MOXDULO 1

Descritivo => Descrever o passado.

Prospectivo => Acompanha os dados.

Preditivo => Adivinhar.

# Descrever a diferença entre dados de lote e dados de streaming

O processamento de dados conforme eles chegam é chamado de streaming. O armazenamento dos dados em buffer e processamento em grupos é chamado de processamento em lotes.

## Conceitos básicos de IaaS e PaaS

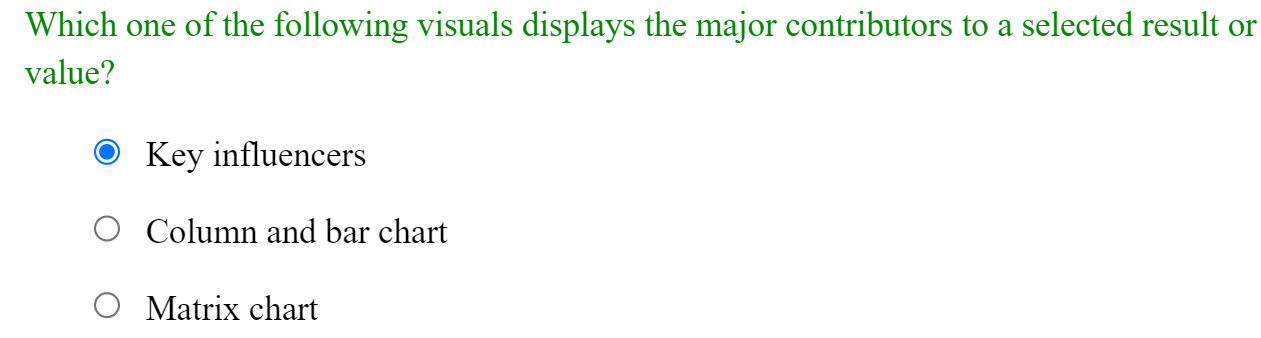
**IaaS** é a abreviação de infraestrutura como serviço. O Azure permite criar uma infraestrutura virtual na nuvem que copia o funcionamento de um data center local.

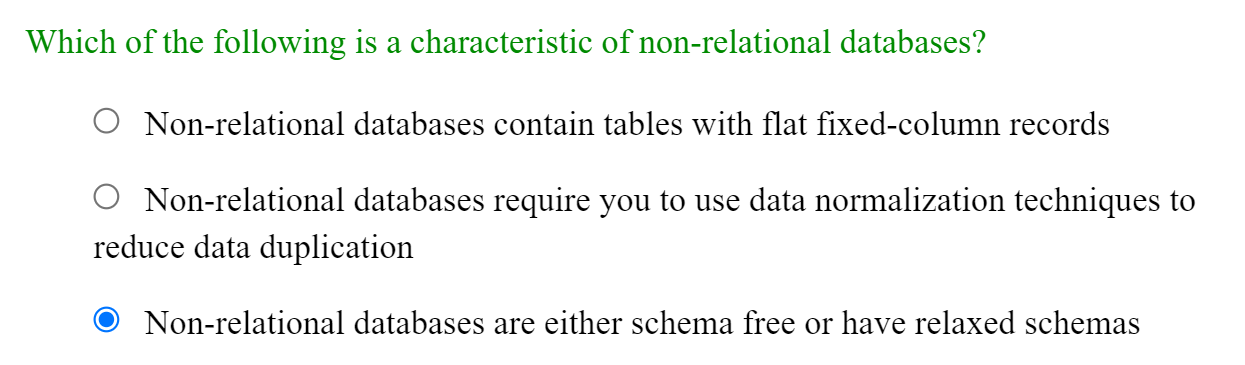
**PaaS** significa plataforma como serviço. Uma solução PaaS pode criar uma infraestrutura virtual e instalar e gerenciar o software de banco de dados para você

ETL significa extração, transformação e carregamento. Os dados brutos são recuperados e transformados antes de serem salvos.  Essa abordagem é adequada para sistemas que exigem apenas modelos simples, com pouca dependência entre os itens.

