

SQL WEEK

RECIFE

Conference - 6ª edição

05/02 até 09/02



APOIO

- **LS Treinamentos**
- **Consulta BD**
- **Lucrécio Paes (CO-HOST)**



Treinamentos



QUEM SOU EU?

- + 12 anos na área de TI/Dados
- Formada em Análise e Desenvolvimento de Sistemas
- Azure Data Engineer na Dataside
- 4 Certificações Microsoft
- Admin e escritora do blog Big Data Navigators
- Hard Skills: Modelagem de dados, BI, Big Data, SQL, Python, Databricks, Azure (Azure Data Factory, Azure Data Lake, Azure Synapse Analytics, Azure DevOps...)



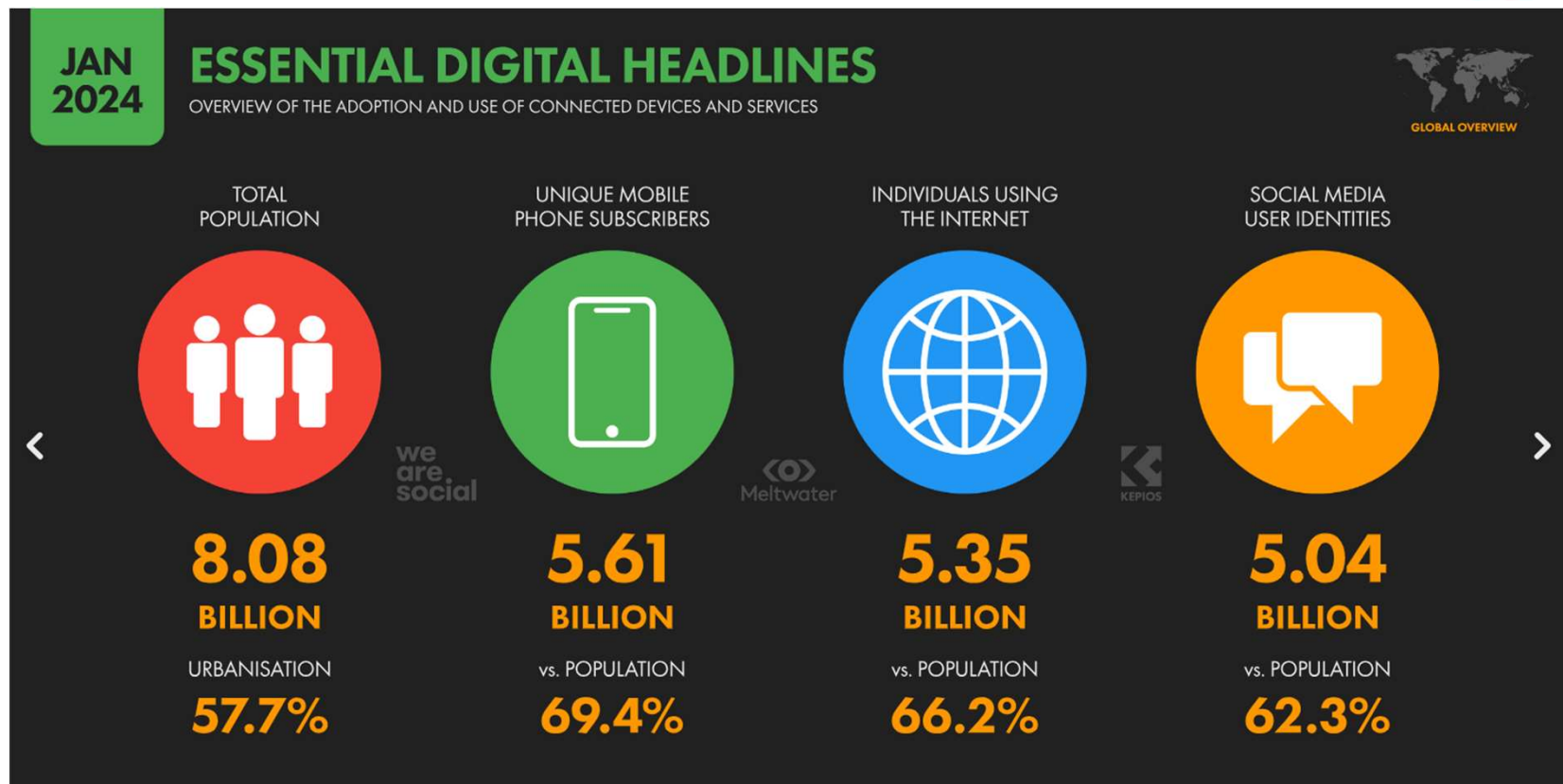
Rayza Wilma





TECNOLOGIAS E FERRAMENTAS UTILIZADAS EM BIG DATA

Algumas análises...



DataReportal – Global Digital Insights

JAN
2024

OVERVIEW OF INTERNET USE

ESSENTIAL INDICATORS OF INTERNET ADOPTION AND USE



INDIVIDUALS
USING THE
INTERNET



5.35
BILLION



INDIVIDUALS USING THE
INTERNET AS A PERCENTAGE
OF TOTAL POPULATION



66.2%
YOY: +0.9% (+60 BPS)



YEAR-ON-YEAR CHANGE IN
THE NUMBER OF INDIVIDUALS
USING THE INTERNET



+1.8%
+97 MILLION



PERCENTAGE OF THE
TOTAL FEMALE POPULATION
THAT USES THE INTERNET



63.5%
YOY: +4.9% (+304 BPS)



PERCENTAGE OF THE
TOTAL MALE POPULATION
THAT USES THE INTERNET



68.8%
YOY: +4.2% (+285 BPS)

AVERAGE DAILY TIME
SPENT USING THE INTERNET
BY EACH INTERNET USER



6H 40M
YOY: +0.8% (+3 MINS)



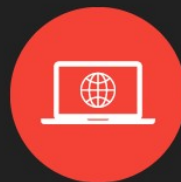
