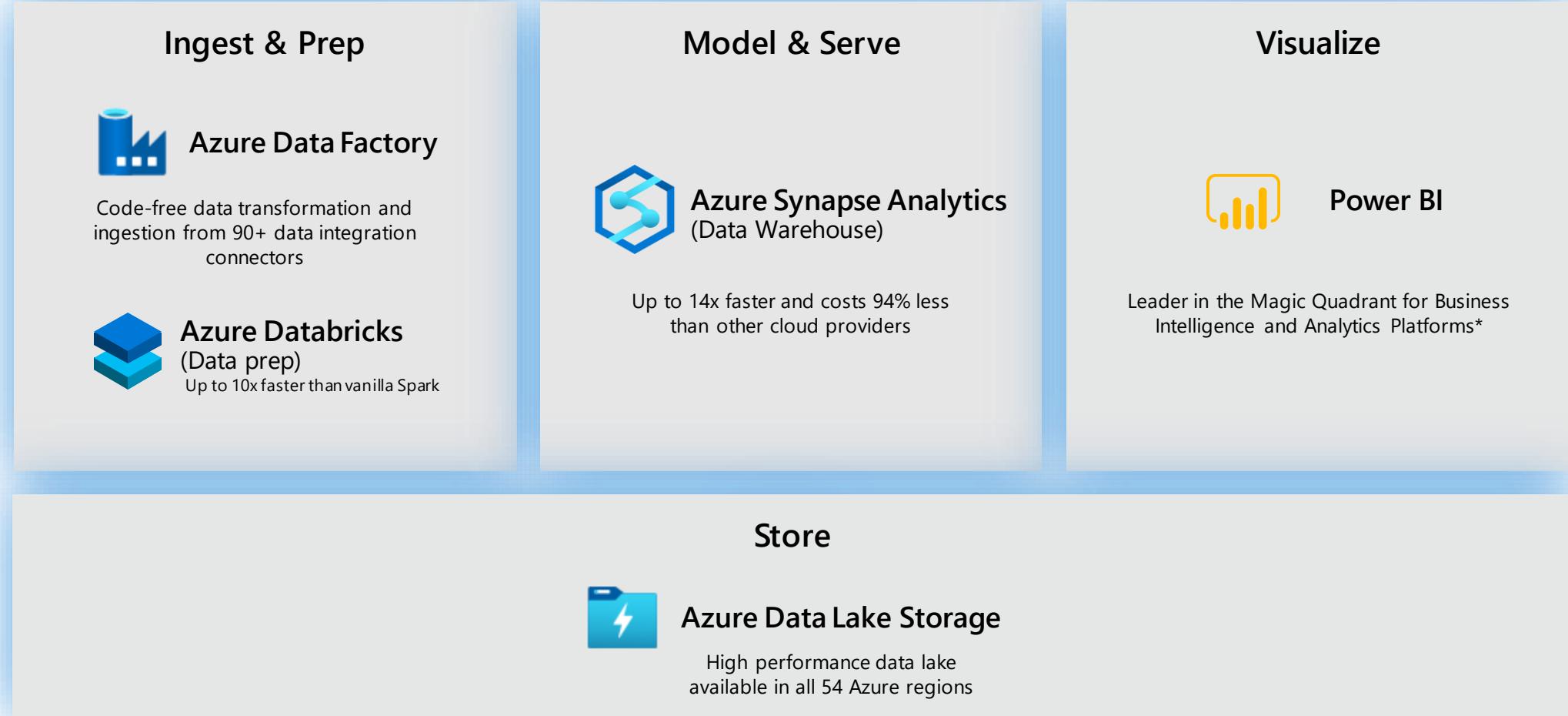


Explore Azure data services in the Azure portal

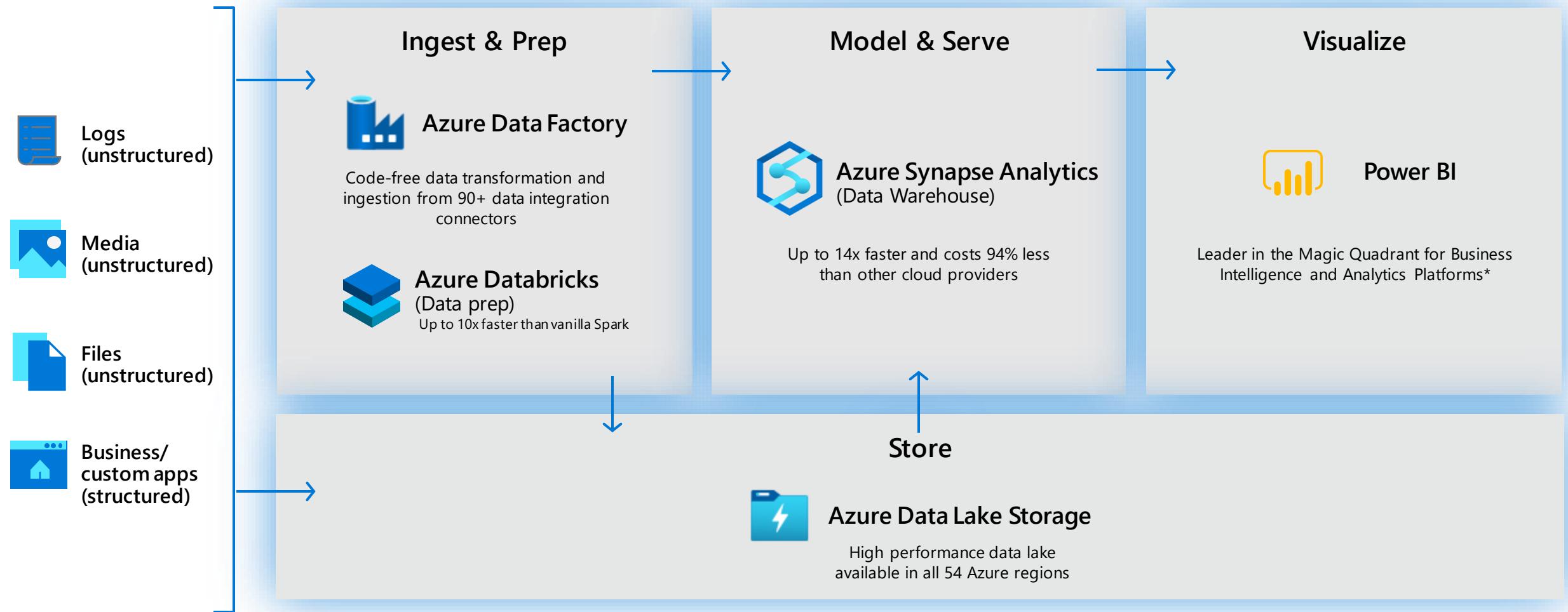
Modern data warehouse components



What is modern data warehousing?



Combine batch and stream processing



