



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

A Revolução Industrial 4.0 já
começou e você pode estar atrasado

O futuro é da Automação e Inteligência Artificial

“Entramos na era da Quarta Revolução Industrial. O Fórum Econômico Mundial sugere que a **automação irá eliminar 75 milhões de postos de trabalhos**, enquanto **133 milhões de novos empregos** com **mais interação entre homem e máquina** serão criados.”



* Fonte: http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf

JOAO ESTEVAN LEONCIO DA SILVA BARBOSA - jestevan12@gmail.com - CPF: 134.982.877-70

Indústria 4.0

Vamos lembrar os últimas revoluções industriais



1ª Revolução Mecanização dos equipamentos

Início: 1760

Principais características:

- Aumento da produtividade através da máquinas
- Invenção da máquina a vapor (indústria textil era a principal)
- Evento mais importante na história da humanidade desde a domesticação de animais e a agricultura

2ª Revolução Eletricidade e Produção em Massa

Início: 1850

Principais características:

- Com o domínio da eletricidade, novos mercados e produtos surgiram
- Novo combustível: Petróleo
- Surgimento dos Carros, TVs, Radios, Telefones, Aviões entre outros

3ª Revolução Produção Flexível e Informática

Início: 1950 (no pós guerra)

Principais características:

- Avanço em pesquisa e desenvolvimento de novas tecnologias: Robótica, Eletrônica e a Informática
- Esgotamento do Fordismo e o início do Toyotismo (produção flexível de acordo com a demanda)

Comum a todas as revoluções: **Empregos serem extintos e novos empregos serem criados!**

Indústria 4.0

E a Quarta Revolução?



Com início estimado entre 2010-2015, **já entramos na 4ª Revolução Industrial**. Veja abaixo algumas de suas características:



Indústria 4.0

Os 5 V's do Big Data



Ciência de Dados

A área de Ciência de Dados engloba uma infinidade de assuntos e conhecimentos. No entanto, é possível resumi-lá como **“a junção das técnicas analíticas, ferramentas computacionais e conhecimento do negócio”** de forma a otimizar o processo decisório das empresas ou mesmo resolver problemas de qualquer espécie com o apoio dos dados.



Técnicas Analíticas

Matemática, Otimização, Estatística Aplicada, Probabilidades, Inferência, Machine Learning.



Ferramentas de dados

Planilhas, linguagens de programação, API's, banco de dados relacionais e não relacionais, ferramentas de Visualização de Dados.



Conhecimento do Negócio

Sem o conhecimento das particularidades do negócio, a prática de Ciência de Dados fica sem sentido e não agrega o valor necessário.



≡ Forbes

The Need For New Skills

According to a [study](#) by the McKinsey Global Institute, "between 400 million and 800 million individuals could be displaced by automation and need to find new jobs by 2030 around the world."

Segundo estudo da McKinsey, milhões de pessoas precisarão se reinventar para encontrar **novos** trabalhos até 2030

Indústria 4.0

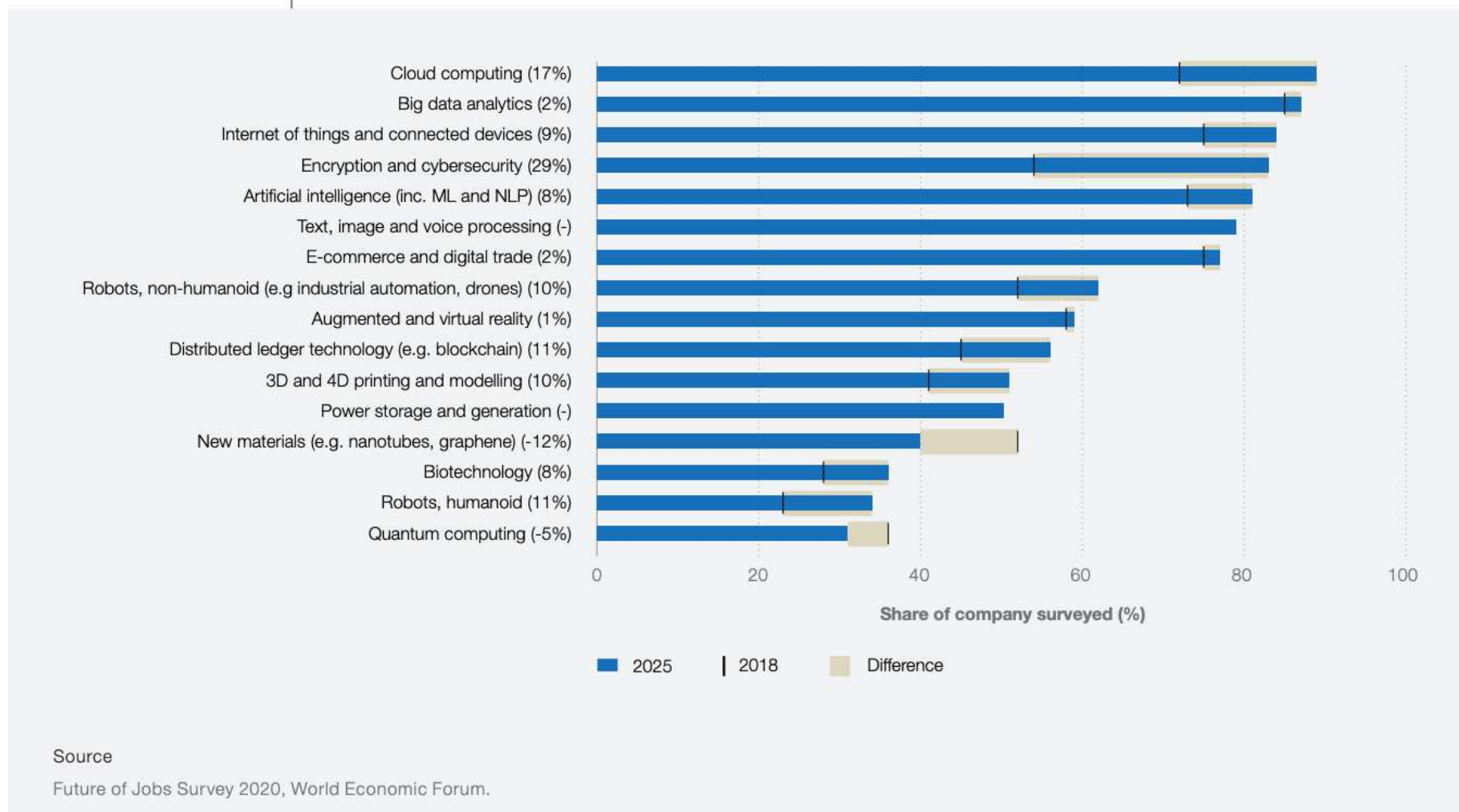
A adoção de novas tecnologias cresce em ritmo acelerado



Preditiva.ai

FIGURE 18

Technologies likely to be adopted by 2025 (by share of companies surveyed)



Como era de se esperar, a **revolução 4.0** nos trouxe um conjunto de novas possibilidades tecnológicas com grande impacto em produtividade e eficiência.

Indústria 4.0

Quais as profissões que estão ganhando ou perdendo relevância?



Preditiva.ai

FIGURE 22

Top 20 job roles in increasing and decreasing demand across industries

↗ Increasing demand

1	Data Analysts and Scientists
2	AI and Machine Learning Specialists
3	Big Data Specialists
4	Digital Marketing and Strategy Specialists
5	Process Automation Specialists
6	Business Development Professionals
7	Digital Transformation Specialists
8	Information Security Analysts
9	Software and Applications Developers
10	Internet of Things Specialists
11	Project Managers
12	Business Services and Administration Managers
13	Database and Network Professionals
14	Robotics Engineers
15	Strategic Advisors
16	Management and Organization Analysts
17	FinTech Engineers
18	Mechanics and Machinery Repairers
19	Organizational Development Specialists
20	Risk Management Specialists

↘ Decreasing demand

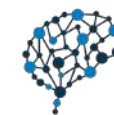
1	Data Entry Clerks
2	Administrative and Executive Secretaries
3	Accounting, Bookkeeping and Payroll Clerks
4	Accountants and Auditors
5	Assembly and Factory Workers
6	Business Services and Administration Managers
7	Client Information and Customer Service Workers
8	General and Operations Managers
9	Mechanics and Machinery Repairers
10	Material-Recording and Stock-Keeping Clerks
11	Financial Analysts
12	Postal Service Clerks
13	Sales Rep., Wholesale and Manuf., Tech. and Sci.Products
14	Relationship Managers
15	Bank Tellers and Related Clerks
16	Door-To-Door Sales, News and Street Vendors
17	Electronics and Telecoms Installers and Repairers
18	Human Resources Specialists
19	Training and Development Specialists
20	Construction Laborers

Source

Future of Jobs Survey 2020, World Economic Forum.

Indústria 4.0

Quais as profissões que estão ganhando ou perdendo relevância?



Preditiva.ai

Carl Frey e Michael Osborne, professores da Universidade de Oxford, avaliaram tarefas cotidianas de mais de 700 ocupações, para identificar o que uma máquina poderá fazer melhor que os humanos nas próximas duas décadas. Chegaram a um índice que varia entre 0 (nenhum risco de substituição) e 100% (risco total).

Site:

<https://willrobotstakemyjob.com/>



Importante: É difícil prever as profissões que realmente vão desaparecer pois tudo muda muito rápido. Porém um bom ponto de partida é: **Quão repetitivo é meu trabalho?**

Accountants and Auditors

94%

[Vote](#) [Comments \(55\)](#)

Will "Accountants and Auditors" be replaced by AI & Robots?

It's highly likely this occupation will be replaced by robots/AI. However, our poll is less clear, and shows a 65% chance of automation within the next couple of decades.



AUTOMATION RISK LEVEL

You are doomed

or 94% probability of automation

PROJECTED GROWTH

11%

by 2024

PEOPLE EMPLOYED

1,280,700

as of 2019

MEDIAN ANNUAL WAGE

\$71,550


or \$34.40 hourly

Indústria 4.0

Apesar disso, parece que não estamos nos preparando muito bem para o futuro do trabalho



Dados: Jul/23

 dados

Brasil

Vagas ▾

Data do anúncio ▾

Nível de experiência ▾

Dados em: Brasil
39.378 resultados

Configurar alerta ☐

Quantidade de oportunidades para pessoas com habilidades em dados

Cenário atual das faculdades

TOP 50% dos cursos com maiores matrículas	Matrículas em 2018	Vagas em 2022	Carga horária
Direito	862.000	695	Quase inexistente
Administração	711.000	*	Baixa
Pedagogia	679.000	642	Quase inexistente
Engenharia Civil	360.000	830	Baixa
Ciências Contábeis	355.000	810	Baixa
Enfermagem	273.000	1793	Quase inexistente
Psicologia	235.000	842	Quase inexistente
Educação Física	185.000	305	Quase inexistente
Arquitetura e Urbanismo	167.000	1194	Quase inexistente
Engenharia de Produção	165.000	896	Média
Total	3.992.000		

* Administração é um curso com múltiplas possibilidades de vagas

Fonte: Censo da Educação Superior 2018 - Inep
Pesquisa de vagas LinkedIn Fev/22

Carga horária de análise de dados

Principais faculdades privadas

Indústria 4.0

Novas profissões já estão aí



Com as novas oportunidades trazidas pela revolução 4.0, novos grupos de funções estão sendo criado. Configura alguns:



Indústria 4.0

A transição de carreira na nova economia

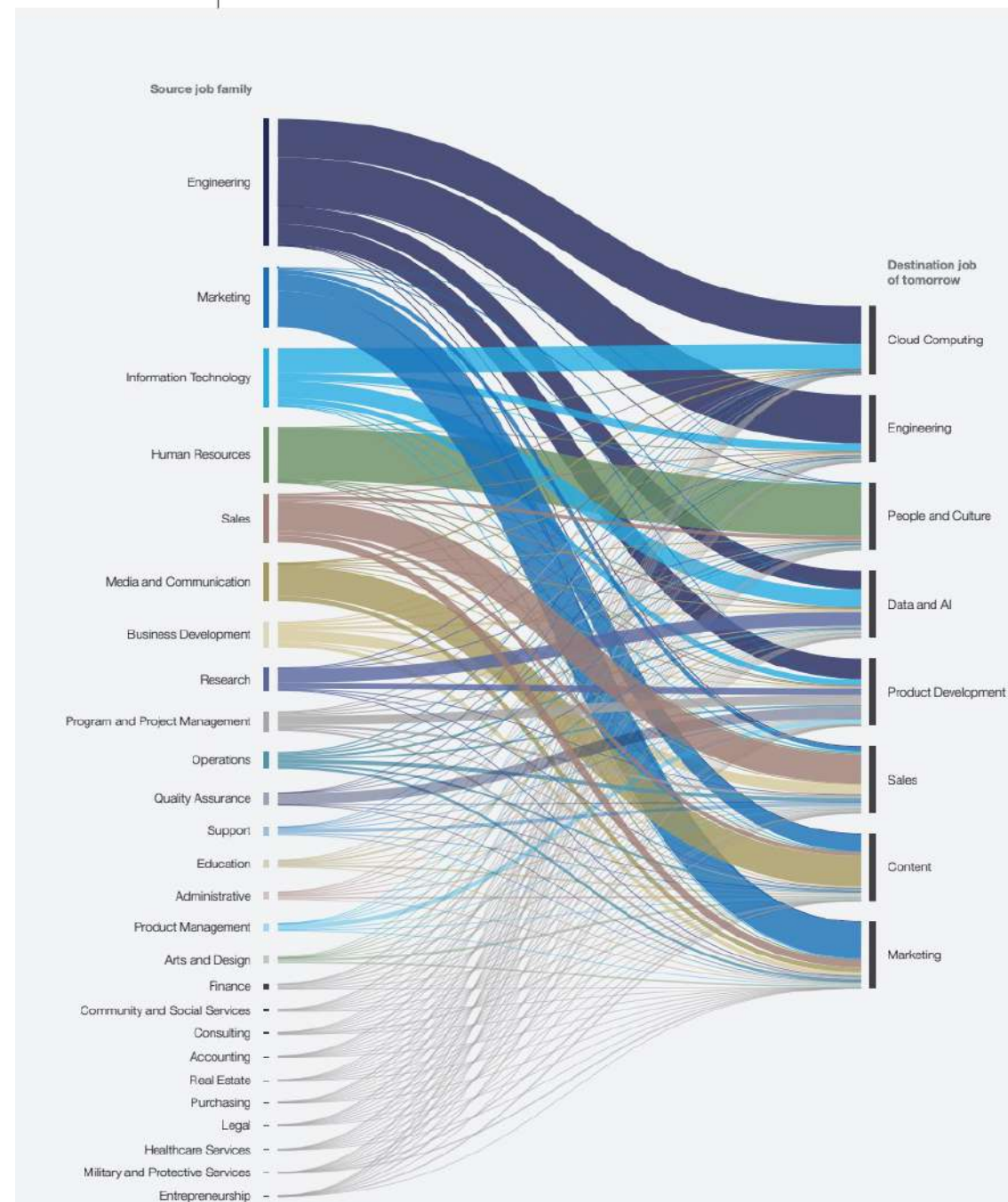
Um dos temas mais falados nos debates sobre o futuro do trabalho é sobre o **Reskilling** (Requalificação dos profissionais).

Isso porque boa parte das novas profissões exigirão novas habilidades, como análise de dados, inteligência emocional, pensamento crítico entre outros.

Na figura ao lado, o **Fórum Econômico Mundial** traz uma análise de habilidades para tentar ajudar os profissionais na afinidade de suas atuais funções com os novas funções dos próximos anos.

FIGURE 25

Transitions into the jobs of the future





Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

Quais as habilidades a serem
desenvolvidas neste ano em diante

Futuro do trabalho



E quais são essas novas habilidades a serem desenvolvidas?

Basicamente, as habilidades para tomada de decisão com o uso de Dados e as habilidades que nos diferenciam dos robôs, ou seja, as habilidades humanas. Vejam as principais:

☐ Flexibilidade

O que é?

Capacidade de se adaptar facilmente a cenários diferentes do habitual.

