



A aula interativa do **Módulo 2 - Bootcamp Analista de Banco de Dados** em breve!

Importante:

- 1) Não se esqueça de acessar a aula com seu e-mail cadastrado no ambiente de aprendizagem do IGTI e seu nome completo. Entrou com os dados errados? Saia da sala e entre novamente com os dados corretos.
- 2) Para sua frequência ser computada, responda a enquete no ambiente de aprendizagem, no horário indicado pelo professor. A enquete ficará disponível por 10 minutos.
- 3) Utilize o Chat para interagir com os colegas durante a aula, a ferramenta Raise Hands para pedir a palavra, e, em caso de dúvidas sobre o conteúdo, utilize o Q&A (perguntas e respostas) para que o professor tutor possa respondê-las.

Informações importantes

- Acesse a aula com o seu nome completo e com o mesmo e-mail utilizado no cadastro do ambiente de aprendizagem do IGTi. Entrou com os dados errados? É só sair da sala e entrar com os dados corretos!
- As presenças das Aulas Interativas são computadas através de uma enquete, que será realizada no ambiente de aprendizagem do IGTi. Para sua frequência ser computada, quando solicitado pelo professor, você deverá ir até a seção “Enquete de presença da Aula Interativa”, localizada na Área Acadêmica da disciplina, e responder a enquete. Essa seção estará logo abaixo da que contiver o link para a Aula Interativa em questão. A enquete ficará no ar por 10 minutos e sua nota estará disponível ao término desse tempo.
- Utilize o chat para interagir com os colegas durante a aula interativa! Em caso de dúvidas sobre o conteúdo, é só postá-las no Q&A que o tutor irá respondê-las! Lembre-se que você ainda pode esclarecer as suas dúvidas nos fóruns disponibilizados no Ambiente de Aprendizagem.

Informações importantes

- Ah! E se você não conseguir assistir a aula interativa, não se preocupe! Sua gravação ficará disponível no Área Acadêmica, juntamente com os slides utilizados pelo professor, em até 24 horas úteis após o término da aula. Você também poderá realizar a atividade de reposição para recuperar os pontos de presença!
- Se você precisar solicitar prorrogação e/ou 2ª oportunidade para entrega de atividades, saiba que isso é realizado somente mediante a apresentação de atestado médico ou de óbito de parentes de 1º grau.
- Para melhor experiência nas aulas interativas, sugerimos que você baixe o aplicativo do Zoom no seu computador.

A Linguagem SQL

Primeira Aula Interativa

Prof. Gustavo Aguilar

Nesta aula



- ☐ Alternativas de SQL Server.
- ☐ Alterações em Schemas Físicos.
- ☐ JOIN x Subconsultas.
- ☐ OPENROWSET.
- ☐ PolyBase.
- ☐ Quiz Premiado.

Alternativas de SQL Server

- Oracle VirtualBox
 - <https://www.virtualbox.org>
 - Windows, Linux, MacOS e Solaris



Alternativas de SQL Server



- Microsoft Azure
 - <https://azure.microsoft.com/pt-br/free>
 - Infraestrutura como Serviço (IaaS)



Alternativas de SQL Server



- Microsoft Azure
 - Azure SQL
 - <https://azure.microsoft.com/pt-br/free/sql-database>



Conta gratuita
(12 meses)