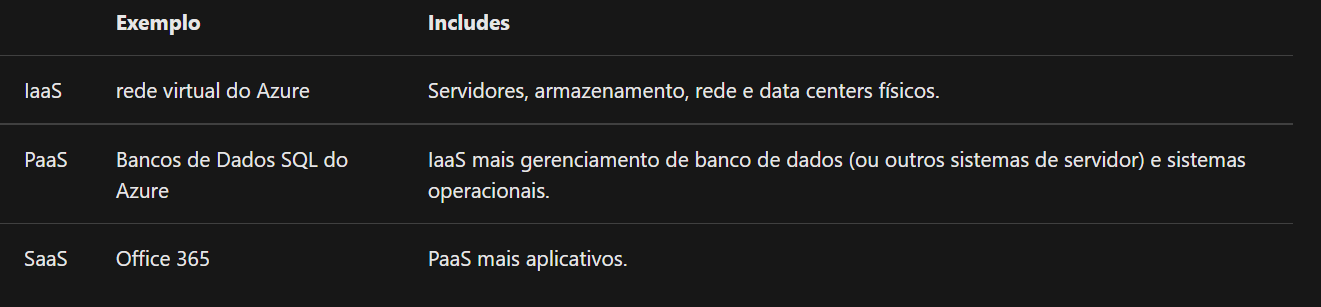
ELT é uma abreviação de extrair, carregar e transformar. O processo difere do ETL, pois os dados são armazenados antes de serem transformados. O ELT é mais adequado para construir modelos complexos que dependem de vários itens do banco de dados, geralmente usando processamento em lotes periódicos.

Usando o Azure Data Factory, é possível criar e agendar fluxos de trabalho orientados a dados (chamados de pipelines)

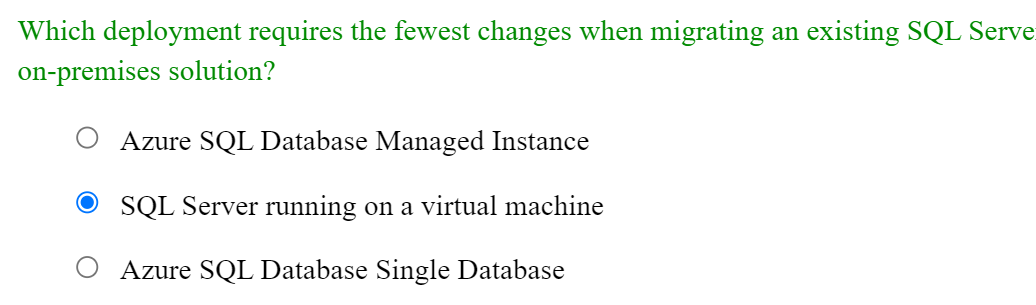




Modulo 2



Os Serviços de Dados do Azure se enquadram na categoria PaaS



MODULO 3

O Armazenamento de Tabelas do Azure implementa o modelo de chave-valor NoSQL.

Uma tabela do Azure permite que você armazene dados semiestruturados.

## Armazenamento de Blobs do Azure?

serviço que permite armazenar grandes quantidades de dados não estruturados (ou blobs) na nuvem.

Blobs de blocos. => Os blobs de blocos são mais bem usados para armazenar objetos binários, grandes e discretos que são alterados com pouca frequência.

Blobs de páginas

Blobs de acréscimo

## O que é provisionamento?

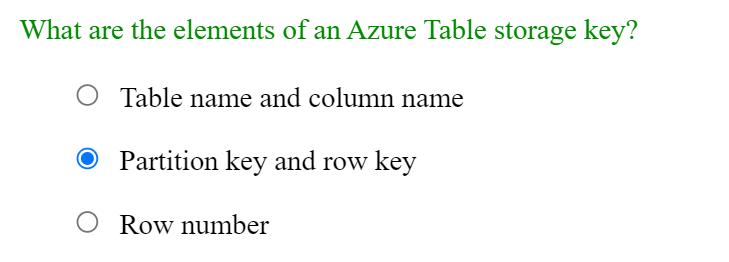
O provisionamento é o ato de executar uma série de tarefas que um provedor de serviços, como o Azure Cosmos DB, realiza para criar e configurar um serviço

