



targettrust
treinamento e tecnologia

Valli (tiago vallinoto)

Especialista de TI - Sicredi

Produtos, Agilidade e Dados

Atuo em movimentos de transformação
digital e mudança de cultura
~20 anos em tech / ~5 dados

Como falar comigo:

- tiago@targettrust.com.br
- linkedin.com/in/vallinoto
- <https://vallinoto.com>

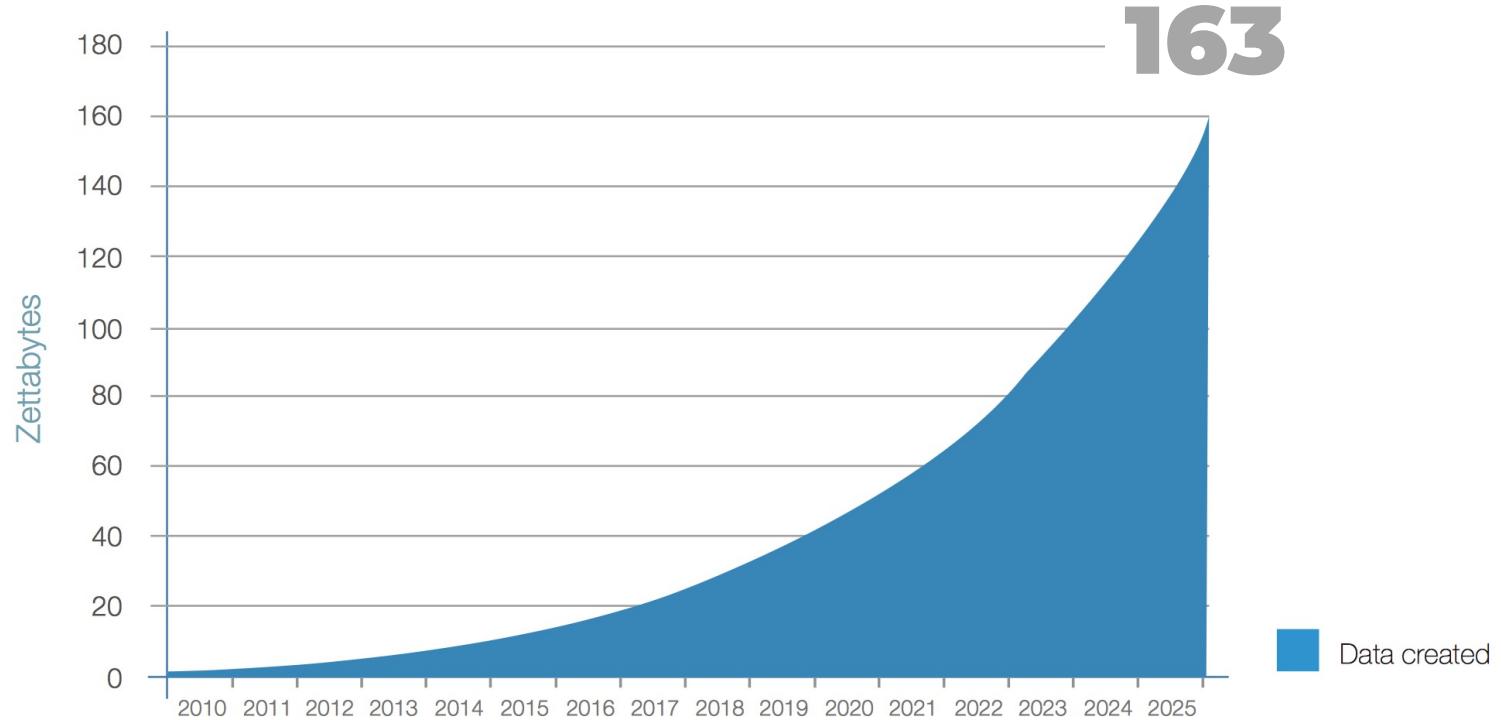




O mercado hoje



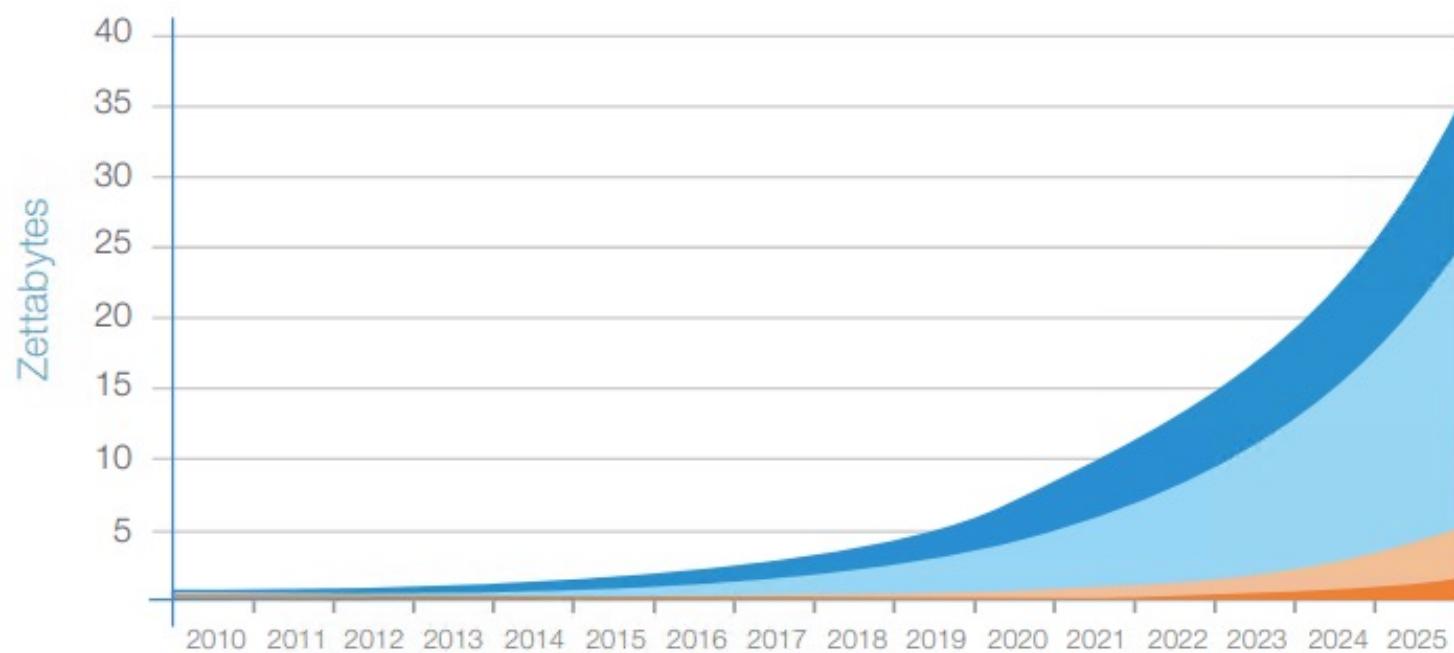
Estamos gerando cada vez + dados



Source: IDC's Data Age 2025 study, sponsored by Seagate, April 2017

<https://www.import.io/wp-content/uploads/2017/04/Seagate-WP-DataAge2025-March-2017.pdf>