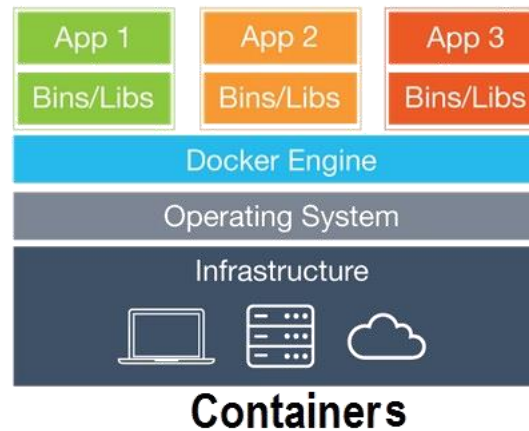
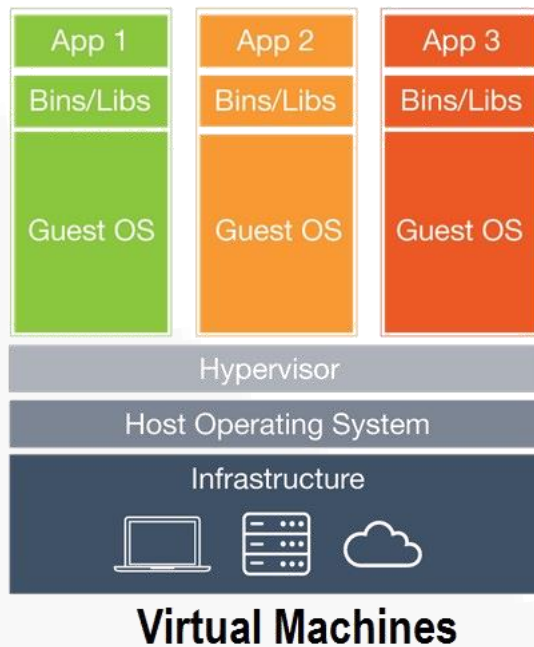
VM (SO + SQL)



Alternativas de SQL Server



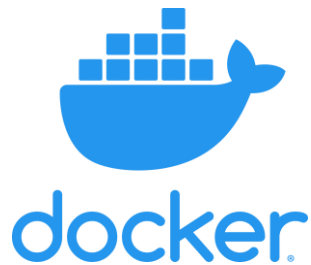
- Container
- Máquinas Virtuais x Container



Alternativas de SQL Server



- Docker Container
 - Docker Desktop
 - <https://www.docker.com/products/docker-desktop>



```
docker pull microsoft/mssql-server-windows-developer
```

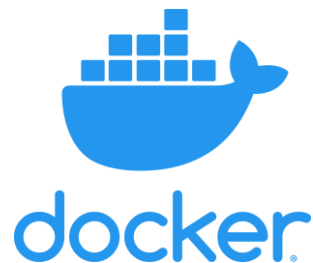
```
docker run -d --name vmsqlbtc -p 1433:1433 -e sa_password=Btcbda@2020 -e ACCEPT_EULA=Y microsoft/mssql-server-windows-developer
```

```
docker inspect --format='{{range .NetworkSettings.Networks}}{{.IPAddress}}{{end}}' vmsqlbtc
```

Alternativas de SQL Server



- Docker Container
 - Nuvem → Play With Docker
 - <https://www.docker.com/play-with-docker>
 - Somente imagens com SQL Server no Linux



```
docker pull mcr.microsoft.com/mssql/server
```

```
docker run --name sqlbtcbda -e 'ACCEPT_EULA=Y' -e 'SA_PASSWORD=gaaBTCBDA@2020' -p 1433:1433 -d mcr.microsoft.com/mssql/server:2019-latest
```

```
docker exec -it sqlbtcbda /opt/mssql-tools/bin/sqlcmd -S localhost -U sa -P 'gaaBTCBDA@2020'
```

Alterações em Schemas Físicos



- Cenário: tabela já criada e populada (com dados)
 - Aumentar o tamanho de uma coluna
 - Alterar o tipo de dados
 - Incluir uma coluna nula
 - Incluir uma coluna não nula
 - Excluir uma coluna



JOIN x Subconsulta

- Títulos sem avaliação?