Por que desenvolver?

Os processos e oportunidades trazidas pela nova economia vão mudar cada vez mais rápido. Os ciclos de mudança diminuíram drasticamente ao longo dos anos.

Futuro do trabalho



E quais são essas novas habilidades a serem desenvolvidas?

Basicamente, as habilidades para tomada de decisão com o uso de Dados e as habilidades que nos diferenciam dos robôs, ou seja, as habilidades humanas. Vejam as principais:

☐ Flexibilidade

☐ Metodologias Ágeis

O que é?

Nomes como MVP: Produto Mínimo Viável, Validação, Sprints, Cultura do Fail Fast (Tentativa e Erro) e Testes A/B fazem parte dessa cultura ágil.

Por que desenvolver?

Porque é a nova forma como os times produtivos trabalham para entregar valor no menor tempo possível.

Futuro do trabalho

E quais são essas novas habilidades a serem desenvolvidas?



Basicamente, as habilidades para tomada de decisão com o uso de Dados e as habilidades que nos diferenciam dos robôs, ou seja, as habilidades humanas. Vejam as principais:

☐ Flexibilidade

☐ Metodologias Ágeis

☒ Inteligência emocional

O que é?

A inteligência emocional é a capacidade de colocar-se no lugar do outro, agindo de maneira empática e também dominar os próprios sentimentos.

Por que desenvolver?

Porque ajuda no relacionamento humano e na gestão das emoções sempre bem desafiadas em cenários de mudanças repentinas.

Futuro do trabalho



E quais são essas novas habilidades a serem desenvolvidas?

Basicamente, as habilidades para tomada de decisão com o uso de Dados e as habilidades que nos diferenciam dos robôs, ou seja, as habilidades humanas. Vejam as principais:

- ☐ Flexibilidade
- ☐ Metodologias Ágeis
- ☐ Inteligência emocional
- ☒ Capacidade de Auto Gestão

O que é?

Esqueça os chefes que fazem micro gerenciamento. Para os próximos anos a gestão do seu trabalho será cada vez mais sua.

Por que desenvolver?

Com o home-office e ambientes cada vez menos hierárquicos, precisamos desenvolver a capacidade de gerir nosso próprio desempenho e produtividade.

Futuro do trabalho

E quais são essas novas habilidades a serem desenvolvidas?



Basicamente, as habilidades para tomada de decisão com o uso de Dados e as habilidades que nos diferenciam dos robôs, ou seja, as habilidades humanas. Vejam as principais:

- ☐ Flexibilidade
- ☐ Metodologias Ágeis
- ☐ Inteligência emocional
- ☐ Capacidade de Auto Gestão
- ☒ Criatividade

O que é?

Capacidade de pensar “fora da caixa” e propor novas formas de solucionar problemas que aparecem a todo instante.

Por que desenvolver?

Os robôs não conseguem fazer isso. Nós conseguimos!

Futuro do trabalho

E quais são essas novas habilidades a serem desenvolvidas?



Basicamente, as habilidades para tomada de decisão com o uso de Dados e as habilidades que nos diferenciam dos robôs, ou seja, as habilidades humanas. Vejam as principais:

- ☐ Flexibilidade
- ☐ Metodologias Ágeis
- ☐ Inteligência emocional
- ☐ Capacidade de Auto Gestão
- ☐ Criatividade
- ☒ Comunicação

O que é?

Habilidade para se comunicar em diversos tipos de cenários, contextos e pessoas diferentes.

Por que desenvolver?

Porque estamos cada vez mais tecnológicos e menos disponíveis presencialmente. Sem comunicação não há interatividade humana.

Futuro do trabalho

E quais são essas novas habilidades a serem desenvolvidas?



Basicamente, as habilidades para tomada de decisão com o uso de Dados e as habilidades que nos diferenciam dos robôs, ou seja, as habilidades humanas. Vejam as principais:

- ☐ Flexibilidade
- ☐ Metodologias Ágeis
- ☐ Inteligência emocional
- ☐ Capacidade de Auto Gestão
- ☐ Criatividade
- ☐ Comunicação
- ☐ Life Long Learning

O que é?

Capacidade de aprender de forma contínua, sem data de término.

Por que desenvolver?

Estaremos em constante formação. Não existirá mais a frase de faculdade “quando eu me formar”. A nova economia trará cada vez mais desafios e skills a serem aprendidos.

Futuro do trabalho

E quais são essas novas habilidades a serem desenvolvidas?



Basicamente, as habilidades para tomada de decisão com o uso de Dados e as habilidades que nos diferenciam dos robôs, ou seja, as habilidades humanas. Vejam as principais:

- ☐ Flexibilidade
- ☐ Metodologias Ágeis
- ☐ Inteligência emocional
- ☐ Capacidade de Auto Gestão
- ☐ Criatividade
- ☐ Comunicação
- ☐ Life Long Learning
- ☒ Empreendedorismo

O que é?

Transformar ideias em produtos tangíveis.

Por que desenvolver?

Vamos viver muito. Como o seu propósito será alcançado ? É bem possível que iremos fazer muito mais transições de carreira do que antes. Como estará o mercado de trabalho para pessoas de maior idade?

Futuro do trabalho

E quais são essas novas habilidades a serem desenvolvidas?



Basicamente, as habilidades para tomada de decisão com o uso de Dados e as habilidades que nos diferenciam dos robôs, ou seja, as habilidades humanas. Vejam as principais:

- ☐ Flexibilidade
- ☐ Metodologias Ágeis
- ☐ Inteligência emocional
- ☐ Capacidade de Auto Gestão
- ☐ Criatividade
- ☐ Comunicação
- ☐ Life Long Learning
- ☐ Empreendedorismo
- ☐ Análise de Dados e Mentalidade Analítica

O que é?

Arelado ao pensamento crítico, é cada vez mais necessário que os profissionais desta nova economia entendam como funciona a tomada de decisão baseada em dados, além de apresentarem as principais técnicas analíticas e as ferramentas que proporcionam aplicar essas técnicas.

Por que desenvolver?

Sobrevivência na nova economia.



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

O que são os "dados" e qual o papel
das técnicas e ferramentas nisso tudo?

Dados, técnicas e ferramentas

Um overview



Ingredientes Brutos



Dados Brutos e Não Estruturados

Preparação



"mise en place"



Ingredientes Organizados



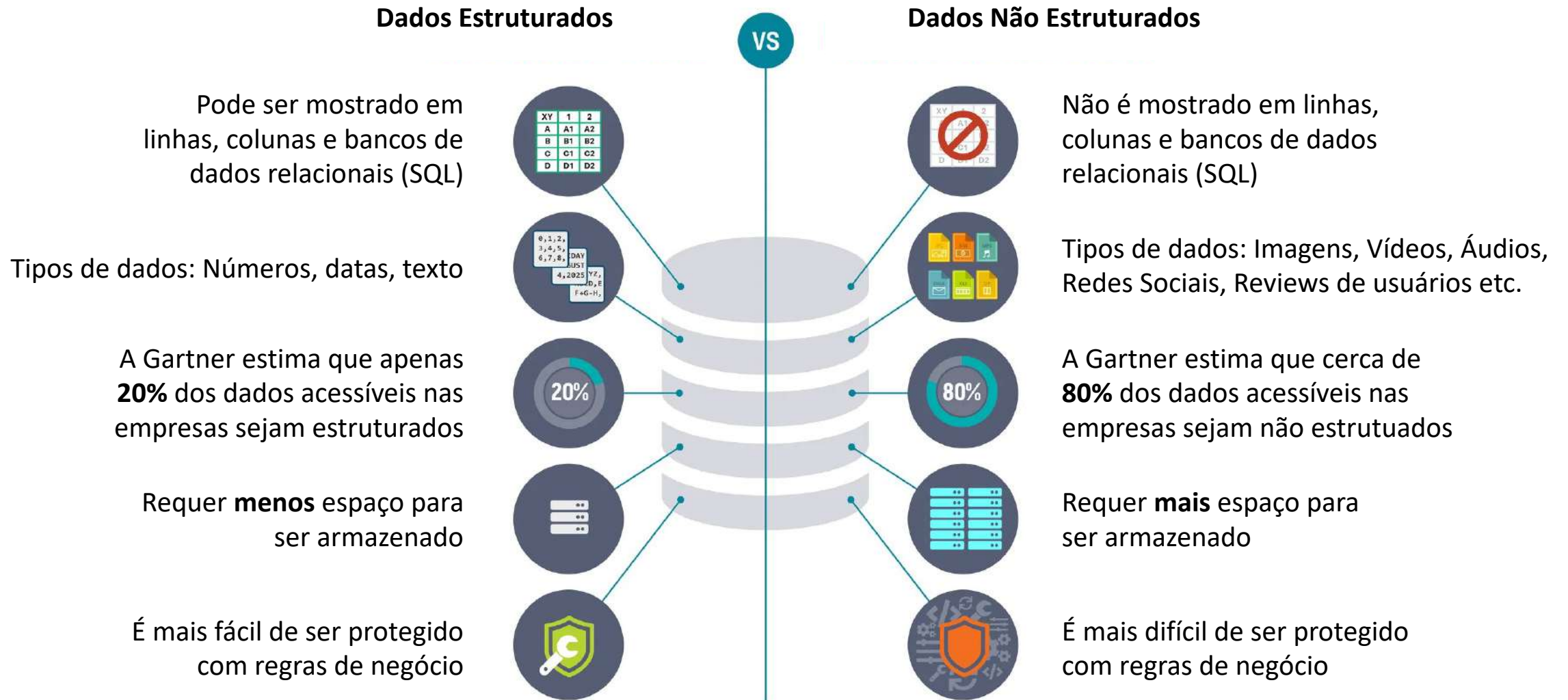
Dados Estruturados



"Data Preparation (Data Prep)"

Dados, técnicas e ferramentas

Um overview



Dados, técnicas e ferramentas

Um overview



Ingredientes



Eletrodomésticos
e utensílios



Receitas



Os pratos



Dados

**Ferramentas
de dados**

**Técnicas e
Metodologias de dados**

**Valor Gerado
pelos Dados**

Dados, técnicas e ferramentas

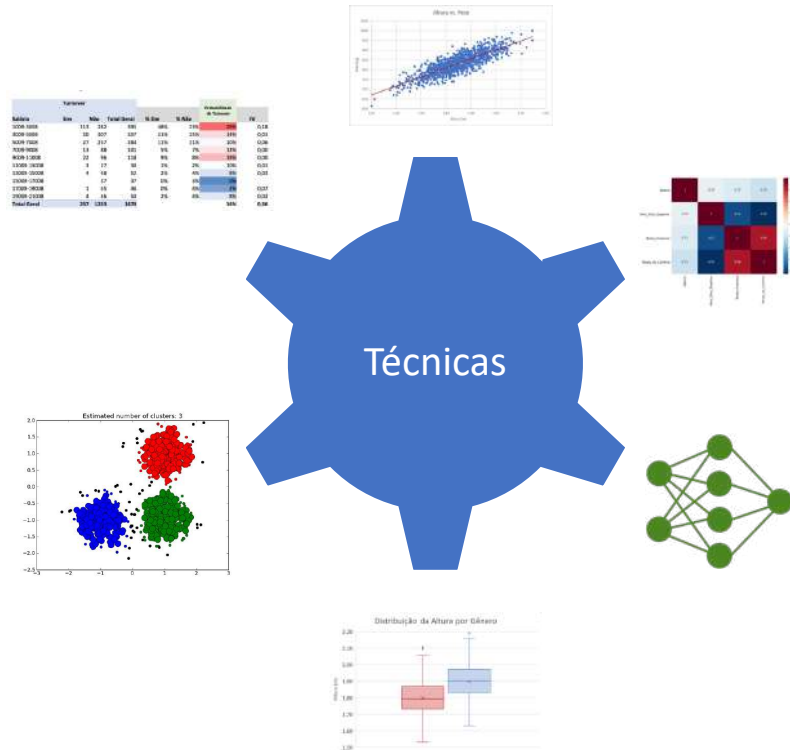
Um overview



		Usou as Ferramentas adequadamente?	Usou as Técnicas e Metodologias?	Resultado	
Gastronomia	Situação 1	Não	Não	Prato ruim e demorado	😡
	Situação 2	Não	Sim	Prato gostoso, mas podia ser mais rápido, neh?	😊
	Situação 3	Sim	Sim	Que experiência fantástica!	🥰
Mundo dos dados	Situação 1	Não	Não	Demorou tanto para um estudo superficial?	😡
	Situação 2	Não	Sim	Que bacana este estudo. Vai ajudar muito. Na próxima seria interessante demorar menos.	😊
	Situação 3	Sim	Sim	Solução que gera muito valor para o negócio	💰

Dados, técnicas e ferramentas

Resumo



Respondem as perguntas
de negócio

Quanto mais técnicas conhecer,
mas perguntas vai responder e
mais valor vai gerar



Geram produtividade para
responder as perguntas

Quanto mais conhecer as
principais ferramentas, mais
produtivo será



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

Overview das principais técnicas de Análise de Dados – Parte 1

Análise de Dados

Overview



Análise de Dados é uma “máquina de solução de problemas”. Mas para que funcione, **precisamos de pelo menos três “engrenagens”**. São elas:



- ✓ Matemática fundamental
- ✓ Técnicas de Otimização
- ✓ Estatística Descritiva
- ✓ Probabilidades
- ✓ Inferência (ex: Testes de Hipótese...)
- ✓ Modelagem Estatística
- ✓ *Machine Learning*
- ✓ Entre outros...



- ✓ Planilhas e Tipos de Arquivos
- ✓ Linguagens de programação
- ✓ Bancos de dados
- ✓ Visualização de Dados
- ✓ Infra estrutura (DWs, Data Lakes etc)
- ✓ ETL / Pipelines de Dados
- ✓ Entre outros...



- ✓ Entendimento dos processos internos
- ✓ Alinhamento com os objetivos da empresa
- ✓ Entrosamento com o time e principais *stakeholders*
- ✓ Análise da concorrência
- ✓ Entre outros...

Análise de Dados

Overview



Análise de Dados é uma “máquina de solução de problemas”. Mas para que funcione, **precisamos de pelo menos três “engrenagens”**. São elas:

Técnicas

- ✓ Matemática fundamental
- ✓ Técnicas de Otimização
- ✓ Estatística Descritiva
- ✓ Probabilidades
- ✓ Inferência (ex: Testes de Hipótese...)
- ✓ Modelagem Estatística
- ✓ *Machine Learning*
- ✓ Entre outros...

Ferramentas



- ✓ Planilhas e Tipos de Arquivos
- ✓ Linguagens de programação
- ✓ Bancos de dados
- ✓ Visualização de Dados
- ✓ Infra estrutura (DWs, Data Lakes etc)
- ✓ ETL / Pipelines de Dados
- ✓ Entre outros...

Conhecimento de Negócio

- ✓ Entendimento dos processos internos
- ✓ Alinhamento com os objetivos da empresa
- ✓ Entrosamento com o time e principais *stakeholders*
- ✓ Análise da concorrência
- ✓ Entre outros...

Overview sobre técnicas

O que temos disponível no mercado?

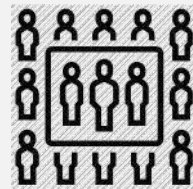


Arsenal analítico



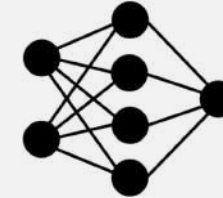
**Estatística
e Probabilidades**

- Resumir uma base de dados
- Entender a relação entre os variáveis



**Inferência
Estatística**

- Estimar valores através de amostras
- Testar hipóteses através de amostras



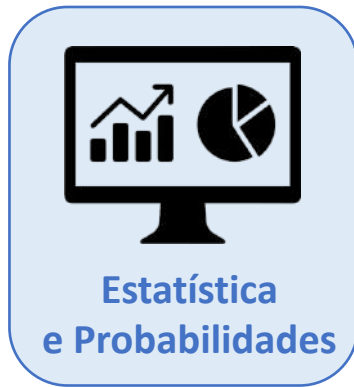
**Machine Learning
e AI**

- Estimar valores de forma automática
- Classificar “coisas” de forma automática
- Gerar dados de forma automática

Overview sobre técnicas

O que temos disponível no mercado?