PERCENTAGE OF USERS
ACCESSING THE INTERNET
VIA MOBILE PHONES



96.5%
YOY: +4.6% (+420 BPS)



PERCENTAGE OF USERS
ACCESSING THE INTERNET
VIA LAPTOPS AND DESKTOPS



61.8%
YOY: -5.8% (-380 BPS)



PERCENTAGE OF THE
TOTAL URBAN POPULATION
THAT USES THE INTERNET



78.8%
YOY: +3.2% (+252 BPS)



PERCENTAGE OF THE
TOTAL RURAL POPULATION
THAT USES THE INTERNET



48.9%
YOY: +7.2% (+340 BPS)

SOURCES: KEPIOS ANALYSIS; ITU; GSMA INTELLIGENCE; EUROSTAT; GOOGLE'S ADVERTISING RESOURCES; CNNIC; KANTAR & IAMA; GOVERNMENT RESOURCES; UNITED NATIONS. TIME SPENT AND MOBILE SHARE DATA FROM GWI (Q3 2023). SEE [GWI.COM](https://www.gwi.com). **NOTES:** GENDER DATA ARE ONLY AVAILABLE FOR "FEMALE" AND "MALE". PERCENTAGE CHANGE FIGURES SHOW RELATIVE YEAR-ON-YEAR CHANGE. "BPS" FIGURES REPRESENT BASIS POINTS, AND SHOW ABSOLUTE YEAR-ON-YEAR CHANGE. **COMPARABILITY:** SOURCE AND BASE CHANGES. ALL FIGURES USE THE LATEST AVAILABLE DATA, BUT SOME SOURCES DO NOT PUBLISH REGULAR UPDATES, SO FIGURES MAY UNDER-REPRESENT ACTUAL USE. SEE [NOTES ON DATA](#).

we
are
social

Meltwater

JAN
2024

DEVICE OWNERSHIP

PERCENTAGE OF INTERNET USERS AGED 16 TO 64 WHO OWN EACH KIND OF DEVICE



ANY KIND OF
MOBILE PHONE



GW.

97.8%

YEAR-ON-YEAR CHANGE
+1.7% (+160 BPS)

SMART
PHONE



we
are
social

97.6%

YEAR-ON-YEAR CHANGE
+1.8% (+170 BPS)

FEATURE
PHONE



GW.