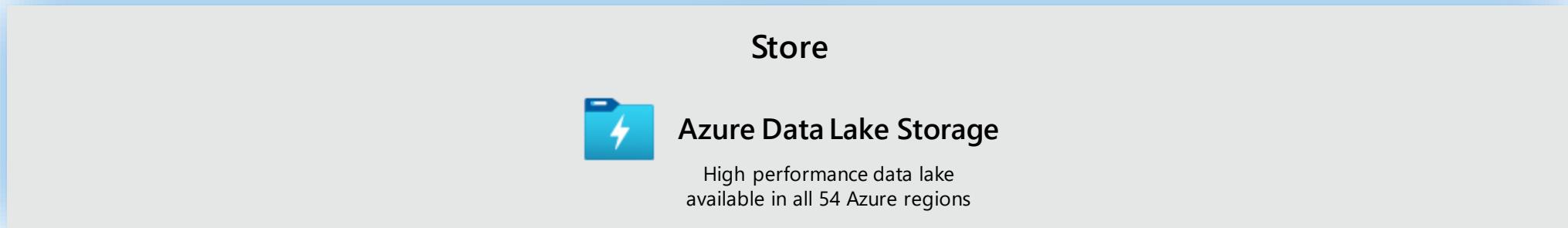
Explore Azure data services for modern data warehousing

[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



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?