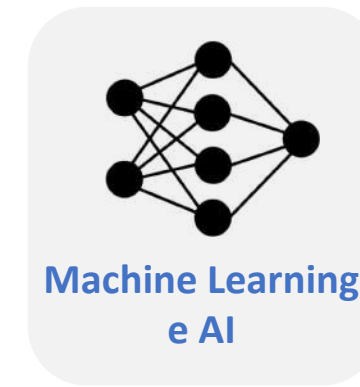
Arsenal analítico



- Resumir uma base de dados
- Entender a relação entre os variáveis



- Estimar valores através de amostras
- Testar hipóteses através de amostras



- Estimar valores de forma automática
- Classificar “coisas” de forma automática
- Gerar dados de forma automática

Estatística e Probabilidades

Estatística Descritiva



O que é?

Conjunto de técnicas para resumir dados de todos os tipos e tamanhos, tanto via **tabelas**, **medidas** ou **gráficos**.

Quando usar?

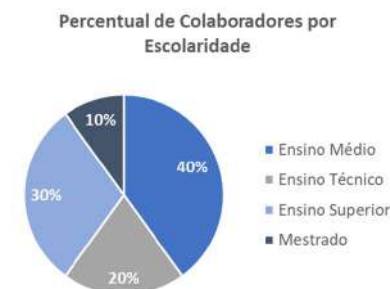
Sempre que quiser **resumir uma ou mais colunas de dados** para interpretar o comportamento.

Técnicas mais usadas

- Tabelas de Frequência
- Medidas Resumo (Média, Mediana, Quartis, Desvio Padrão etc)
- Gráficos analíticos (Barras, Linhas, Pareto, Histograma, Box-Plots etc)

Exemplos

Escolaridade	Frequência absoluta	Frequência relativa
Ensino Médio	4	40%
Ensino Técnico	2	20%
Ensino Superior	3	30%
Mestrado	1	10%
Total	10	100%



Estatística e Probabilidades

Estatística Descritiva

O que é?

Conjunto de técnicas para resumir dados de todos os tipos e tamanhos, tanto via **tabelas**, **medidas** ou **gráficos**.

Quando usar?

Sempre que quiser **resumir uma ou mais colunas de dados** para interpretar o comportamento.

Técnicas mais usadas

- Tabelas de Frequência
- Medidas Resumo (Média, Mediana, Quartis, Desvio Padrão etc)
- Gráficos analíticos (Barras, Linhas, Pareto, Histograma, Box-Plots etc)



Onde estudar?

Módulo

Resumindo os dados

Tabelas, Medidas,
Gráficos Analíticos



2) Como resumir e interpretar dados?

Transforme dados em conhecimento, gerando poderosos insights e tomando as melhores decisões utilizando técnicas simples e práticas de Estatística Descritiva.

Estatística e Probabilidades

Análise de Associação



O que é?

Conjunto de técnicas para analisar o comportamento conjunto de **duas** variáveis

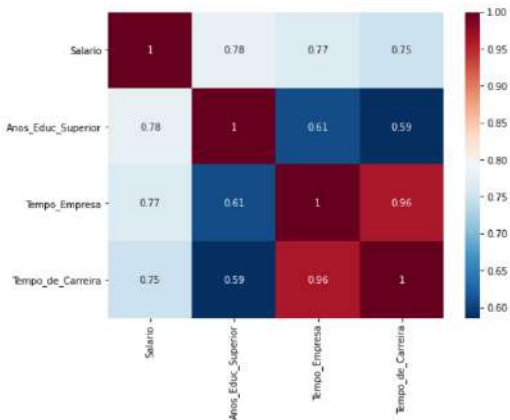
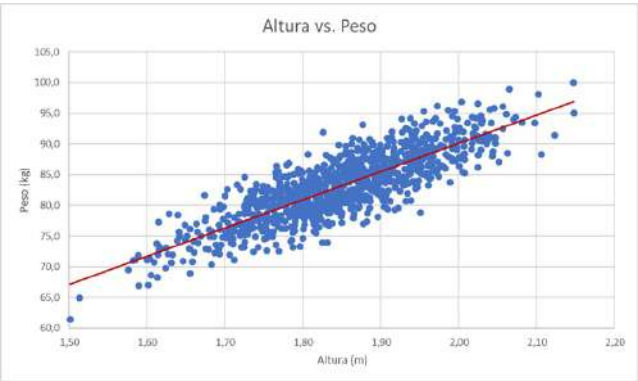
Quando usar?

Sempre que quiser **resumir simultaneamente duas colunas** e interpretar o comportamento e associação entre elas.

Técnicas mais usadas

- Correlação de Pearson (quando dados lineares)
- Coeficiente de Determinação (R Quadrado)
- Information Value (IV)
- Estudo de Causalidade

Exemplos



Turnover						Probabilidade de Turnover	IV
Salário	Sim	Não	Total Geral	% Sim	% Não		
1009-3008	113	282	395	48%	23%	29%	0,18
3009-5008	50	307	357	21%	25%	14%	0,01
5009-7008	27	257	284	11%	21%	10%	0,06
7009-9008	13	88	101	5%	7%	13%	0,00
9009-11008	22	96	118	9%	8%	19%	0,00
11009-13008	3	27	30	1%	2%	10%	0,01
13009-15008	4	48	52	2%	4%	8%	0,02
15009-17008		37	37	0%	3%	0%	
17009-19008	1	45	46	0%	4%	2%	0,07
19009-21008	4	46	50	2%	4%	8%	0,02
Total Geral	237	1233	1470			16%	0,36

Estatística e Probabilidades

Análise de Associação

O que é?

Conjunto de técnicas para analisar o comportamento conjunto de **duas** variáveis

Quando usar?

Sempre que quiser **resumir simultaneamente duas colunas** e interpretar o comportamento e associação entre elas.

Técnicas mais usadas

- Correlação de Pearson (quando dados lineares)
- Coeficiente de Determinação (R Quadrado)
- Information Value (IV)
- Estudo de Causalidade



Onde estudar?

Módulo

Análise de Correlação

Pearson, R Quadrado,
IV, Causalidade

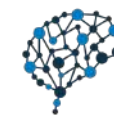


3) Análise de Correlação e Associação

Entenda como cada variável de sua base de dados se relaciona usando técnicas de correlação e associação. Explore esses insights do jeito certo.

Estatística e Probabilidades

Probabilidade



Preditiva.ai

O que é?

Conjunto de técnicas para **prever** o acontecimento de algum evento.

Quando usar?

Quando seu trabalho envolver projeções futuras de indicadores e eventos de interesse. (ex: Inadimplência de clientes, Churn, TurnOver, Vendas etc.)

Técnicas mais usadas

- Teoria Frequentista
- Modelos de Probabilidade (Uniforme, t-student, Normal etc)
- Machine Learning

Exemplos

Funil para um E-commerce



Estatística e Probabilidades

Probabilidade

O que é?

Conjunto de técnicas para **prever** o acontecimento de algum evento.

Quando usar?

Quando seu trabalho envolver projeções futuras de indicadores e eventos de interesse. (ex: Inadimplência de clientes, Churn, TurnOver, Vendas etc.)

Técnicas mais usadas

- Teoria Frequentista
- Modelos de Probabilidade (Uniforme, t-student, Normal etc)
- Machine Learning



Onde estudar?

Módulo

Prevendo o futuro

Teoria Frequentista, Modelos de Probabilidade

4) Prevendo o futuro

Não temos bola de cristal, porém conseguimos dar ótimos palpites com o uso correto das técnicas de Probabilidade e Análise Preditiva. Entenda como utilizar Probabilidades.



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

Overview das principais técnicas de Análise de Dados – Parte 2

Overview sobre técnicas

O que temos disponível no mercado?

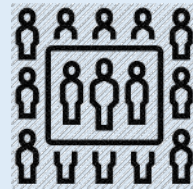


Arsenal analítico



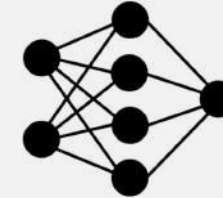
**Estatística
e Probabilidades**

- Resumir uma base de dados
- Entender a relação entre os variáveis



**Inferência
Estatística**

- Estimar valores através de amostras
- Testar hipóteses através de amostras



**Machine Learning
e AI**

- Estimar valores de forma automática
- Classificar “coisas” de forma automática
- Gerar dados de forma automática

Inferência Estatística

Estimação



O que é?

Conjunto de técnicas para **estimar valores** de interesse **através de amostras de dados**.

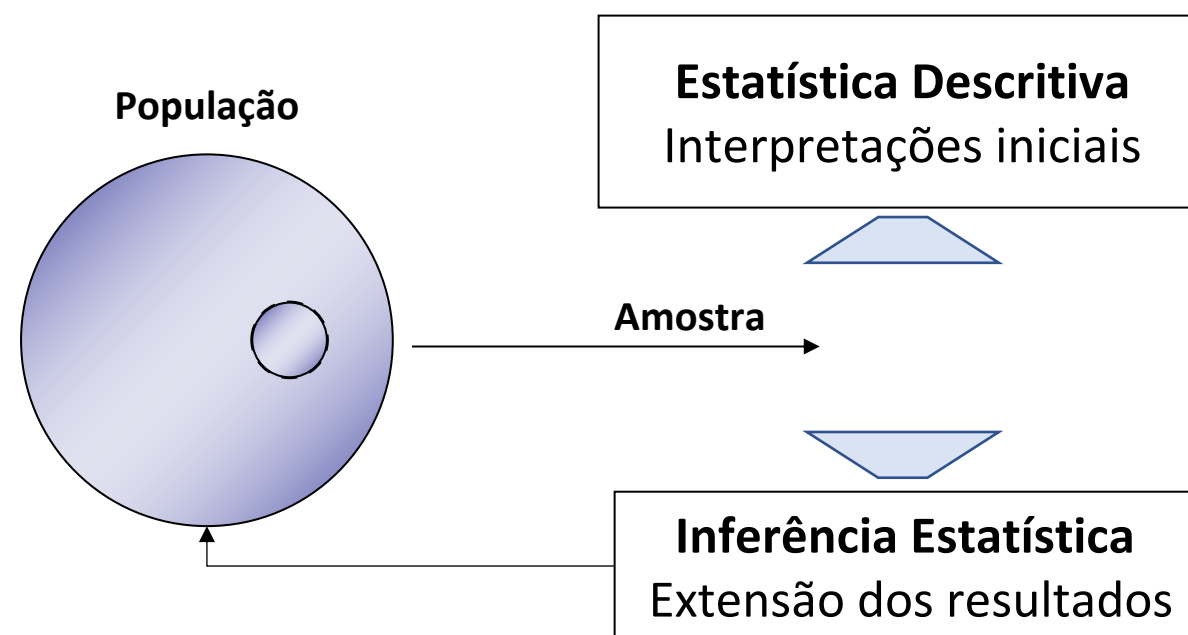
Quando usar?

Sempre que **não tiver acesso à base completa** (população) do evento de interesse. Ex: Pesquisas de Opinião, Pesquisas de Satisfação etc.

Técnicas mais usadas

- Extração de amostras randomizadas e com o menor viés possível
- Estimação por Intervalos de Confiança
 - Cálculo do tamanho da Amostra
 - Controle da Margem de Erro

Diagrama



Inferência Estatística

Estimação

O que é?

Conjunto de técnicas para **estimar valores** de interesse **através de amostras de dados**.

Quando usar?

Sempre que **não tiver acesso à base completa** (população) do evento de interesse. Ex: Pesquisas de Opinião, Pesquisas de Satisfação etc.

Técnicas mais usadas

- Extração de amostras randomizadas e com o menor viés possível
- Estimação por Intervalos de Confiança
 - Cálculo do tamanho da Amostra
 - Controle da Margem de Erro



Onde estudar?

Módulo

Como trabalhar com amostras de dados?

Intervalos de confiança, margem de erro, amostras



5) Como trabalhar com amostras de dados

Na maioria da vezes não temos acesso à todos os clientes/dados. Desta forma, precisamos trabalhar com amostras. Descubra como!

Inferência Estatística

Testes de Hipóteses

O que é?

Técnica para se **testar** valores da população **através de amostras de dados**.

Quando usar?

Sempre que tiver uma hipótese sobre a população e quiser ter um **suporte estatístico para a tomada de decisão**.

Técnicas mais usadas

- Experimentos (Grupos de Controle e Teste)
- Testes de Hipóteses para médias e proporções
- Testes de Hipóteses para Grupos (Teste A/B)
- Análise do p-valor

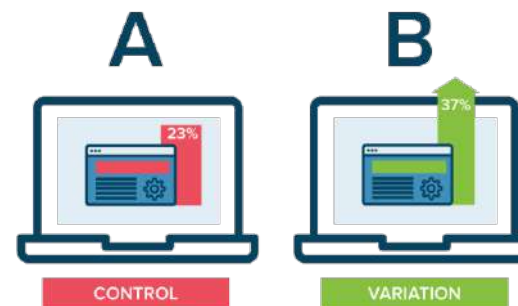


Preditiva.ai

Exemplos

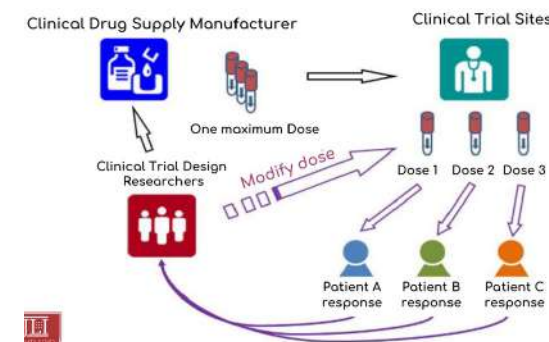
Testes A/B

Comparação do efeito de duas campanhas diferentes de marketing.



Estudos Clínicos

Medicamento vs. Placebo, Tratamento A vs. Tratamento B.



Fonte: <https://www.iliedu.com/blog/adaptive-design-clinical-trials/>

Inferência Estatística

Testes de Hipóteses

O que é?

Técnica para se **testar** valores da população **através de amostras de dados**.

Quando usar?

Sempre que tiver uma hipótese sobre a população e quiser ter um **suporte estatístico para a tomada de decisão**.

Técnicas mais usadas

- Experimentos (Grupos de Controle e Teste)
- Testes de Hipóteses para médias e proporções
- Testes de Hipóteses para Grupos (Teste A/B)
- Análise do p-valor



Onde estudar?

Módulo

Trabalhando com Testes A/B

Testes de hipóteses, p-valor, médias e proporções



1) Testes A/B (Testes de Hipóteses)

Entenda como fornecer significância estatística a qualquer estudo analítico com o uso dos testes de hipóteses. Importante para avaliar os experimentos que realizamos no dia a dia.