6.9%

YEAR-ON-YEAR CHANGE
-12.7% (-100 BPS)

LAPTOP OR
DESKTOP COMPUTER



GW.

57.7%

YEAR-ON-YEAR CHANGE
-0.5% (-30 BPS)

TABLET
DEVICE



30.9%

YEAR-ON-YEAR CHANGE
-8.3% (-280 BPS)

GAMES
CONSOLE



Meltwater

19.1%

YEAR-ON-YEAR CHANGE
-5.9% (-120 BPS)

SMART WATCH OR
SMART WRISTBAND



GW.

30.1%

YEAR-ON-YEAR CHANGE
+0.7% (+20 BPS)

TV STREAMING
DEVICE



KEPIOS

15.7%

YEAR-ON-YEAR CHANGE
-4.8% (-80 BPS)

SMART HOME
DEVICE



GW.

16.2%

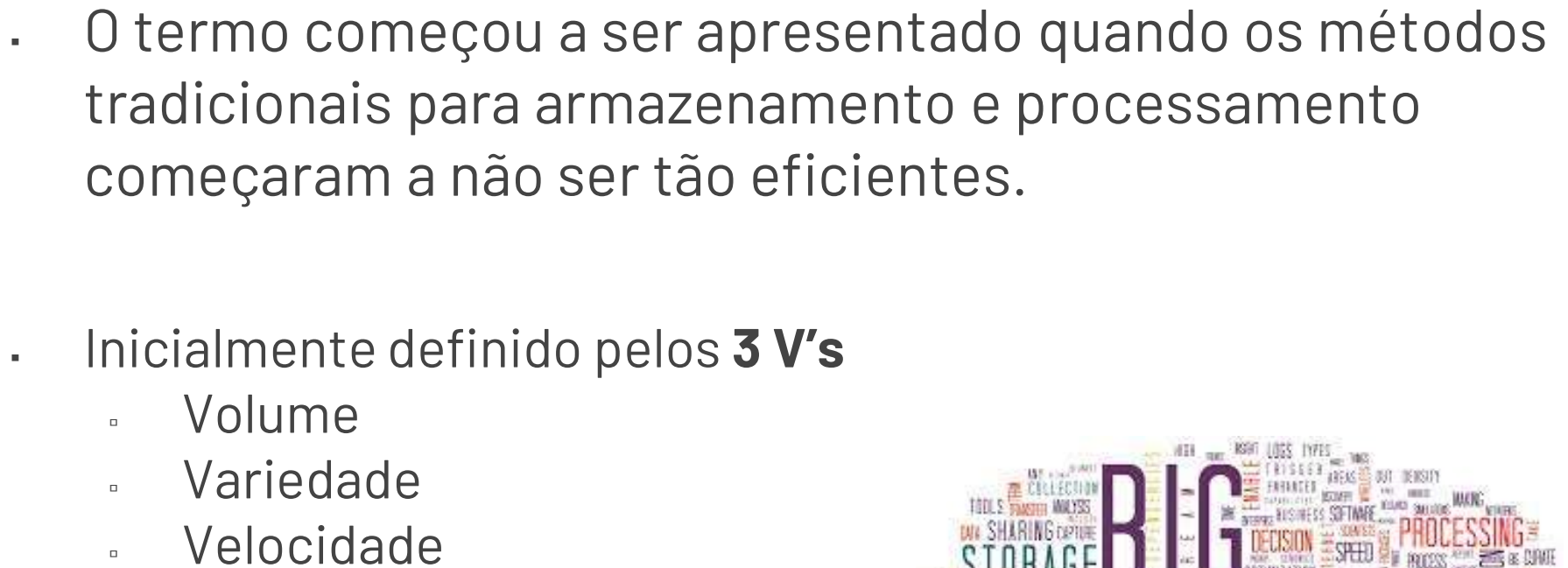
YEAR-ON-YEAR CHANGE
-1.2% (-20 BPS)

VIRTUAL REALITY
DEVICE



4.4%

YEAR-ON-YEAR CHANGE
-21.4% (-120 BPS)



5Vs do Big Data



Volume

Cada vez mais, produzimos informações em maior quantidade

Variedade

Bancos de dados, json, xml, mídias sociais, e-mails, IoT, Áudio e vídeo, arquivos de logs, sensores...