Synapse
Pipelines



Synapse
SQL



Azure Synapse
Analytics

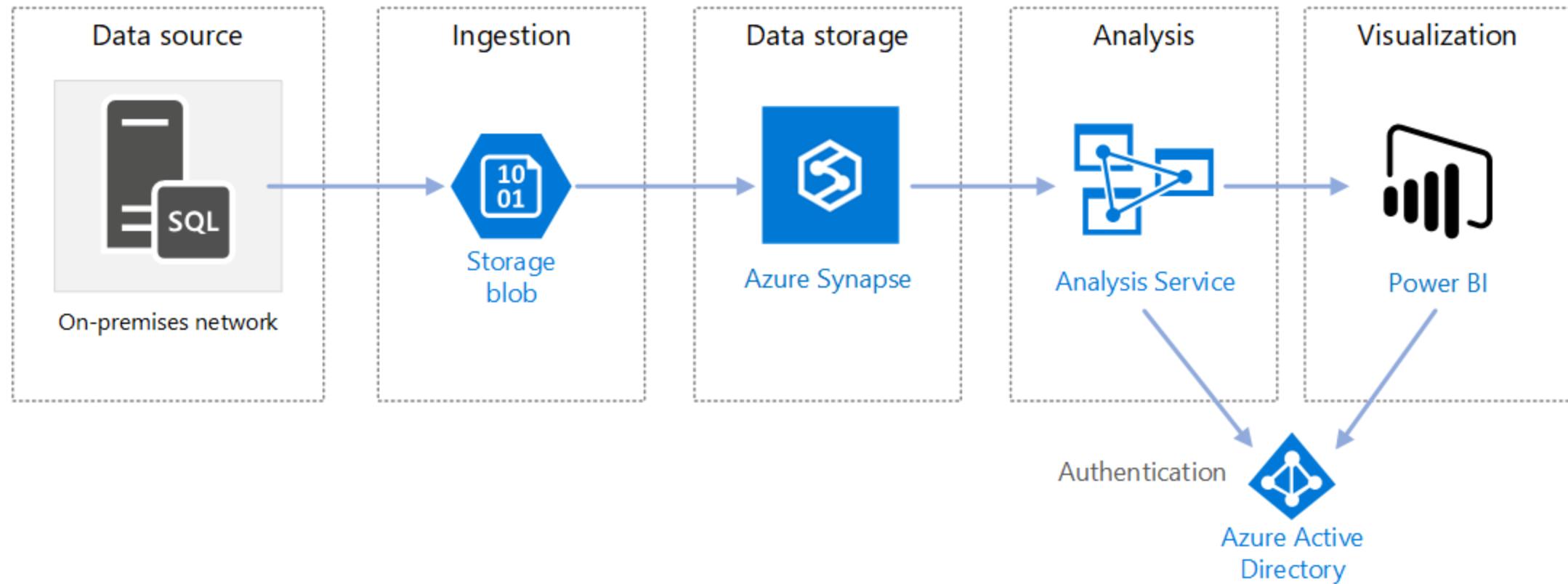


Synapse
Studio

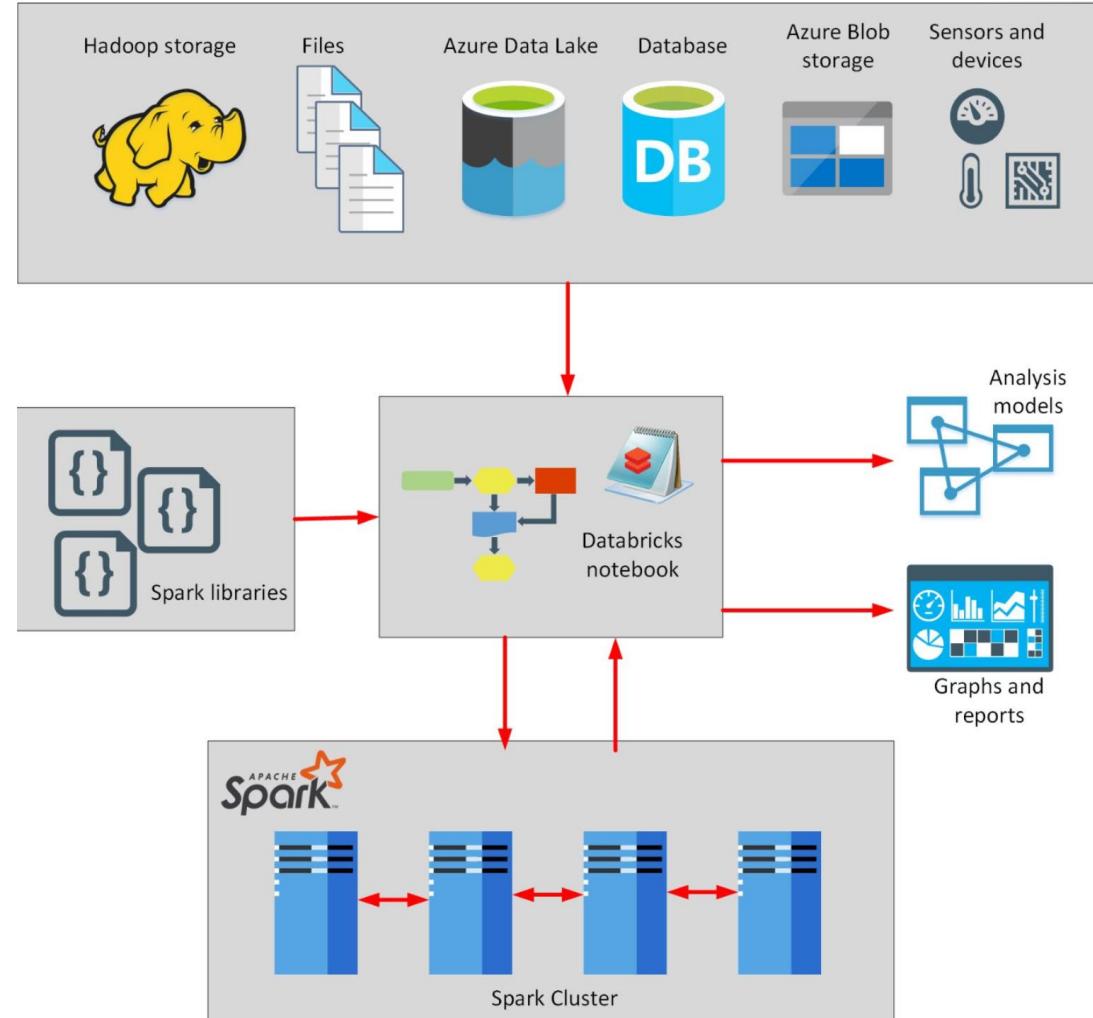


Synapse
Spark

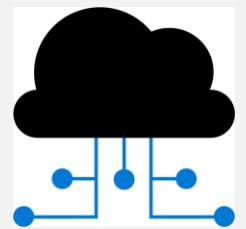
What is Azure Analysis Services?



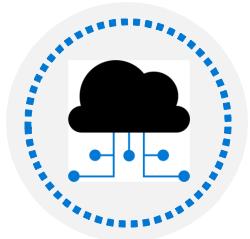
What is Azure HDInsight?