Testes de Hipóteses

Exemplo em *Growth Marketing*

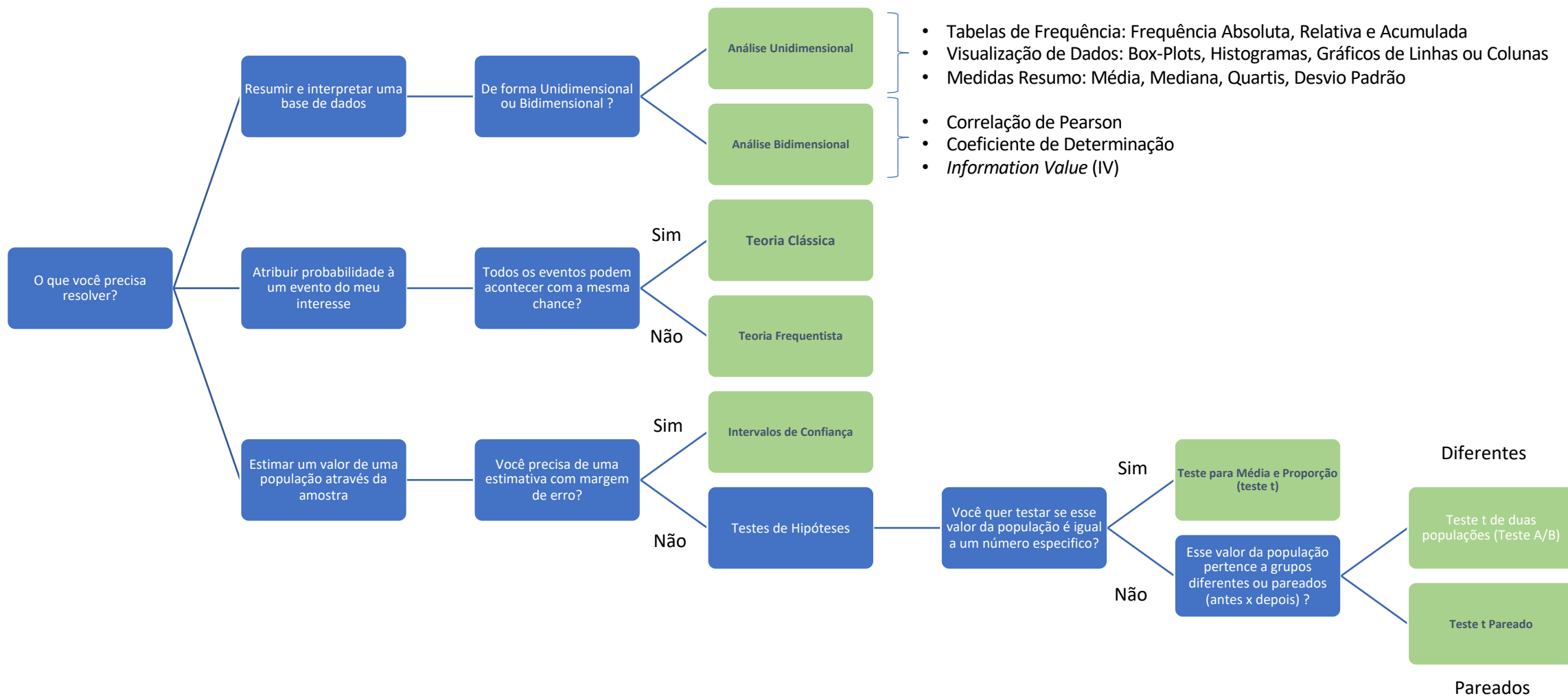


Exemplo de **Teste A/B com mudanças (variantes) em uma página de internet**. Ferramentas como Google Optimize ajudam muito na aplicação das técnicas.

	DADOS OBSERVADOS			
	Sessões de experimentos	Conversões do experimento	Taxa de conversão calculada	Taxa de conversão estimada
Original	175	8	4,57%	2,0% 9,9%
Variante 1	173	7	4,05%	2,1% 9,7%
Variante 2	142	14	9,86%	5,9% 17,8%
Variante 3	186	6	3,23%	1,4% 8,0%

Resumo

Resumo das principais técnicas estatísticas até aqui





Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

Overview das principais técnicas de Análise de Dados – Parte 3

Overview sobre técnicas

O que temos disponível no mercado?

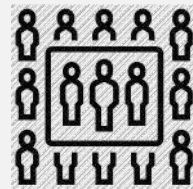


Arsenal analítico



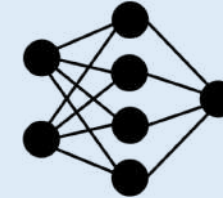
**Estatística
e Probabilidades**

- Resumir uma base de dados
- Entender a relação entre os variáveis



**Inferência
Estatística**

- Estimar valores através de amostras
- Testar hipóteses através de amostras



**Machine Learning
e AI**

- Estimar valores de forma automática
- Classificar “coisas” de forma automática
- Gerar dados de forma automática

Machine Learning

Modelos Supervisionados



O que é?

Conjunto de técnicas que estimam valores numéricos (**regressores**) ou classificam “eventos” de forma automática (**classificadores**). Obs: É **obrigatório** ter a variável resposta (target) mensurada.

Quando usar?

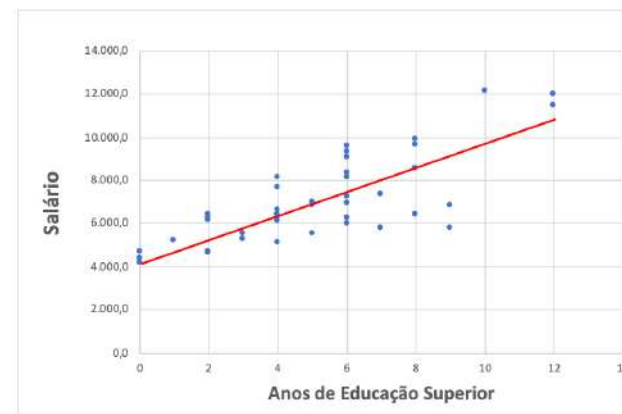
Sempre que precisar tomar uma **decisão estatística de forma automatizada**, via sistemas, algoritmos ou algum produto de dados.

Técnicas mais usadas

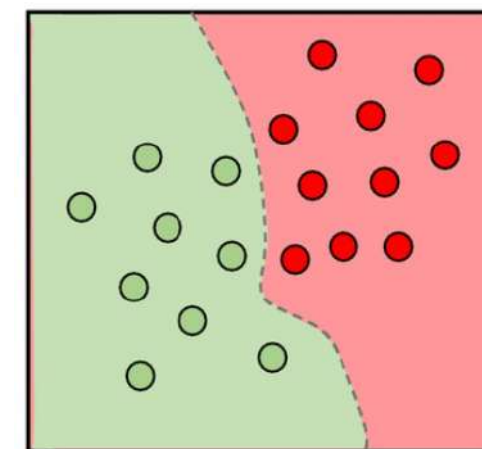
- Regressões
- Árvores de Decisão
- Redes Neurais Artificiais
- Modelos combinados (ensemble)

Exemplos

Estimação de Salário (Regressão)



Predição de Churn (Classificação)



Valor da
assinatura

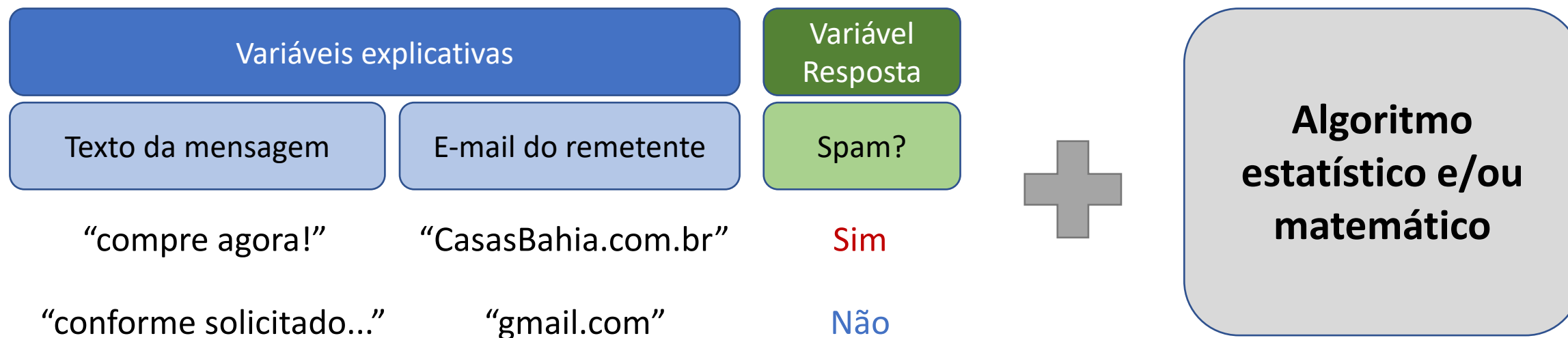
Número de reclamações

Machine Learning

De forma geral



Como a máquina “**aprende**”? Exemplo de modelo para detectar spam (modelo classificador):



**Modelo de Machine Learning
para detecção de Spam**



Machine Learning

Criação de modelos que explicam uma infinidade de perguntas, como por exemplo:

Tecnologia

- Detecção de SPAM
- Carros autônomos
- Drones
- Internet das Coisas
- Compressão de arquivos

Finanças

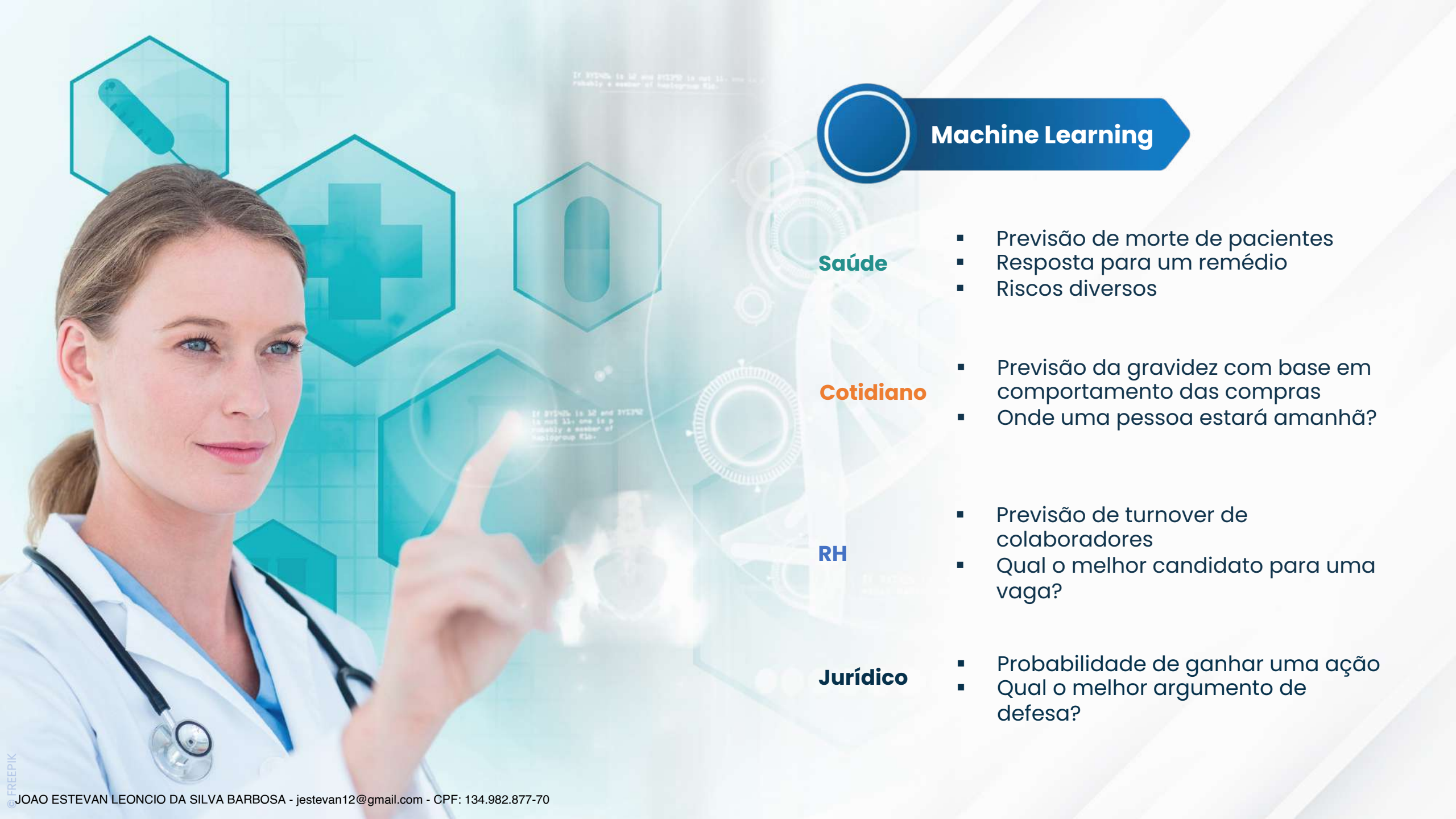
- Probabilidade de “calote” de clientes
- Trading algorítmico
- Precificação de seguros

Atendimento

- Robôs de autoatendimento
- Maximização de NPS
- Monitoria da Qualidade automática

Marketing e CRM

- Qual o cliente mais propenso a comprar?
- Recomendação de produtos
- Previsão de Churn



Machine Learning

Saúde

- Previsão de morte de pacientes
- Resposta para um remédio
- Riscos diversos

Cotidiano

- Previsão da gravidez com base em comportamento das compras
- Onde uma pessoa estará amanhã?

RH

- Previsão de turnover de colaboradores
- Qual o melhor candidato para uma vaga?

Jurídico

- Probabilidade de ganhar uma ação
- Qual o melhor argumento de defesa?

Machine Learning

Modelos Supervisionados



O que é?

Conjunto de técnicas que estimam valores numéricos (**regressores**) ou classificam “eventos” de forma automática (**classificadores**). Obs: É **obrigatório** ter a variável resposta (target) mensurada.

Quando usar?

Sempre que precisar tomar uma **decisão estatística de forma automatizada**, via sistemas, algoritmos ou algum produto de dados.

Técnicas mais usadas

- Regressões
- Árvores de Decisão
- Redes Neurais Artificiais
- Modelos combinados (ensemble)

Onde estudar?

Módulo Regressão Linear Múltipla Modelagem, Estimação, R^2 , Variáveis Dummy 	Módulo Regressão Logística Conceitos, aplicação e interpretação do modelo 	Módulo Árvores de decisão Árvores de regressão e classificação 
2) Regressão Linear Múltipla Entenda como estimar medidas com o uso simultâneo de dezenas de variáveis de sua base de dados. Aprenda como criar modelos estatísticos.	6) Regressão Logística Desenvolva modelos que te permitirão classificar os dados quando sua variável resposta for do tipo binária (Ex: Cliente Comprou ou Não Comprou, Churn ou Não Churn entre outros).	8) Árvores de Decisão Aprenda um dos algoritmos mais robustos de Machine Learning. Com as árvores de decisão, você conseguirá prever variáveis numéricas ou qualitativas.
Módulo Ensemble de modelos Blending, Bagging e, Boosting de modelos 1+1 > 2	Módulo Redes Neurais Artificiais Conceitos, aplicação e interpretação do modelo 	Módulo Diagnóstico de Modelos Métricas, Validação Cruzada, Interpretabilidade de modelos 
9) Ensemble de Modelos no Python Utilize técnicas de combinação de modelos para obter um desempenho maior no desenvolvimento de modelos supervisionados. Algoritmos “estado da arte” são produzidos assim.	10) Redes Neurais Artificiais no Python Aprenda a utilizar uma das técnicas mais poderosas e complexas do grupo de Machine Learning.	7) Diagnóstico e Interpretação de Modelos Garantindo a qualidade dos modelos desenvolvidos avaliando o desempenho, generalização e sua interpretação de negócio.

Machine Learning

Modelos Não Supervisionados



O que é?

Conjunto de técnicas buscam padrões nos dados sem depender de uma variável resposta (target) mensurada.

Quando usar?

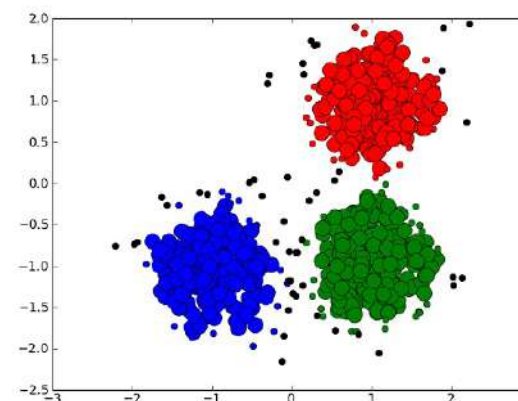
Sempre que quiser entender concentrações, dados fora de padrão (outliers) e realizar visualizações de dados complexas com múltiplas variáveis.