Estamos gerando cada vez + dados



Source: IDC's Data Age 2025 study, sponsored by Seagate, April 2017

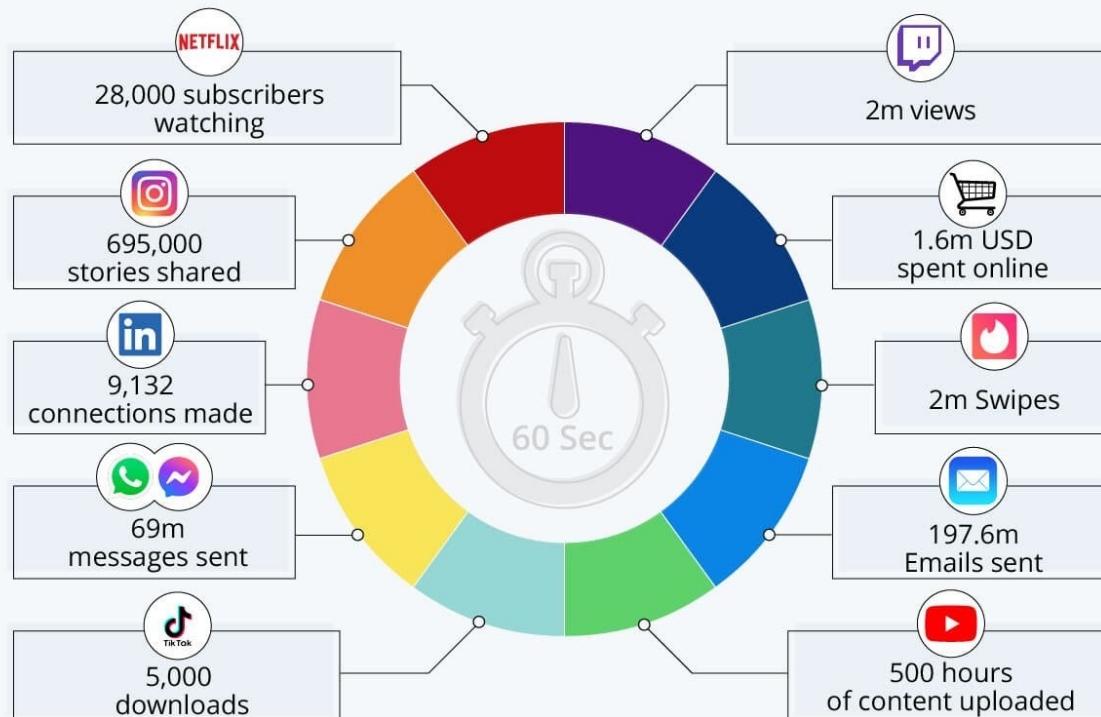
O IDC estima que até o final de 2025, apenas 15% dos dados na esfera de dados global será marcada e apenas um quinto disso será realmente ser analisado.