Valor

Vantagem competitiva,
aumento de receita,
economia de custos, satisfação de clientes internos e externos...

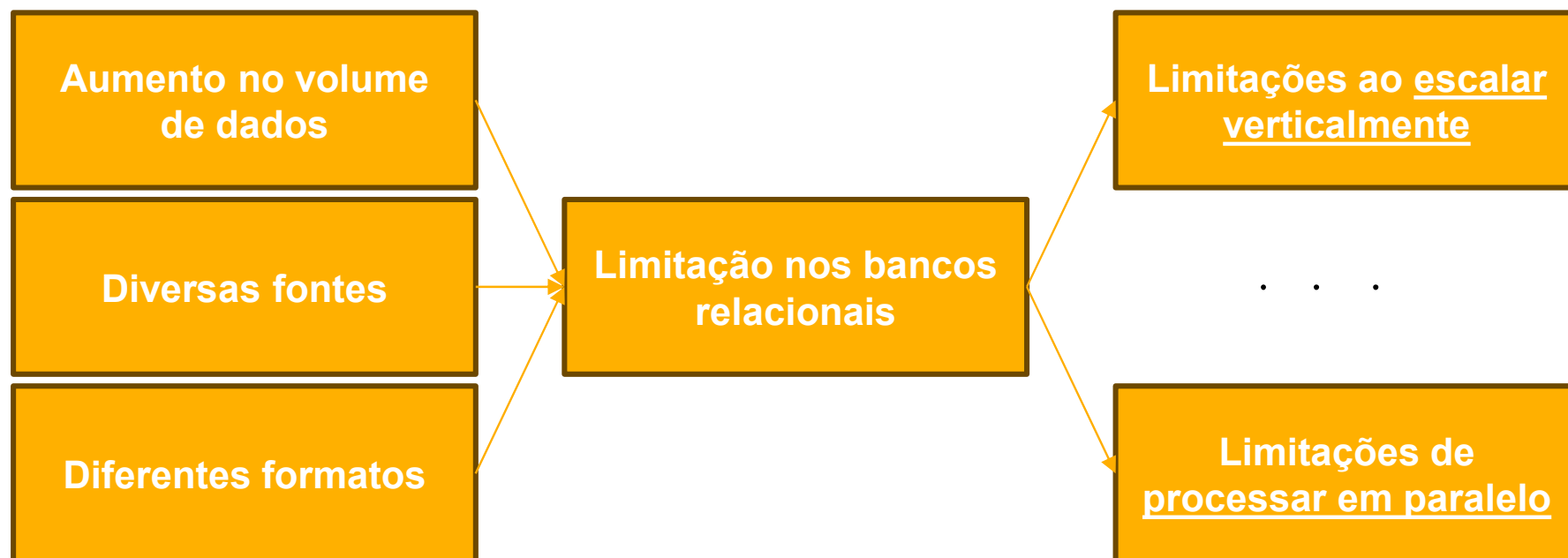
Velocidade

Com o avanço das tecnologias, a geração dos dados é mais rápida, logo, a tomada de decisão necessita ser mais rápida

Veracidade

Confiabilidade dos dados. Devem ter qualidade, consistência, origens conhecidas e não inventados (opinião)...

Algumas análises...



Logo, nos perguntamos...



- Onde **armazenar** esses dados?
- Como **processar** esses dados?
- Que **tecnologias** e **ferramentas** utilizar?
- Como conseguir explorar e gerar **valor** através desses **dados**?