Técnicas mais usadas

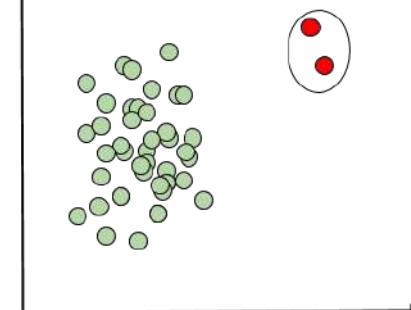
- Clusterizações (método k-means, DBScan etc)
- Detecção de Anomalias
- PCA (componentes principais)
- t-SNE

Exemplos

Criação de perfis de clientes



Detecção de Anomalias em Fraudes



Machine Learning

Modelos Não Supervisionados



O que é?

Conjunto de técnicas buscam padrões nos dados sem depender de uma variável resposta (target) mensurada.

Quando usar?

Sempre que quiser entender concentrações, dados fora de padrão (outliers) e realizar visualizações de dados complexas com múltiplas variáveis.

Técnicas mais usadas

- Clusterizações (método k-means, DBScan etc)
- Detecção de Anomalias
- PCA (componentes principais)
- t-SNE

Onde estudar?

Módulo Clusterização: K-means

Conceito, aplicação e interpretação do algoritmo



11) Clusterização

Com essa técnica é possível identificar grupos de observações que tem características semelhantes apenas considerando as variáveis explicativas.

Módulo Redução de Dimensionalidade

Algoritmos PCA, t-SNE e AutoEncoder



12) Redução de Dimensionalidade

Em algumas situações, reduzir o número de variáveis disponíveis traz muitos benefícios, principalmente do ponto de vista computacional.

Módulo Detecção de Outliers (dados discrepantes)

Algoritmos Isolation Forest e AutoEncoder



13) Detecção de Outliers (anomalias)

Identifique outliers (dados discrepantes ou anomalias) utilizando técnicas sofisticadas como Isolation Forest e AutoEncoder.

Vini, vou ter que aprender tudo isso?

Não! Tudo vai depender do seu objetivo de carreira (ser um analista ou gestor com habilidades em dados ou ser um analista ou cientista de dados especialista)

A dica é usar os níveis de especialização da plataforma como norte e ir aprendendo outros assuntos quando tiver necessidade na sua carreira.



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

Excel para Análise de Dados

Demonstração do Excel



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

SQL para Análise de Dados

Ferramentas de dados

SQL



Até hoje, a forma mais utilizada para acesso aos dados é através de uma **Linguagem de Consulta Estruturada**, ou do inglês, **Structured Query Language – SQL** (lê-se “Siquou”).

Por exemplo:

```
SELECT
    Clientes
FROM
    Tabela_de_Clientes
WHERE
    Idade > 18
    and Produto = 'Samsung Galaxy S'
```

O SQL foi desenvolvido originalmente no início dos anos 70 pela IBM e hoje, após várias evoluções e padronizações, se tornou a linguagem padrão para banco de dados. Exemplos de fornecedores e versões do SQL:



Demonstração do SQL



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

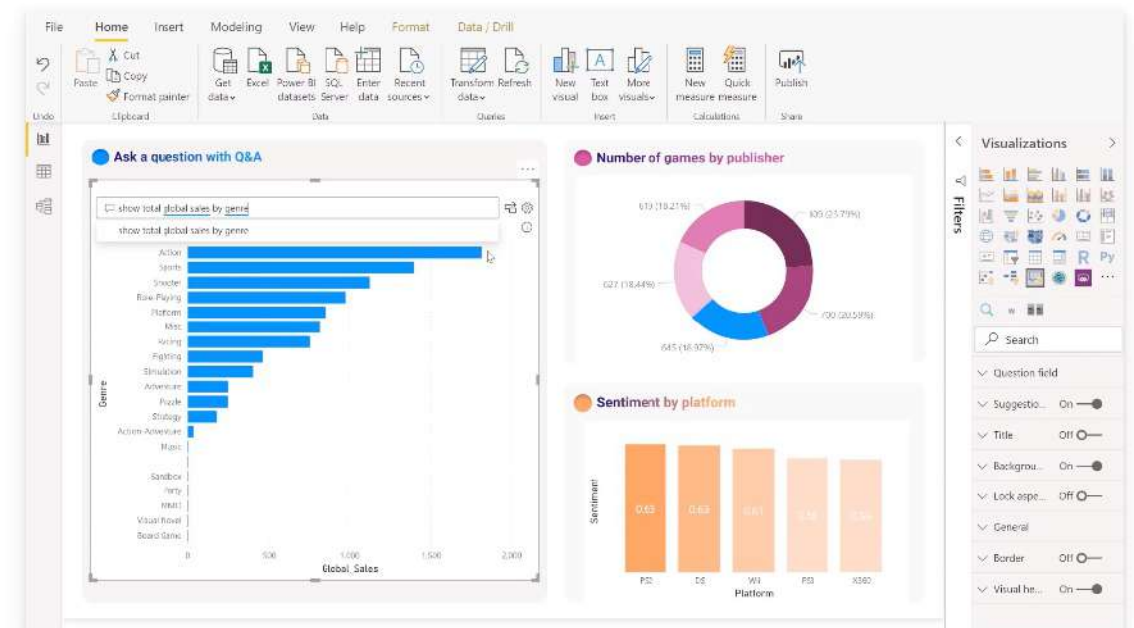
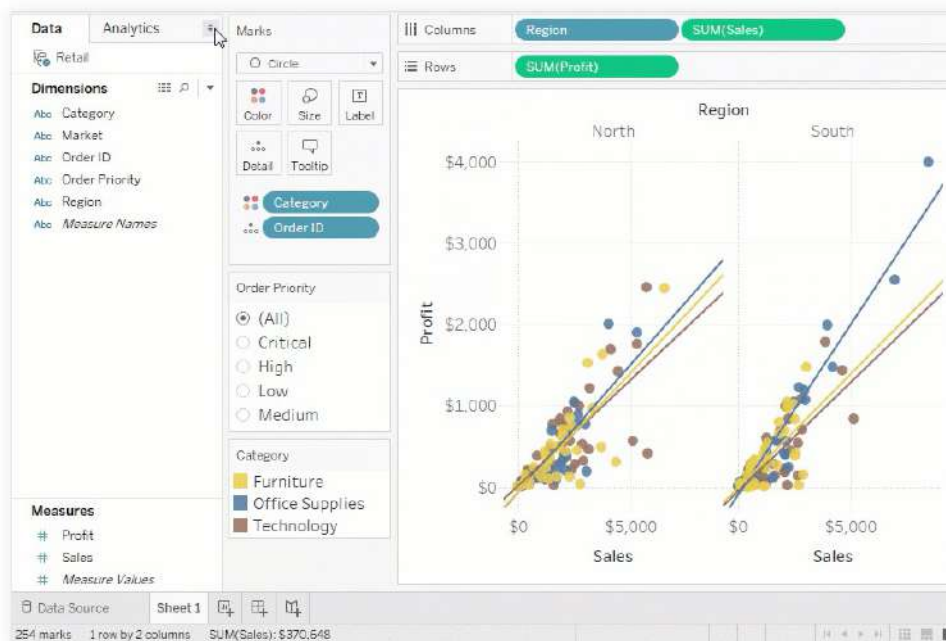
Power BI para Análise de Dados

Ferramentas de dados

Power BI



Embora muito útil e flexível, o Excel não tem capacidades de entrega robustas para entregar relatórios em produção nas empresas. Desta forma é importante partir para alternativas mais eficazes. Nesta linha, as principais ferramentas de visualização de dados do mercado são o **Tableau** e o **Power BI**.



Demonstração do Power BI



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

Python para Análise de Dados

Ferramentas de dados

Programar ou não programar? Eis a questão...

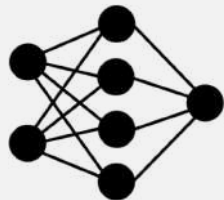


Ferramentas No-Code ou Low-Code

Ferramentas que usam código



Analytics e
Dashboards



Machine Learning
e AI



Demonstração do Python



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

Qual ferramenta usar?

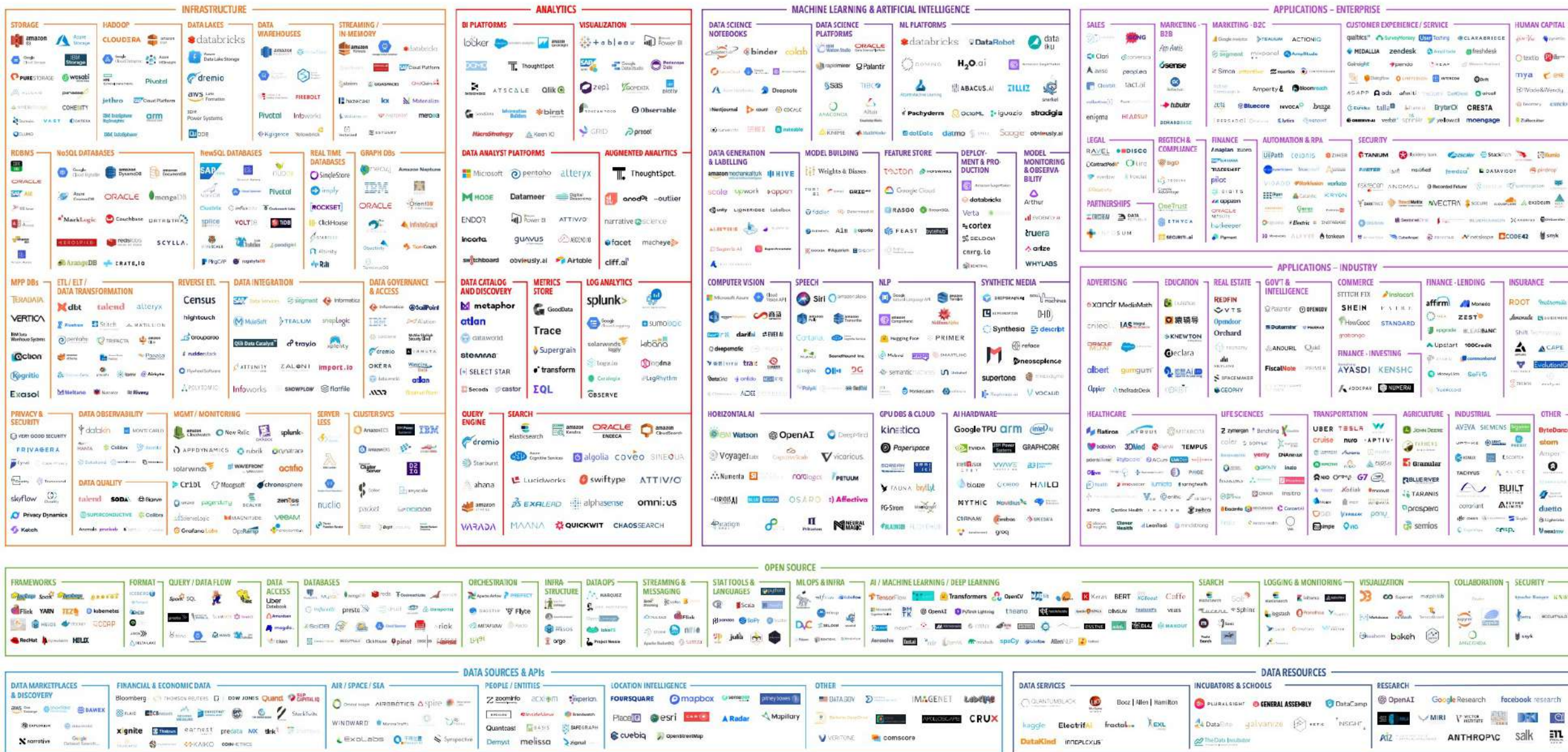
Dúvidas que recebemos constantemente

Sobre ferramentas de dados...



- ☐ Preciso aprender **Python** para analisar dados?
- ☐ Preciso aprender **Power BI** para analisar dados?
- ☐ **Excel** não está obsoleto? Não é melhor focar no Python?
- ☐ **SQL** é complicado? Não tenho experiência com linguagem de programação...

MACHINE LEARNING, ARTIFICIAL INTELLIGENCE, AND DATA (MAD) LANDSCAPE 2021



Version 3.0 - November 2021

© Matt Turck (@mattturck), John Wu (@john_d_wu) & FirstMark (@firstmarkcap)

mattturck.com/data2021

FIRSTMARK
EARLY STAGE VENTURE CAPITAL

Comparativo das Ferramentas de Dados

Overview das ferramentas analíticas



Preditiva.ai



Análises Estatísticas Fundamentais

Sim

Poucas possibilidades

Sim

Sim

Análises Estatísticas Mais Avançadas

Sim, de forma manual

Não

Não

Sim

Processamento e Performance

Baixa

Alta

Médio

Médio (localmente),
Alto (Na cloud)

Facilidade de Uso

Alta

Média

Alta

Baixa

Flexibilidade

Alta

Baixa

Média

Alta

Resolvem qualquer problema

Resolvem qualquer problema
com produtividade e automação



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

A importância da Metodologia em Projetos de Dados

Metodologia

Overview



Análise de Dados é uma “máquina de solução de problemas”. Mas para que funcione, **precisamos de pelo menos três “engrenagens”**. São elas:



- ✓ Matemática fundamental
- ✓ Técnicas de Otimização
- ✓ Estatística Descritiva
- ✓ Probabilidades
- ✓ Inferência (ex: Testes de Hipótese...)
- ✓ Modelagem Estatística
- ✓ *Machine Learning*
- ✓ Entre outros...



- ✓ Planilhas e Tipos de Arquivos
- ✓ Linguagens de programação
- ✓ Bancos de dados
- ✓ Visualização de Dados
- ✓ Infra estrutura (DWs, Data Lakes etc)
- ✓ ETL / Pipelines de Dados
- ✓ Entre outros...



- ✓ Entendimento dos processos internos
- ✓ Alinhamento com os objetivos da empresa
- ✓ Entrosamento com o time e principais *stakeholders*
- ✓ Análise da concorrência
- ✓ Entre outros...



Analista de Dados | Desenvolvedor...
9 h • Editado •

ANALISTA DE DADOS E NÃO ESTILISTA DE MODA

Essa semana minha gestora solicitou um novo Dashboard para mim. Ela disse o que precisava, eu fui fazendo anotações e já imaginando como eu poderia fazer.

Quando eu me dispus em frente ao laptop, abri o PowerBI, comecei a colocar os gráficos, criar medidas e tudo mais. A coisa é que eu imaginei algo super automatizado, quase que um "gerador de resultados", já ordenando as variáveis e mostrando os departamentos conforme a regra de negócio. Só que eu estava com dificuldades em fazer aquele "DashMonster" e imaginei de repente que desenvolver uma solução em Python seria mais adequado. Passei o dia pensando como seria essa integração, como os usuários iriam interagir e muitos etc.

Um pouco cansado de planejar, levantei, fui à máquina de capuccino, dei um "oi" para uns colegas do caminho e voltei para minha mesa. Ao voltar, olhei as anotações, olhei para a tela, olhei para minha xícara, olhei outra vez para a tela e concluí: Estou inventando moda.

A solução que eu tinha imaginado para aquele problema não precisava de robôs extratores como as outras necessidades pediam. Não precisavam de super automações e ordenação automática - se eu tivesse implantado isso, talvez até atrapalharia a análise. Percebi que o problema só exigia um leiaute mais adequado e que eu estava querendo - na melhor das intenções, lógico - criar algo mirabolante, quando bastava só um pouco de adaptação do que eu já tinha feito.

Às vezes quem trabalha com dados pode querer ser "estilista de moda", inventando um monte de coisas e criando tendências. Querendo aplicar todas as tecnologias e conhecimentos que vimos nos cursos, acabamos por deixar nosso projeto demasiado lento e complexo, sendo que algo bem mais simples poderia ter sido feito. Melhorias serão implementadas aos poucos, com certeza. Mas talvez essa coluna calculada dentro de um segmentador, ou algumas VAR dentro das medidas podem ser eliminadas se o projeto fosse mais simples. E como é difícil ser simples e direto.