Avaliacao

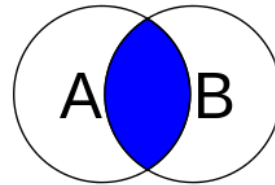
cod_titulo: int NOT NULL (FK)
classificacao_media: int NOT NULL
qtd_votos: int NOT NULL

FK_Avaliacao_Titulo

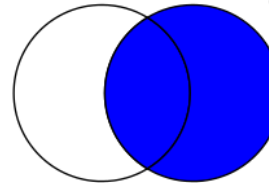
Titulo

 cod_titulo: int NOT NULL
nom_titulo: varchar(1000) NULL

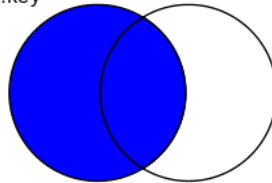
```
SELECT <fields>
FROM TableA A
INNER JOIN TableB B
ON A.key = B.key
```



```
SELECT <fields>
FROM TableA A
RIGHT JOIN TableB B
ON A.key = B.key
```

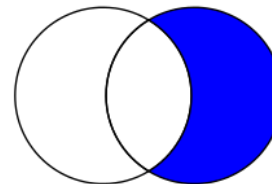


```
SELECT <fields>
FROM TableA A
LEFT JOIN TableB B
ON A.key = B.key
```

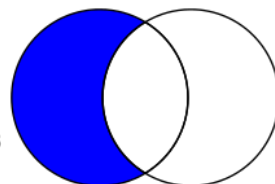


SQL JOINS

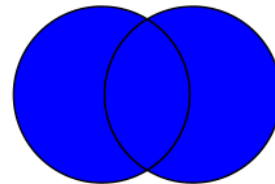
```
SELECT <fields>
FROM TableA A
RIGHT JOIN TableB B
ON A.key = B.key
WHERE A.key IS NULL
```



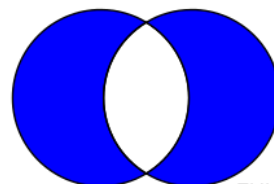
```
SELECT <fields>
FROM TableA A
LEFT JOIN TableB B
ON A.key = B.key
WHERE B.key IS NULL
```



```
SELECT <fields>
FROM TableA A
FULL OUTER JOIN TableB B
ON A.key = B.key
```



```
SELECT <fields>
FROM TableA A
FULL OUTER JOIN TableB B
ON A.key = B.key
WHERE A.key IS NULL
OR B.key IS NULL
```



OPENROWSET

- Permite ler arquivos de diversas fontes:
 - Para acessar outros bancos de dados → preferível usar linked server

```
SELECT a.*
FROM OPENROWSET('SQLNCLI', 'Server=Seattle1;Trusted_Connection=yes;',
'SELECT GroupName, Name, DepartmentID
FROM AdventureWorks2012.HumanResources.Department
ORDER BY GroupName, Name') AS a;
```

```
USE Northwind;
GO
SELECT c.*, o.*
FROM Northwind.dbo.Customers AS c
INNER JOIN OPENROWSET('Microsoft.Jet.OLEDB.4.0',
'C:\Program Files\Microsoft Office\OFFICE11\SAMPLES\Northwind.mdb';'admin';'', Orders)
AS o
ON c.CustomerID = o.CustomerID ;
```

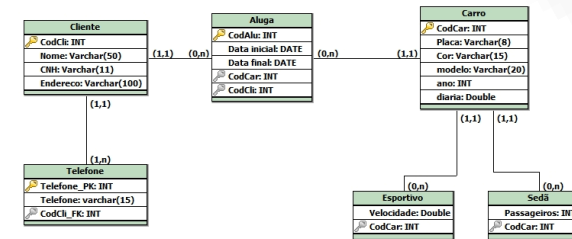
Tipos de Dados

■ Dados Estruturados

- Organizados e representados com uma estrutura rígida;
- Definição prévia de um modelo de dados.

■ Dados Não Estruturados

- Não são estruturados por meio de modelos ou esquemas de dados predefinidos;
- Arquivos de imagens, áudio, vídeo, etc.