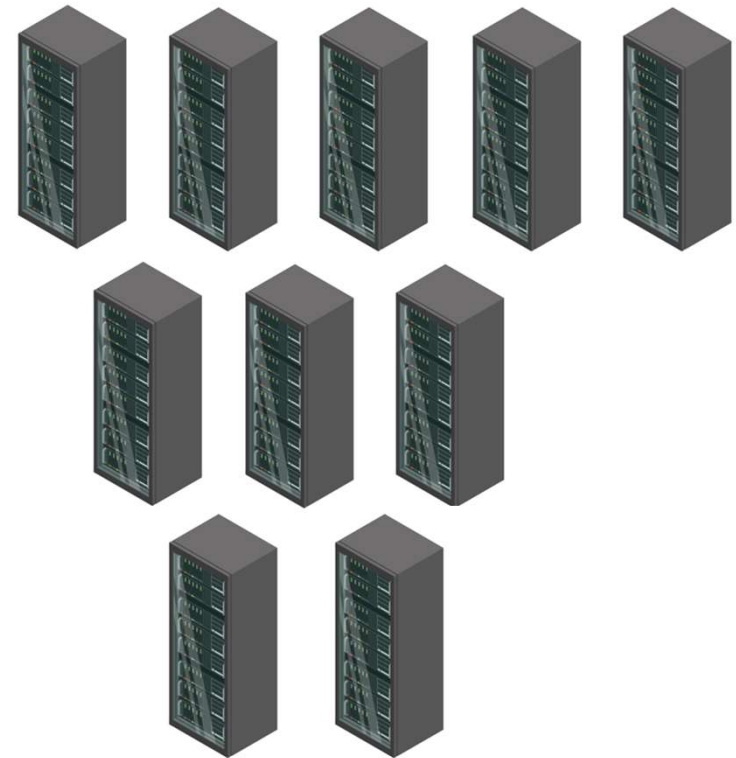
Escalabilidade Vertical x Horizontal



"Máquina Poderosa"



"Conjunto de máquinas realizando processamentos em paralelo"



Tecnologias em Big Data



- Conjunto de sistemas, práticas e tecnologias para lidar com grandes volumes de dados.
- Inclui infraestruturas de hardware (por exemplo, clusters de servidores), sistemas de armazenamento distribuído (como Hadoop Distributed File System - HDFS), sistemas de processamento distribuído (como Apache Spark) e outras tecnologias subjacentes.
- As tecnologias de big data fornecem a **base** para lidar com os 3Vs.

E as ferramentas em Big Data?



- São softwares ou aplicativos específicos que operam **sobre** as tecnologias de big data para realizar tarefas específicas de **análise**, **processamento** ou **visualização de dados**.

Tecnologia X Ferramentas em Big Data



As **tecnologias** de Big Data são como os materiais de construção, como tijolos, madeira e concreto.

As **ferramentas** de Big Data são como as ferramentas utilizadas para construir a casa, como martelos, serras e chaves de fenda.



Cloud Computing (Computação em nuvem)



- Oferece diferentes tipos de serviços baseados no provisionamento e disponibilização de recursos.
- Exemplos de provedores de serviços de nuvem
 - Amazon
 - Google
 - Microsoft



Cloud Computing (Computação em nuvem)



- Provisionamento rápido dos recursos
- Custo – Pay-per-use / Pay As You Go
- Acesso simplificado
- Manutenção simplificada
- Confiabilidade
- Escalabilidade e elasticidade virtualmente e ampla
 - Escalabilidade – lidar com mais demanda
 - Elasticidade – ajustar-se automaticamente à quantidade de demanda que está acontecendo no momento, aumentando ou diminuindo recursos conforme necessário.

Sistema de Arquivos Distribuídos



- É uma infraestrutura de armazenamento de dados projetada para **armazenar** e **acessar arquivos** em uma rede de computadores **distribuída**.
- Em vez de armazenar todos os arquivos em um **único servidor centralizado**, um sistema de arquivos distribuídos **divide** os arquivos em várias partes e as armazena em vários **nós** (ou servidores) dentro da rede.