[Ver tradução](#)



50

1 comentário • 2 compartilhamentos

Metodologia

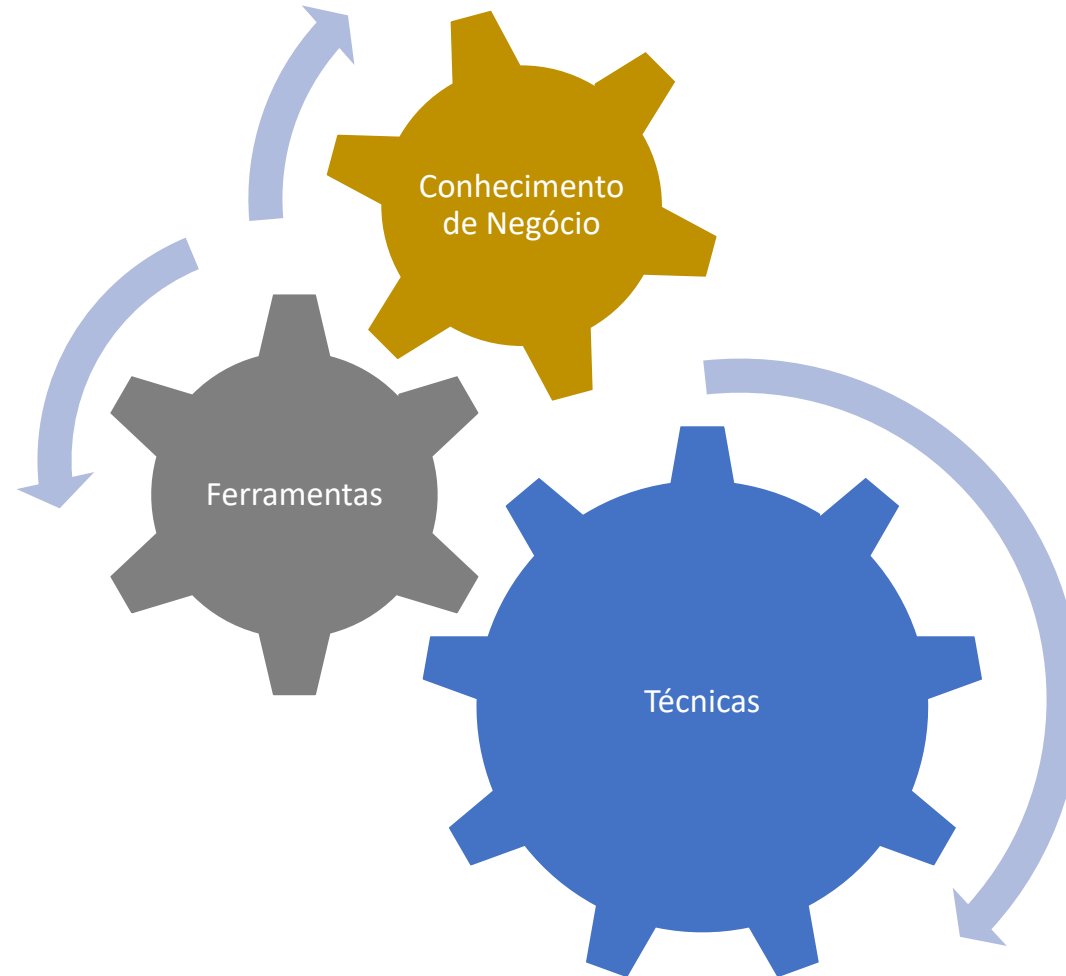
Overview



No entanto, essas “engrenagens” funcionando sozinhas **não** entregam o real valor para as áreas das empresas como deveriam.

Felizmente, existe uma forma eficiente para que as engrenagens funcionem em sinergia:

Aplicando-se uma
Metodologia de Projetos de Dados.



Metodologia

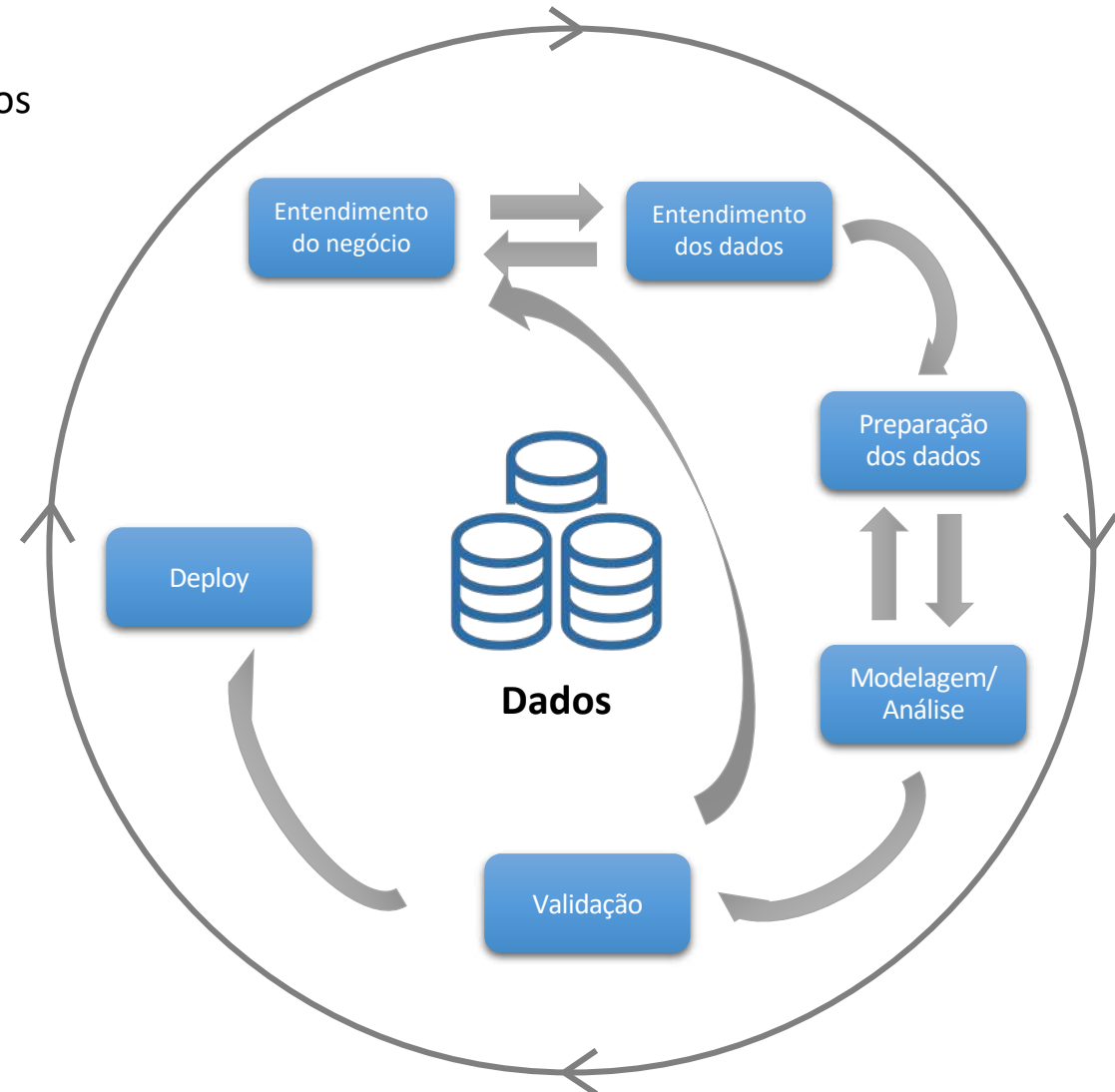
A metodologia CRISP-DM



Uma metodologia muito utilizada por profissionais de Dados é a chamada **CRISP-DM**. Sua abreviação vem de **Cross Industry Standard Process for Data Mining**.

A metodologia foi criada em unindo-se as melhores práticas no processo de análise de dados.

O **CRISP-DM** tem **6 grandes passos** conforme mostrado na imagem ao lado.



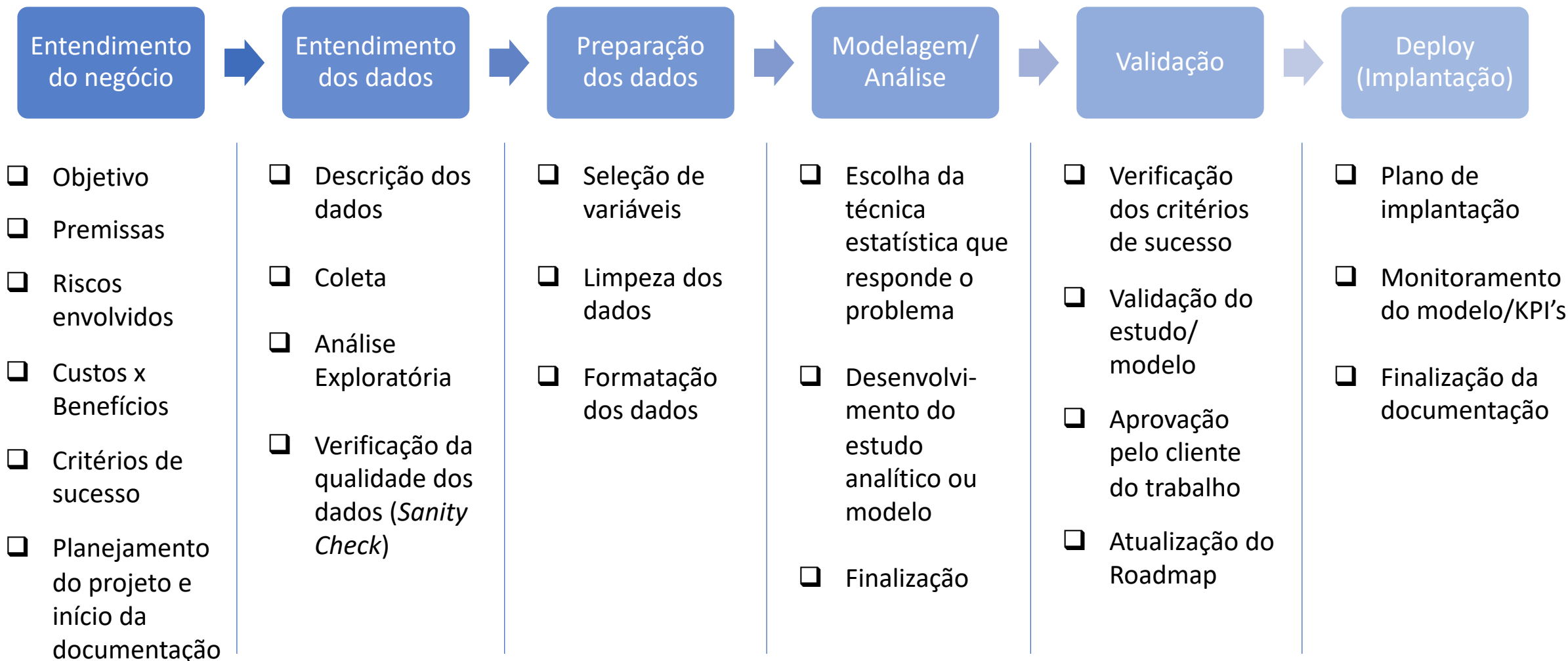
A metodologia CRISP-DM

Metodologia

A metodologia CRISP-DM



O **CRISP-DM** é praticado em 6 grandes passos. Veja:



“ Antes de olhar um resultado de um trabalho, olhe como ele foi construído. Só assim é possível confiar no resultado. Essa é a **importância da metodologia.** ”



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

A tal da Cultura Data Driven

Pergunta 1

Responda este exemplo



Você pede para um analista um relatório de vendas de um determinado produto. A ideia é saber se o gênero explica a compra desse produto. O analista te entrega isso:

Gênero	Unidades Vendidas	Frequência de Compra (%)
Masculino	400	20%
Feminino	1600	80%

2000

Conclusões mais comuns:

Produto claramente voltado para o público feminino.

Pergunta 1

Responda este exemplo



Essa análise simples não considera o fator **contrafactual**, ou seja, não considera as compras **não** realizadas por cada um dos gêneros. Teria a mesma proporção? Vejamos:

Gênero	Unidades enviadas para o carrinho de compras	Frequência de Compra (%)	Frequência de Não Compra (%)	Diferença
Masculino	3600	20%	19%	1%
Feminino	7200	80%	81%	-1%

6000

E agora? A conclusão muda?

Pergunta 2

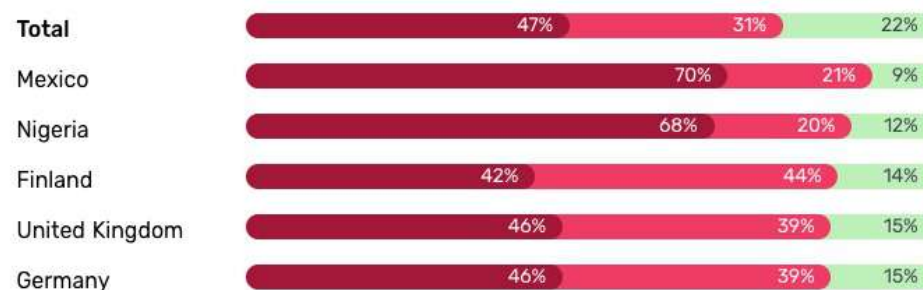
Agora responda esta aqui:

Qual o percentual da população mundial **não** tem comida suficiente para suas necessidades diárias?

- A) Cerca de 11%
- B) Cerca de 23%
- C) Cerca de 37%

Survey Results

Of the people we have tested, 78% got this question wrong.



[More](#) ▾

O que acontece?

Mas por que em geral erramos questões como essa?



Pergunta 1:

Erro:



Gênero	Unidades Vendidas	Frequência de Compra (%)
Masculino	400	20%
Feminino	1600	80%

2000

Porque **não conhecemos as técnicas analíticas** e quais os erros a serem evitados em seu uso.

Pergunta 2:

Qual o percentual da população mundial **não** tem comida suficiente para suas necessidades diárias?

- A) Cerca de 11%
- B) Cerca de 23%
- C) Cerca de 37%

Erro:

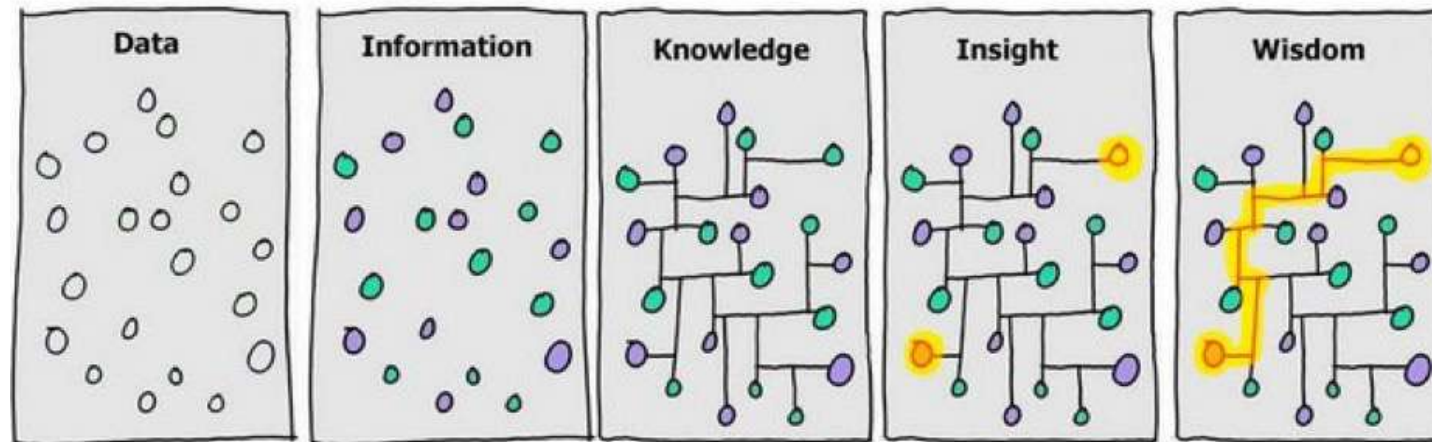


Porque em geral tendemos a dar opiniões baseadas em **achismos** e não em **dados e fatos**.

Cultura Data-Driven

O que é essa cultura?