Lesson 2: Explore data ingestion in Azure



Lesson 2 objectives



Describe data ingestion in Azure



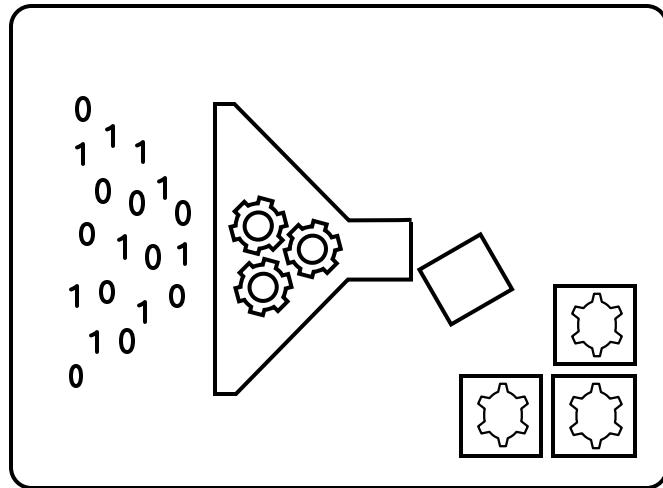
Describe components of Azure Data Factory



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

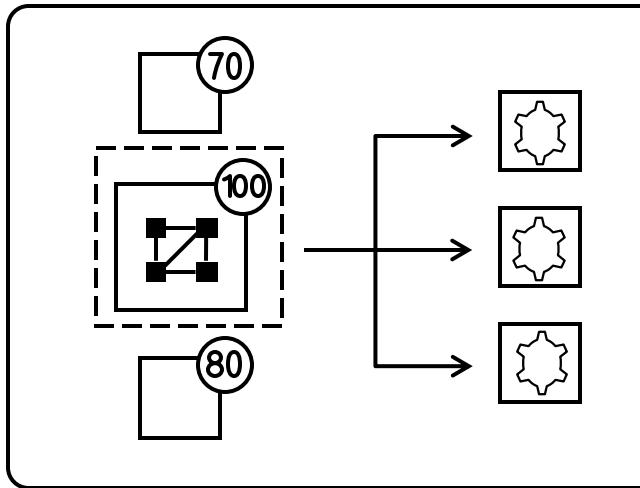
Describe data ingestion in Azure

ADF



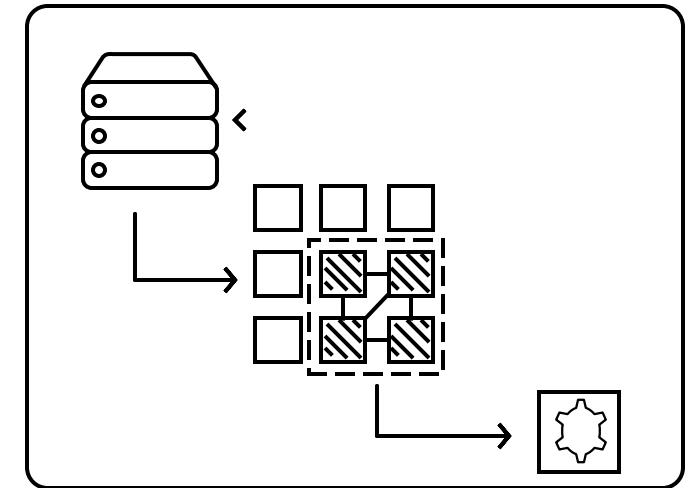
Heterogenous

PolyBase



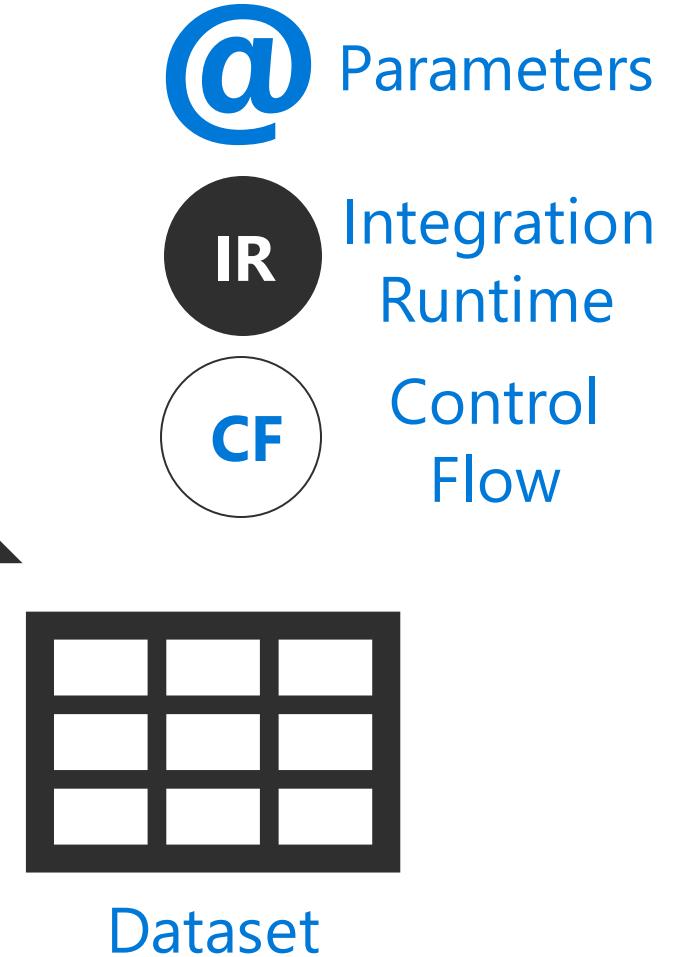
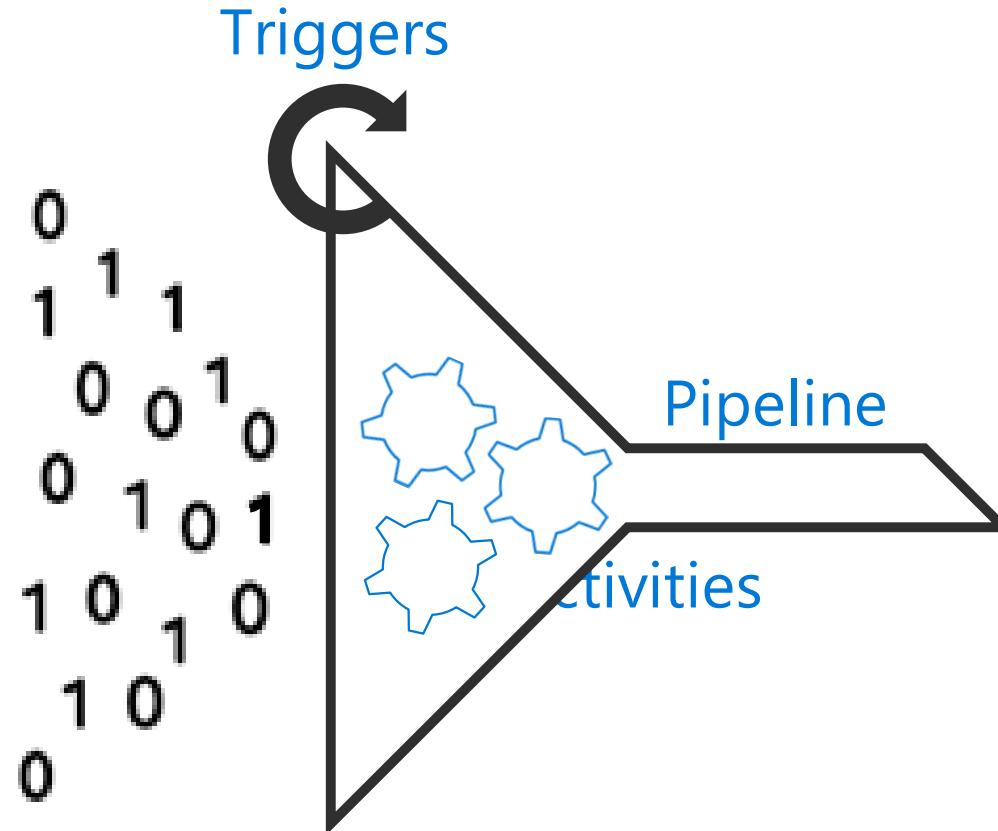
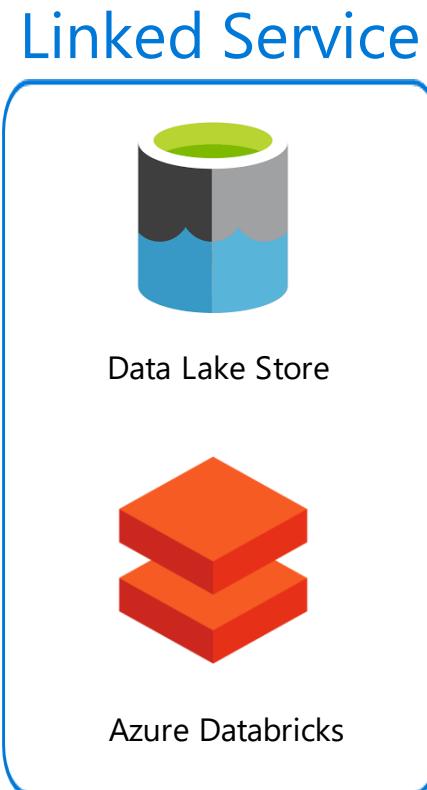
File based