Sistemas de Armazenamento Distribuídos e Escaláveis



- Arquivos são distribuídos em vários nós da rede
- Armazenamento é escalado horizontalmente
- Redundância e tolerância a falhas
- Suporta acesso concorrente aos arquivos por vários usuários e aplicativos
- Distribuição focada em minimizar tráfego de rede, otimizando comunicação entre nós e latência de acesso aos dados



Azure Data Lake Storage Gen2

Processamento Paralelo - APACHE SPARK



- É um framework (estrutura/conjunto de ferramentas , bibliotecas, padrões) de **computação distribuída**, de **código aberto** projetado para **processamento de dados** em **larga escala**
- Foi desenvolvido pelo grupo Apache
- Está em constante evolução
- Realiza processamento em memória sempre que possível, minimizando a necessidade de leitura e gravação em disco
- É possível utilizar para processamento tanto em lote (batch), como em tempo real (streaming)



DATABRICKS



- Foi projetado pelos fundadores do Apache Spark
- Plataforma de análise de dados unificada e baseada na nuvem
- Construída em torno do Apache Spark
- Integração com Azure, AWS, GCP
- Análise de dados, ingestão, preparação, visualização, construção de machine learning
- SQL, Python, Scala, R
- Unity Catalog



databricks

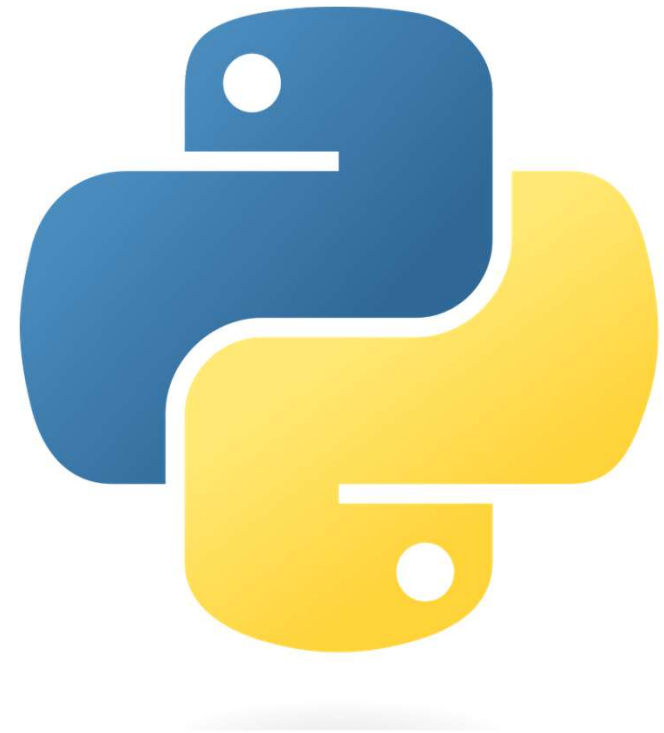
SNOWFLAKE



- Plataforma de armazenamento de dados e análise na nuvem
- Integração com Azure, AWS, GCP
- Análise de dados, ingestão, preparação, visualização, construção de machine learning
- SQL, Python, Java, Javascript



SQL x PYTHON



PYSPARK



- É uma api Python que fornece uma interface de programação que permite interagir com o Spark
- É uma camada de abstração sobre o Apache Spark
- É uma das principais bibliotecas para processamento de Big Data e Machine Learning
- python > processamento sequencial
- pyspark > clusters



ETL e Orquestração De Dados



Azure
Data Factory



Google
Cloud Composer



databricks



Apache
Airflow

Inteligência Artificial



- Algoritmos de Machine Learning (sistemas de recomendação)
- Chatbots
- Reconhecimento facial, digital, voz
- Detecção de sentimentos
- Detecção de anomalias na área de medicina
- Automação de Processos Robóticos (RPA)

E COMO ESTÁ O MERCADO?



SQL em: Brasil

92.282 resultados

Configurar alerta

CATERPILLAR

IT Analyst Database Administration

Caterpillar Inc.

Piracicaba, São Paulo, Brasil

27 ex-alunos trabalham aqui

Promovida • 23 candidaturas

Data Migration

Wipro Limited

Brasil (Remoto)

Promovida •

Candidatura simplificada

wipro

Data Migration

Wipro

Curitiba, Paraná, Brasil (Remoto)

1 conexão trabalha aqui

Promovida •

Candidatura simplificada

Landis Gyr

Database Developer

Landis+Gyr

Curitiba, Paraná, Brasil (Híbrido)

Promovida

Analista de TI - Recife

Mercado Livre

Recife, Pernambuco, Brasil (Presencial)

203 ex-alunos trabalham aqui

Promovida •

Candidatura simplificada

Analista de Migração de Dados

Group Software

Belo Horizonte, Minas Gerais, Brasil (Presencial)