Cultura Data-Driven (Dados gerando sabedoria)



Uma empresa com uma **Cultura Data-Driven** utiliza os dados em todo processo decisório. A empresa trata os dados como um **ativo estratégico** e busca disseminar o acesso aos dados de forma ampla para todos os colaboradores. Este processo foca em capturar, limpar e armazenar dados relevantes. A empresa promove a experimentação e aprendizado constantes através dos dados.

Cultura Data-Driven

Os passos para implantação



Mudança de Mentalidade

- ☐ Migração de sistemas e processos “**legados**”. Desapegue!
- ☐ A alta administração deve colocar os **dados** no centro do **processo decisório**.
- ☐ Dados devem passar a serem vistos como um **ativo** da empresa e não como um sub-produto.

Fornecimento de Dados

- ☐ Todos os colaboradores devem utilizar **um único repositório de dados na empresa**.
- ☐ Esse repositório (Data Lake etc) deve passar por um processo de governança que assegure sua **documentação e qualidade**.

Alfabetização de Dados

- ☐ Os colaboradores precisam ser capazes de explorar os dados da empresa para resolver problemas de negócio. Isso envolve **muito treinamento**.
- ☐ Nem todos precisam ser Analistas ou Cientistas de Dados, mas todos precisam **ser alfabetizados**.

Mas afinal, o que é Alfabetização de Dados ?

Introdução ao Mundo dos Dados

Os desafios na implantação da
Cultura Data Driven e a
Alfabetização de Dados



Preditiva.ai
Especialista em Analytics

Mas afinal, o que é Alfabetização de Dados ?

A [Gartner](#) define “Alfabetização de Dados” como a habilidade para **ler, escrever e comunicar os dados** em seu devido contexto, incluindo um **entendimento de fontes de dados** e suas transformações, dos **métodos e técnicas analíticas aplicadas**, descrevendo assim as aplicações dessas técnicas no problema de negócio e como espera-se obter valor disso (resultado tangível e esperado).

No fundo, a partir de agora, a pergunta é:

“Do you speak data?”

Cultura Data-Driven

O Processo de Alfabetização de Dados - *Data Literacy*



Preditiva.ai

Os níveis de
**Alfabetização de
Dados** são
alarmantes em
todo o mercado!



Cultura Data-Driven

O Processo de Alfabetização de Dados - *Data Literacy*



Os níveis de
**Alfabetização de
Dados** são
alarmantes em
todo o mercado!



G1

EDUCAÇÃO

Brasil cai em ranking mundial de educação em matemática e ciências; e fica estagnado em leitura

País teve novamente um dos 10 piores desempenhos do mundo em matemática no Pisa 2018, a avaliação mundial de educação.

Por Ana Carolina Moreno e Elida Oliveira, G1

03/12/2019 05h01 · Atualizado há um ano



Déficit educacional do país sempre foi alarmante, principalmente em temas de Ciências Exatas.

Cultura Data-Driven

Mas por que não somos alfabetizados em Dados?



Pisa 2018 - resultados do Brasil

País conseguiu avançar alguns pontos entre as edições 2015 e 2018 da prova, mas ainda não subiu de patamar e segue longe do desempenho dos países desenvolvidos

BRASIL	Leitura	Matemática	Ciências
Nota média 2018	413	384	404
Margem de erro	2	2	2
Variação 2015-2018	6	6	3
Posição no ranking	58-60	72-74	66-68

Fonte: OCDE/Pisa 2018

China é referência em Data Science no mundo.

Será coincidência?

Pisa 2018 - TOP 10

Regiões da China continental superaram Singapura, a campeã da edição anterior, nas três provas

Posição	LEITURA	MATEMÁTICA	CIÊNCIA
1º	China (P-X-J-G)* (555)	China (P-X-J-G)* (591)	China (P-X-J-G)* (590)
2º	Singapura (549)	Singapura (569)	Singapura (551)
3º	Macau (China) (525)	Macau (China) (558)	Macau (China) (544)
4º	Hong Kong (China) (524)	Hong Kong (China) (551)	Estônia (530)
5º	Estônia (523)	Taipei (China) (531)	Japão (529)
6º	Canadá (520)	Japão (527)	Finlândia (522)
7º	Finlândia (520)	Coreia do Sul (526)	Coreia do Sul (519)
8º	Irlanda (518)	Estônia (523)	Canadá (518)
9º	Coreia do Sul (514)	Países Baixos (519)	Hong Kong (China) (517)
10º	Polônia (512)	Polônia (516)	Taipei (China) (516)

*Pequim-Xangai-Jiangsu-Guangdong

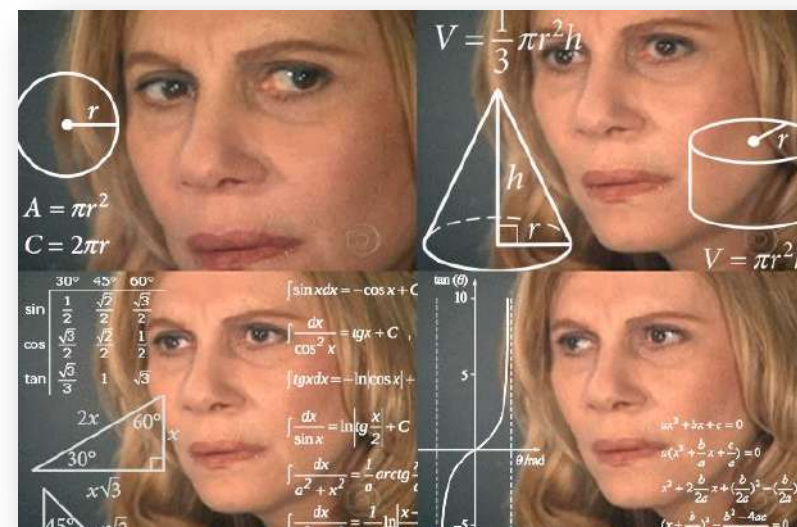
Fonte: OCDE/Pisa 2018

Fonte: <https://g1.globo.com/educacao/noticia/2019/12/03/brasil-cai-em-ranking-mundial-de-educacao-em-matematica-e-ciencias-e-fica-estagnado-em-leitura.ghtml>

A maioria das empresas **não** têm iniciativas sólidas de Ciência de Dados

Principais desafios:

- a) Cultura de coleta de Dados
- b) Governança e Qualidade dos Dados
- c) Alfabetização de Dados
- d) Dados no centro do processo decisório



a) Cultura de coleta de Dados

- Só se faz análise de dados com... veja só: **“com dados”!**
- Desta forma, muitas vezes **projetos não vão pra frente devido à falta de acesso e/ou existência de dados** armazenados.
- O dados deve ser pensado desde o “day one” de um novo processo e/ou produto. **Quais dados serão armazenados? Onde serão armazenados? Por qual período? Em qual formato?**

b) Governança e Qualidade dos Dados

- Os **dados** devem ser **consistentes** em toda a empresa.
- Conceitos utilizados no **tratamento e transformação** dos dados devem ser **documentados** e estarem acessíveis.
- Considerar a **LGPD** é fundamental!



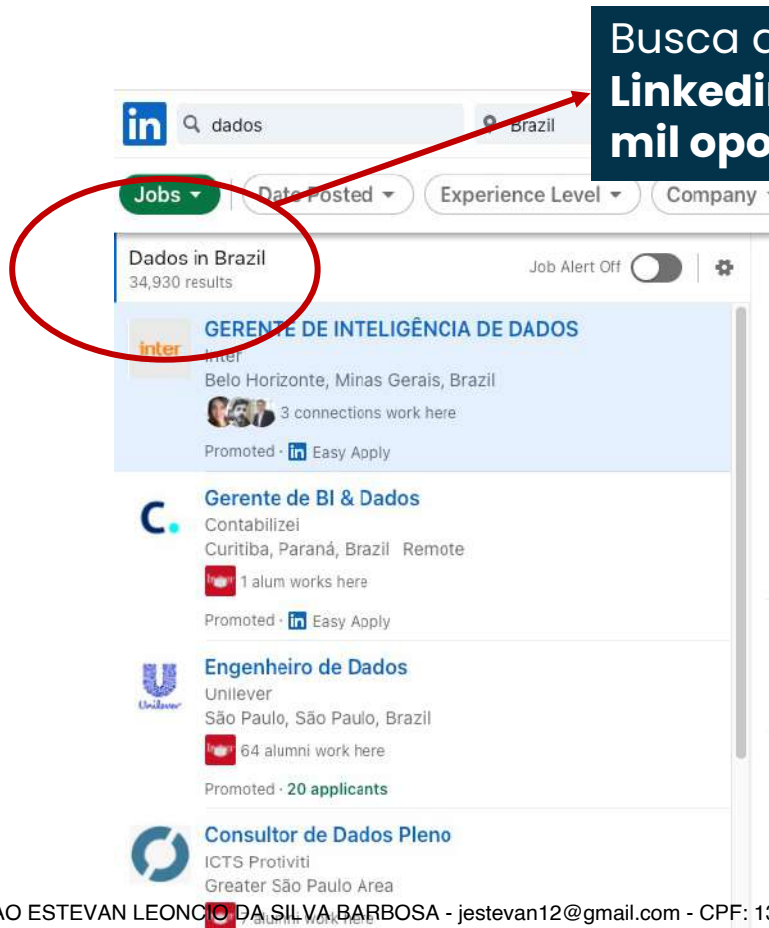
Cultura Data-Driven

Principais desafios na implantação dessa cultura



c) Alfabetização de Dados

Busca da palavra “**dados**” no **LinkedIn** retornou mais de **34 mil oportunidades** abertas.



- Escassez de profissionais qualificados e alfabetizados em dados em diversos níveis.
- Formação de talentos *in company* tem sido uma boa estratégia adotada pela maioria das empresas.

Cultura Data-Driven

Principais desafios na implantação dessa cultura



Do estagiário ao CEO: Itaú faz da IA o novo mantra para toda a empresa

Nova fase de transformação, focada em inteligência artificial e computação quântica, inclui capacitação para todos os colaboradores, como explica Moisés Nascimento, CDAO do banco



Luiz Gustavo Pacete

3 de agosto de 2023

FB – Quais outros grupos, além dos C-Levels, têm sido estratégicos nesse letramento de inteligência artificial e dados?

Moisés – Hoje, no banco, todas as comunidades usam IA e dados. Mas um grupo importante em que focamos foi o de product management. Os líderes de produtos, para que eles possam inserir a tecnologia como parte do design e áreas mais avançadas, dentro de cada produto, desenhar features e outras possibilidades. Melhorar, por exemplo, navegação dos apps. Isso melhora o que já temos e cria novas possibilidades. Isso vale para IA tradicional como também para IA generativa. Permite encontrar nichos de mercado e oportunidades para oferecer soluções.

Leia mais em: <https://forbes.com.br/forbes-tech/2023/08/do-estagiario-ao-ceo-itaui-faz-da-ia-o-novo-mantra-para-toda-a-empresa/>

Cultura Data-Driven

Principais desafios na implantação dessa cultura



Predictiva.ai



#RADARMARTHAGABRIEL

DESAFIOS PARA IMPLANTAR IA NOS NEGÓCIOS

Twitter LinkedIn Instagram @MARTHAGABRIEL



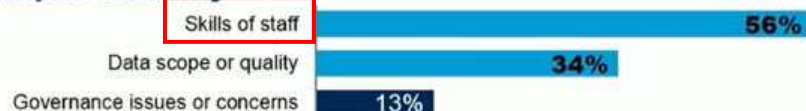
Nov/2021

FONTE: CMO Council / Deloitte Digital / Marketing Charts

Top 3 Challenges to AI/ML Adoption

Sum of 1-3 rank

Enterprise maturity



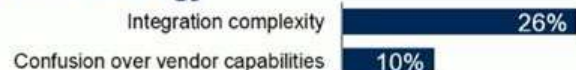
Fear of unknown



Finding a starting point



Vendor strategy



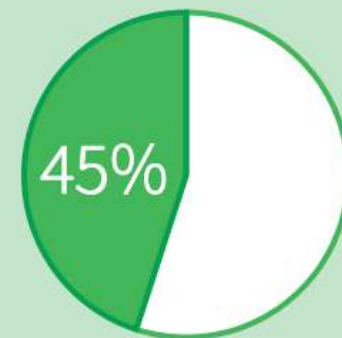
Gartner

2019

d) Dados no centro do processo decisório

- **Dados** coletados, organizados e validados devem ser utilizados para a **tomada de decisão!**
- Experiências passadas precisam cada vez mais estar acompanhadas de **análise de dados.**

45% dos diretores entrevistados ainda tomam decisões no *feeling* em vez de *insights* trazidos por dados *



- ☐ Sua empresa pode não ser Data Driven, mas a maioria está tentando. Não acontece da noite para o dia. Faça parte dessa mudança.
- ☐ Policie-se para **evitar os achismos**. Caso não sabia, peça um tempo para analisar e extrair dados que suportem sua conclusão.
- ☐ Foque em sua **Alfabetização e Letramento em Dados**. O mercado já tem muito apertador de botão (ferramenteiros).
- ☐ Você não precisa ser um analista ou cientista de dados para prosperar nesta nova economia. **Se você já é bom no que faz, alfabetize-se em dados** e busque oportunidades de melhoria na própria carreira antes de optar por uma transição.



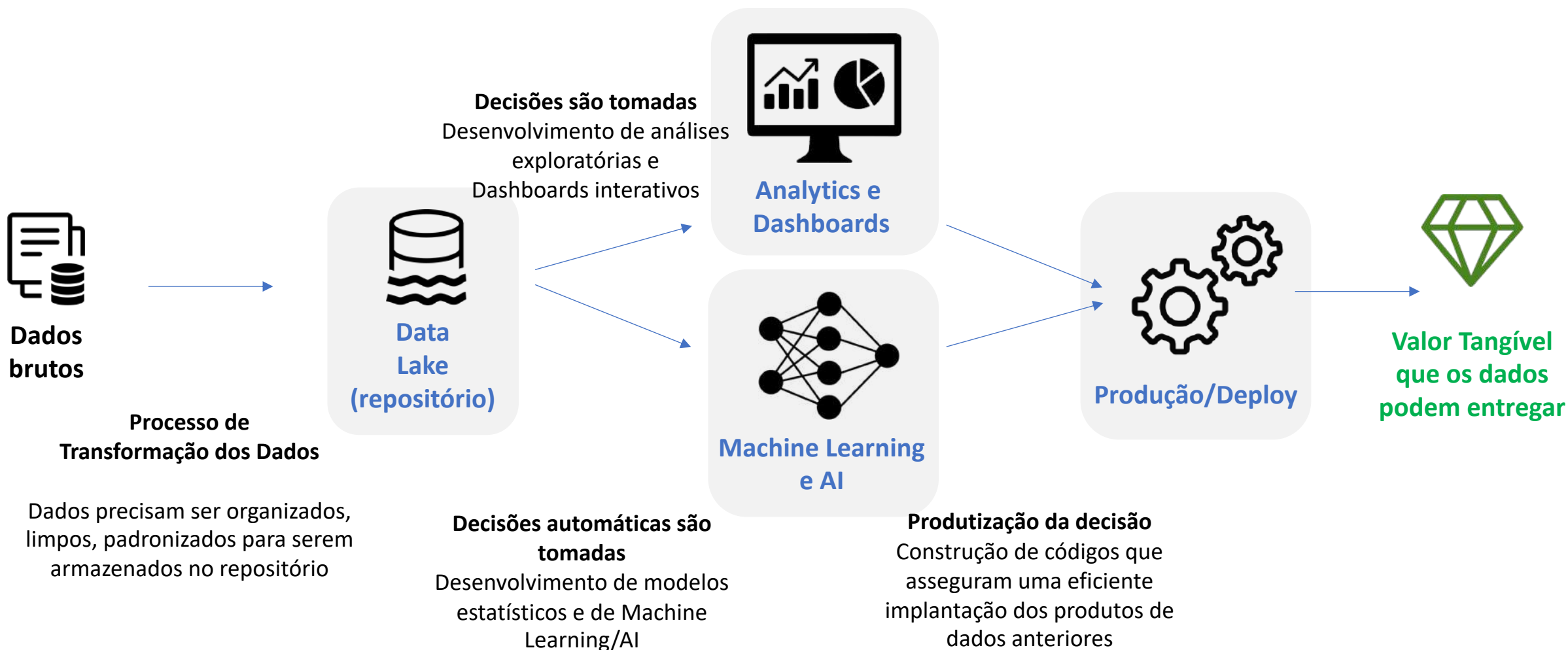
Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