- Useful if Tagged
- Tagged
- Analyzed
- Touched by Cognitive Analytics

Estamos gerando cada vez + dados

A Minute on the Internet in 2021

Estimated amount of data created
on the internet in one minute



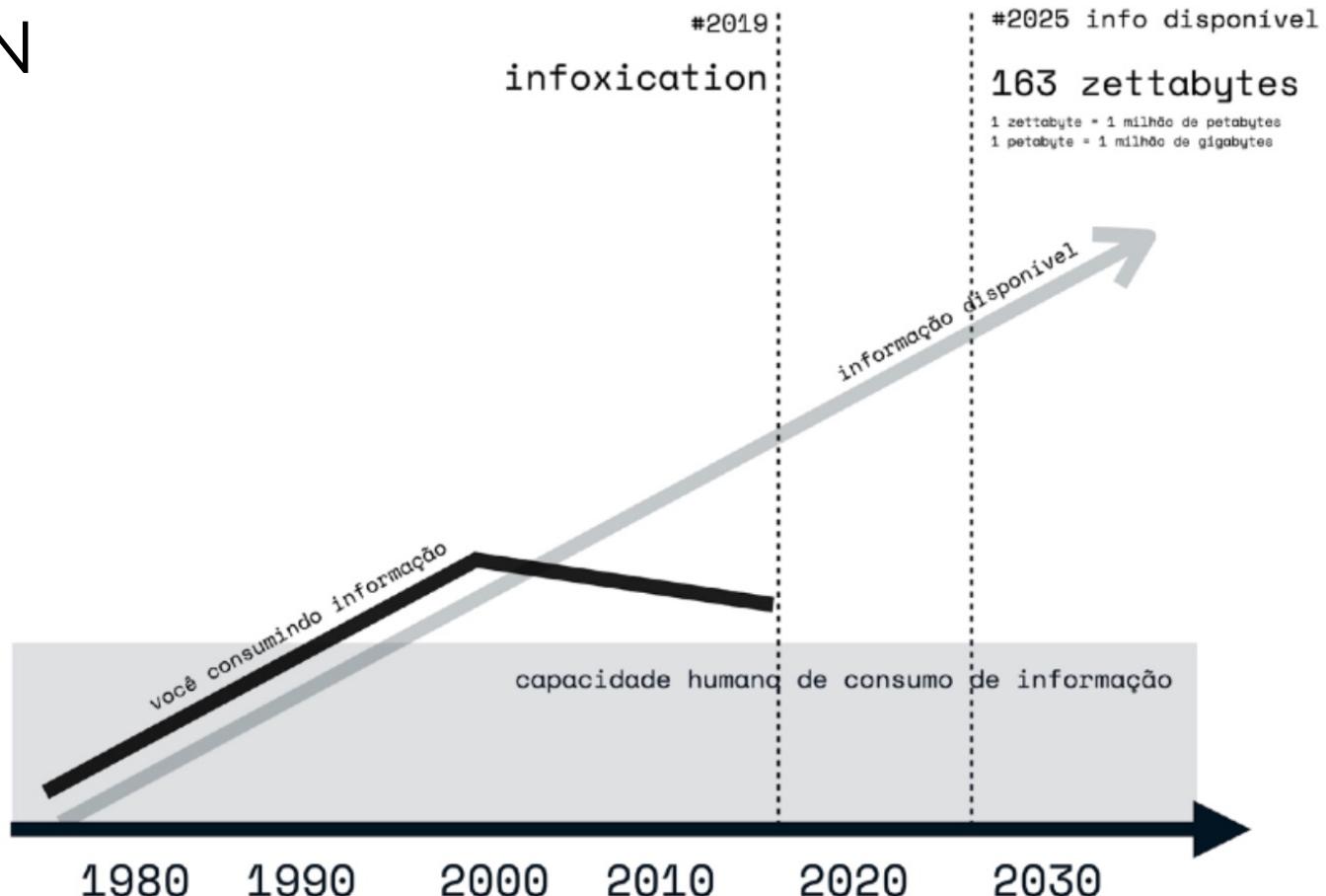
Source: Lori Lewis via AllAccess



E nossa capacidade cognitiva é limitada

<https://www.cappra.institute/2019>

INFOXICATION



>> infográfico relação da quantidade de informação disponível com a capacidade humana de consumo de informação

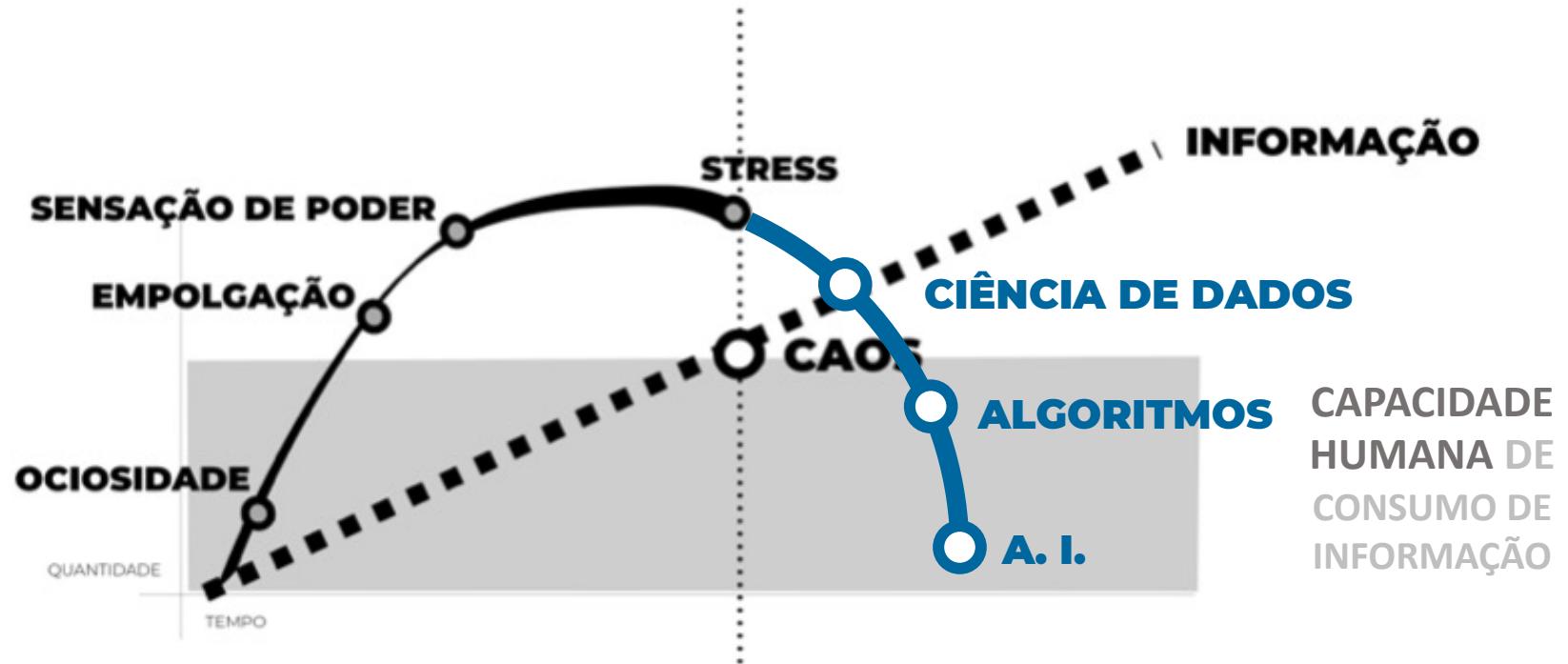
Classificação da informação: Uso Interno

“

Chegamos aos nossos limites físicos e biológicos como seres humanos. Não conseguimos mais consumir, processar e analisar a quantidade de informação que temos à nossa disposição.

