

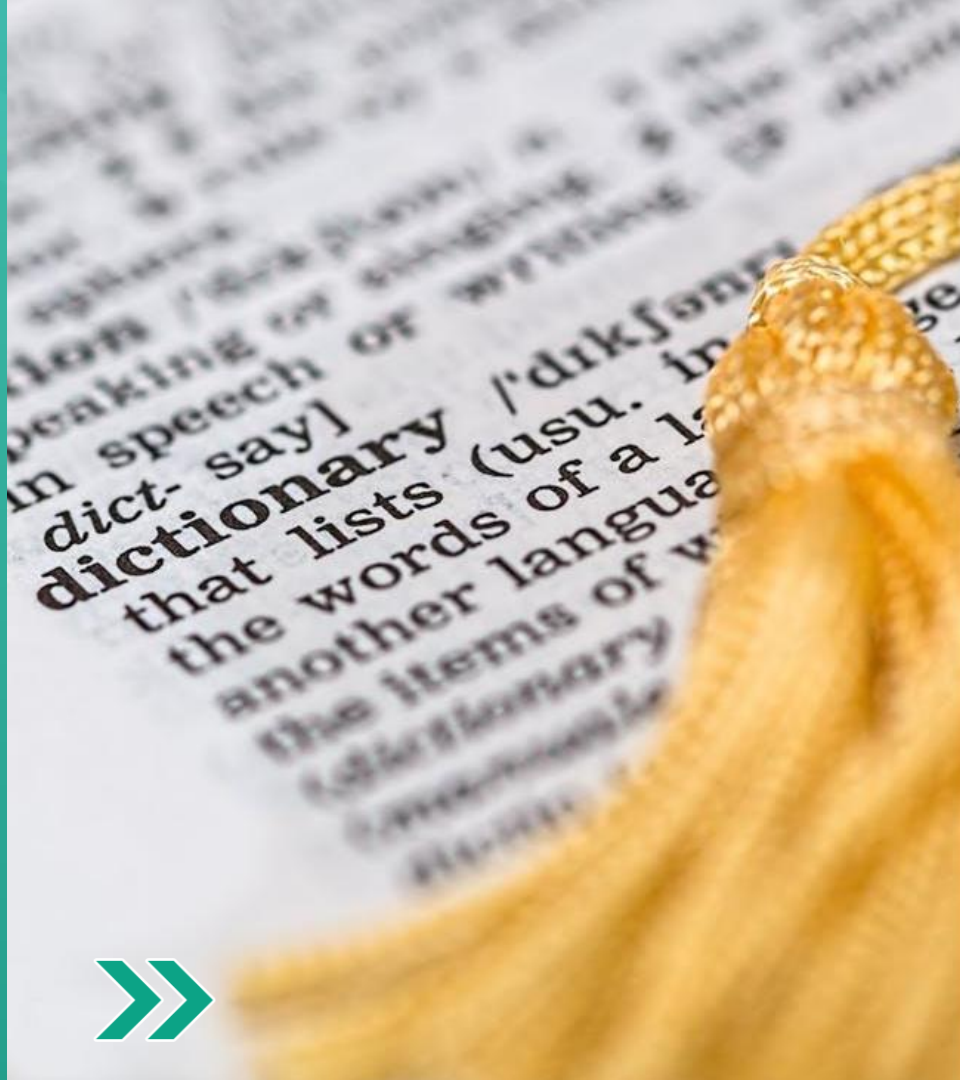


escola
britânica de
artes criativas
& tecnologia

Profissão: Analista de dados



GLOSSÁRIO



Big Data II – Processamento



Dica: para encontrar rapidamente a palavra que procura aperte o comando CTRL+F e digite o termo que deseja achar.

- Banco de dados colunares
- Chave de partição
- Escalabilidade horizontal
- Particionamento horizontal
- Particionamento vertical
- Processamento paralelo



• Banco de dados colunares

É um tipo de sistema de gerenciamento de banco de dados (SGBD) que armazena os dados de forma organizada por colunas, em vez de por linhas. Isso significa que, em vez de armazenar todos os valores de uma tupla (linha) juntos, um banco de dados colunar armazena os valores de cada coluna juntos.

• Escalabilidade horizontal

É a capacidade de aumentar a capacidade de um sistema adicionando mais recursos, como servidores, sem que isso afete o desempenho ou a disponibilidade do sistema.

• Chave de partição

É um identificador único usado para determinar a localização dos dados em um sistema distribuído. É usado para particionar os dados em diferentes nós ou servidores e garantir que os dados sejam distribuídos uniformemente.

• Particionamento horizontal

É uma técnica de particionamento de dados em que os dados são divididos em partes menores e distribuídos em diferentes nós em um cluster ou em diferentes servidores em uma rede. Essa técnica é usada para permitir que um banco de dados possa lidar com grandes quantidades de dados e aumentar a escalabilidade horizontal do sistema.



● **Particionamento vertical**

É uma técnica de particionamento de dados em que os dados são divididos em colunas, em vez de linhas, e distribuídos em diferentes tabelas ou servidores. Essa técnica é usada para melhorar o desempenho em operações que envolvem apenas um subconjunto de colunas, reduzindo o tamanho da tabela e permitindo que as consultas sejam executadas mais rapidamente.

● **Processamento paralelo**

É a execução simultânea de múltiplas tarefas em um sistema distribuído ou em um cluster de computadores para processar grandes volumes de dados de maneira mais rápida e eficiente. Cada tarefa é executada em um nó ou servidor diferente, e os resultados são combinados no final para produzir o resultado final.



Bons estudos!