Tipos de Dados

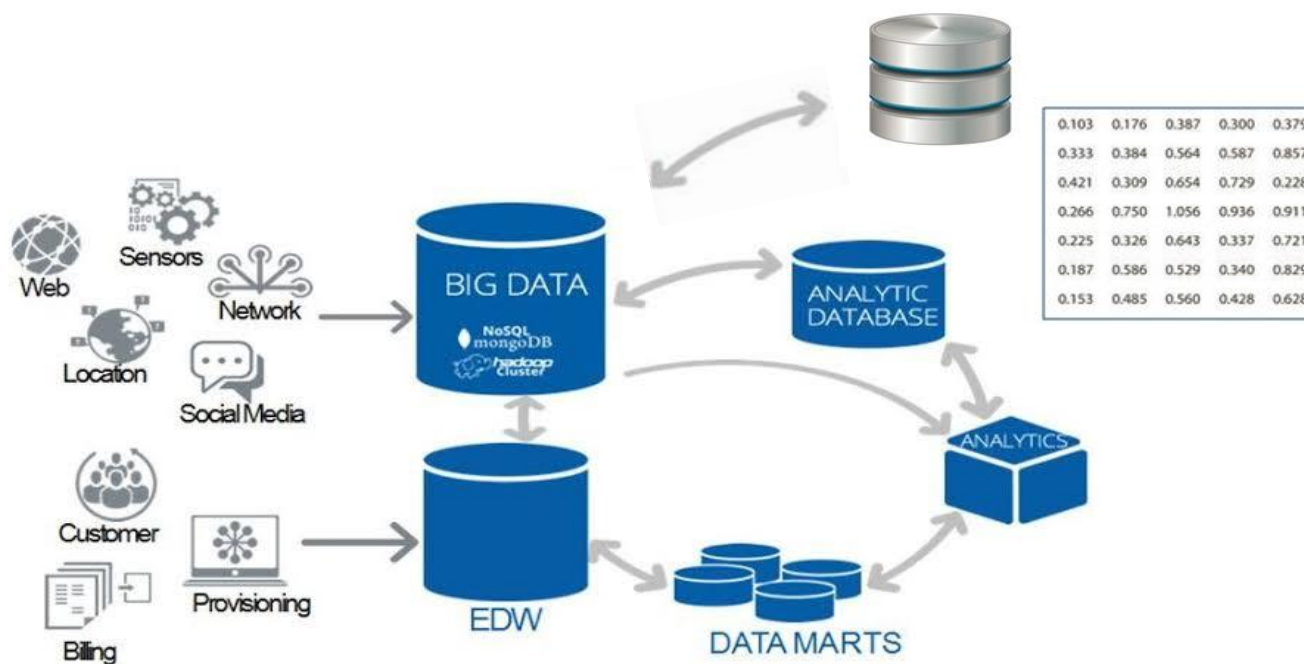
■ Dados Semiestruturados

- Não contém toda a rigidez requerida na definição dos tipos de dados estruturados;
- Procuram manter certa uniformidade no armazenamento das informações
 - Mantém tags e marcações internas que identificam elementos de dados separados.

```
{  
  name: "sue",  
  age: 26,  
  status: "A",  
  groups: [ "news", "sports" ]  
}
```