Ricardo Capra

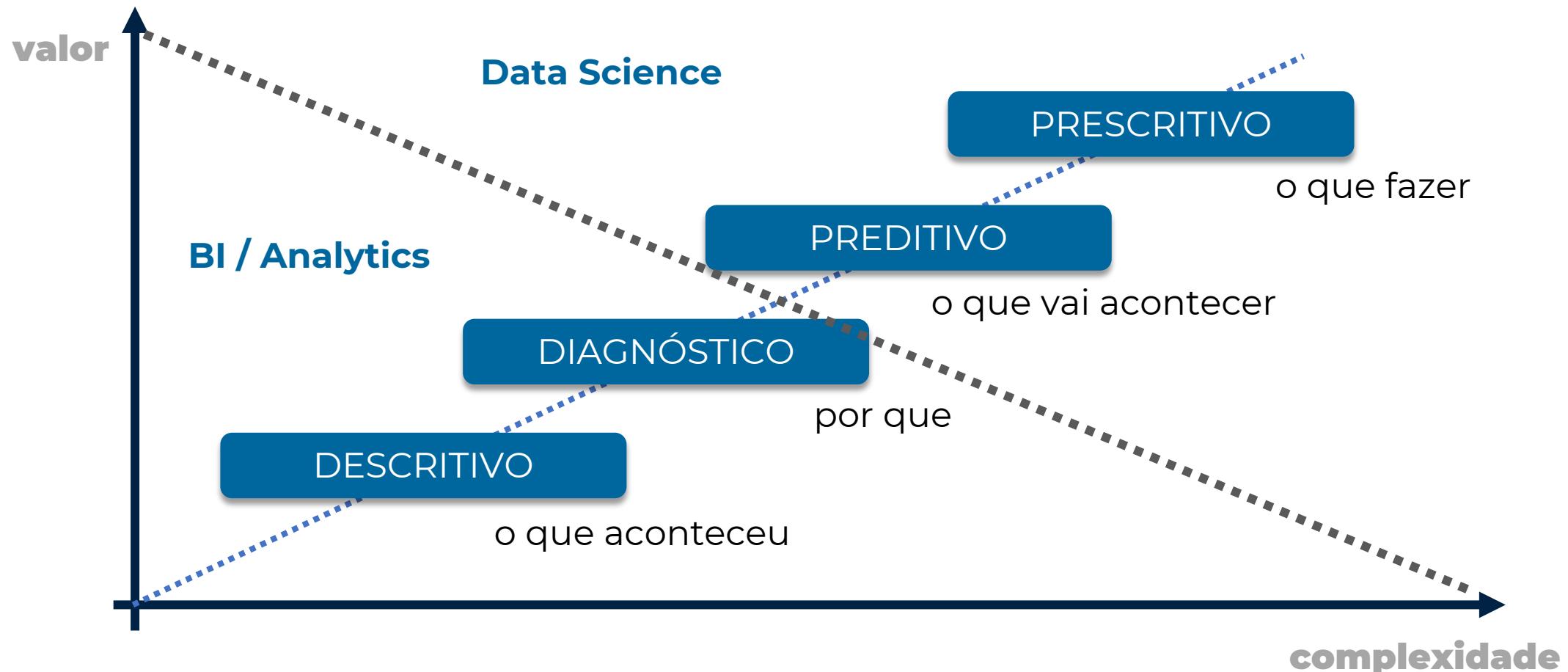
E nossa capacidade cognitiva é limitada



A geração do valor é no topo



A complexidade cresce também





DATA

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE

[SUMMARY](#) [SAVE](#) [SHARE](#) [COMMENT](#) [TEXT SIZE](#) [PRINT](#) [\\$8.95 BUY COPIES](#)

When Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't seeking out

g1

GLOBO REPÓRTER

Globo Repórter revela quais são as profissões do futuro

Nesta sexta-feira (28), você vai conhecer a rotina de um superprofissional: o homem das mil habilidades.

Por Globo Repórter

25/06/2019 18h43 · Atualizado há 2 anos



fora o hype



O PROGRAMA ▾ NOTÍCIAS ▾ VÍDEOS ▾ FALE COM A PRODUÇÃO ▾ PODCAST REDES SOCIAIS

Carreira em ciência de dados promete salários de R\$ 22 mil

Pouco conhecida no Brasil, a profissão é promissora e muito procurada pelas empresas. Saiba mais sobre o trabalho desenvolvido por quem escolhe atuar nessa área

FALA BRASIL | Do R7

16/08/2018 - 01H00 (ATUALIZADO EM 16/08/2018 - 09H44)

COMPARTILHE: [f](#) [t](#) [w](#) [...](#)

ciência de dados | Brasil | Pesq.

Vagas | Último mês | Nível de experiência | Empres.

Ciência de dados em: Brasil 1.165 resultados

Gerente de Ciência de Dados
Dasa
Pinheiros, Espírito Santo, Brasil
2 ex-funcionários da empresa trabalham aqui
Promovida • 17 candidaturas

Especialista em Ciência de Dados - IT Data Analytics
BTG Pactual
São Paulo, São Paulo, Brasil
9 ex-funcionários da empresa trabalham aqui
Promovida

Analista de Dados III - Marketing Insights
Grupo Boticário
Brasil (Remoto)
1 conexão trabalha aqui
Há 3 horas

Data Visualization Analyst Senior
CINT
Brasil (Remoto)
2 conexões trabalham aqui
Há 3 semanas

Engenheiro de dados
Keep Simple
Brasil (Remoto)
Recrutando agora
Há 4 horas • Candidatura simplificada

Engenheiro(a) de Dados PL
Geofusion
Brasil (Remoto)
Recrutando agora
Há 5 horas

analista de dados | Brasil | Pesq.

Vagas | Último mês | Nível de experiência | Empres.

Analista de dados em: Brasil 3.673 resultados

Analista Engenharia de Dados PL/SR
Accenture Brasil
Rio de Janeiro, Rio de Janeiro, Brasil
1 conexão trabalha aqui
Promovida • 7 candidaturas

Analista de Dados Sênior
PicPay
São Paulo, São Paulo, Brasil
1 conexão trabalha aqui
Promovida

Analista de Dados (Data Engineer) - Asset Management
BTG Pactual
São Paulo, São Paulo, Brasil
10 ex-alunos trabalham aqui
Promovida

Analista de Dados (Foco em Modelagem) | BANCO DE TALENTOS
Stone
Rio de Janeiro, Rio de Janeiro, Brasil
23 ex-funcionários da empresa trabalham aqui
Promovida

Analista de dados
LUMA Ensino Individualizado
Vitória, Espírito Santo, Brasil (Remoto)
Seu perfil corresponde a esta vaga
Há 1 hora

Analista de Dados
Business Partners Consulting
São Paulo, Brasil (Presencial)
Recrutando agora

engenheiro de dados | Brasil | Pesq.

Vagas | Último mês | Nível de experiência | Empres.

Engenheiro de dados em: Brasil 1.306 resultados

Engenheiro(a) de Dados Sênior - Big Data
EY
Recife, Pernambuco, Brasil
1 conexão trabalha aqui
Promovida • 3 candidaturas

Engenheiro de Dados - Risco
Stone
Brasil (Remoto)
23 ex-funcionários da empresa trabalham aqui
Promovida • Candidatura simplificada

Engenheiro de Dados- Hadoop
Accenture Brasil
Rio de Janeiro, Rio de Janeiro, Brasil
1 conexão trabalha aqui
Promovida • 2 candidaturas

Engenheiro de Dados Sênior
Cartão Elo
São Paulo, Brasil (Híbrido)
2 ex-alunos trabalham aqui
Promovida • 5 candidaturas

Engenheiro de dados
GAVB
Canoas, Rio Grande do Sul, Brasil (Remoto)
7 conexões trabalham aqui
Há 1 dia • 18 candidaturas

Engenheiro de dados
Hrsoul
Brasil (Remoto)
Recrutando agora
Há 5 minutos • Candidatura simplificada

analytics | Brasil | Pesq.

Vagas | Último mês | Nível de experiência | Empres.

