**Curso – Microsoft Azure Data Fundamentals – DP-900 (Aula 1 )**

**\*Definição de Dados**

Coleção de fatos, números, descrições, objetos, armazenados de maneira estruturada, semiestruturada, não estruturada.

Exemplos:

Dados estruturados: Tabelas, linhas e colunas.

Dados semiestruturados: Informações que não residem em um banco de dados relacionais. Exemplos através de Jaison.

Dados Não estruturados: Através de arquivos e não mantém uma padronização.

**\*Armazenamento de dados transacional versus analítico**

Processamento transacional on-line (OLTP) = Os dados são armazenados em uma transação por vez.

Processamento analítico on-line (OLAP) = Os dados são periodicamente carregados, agregados e armazenados em cubo.

**Cargas de trabalho transicionais**

Os dados transacionais são informações que rastreiam as interações relacionadas às atividades de uma organização.

* **Atomicidade** – cada transação é tratada como uma unidade independente que resulta em sucesso completo ou falha completa.
* **Consistência** – as transações só podem conduzir os dados do banco de dados de um estado válido para outro estado válido.
* **Isolamento** – a execução concorrente de transações deixa o banco de dados no mesmo estado.
* **Durabilidade** – assim que uma transação tiver sido confirmada, permanecerá assim.

**Cargas de trabalho analíticas**

As cargas de trabalho analíticas são usadas para análise de dados e tomada de decisões.

* Resumos
* Tendências
* Informações comerciais

Processamento de dados

O processamento de dados é a conversão de dados brutos em informações relevantes por meio de um processo.

Processamento em lotes: Os elementos de dados são reunidos em um grupo. Então, o grupo inteiro é processado em momento futuro como um lote.

Processamento de streaming: cada novo dado é processado quando chega.

**Explorar funções e responsabilidades no mundo dos dados**

Funções em dados

**\*Administrador de Banco de Dados**

Gerenciamento de banco de dados

Segurança de dados de implementos

Backups

Acesso do usuário

Desempenho de monitores

**\*Engenheiro de dados**

Pipelines e processos de dados

Armazenamento de ingestão de dados

Preparar dados para análise

Preparar dados para processamento de análise

**\*Analista de dados**

Oferece insights sobre os dados

Geração de relatórios visuais

Modelagem de dados para análise

Combina dados para visualização e análise

**Ferramentas comuns – Administrador de banco de dados**

**\*Azure Data Studio**

Interface gráfica para gerenciar serviços de dados no local e baseados na nuvem

Funciona no Windows, no macOS e no Linux

**\*SQL Server Management Studio**

Interface gráfica para gerenciar serviços de dados no local e baseados na nuvem

Funciona no Windows

Ferramenta de administração de banco de dados abrangente

**\*Portal/CLI do Azure**

Ferramentas para gerenciamento e provisionamento de serviços de dados do Azure

Manual e automação de scripts usando o Azure Resource Manager ou criação de scripts em interface de linha de comando

**Ferramentas comuns – Engenharia de dados**

**\*Azure Synapse Studio**

Portal do Azure integrado para gerenciar o Azure Synapse

Ingestão de dados (Azure Data Factory)

Gerenciamento de ativos do Azure Synapse (Pool de SQL/Pool do Spark)

**\*SQL Server Management Studio**

Interface para gerenciar serviços de dados no local e baseados na nuvem

Ferramenta de administração de banco de dados abrangente

**\*Portal/CLI do Azure**

Ferramentas para gerenciamento e provisionamento de recursos do Azure

Manual e automação de scripts usando o Azure Resource Manager ou criação de scripts em interface de linha de comando

**Ferramentas comuns – Analista de dados**

**\*Power Bi Desktop**

Ferramenta de visualização de dados

Modele e visualiza dados

Gerenciamento de ativos do Azure Synapse (Pool de SQL/Pool do Spark)

**\*Power Bi Portal/serviço do Power Bi**

Criação e gerenciamento de relatórios do Power BI

Compartilhe relatórios/conjuntos de dados

**\*Power BI Report Builder**

Ferramenta de visualização de dados para relatórios paginados

Modele e visualiza relatórios paginados

**Curso – Microsoft Azure Data Fundamentals – DP-900 (Aula 2)**

Descrever os conceitos de dados relacionais

**\*Identificar casos de uso do banco de dados relacional**

IoT: Embora sejam normalmente considerados como não relacionais, os dados de dispositivos IoT poderiam ser estruturados e consistentes

Processamento de transação on-line: Por exemplo, sistemas de pedidos que realizam muitas atualizações transacionais pequenas

Data warehousing: Grande quantidade de dados podem ser importadas de várias fontes e estruturadas para permitir consultas de alto desempenho

Funcionamento de banco de dados relacional

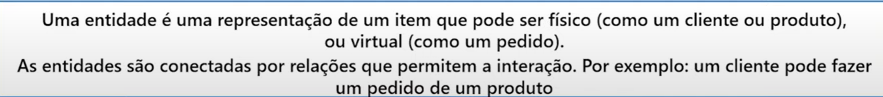
Os dados são armazenados em uma tabela

A tabela consiste em linhas e colunas

Todas as linhas têm o mesmo número de colunas

Cada coluna é definida por um tipo de dados

Observação:



Normalização

**\*Os dados são normalizados para:**

Reduzir armazenamento Evitar a duplicação de dados Melhorar a quantidade dos dados

**\*Em um esquema de bancos de dados normalizado:**

As chaves primárias e estrangeiras são usadas para definir relacionamentos

Não há nenhuma duplicação de dados (a não ser os valores chave no 3º formulário norma (3FN)

Os dados são recuperados unindo as tabelas em uma consulta

**\*Um índice**

Otimiza consultas de pesquisa para recuperação mais rápida de dados

Reduz a quantidade de páginas de dados que precisam ser lidas para recuperar os dados em uma Declaração SQL

Os dados são recuperados unindo as tabelas em uma consulta

**\*Exibição**

Uma exibição é uma tabela virtual com base no conjunto de resultados de uma consulta:

As exibições são criadas para simplificar a consulta

Combine dados relacionados em uma exibição de painel simples

Descrever os conceitos de dados não relacionais

Características dos dados não relacionais

**\*Coleções não relacionais podem ter:**

Várias entidades na mesma coleção ou contêiner com diferentes campos

Têm um esquema diferente, não tabular

São frequentemente definidos por identificar cada campo com o nome que ele representa

**\*Identificar casos de uso do banco de dados não relacional**

IoT e telemática: Com frequência exigem ingerir grandes quantidades de dados em frequentes explosões de atividades, os dados são semiestruturados ou estruturados, com frequência exigem processamento em tempo real

Varejo e marketing: Cenários comuns para dados distribuídos globalmente, armazenamento de documentos

Jogos: Estatísticas de jogo, integração com mídias sociais, quadros de líderes, aplicativos de baixa latência

Web e móvel: Normalmente usados com análises de cliques na web, aplicativos modernos incluindo bots