← field: value
← field: value
← field: value
← field: value

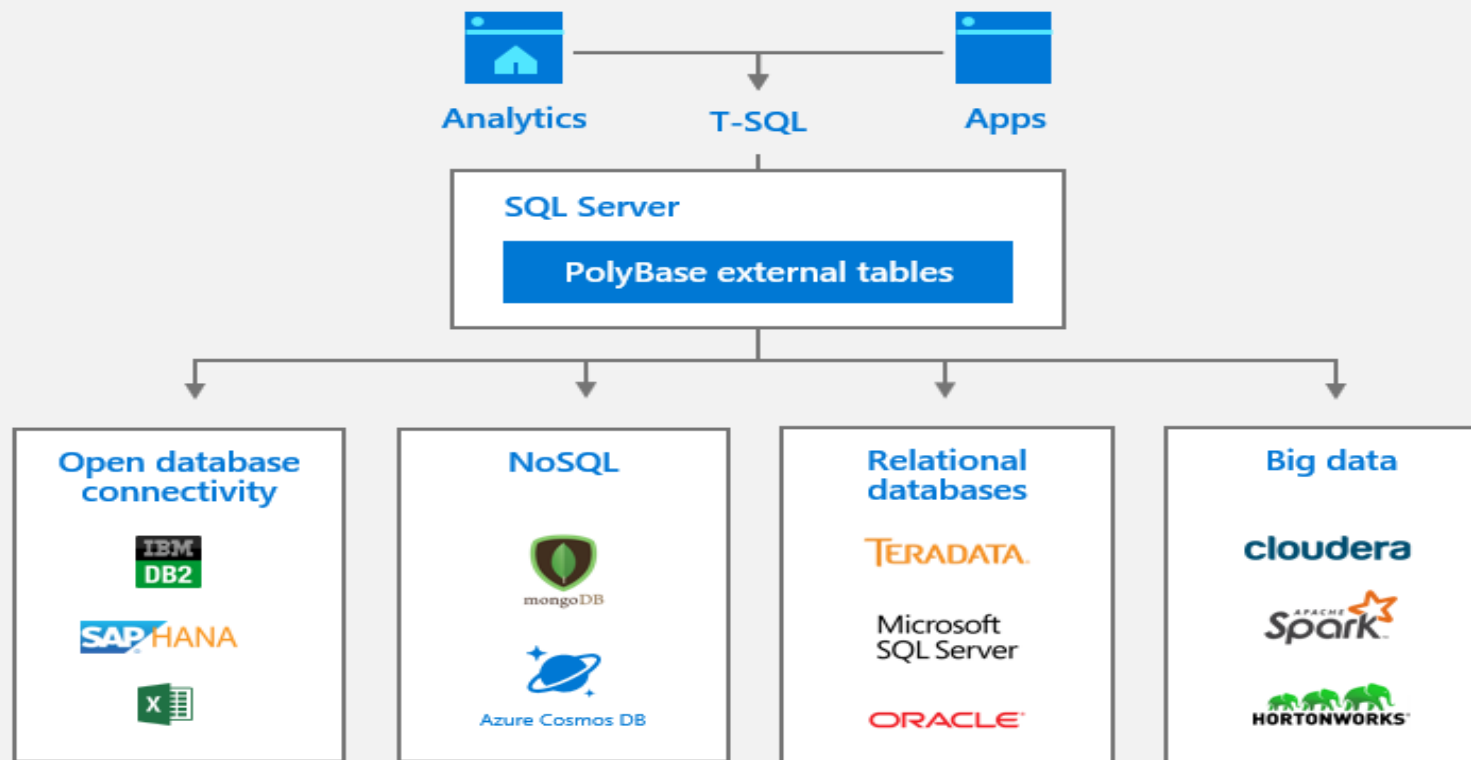
Como fazer a integração de dados estruturados com dados não estruturados da forma mais simples, econômica e ágil possível?