Analytics em: Brasil 5.519 resultados

Especialista em Engenharia de Dados - IT Data Analytics
BTG Pactual
São Paulo, São Paulo, Brasil
10 ex-alunos trabalham aqui
Promovida

Insight and Analytics Global Lead, People with Disabilities
Google
São Paulo, São Paulo, Brasil
Recrutando agora
Promovida • 4 candidaturas

Marketplace Analytics Operations Manager, Marketplace Central Ops LatAm
Uber
Osasco, São Paulo, Brasil
1 conexão trabalha aqui
Promovida • 18 candidaturas

Business Analyst - Marketing Analytics
Toptal
Porto Alegre, Rio Grande do Sul, Brasil (Remoto)
6 ex-alunos trabalham aqui
Promovida • 14 candidaturas

Gerente de Dados
Business Partners Consulting
São Paulo, Brasil (Presencial)
Recrutando agora
Há 41 minutos

Analista de Performance e BI
JB3 Investimentos
Florianópolis, Santa Catarina, Brasil (Presencial)

Top Trends in Data and Analytics For 2021



Accelerating Change

- 1 Smarter, Responsible, Scalable AI
- 2 Composable Data and Analytics
- 3 Data Fabric is The Foundation
- 4 From Big to Small and Wide Data



Operationalizing Business Value

- 5 XOps
- 6 Engineering Decision Intelligence
- 7 D&A as a Core Business Function



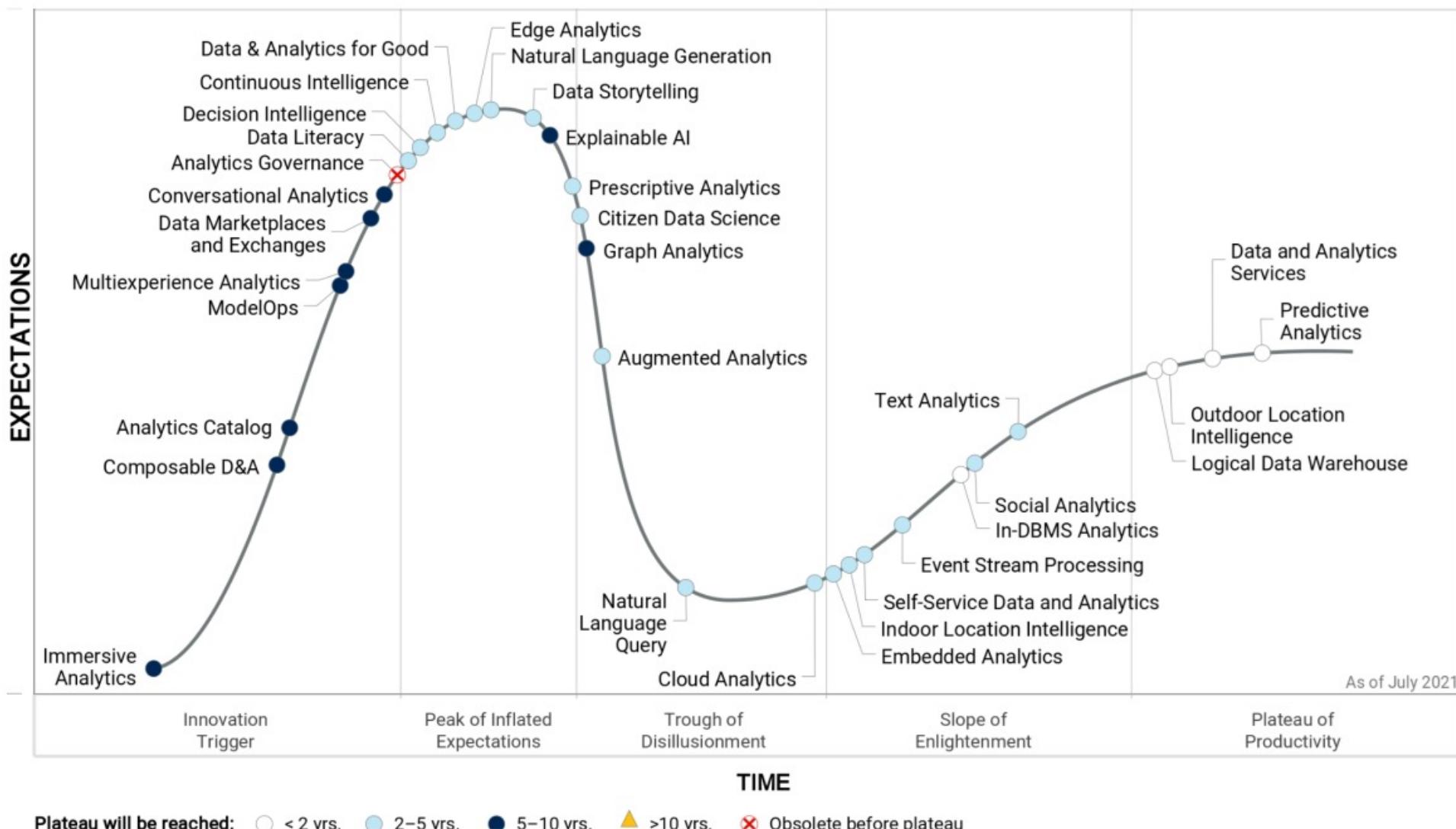
Distributed Everything

- 8 Graph Relates Everything
- 9 The Rise of the Augmented Consumer
- 10 D&A At the Edge

Source: Gartner

729348_C

Hype Cycle for Analytics and Business Intelligence, 2021



Source: Gartner (July 2021)

747544



Papéis envolvendo
dados

Papéis envolvidos

- Quando falamos sobre trabalhar com dados, geralmente, falamos de atribuições de mais de um papel.
- Unicórnios, no contexto de dados, não existem :)