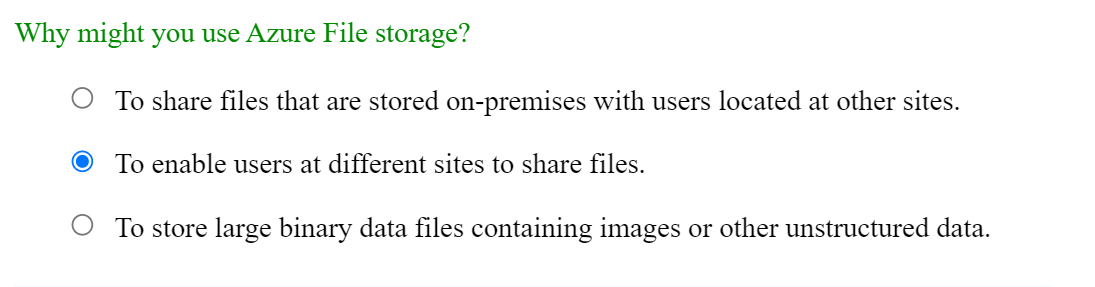
#### Azure File Storage

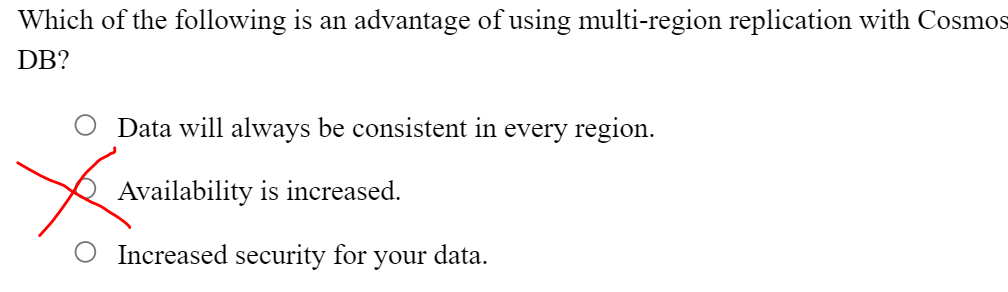
permite que você crie compartilhamentos de arquivo na nuvem e acesse esses compartilhamentos de arquivo de qualquer lugar com uma conexão com a Internet

Cosmos DB => sistema de gerenciamento de banco de dados NoSQL de vários modelos.

Um documento pode conter até 2 MB de dados, incluindo pequenos objetos binários. Se você precisar armazenar blobs maiores como parte de um documento, use o Armazenamento de Blobs do Azure e adicione uma referência ao blob no documento.







MODULO 4

Azure Synapse Analytics => To perform very complex queries and aggregations.

serviço de análise que reúne data warehouse corporativo e análise de Big Data. O pool de SQL dedicado (antigo SQL DW) refere-se aos recursos de data warehouse empresariais que estão disponíveis no Azure Synapse Analytics.

data ingestion => o capture data flowing into a data warehouse system as quickly as possible

 data lake => Raw data

Data WareHouse => **structured information**

Azure Data Factory

* serviço de integração de dados sem servidor totalmente gerenciado. Integre visualmente as fontes de dados usando mais de 90 conectores internos livres de manutenção sem custo adicional. Construa processos ETL e ELT sem código com facilidade em um ambiente visual intuitivo ou escreva seu próprio código. Em seguida, forneça dados integrados ao Azure Synapse Analytics para desbloquear insights de negócios.

Which component of an Azure Data Factory can be triggered to run data ingestion tasks? Pipeline

PolyBase => permite que sua instância SQL Server consulte dados com o T-SQL diretamente do SQL Server, Oracle, Teradata, MongoDB, clusters Hadoop, Cosmos DB sem instalar separadamente o software de conexão cliente.

Azure DataBrink => ambiente do Apache Spark em execução no Azure para fornecer processamento de Big Data, streaming e machine learning.

## Azure Analysis Services?

 permite que você crie modelos de tabela para dar suporte a consultas OLAP (processamento analítico online).

## Azure HDInsight

serviço de processamento de Big Data, que fornece a plataforma para tecnologias como o Spark em um ambiente do Azure.