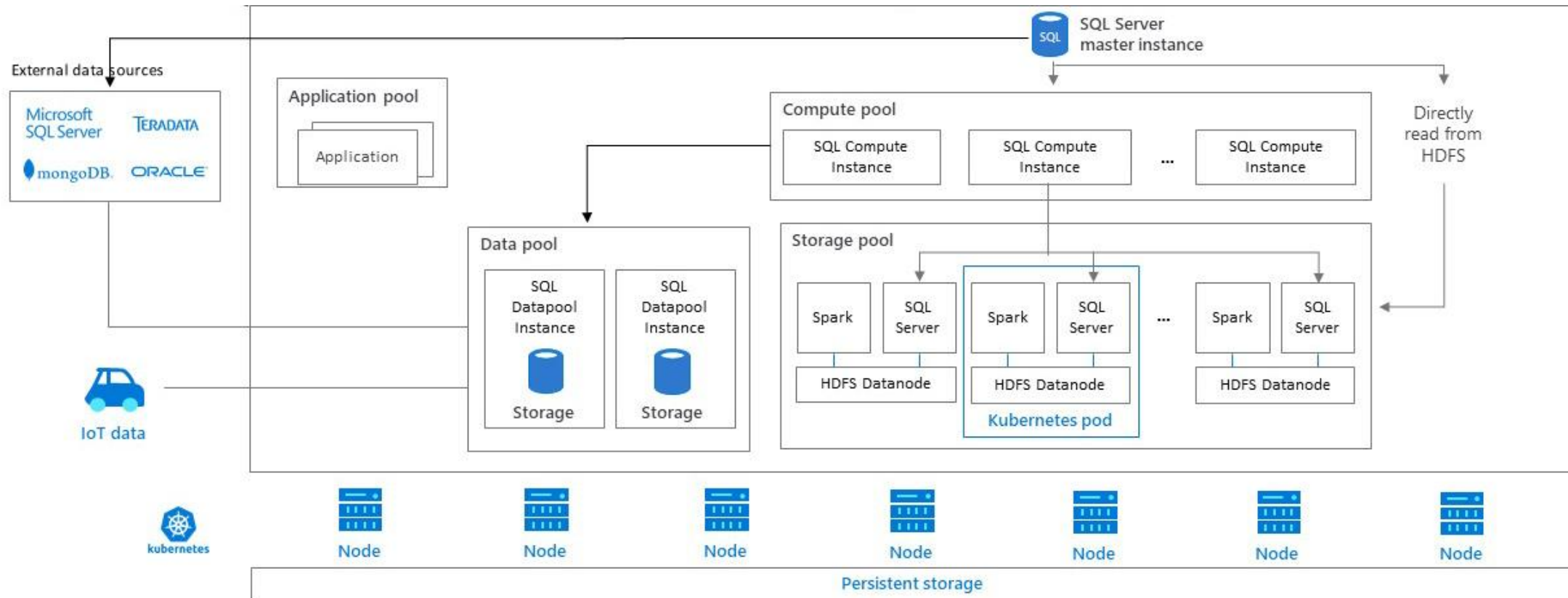
PolyBase



- Recurso que permite que uma instância do SQL Server processe consultas escritas em sua linguagem nativa (Transact-SQL), mas que leem dados de fontes de dados externas.



SQL Server Big Data Cluster



SQL Server Big Data Cluster



```
Windows PowerShell
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

PS C:\Users\gusta> azdata login --cluster-name mssql-cluster
Controller Username: admin
Controller Password:
Logged in successfully to `https://13.72.108.63:30080`
PS C:\Users\gusta> azdata bdc status show

Mssql-cluster: ready
=====
Services: ready
-----


| ServiceName | State | Healthstatus | Details |
|-------------|-------|--------------|---------|
| sql         | ready | healthy      | -       |
| hdfs        | ready | healthy      | -       |
| spark       | ready | healthy      | -       |
| control     | ready | healthy      | -       |
| gateway     | ready | healthy      | -       |
| app         | ready | healthy      | -       |



Sql Services: ready
-----


| Resourcename | State | Healthstatus | Details                          |
|--------------|-------|--------------|----------------------------------|
| master       | ready | healthy      | StatefulSet master is healthy    |
| compute-0    | ready | healthy      | StatefulSet compute-0 is healthy |
| data-0       | ready | healthy      | StatefulSet data-0 is healthy    |
| storage-0    | ready | healthy      | StatefulSet storage-0 is healthy |



Hdfs Services: ready
-----


| Resourcename | State | Healthstatus | Details                          |
|--------------|-------|--------------|----------------------------------|
| nmnode-0     | ready | healthy      | StatefulSet nmnode-0 is healthy  |
| storage-0    | ready | healthy      | StatefulSet storage-0 is healthy |
| sparkhead    | ready | healthy      | StatefulSet sparkhead is healthy |



Spark Services: ready
-----


| Resourcename | State | Healthstatus | Details                          |
|--------------|-------|--------------|----------------------------------|
| sparkhead    | ready | healthy      | StatefulSet sparkhead is healthy |
| storage-0    | ready | healthy      | StatefulSet storage-0 is healthy |