analista de negócios



analista de dados



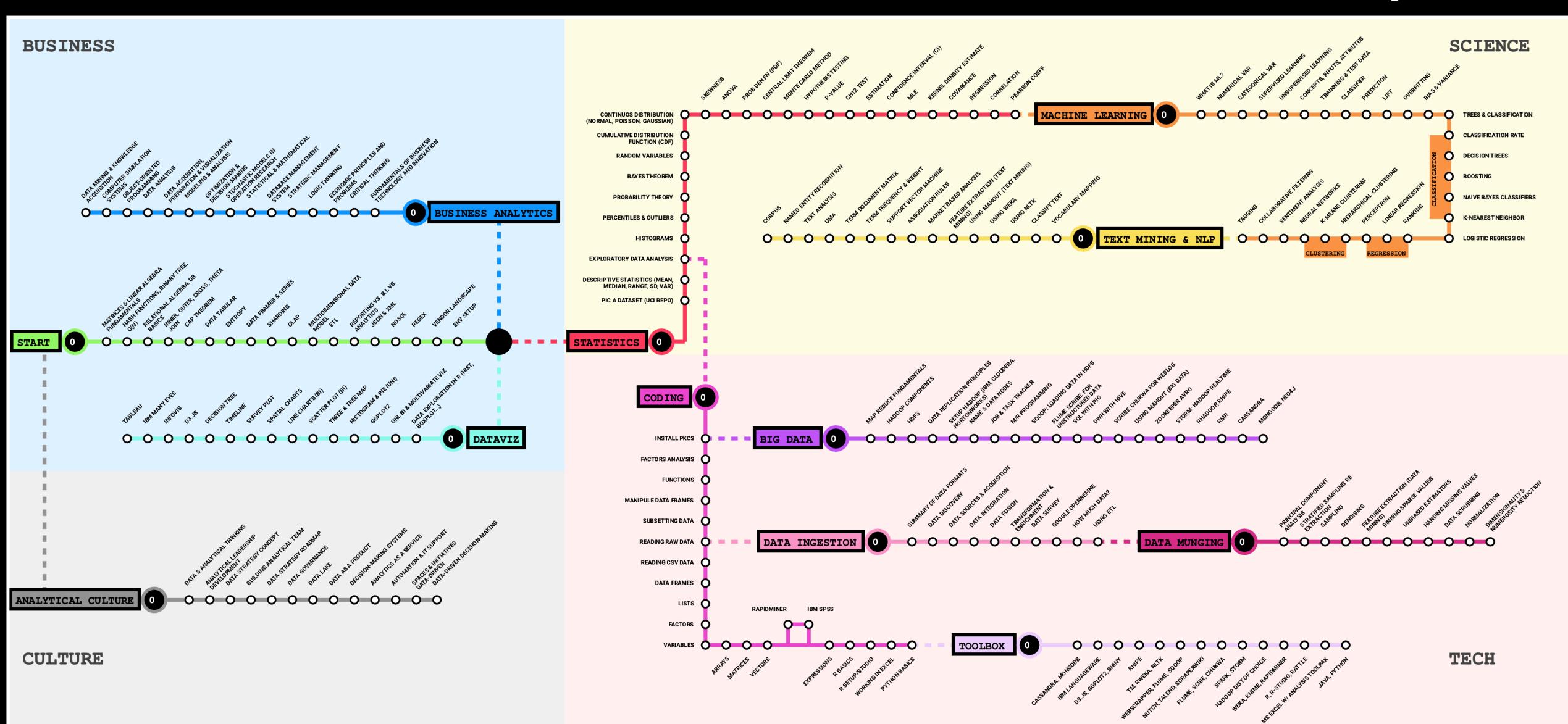
engenheiro de dados



cientista de dados



**administrador
de banco de dados**



Analista de negócios



- Embora existam algumas semelhanças entre um analista de dados e um analista de negócios, a principal diferença entre as duas funções é o que eles fazem com os dados.
- Um analista de negócios está mais perto da empresa e é especialista em interpretar os dados provenientes da visualização.
- Muitas vezes, as funções do analista de dados e o analista de negócios podem ser a responsabilidade de uma só pessoa.

Analista de dados



- Foco em maximizar o valor dos ativos de dados deles por meio de ferramentas de visualização e relatórios.
- Responsáveis pela criação de perfil, limpeza e transformação de dados. Inclui a criação de modelos de dados escalonáveis e eficientes e a habilitação e implementação de funcionalidades analíticas avançadas em relatórios para análise. Identifica os requisitos apropriados e necessários de dados e de relatórios, depois tem a tarefa de transformar dados brutos em insights relevantes e significativos.
- Os analistas de dados trabalham com engenheiros de dados para determinar e localizar fontes de dados apropriadas que atendam aos requisitos de stakeholder.

Engenheiro de dados



- Os engenheiros de dados provisionam e configuram tecnologias de plataforma de dados locais e na nuvem. Eles gerenciam e protegem o fluxo de dados estruturados e não estruturados de várias fontes. Os engenheiros de dados também garantem que os serviços de dados se integrem de maneira segura e direta uns aos outros.
- As principais responsabilidades de engenheiros de dados incluem o uso de serviços de dados locais e na nuvem e ferramentas para ingestão, saída e transformação de dados de várias fontes.
- Engenheiros de dados agregam um valor imenso a projetos de *business intelligence* e ciência de dados, ao reunir, estruturar e ingerir dados.

Cientistas de dados



- Os cientistas de dados realizam a análise avançada para extrair valor dos dados. Seu trabalho pode variar de análise descritiva a análise preditiva. A análise descritiva avalia os dados por meio de um processo conhecido como EDA (análise exploratória de dados).
- A análise preditiva é usada no machine learning para aplicar técnicas de modelagem que possam detectar anomalias ou padrões. Essas análises são partes importantes dos modelos de previsão.
- As análises descritiva e preditiva são apenas aspectos parciais do trabalho dos cientistas de dados. Alguns cientistas de dados podem trabalhar na área de aprendizado profundo, realizando experimentos iterativos para resolver um problema de dados complexo com o uso de algoritmos personalizados.
- A maior parte do trabalho em um projeto de ciência de dados é gasto em estruturação de dados e engenharia de recursos.
- À primeira vista, pode parecer que um cientista de dados e um analista de dados fazem trabalhos completamente diferentes, mas não é verdade.

Administrador de banco de dados

- Um administrador de banco de dados implementa e gerencia os aspectos operacionais de soluções de plataformas de dados híbridas e nativas de nuvem
- Um administrador de dados é responsável pela disponibilidade geral e pelo desempenho e otimizações consistentes das soluções de banco de dados.
- Um administrador de banco de dados monitora e gerencia a integridade geral de um banco de dados e o hardware no qual ele reside, enquanto um engenheiro de dados está envolvido no processo de estruturação de dados, em outras palavras, ingestão, transformação, validação e limpeza de dados para atender às necessidades e requisitos empresariais.
- O administrador de banco de dados também é responsável por gerenciar a segurança geral dos dados, concedendo e restringindo o acesso e os privilégios dos usuários aos dados, conforme determinado pelas necessidades e requisitos empresariais.



E como a **targettrust** pode ajudar?



Formação **Data Science**

A Formação Data Science é para você que deseja se especializar na Manipulação e Extração de Dados! Composta por 3 cursos essenciais (96h) , é ideal para quem busca desenvolver uma nova Skill e quer aprender na prática a criar aplicações que se destaquem por sua inteligência artificial. Nossa formação simula todos os desafios reais enfrentados por um cientista de dados. Por isso prepare-se para muito conhecimento e dedicação!



Formação **Data Analytics**

Quem deve fazer essa Jornada: pessoas que queiram entrar ou evoluir nas carreiras de Análise de Dados, Ciência de Dados, Engenharia de Dados, além de Analistas de Negócios, Analistas de Marketing, P.Os, P.Ms, que queiram apurar dados, gerar métricas, fazer análises e demonstrações mais assertivas do seu trabalho e resultados. Enfim, qualquer pessoa que queira conhecer a cultura Data Driven!