SSIS



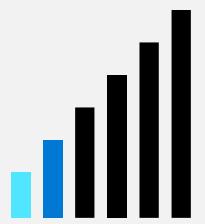
Heterogenous

Describe components of Azure Data Factory



Demo: Load data into Azure Synapse Analytics

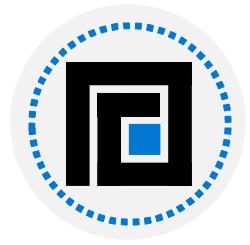
Lesson 3: Explore data storage and processing in Azure



Lesson 3 objectives



Describe data processing options for performing analytics in Azure



Explore Azure Synapse Analytics

Data processing options for performing analytics in Azure



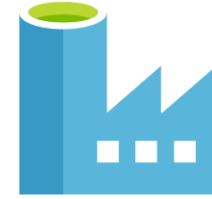
Azure Synapse
Analytics



Azure Databricks



Azure HDInsight



Azure Data Factory



Data Lake Store

Explore Azure Synapse Analytics



Synapse
Pipelines



Synapse
SQL



Synapse
Link



Azure Synapse
Analytics

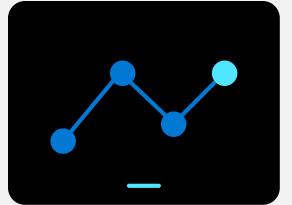


Synapse
Studio



Synapse
Spark

Lesson 4: Get started building with Power BI



Lesson 4 objectives



Learn how Power BI services and applications work together

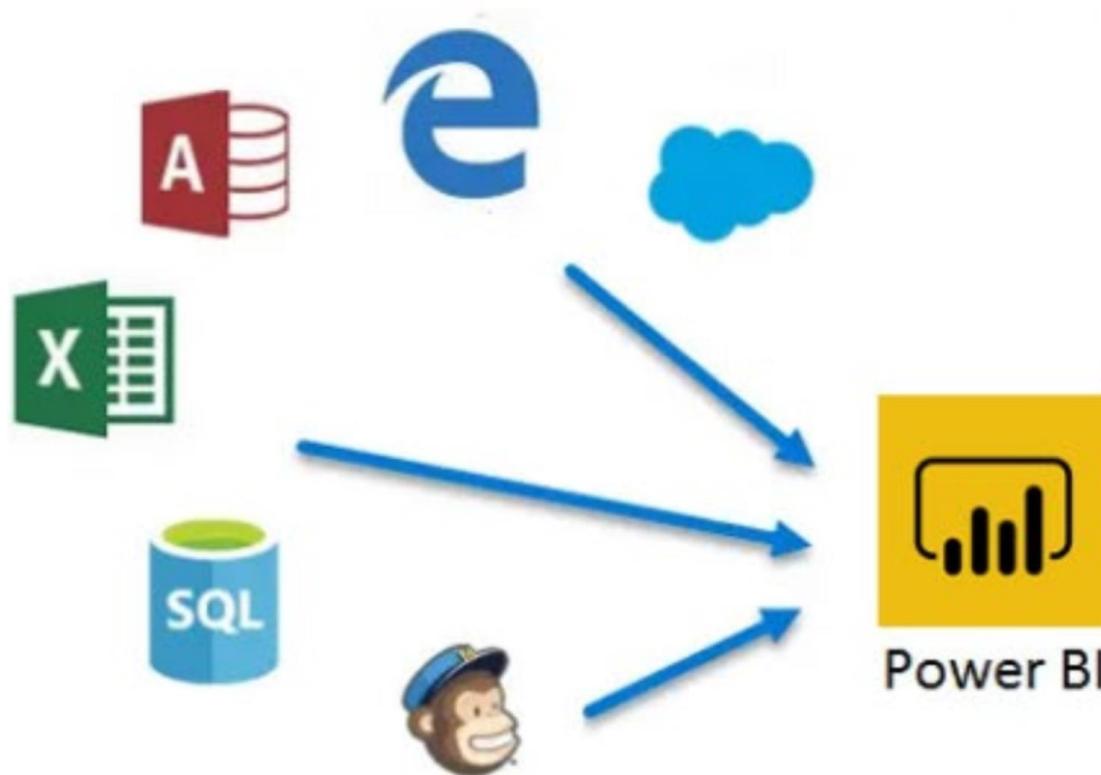


Explore how Power BI can make your business more efficient



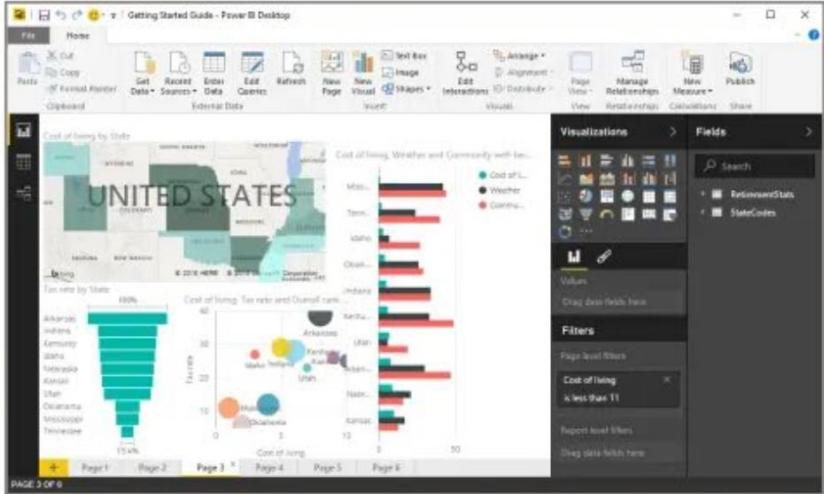
Learn how to create compelling visuals and reports.

Learn how Power BI services and applications work together



Explore how Power BI can make your business more efficient

Power BI Desktop



Power BI service



Power BI Mobile



Learn how to create compelling visuals and reports.