```

PolyBase



- **External Table:** objeto que permite que uma mesma consulta acesse dados externos ao SQL.
- Possível fazer joins entre external tables e tabelas no banco de dados relacional SQL Server, Oracle, MySQL etc.

```
CREATE EXTERNAL FILE FORMAT formato_csv
WITH (
    FORMAT_TYPE = DELIMITEDTEXT,
    FORMAT_OPTIONS (
        FIELD_TERMINATOR = ';',
        STRING_DELIMITER = '"',
        FIRST_ROW = 2,
        USE_TYPE_DEFAULT = TRUE
    )
);
```

```
CREATE EXTERNAL DATA SOURCE FonteHDFS
WITH (LOCATION = 'sqlhdfs://controller-svc/default');
```

```
CREATE EXTERNAL TABLE HDFS_clickstreams
(
    "data_do_clique_no_item"          BIGINT,
    "horario_do_clique_no_item"      BIGINT,
    "codigo_do_item"                  BIGINT,
    "endereco_da_pagina"              BIGINT,
    "nome_usuario"                    BIGINT
)
WITH ( DATA_SOURCE = FonteHDFS,
        LOCATION = '/clickstream_data',
        FILE_FORMAT = formato_csv );
```

PolyBase



• SQLQuery_1 - 13.82.56.237,31433:sales (sa) - Azure Data Studio - Insiders

13.82.56.237,31433:sales | SQLQuery_1 - 13.82....es (sa) | deploy-bdc-aks-1 | deploy-bdc

Run Cancel Disconnect Change Connection sales Explain Enable SQLCMD Mod

```
1 SELECT TOP (100) *
2 FROM HDFS_clickstreams
3 ORDER BY data_do_clique_no_item DESC, horario_do_clique_no_item DESC;
4
```

Messages Results

	data_do_clique_no_item	horario_do_clique_no_item	codigo_do_item	endereco_da_pagina	nome_usuario
1	38697	86380	NULL	12560	2
2	38697	86305	NULL	700	12
3	38697	86280	NULL	7489	2
4	38697	86124	NULL	15412	5
5	38697	86120	NULL	8601	57
6	38697	86040	NULL	14377	2
7	38697	86023	NULL	8978	54
8	38697	85941	NULL	1724	19
9	38697	85906	NULL	8352	1
...	38697	85772	NULL	4544	34
...	38697	85762	NULL	8053	16
...	38697	85757	NULL	11737	57
...	38697	85741	NULL	17176	45
...	38697	85728	NULL	12951	53

PolyBase



SQLQuery_1 - 13.82.56.237,31433.sales (sa) - Azure Data Studio - Insiders

13.82.56.237,31433:sales SQLQuery_1 - 13.82....es (sa) deploy-bdc-aks-1

Run Cancel Disconnect Change Connection sales Explain E

```
1 SELECT I.nome_categoria AS 'Categoria',  
2      COUNT (C.codigo_do_item) AS 'Quantidade de Cliques'  
3 FROM itens I JOIN HDFS_clickstreams C  
4 ON I.codigo_do_item = C.codigo_do_item  
5 GROUP BY I.nome_categoria  
6 ORDER BY 'Quantidade de Cliques';  
7  
8  
9
```

Messages Results

	Categoria	Quantidade de C...
1	Toys & Games	2853
2	Tools & Home Improvement	3136
3	Sports & Outdoors	3502
4	Home & Kitchen	3564
5	Clothing & Accessories	4583
6	Electronics	5138
7	Movies & TV	9781
8	Books	40080
9	Music	61333



Quiz Premiado

- Vale-presente (gift card) da **Americanas.com**, no valor de **R\$25,00**, para o **primeiro colocado**.
- Acesse o site **www.kahoot.it** ou o app **Kahoot!**
- Digite o Game PIN: **??????**

Kahoot!