E como a **targettrust** pode ajudar?

Lóg. Programação e Ferramentas p/ trabalhar com Dados

Duração: 27h

Através da resolução de problemas, você aprenderá lógica de programação juntamente com a sintaxe da linguagem Python. Além de desenvolver o raciocínio lógico aplicado à resolução de problemas em nível computacional, capacitar-se a trabalhar com estruturas de dados identificando a aplicação destas na solução de problemas reais e entender as boas práticas de programação e organização de código.



**Formação
Data
Science**

Extração e Manipulação de Dados com Python

Duração: 30h

Aqui, você vai ser preparado para organizar, preparar e entender os dados manipulando-os com Python. Compreender os fundamentos da ciência de dados e da aprendizagem de máquina. Conhecer o processo de coleta e análise de dados aplicando métodos e ferramentas para responder a questões úteis à tomada de decisão. Apresentar os conceitos envolvidos na concepção de técnicas de visualização de informações, trabalhando com as principais bibliotecas de manipulação e visualização de informações.

Técnicas de Machine Learning aplicadas no Mercado

Duração: 39h

Ao final deste módulo você será um cientista de dados que além de organizar, preparar e entender, vai manipular dados com Python, compreender os fundamentos da ciência de dados e da aprendizagem de máquina, aplicando algoritmos de Machine Learning na prática e criando aplicações que se destaquem por sua inteligência artificial! Tomada de decisões com base na análise prévia de dados que, disponíveis em grande volume, conseguem detectar tendências de comportamento ao acompanhar padrões registrados.

E como a **targettrust** pode ajudar?



**Formação
Data Analytics**

Análise e Visualização de Dados com SQL e Power BI

Duração: 18h

Torne-se um Analista de Dados capaz de organizar e preparar os dados para análise, descobrir padrões e insights, obter e expor conclusões significativas, comunicar claramente suas descobertas através de visualizações usando o SQL para acessar dados e o Power BI.

Data Stories - contextualizando dados

Duração: 12h

Aprenda a construir histórias que combinam dados confiáveis e análises, dando significado aos dados, fornecendo visão, contexto e explicação, tornando a análise instantaneamente mais relevante, interessante e compreendida, gerando insights de dados de forma eficiente e eficaz.

Bora estudar?



Lógica de
Programação e
Ferramentas para
trabalhar com Dados

Quem são vcs?

- Nome / Empresa
- Experiência com o tema
- Expectativas com o treinamento



Houston, we have a problem



problema

substantivo masculino

1. assunto controverso, que pode ser objeto de pesquisas científicas ou discussões acadêmicas.
"o p. do descobrimento do Brasil"
2. questão social que traz transtornos e que exige grande esforço e determinação para ser
solucionado.
"o p. da seca no Nordeste brasileiro"
3. obstáculo, dificuldade que desafia a capacidade de solucionar de alguém.
"os moradores listaram os p. mais graves do bairro"
4. situação difícil; conflito emocional.
"ela vive cheia de problemas"
5. distúrbio, disfunção orgânica.
"p. digestivo"
6. pessoa, coisa ou situação incômoda, preocupante, fora de controle etc.
"essa menina é um p."
7. **MATEMÁTICA**
tarefa de calcular uma ou várias quantidades desconhecidas (*incógnitas*) relacionadas a outras
conhecidas (*dados*).

Origem

○ ETIM lat. *problēma, ātis* 'id.', do gr. *próblerma, atos* 'o que se tem diante de si; obstáculo; questão'

Questão DIFÍCIL, delicada,
suscetível de DIVERSAS
SOLUÇÕES

- Trocar a resistência de
um chuveiro
- Trocar o pneu do carro
- Definir onde almoçar
- Entregar um trabalho

Houston, we have a problem

E quando nos deparamos com um PROBLEMA?
Buscamos uma SOLUÇÃO!

Mas COMO?
Elaborando um PROCEDIMENTO

Ou seja, uma LISTA DE PASSOS

Houston, we have a problem

Procedimento para TROCAR RESISTÊNCIA DE UM CHUVEIRO:

1. Adquirir uma resistência nova
2. Abrir o chuveiro
3. Troca a resistência
4. Fechar o chuveiro
5. Testar o chuveiro

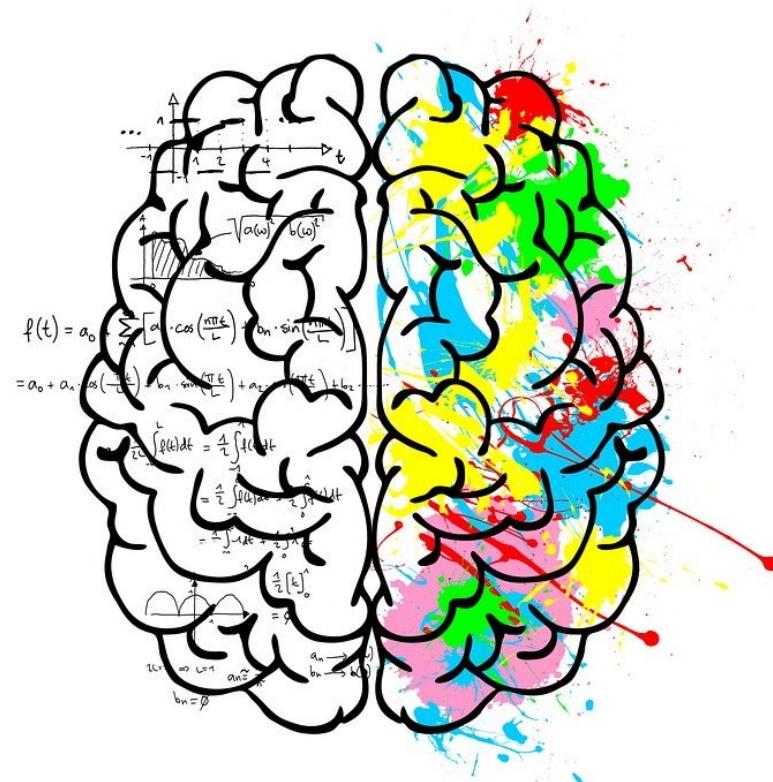
Tem certa **lógica**, né?

E o que é Lógica?

Ramo da FILOSOFIA e MATEMÁTICA
que tem como objetivo definir se
RACIOCÍNIOS são VÁLIDOS ou não.

Primordial para SOLUÇÃO DE
PROBLEMAS, pois permite alcançar
EFICIÊNCIA e EFICÁCIA

Imagine tentar trocar a resistência
de um chuveiro sem abri-lo ou sem
desliga-lo.