Há 7 horas

Python em: Brasil

58.210 resultados

Configurar alerta

DBA de Oracle Junior

DBACorp - Inteligência Aplicada a TI

São Paulo, São Paulo, Brasil (Remoto)

Seu perfil corresponde a esta vaga

Há 3 semanas

Senior DevOps Cloud Engineer

EngSearch | engineering people

Porto Alegre, Rio Grande do Sul, Brasil (Híbrido)

Seu perfil corresponde a esta vaga

Promovida •

Candidatura simplificada

Analista de TI – Canoas – RS

Divulga Vagas - Consultoria

Canoas, Rio Grande do Sul, Brasil

Há 8 horas

Meta

Engenheiro(a) de Dados II

Meta

Brasil (Remoto)

18 ex-alunos trabalham aqui

Há 4 horas

ESTÁGIO EM TI / INFORMATICA

WILLISA

São Paulo, São Paulo, Brasil (Híbrido)

Há 6 dias

cloud computing em: Brasil

7.175 resultados

Configurar alerta

Senior DevOps Cloud Engineer

EngSearch | engineering people

Porto Alegre, Rio Grande do Sul, Brasil (Híbrido)

Seu perfil corresponde a esta vaga

Promovida •

Candidatura simplificada

rw

Engenheiro de Dados Especialista em Cloud Snow Flake

Robert Walters

São Paulo, São Paulo, Brasil (Remoto)

Promovida • 10 candidaturas

BRQ

Profissional de Engenharia de Dados Azure - SR

BRQ Digital Solutions

São Paulo e Região (Híbrido)

2 conexões trabalham aqui

Promovida • 14 candidaturas •

Candidatura simplificada

ESPECIALISTA DE INFRAESTRUTURA CLOUD

Brasilprev

São Paulo, Brasil (Híbrido)

1 conexão trabalha aqui

Promovida • 19 candidaturas

arquivar

Analista de Infra em Cloud Azure (TI)

ARQUIVAR

Belo Horizonte, Minas Gerais, Brasil (Remoto)

3 ex-alunos trabalham aqui

Há 1 mês

E COMO ESTÁ O MERCADO?



SQL em: Brasil
92.282 resultados Configurar alerta

IT Analyst Database Administration
Caterpillar Inc.
Piracicaba, São Paulo, Brasil
27 ex-alunos trabalham aqui
Promovida • 23 candidaturas

Data Migration
Wipro Limited
Brasil (Remoto)
Promovida • Candidatura simplificada

Data Migration
Wipro
Curitiba, Paraná, Brasil (Remoto)
1 conexão trabalha aqui
Promovida • Candidatura simplificada

Database Developer
Landis+Gyr
Curitiba, Paraná, Brasil (Híbrido)
Promovida

Analista de TI - Recife
Mercado Livre
Recife, Pernambuco, Brasil (Presencial)
203 ex-alunos trabalham aqui
Promovida • Candidatura simplificada

Analista de Migração de Dados
Group Software
Belo Horizonte, Minas Gerais, Brasil (Presencial)
Há 7 horas

Python em: Brasil
58.210 resultados Configurar alerta

DBA de Oracle Junior
DBACorp - Inteligência Aplicada a TI
São Paulo, São Paulo, Brasil (Remoto)
Seu perfil corresponde a esta vaga
Há 3 semanas

Senior DevOps Cloud Engineer
EngSearch | engineering people
Porto Alegre, Rio Grande do Sul, Brasil (Híbrido)
Seu perfil corresponde a esta vaga
Promovida • Candidatura simplificada

Analista de TI – Canoas – RS
Divulga Vagas - Consultoria
Canoas, Rio Grande do Sul, Brasil
Há 8 horas

Engenheiro(a) de Dados II
Meta
Brasil (Remoto)
18 ex-alunos trabalham aqui
Há 4 horas

ESTÁGIO EM TI / INFORMATICA
WILLISA
São Paulo, São Paulo, Brasil (Híbrido)
Há 6 dias

cloud computing em: Brasil
7.175 resultados Configurar alerta

Senior DevOps Cloud Engineer
EngSearch | engineering people
Porto Alegre, Rio Grande do Sul, Brasil (Híbrido)
Seu perfil corresponde a esta vaga
Promovida • Candidatura simplificada