Algoritmos

SEQUÊNCIA DE PASSOS BEM DEFINIDA que deve ser seguida para realização de uma tarefa ou SOLUÇÃO DE UM PROBLEMA.

Para tanto, utilizamos a LÓGICA para LISTAR PASSOS que resultam na SOLUÇÃO de um determinado PROBLEMA

Exemplos:

- Resolver uma equação de segundo grau
- Somar vetores
- Preparar uma bebida



Exercício

Escreva um algoritmo em linguagem natural para calcular a média de três valores.



5 min

Algoritmo: MÉDIA TRÊS VALORES

1. Adicione o **VALOR1** ao **VALOR2**
2. Armazene o resultado da soma em **SOMA1**
3. Adicione o **VALOR3** ao **SOMA1**
4. Armazene o resultado da soma em **SOMA2**
5. Divida **SOMA2** por 3
6. Armazene o resultado da divisão em **MEDIA**

Fim_Algoritmo

Lógica de Programação

Estudo da estrutura da linguagem das máquinas

entrada

Forma de comunicação com computadores

Passos necessários para que os computadores precisam executar para resolver determinado problema ou atingir determinado objetivo

processamento

Essa comunicação é através de algoritmos (passos finitos e organizados)

saída

Lógica de Programação

- Variáveis e Constantes
 - Espaço reservado para guardar um tipo de dado. Um de cada vez.
Podendo ser numérico, caractere, alfanumérico ou lógico.
- Operadores
 - Trata a forma como comparamos, avaliamos, incrementamos ou decrementamos dados
- Estruturas de Controle (se)
- Estruturas de Repetição



Exercício

Escreva um algoritmo em linguagem natural para fazer um purê de batata com o que existe na geladeira seguindo a receita que vc tem.



5 min



Python

—

Linguagem Python

- Criada em 1991
 - Guido van Rossum no Instituto de Pesquisa Nacional para Matemática e Ciências da Computação (Países Baixos)
- Python é uma linguagem de programação de alto nível, versátil
- Projetada para ser legível, de fácil manutenção e com suporte avançado a mecanismos de reutilização de software.
- Filosofia por trás: esforço do programador sobre o esforço computacional. Prioriza a legibilidade do código sobre a velocidade.



Linguagem Python

Principais motivos listados por quem desenvolve em Python:

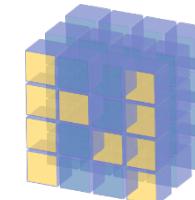
- Fácil aprendizado;
- Simples de programar;
- Sintaxe intuitiva;
- Open Source;
- Modularização;
- Multiplataforma;
- Grande quantidade de bibliotecas disponíveis;
- Grande comunidade de usuários;
- Documentação extensa;
- Número de oportunidades no mercado de trabalho.

Linguagem Python

Principais motivos listados por quem desenvolve em Python:

- Fácil aprendizado;
- Simples de programar;
- Sintaxe intuitiva;
- Open Source;
- Modularização;
- Multiplataforma;
- Grande quantidade de bibliotecas disponíveis;
- Grande comunidade de usuários;
- Documentação extensa;
- Número de oportunidades no mercado de trabalho.

Poder x Usabilidade x Comunidade



NumPy



Folium



Pandas



TensorFlow



Keras

Google Colab

- Serviço em nuvem, gratuito;
- Similar ao Jupyter Notebook (comum entre o pessoal que trabalha com Data Science);
- Suporta comandos bash;
- Vasta variedade de bibliotecas instaladas;
- Integração com Drive e Git;
- Vc só precisa de uma conta no Google :)

<https://colab.research.google.com/>

Olá, este é o Colaboratory

Arquivo Editar Ver Inserir Ambiente de execução Ferramentas Ajuda

Compartilhar gear icon T

Índice

- + Código
- + Texto
- Copiar para o Drive

Conectar dropdown | pencil icon Editar | up arrow icon

Primeiros passos

Ciência de dados

Machine learning

Mais recursos

Exemplos de machine learning

Seção

• Experimentos com TPUs

• Divulgação de pesquisas em IA

• Criação de tutoriais

Para ver notebooks do Colab que demonstram aplicações de machine learning, consulte os [exemplos de machine learning](#) abaixo.

Mais recursos

Como trabalhar com Notebooks no Colab

- [Visão geral do Colaboratory](#)
- [Guia sobre Markdown](#)
- [Importar bibliotecas e instalar dependências](#)
- [Salvar e carregar notebooks no GitHub](#)
- [Formulários interativos](#)
- [Widgets interativos](#)
- [TensorFlow 2 no Colab](#)

Como trabalhar com dados

Uma **variável** é uma posição de memória capaz de armazenar um **valor**
Uma **constante** é uma variável com valores fixos

Tipo de Dado	Descrição
str	Caracteres
int	Número inteiro
float	Número real
complex	Números complexos
bool	Booleano (Verdadeiro ou Falso)

Existem outros tipos de dado com representatividade e precisão maior, sendo que, cada tipo possui um intervalo de valores aceitos.

Importante: Variáveis só existem em tempo de execução!

```
# CONSTANTES
PI = 3.14159265359
DIAS_DA_SEMANA = 7

# VARIÁVEIS
alunos = 13
nota = 9.5
nome = 'Matheus'

print(type(PI))
print(type(DIAS_DA_SEMANA))
print(type(alunos))
print(type(nota))
print(type(nome))
```

*Conheça os padrões de nomes de variáveis utilizados em diferentes linguagens

<https://medium.com/better-programming/string-case-styles-camel-pascal-snake-and-kebab-case-981407998841>

A linguagem Python oferece uma gama variada de operadores

- Operadores aritméticos
- Operadores de atribuição
- Operadores relacionais
- Operadores lógicos

Operadores matemáticos usuais:

Operador	Nome	Ação que executa
+	Adição	Soma dois valores
-	Subtração	Subtrai dois valores
*	Multiplicação	Multiplica dois valores
**	Potenciação	Potência entre dois valores
/	Divisão	Divide dois valores(resultado é um real)
//	Divisão Inteira	Divide dois valores (resultado é um inteiro)
%	Módulo	Retorna o resto da divisão de um valor por outro

```
adicao = 1 + 4  
subtracao = 5 - 3  
multiplicacao = 2 * 5  
potenciacao = 2 ** 3  
divisao_real = 5 / 2  
divisao_inteira = 5 // 2  
modulo = 10 % 3
```

```
print(adicao)  
print(subtracao)  
print(multiplicacao)  
print(potenciacao)  
print(divisao_real)  
print(divisao_inteira)  
print(modulo)
```