Engenheiro de Dados Especialista em Cloud Snow Flake
Robert Walters
São Paulo, São Paulo, Brasil (Remoto)
Promovida • 10 candidaturas

Profissional de Engenharia de Dados Azure - SR
BRQ Digital Solutions
São Paulo e Região (Híbrido)
2 conexões trabalham aqui
Promovida • 14 candidaturas • Candidatura simplificada

ESPECIALISTA DE INFRAESTRUTURA CLOUD
Brasilprev
São Paulo, Brasil (Híbrido)
1 conexão trabalha aqui
Promovida • 19 candidaturas

Analista de Infra em Cloud Azure (TI)
ARQUIVAR
Belo Horizonte, Minas Gerais, Brasil (Remoto)
3 ex-alunos trabalham aqui
Há 1 mês

E COMO ESTÁ O MERCADO?



Profissional de Engenharia de Dados BRQ Digital Solutions - São Paulo

Candidatura simplificada

Salvar

...

- Ei Profissional de Engenharia de Dados Databricks - SR, venha fazer parte de uma das maiores empresas de transformação digital do país!

- O que procuramos?

Buscamos um #fera com espírito empreendedor e que sabe que é o verdadeiro protagonista do seu sucesso! Queremos em nosso time pessoas motivadas por desafios, constante evolução e, claro, que são apaixonadas por tecnologia e transformação!

- O que você deve ter?

- Experiência com os principais serviços da Microsoft Azure, incluindo o Azure Data Factory, VMs(Maquinas virtuais), Azure StorageAccounts, BlobStorage, Resource groups, Azure Databricks, AKS
- Experiência em Pyspark e Python e Kedro
- Conhecimento em DevOps
- Desejável Experiência anterior com a plataforma Cloudera ou outras soluções on-premises para big data, incluindo Hadoop, HBase e Hive

Analista de TI Cloud

VIASOFT - Pato Branco, Paraná, Brasil (Presencial)

Aplicar ↗

Salvar

...

Requisitos e qualificações

- Graduação ou andamento em Análise e desenvolvimento de Sistemas, Tecnologia da informação, ou áreas correlatas;
- Conhecimentos em soluções de virtualização;
- Conhecimento em soluções de Cloud (Oracle Cloud, Azure, AWS);
- Resiliência e trabalho em equipe;
- Conhecimento em Computação em nuvem (IaaS, PaaS, SaaS);
- Proficiência em elementos de Infraestrutura de TI (Servidores, Redes, SO, etc);
- Conhecimento em Segurança da Informação;
- Conhecimento gerenciamento de serviços de TI (ITIL, Cobit5);
- Certificações na área serão consideradas um diferencial;
- Infraestrutura como código (CI/CD, Terraform);
- Conhecimento em ferramentas de automação;

PARA REFLETIR...



Provérbio Japonês

なな ころ び や お き
七 転 び 八 起 き

Se cair sete vezes, levante oito.

Obrigada!



in

[rayzawilma](#)



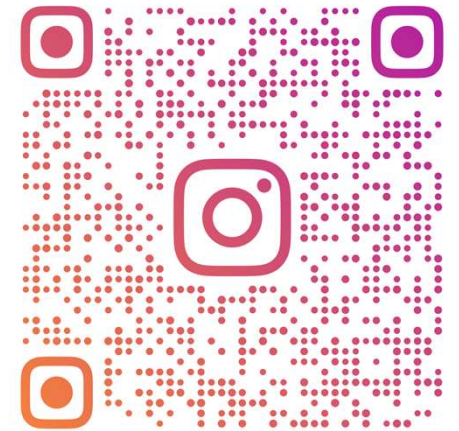
[bigdatanavigators](#)



[bigdatanavigators](#)



<https://bigdatanavigators.com/>



BIGDATANAVIGATORS

SQL WEEK RECIFE

Conference



**O MAIOR EVENTO
ONLINE EM LINHA
RETA DO NORDESTE.**