O papel de cada profissional na Jornada dos Dados

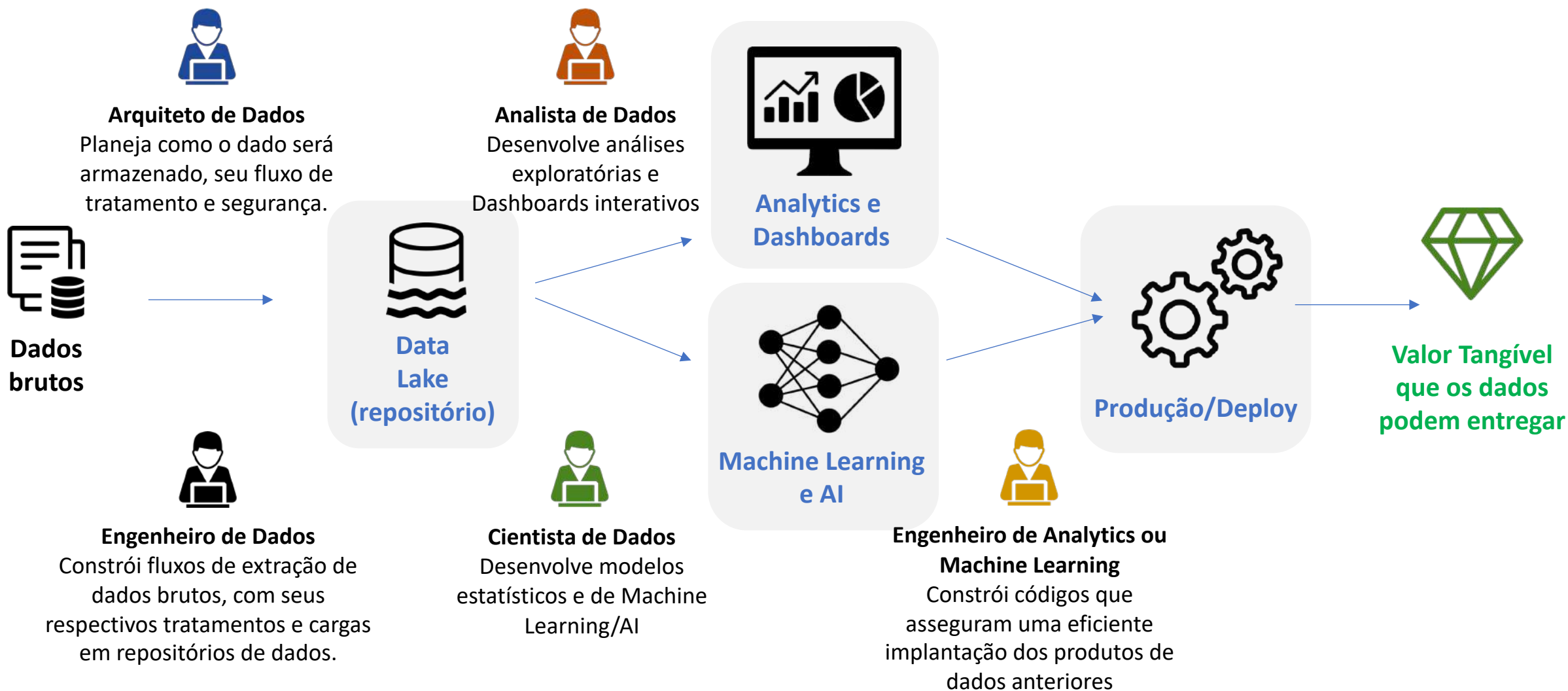
Jornada dos dados

Do dado bruto ao valor “tangível”



Jornada dos dados

As principais funções analíticas na jornada dos dados



Jornada dos dados

Funções por hierarquia de necessidades



THE DATA SCIENCE HIERARCHY OF NEEDS

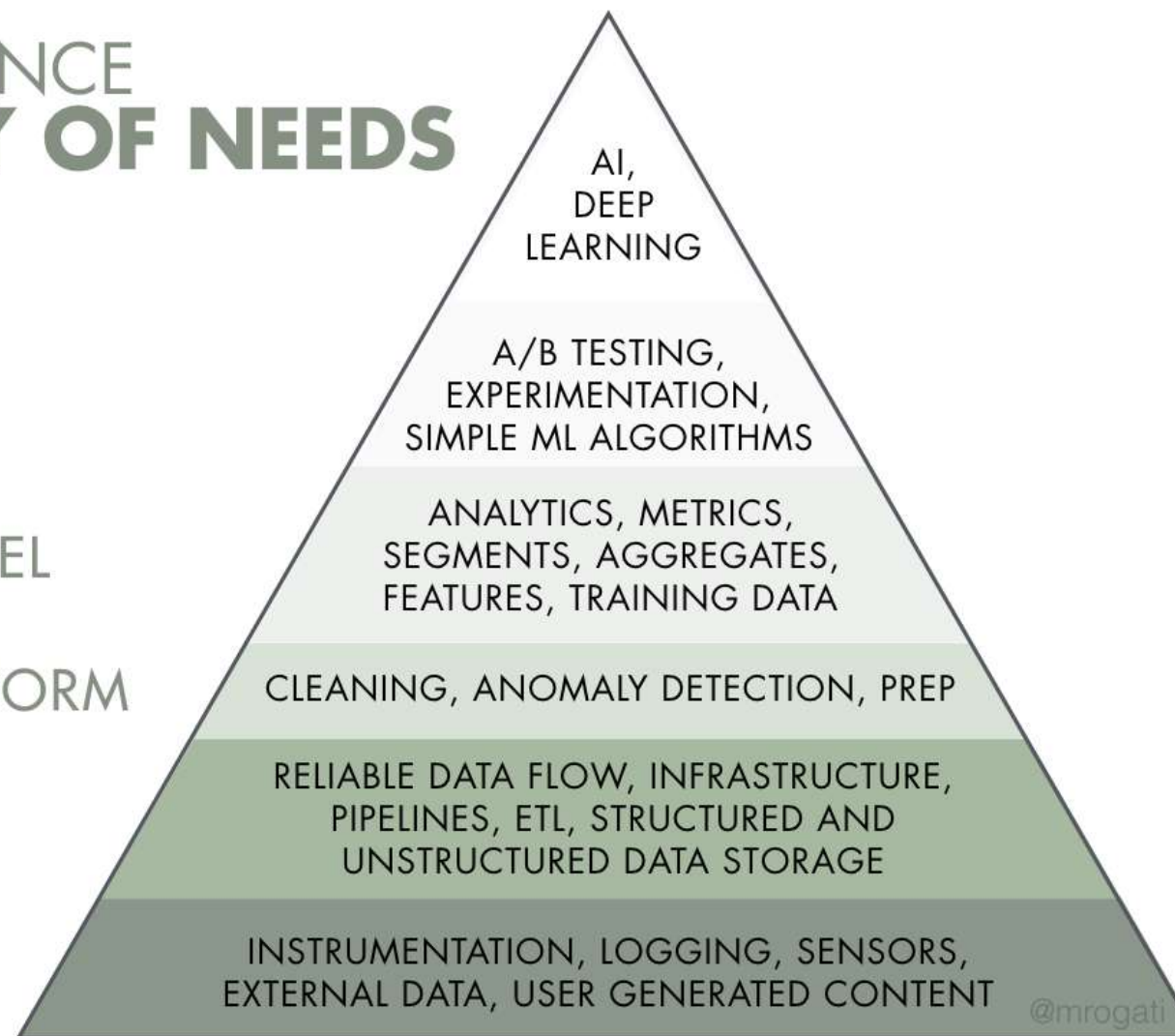
LEARN/OPTIMIZE

AGGREGATE/LABEL

EXPLORE/TRANSFORM

MOVE/STORE

COLLECT



Chief Data Scientist
Pesquisador de Dados
Engenheiro de Machine Learning

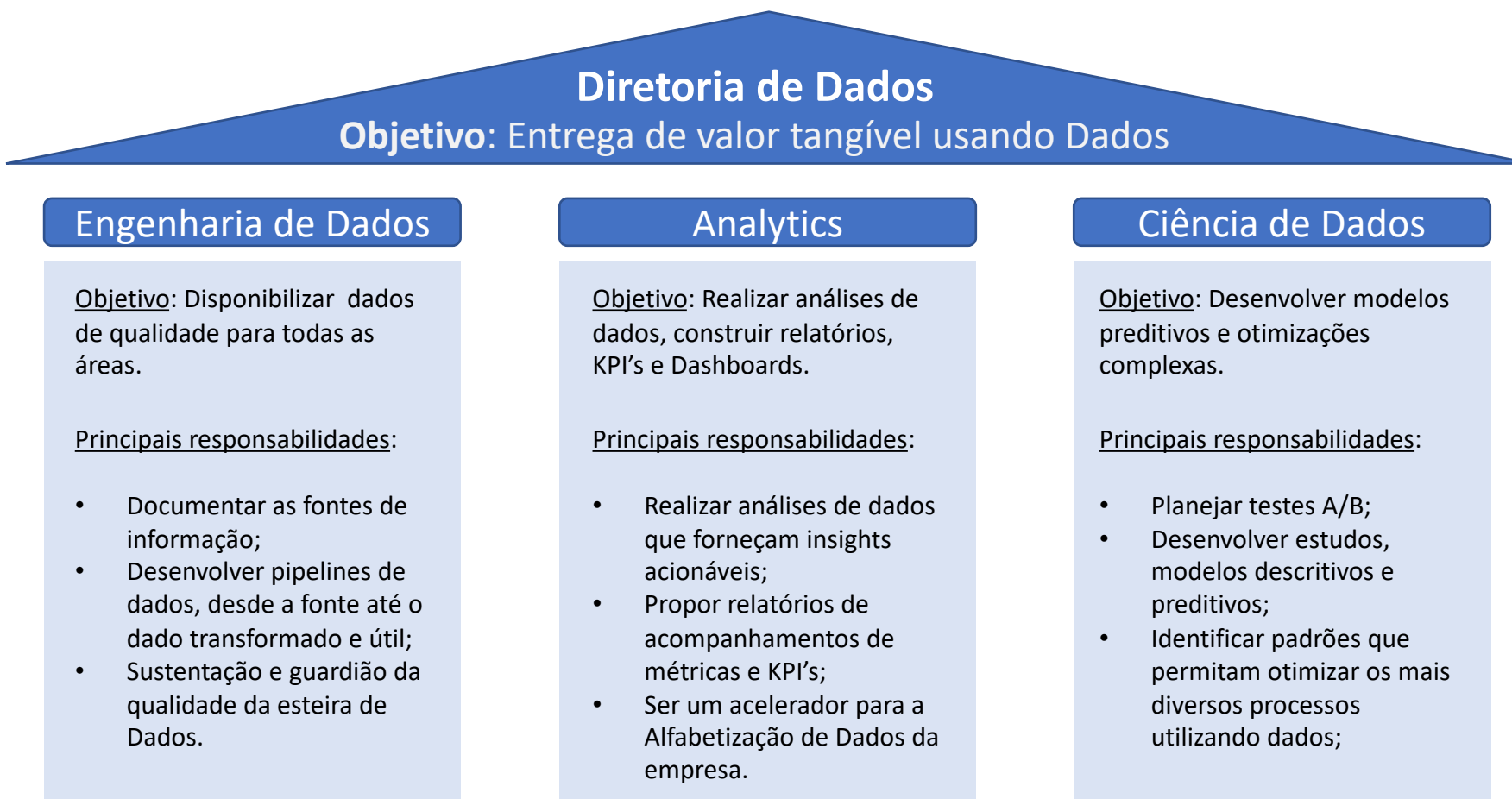
Analistas de Informação
Analistas de Dados
Analistas de Negócio
Cientistas de Dados

Engenheiros de Dados

Engenheiros, Cientistas e
Desenvolvedores de Software

Jornada dos dados

Estrutura organizacional de um CDO (Chief Data Officer)



Essa é a chamada estrutura **“centralizada”**. Existem outras formas de organizar os times, cada uma com seus prós e contras.

As estruturas acima prestam serviços a todas as áreas e/ou Squads da empresa

- ❑ A quantidade de funções vai depender do tamanho da empresa e **maturidade analítica** (estágio da cultura data driven implantada). Tem empresas que têm analistas de dados mas não tem engenheiros, por exemplo. O que é um grande erro, como vimos na pirâmide de necessidades.
- ❑ Existem **vagas que pedem** que um profissional de dados **faça o trabalho de mais de um** profissional. Exemplo:
 1. ETL (trabalho de Engenheiro);
 2. Análise de Dados e Dashboards (trabalho de um Analista de Dados) e;
 3. Modelos de Machine Learning (trabalho de um Cientista de Dados).
- ❑ Nestes casos, só aceite se estiver em transição de carreira e precisa ganhar experiência. Caso contrário, nossa sugestão é que busque outras vagas. Grandes são as chances desta empresa **não ter uma cultura analítica definida ou praticada**.



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

A importância de uma boa
gestão de carreira

Gestão de carreira

Passo a passo para conquistar a carreira de seus sonhos



Estudo

Estatística descritiva

Probabilidades

Inferência

**Modelos e Machine
Learning**

Excel

SQL

Power BI

Python e API's

Metodologia

Conhecimento de negócio

Linkedin

Entender o funil

Taxa de Empregabilidade

Otimização

Portfólio

Posicionamento online

Processo seletivo

Entrevista com o RH

**Entrevista com o
Gestor da vaga**

Projeto prático



**Emprego em
dados**

Gestão de carreira



Dicas para alavancar suas chances de crescimento

Uma vez dentro da empresa, é sua responsabilidade **administrar a própria carreira**.

- ☐ **Mapeie as oportunidades** de melhoria em sua área. Nem sempre o cliente vai fazer isso. Tome a iniciativa.
- ☐ **Calcule o ROI de seus projetos**. Use isso nas reuniões de feedback para planejar a carreira na empresa.
- ☐ Tenha uma **conversa de carreira** (1x1) a cada 3 meses com seu gestor.
- ☐ Não se esqueça das **pessoas** e da cultura da empresa.



Gestão de carreira

Consulte estes materiais



A Preditiva tem os melhores materiais de carreira em Dados do mercado. Não perca esta oportunidade.

Bônus: Gestão de Carreira (construindo e administrando uma carreira de sucesso) [Ver tudo →](#)

Módulo

Conceitos fundamentais para uma boa gestão de carreira



1) Entenda como aumentar sua taxa de empregabilidade

Não adianta analisar dados se você não for eficiente nas suas escolhas de carreira. Reunimos todos os principais fatores que vão lhe trazer muitas oportunidades.

Módulo

Os processos seletivos nas áreas de dados



2) Processos seletivos em dados

Ninguém gosta de processo seletivo. Nem a gente. Mas aprendemos o mais importante para aumentar suas chances. Confira aqui.

Módulo

Crescendo na carreira



3) Crescendo na carreira: O que fazer?

Parabéns por conquistar sua vaga. Agora o jogo da carreira (e vida) começam. Quais suas metas? Como alcançá-las? Crescer na empresa ou trocar de empresa? Isso e muito mais você descobre aqui.



Live #36 - As principais perguntas e respostas em processos seletivos em Dados



Live #31 - 10 situações de carreira que profissionais de dados devem superar



Live #30 - Dicas de Carreira com os Professores da Preditiva



Live #23 - Job Hacking - Como otimizar a sua busca de emprego?

E muito mais...



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

O processo de aprendizado das
habilidades em dados

O processo de aprendizado

Cuidado para não se perder no caminho



Problema 1) Ele não considera a senioridade do profissional

Acredito que nem o autor do mapa conheça a maioria do que é proposto ali. Não coloque uma meta irreal para você mesmo.

Problema 2) O mapa não considera a Lei de Pareto para Ciência de Dados

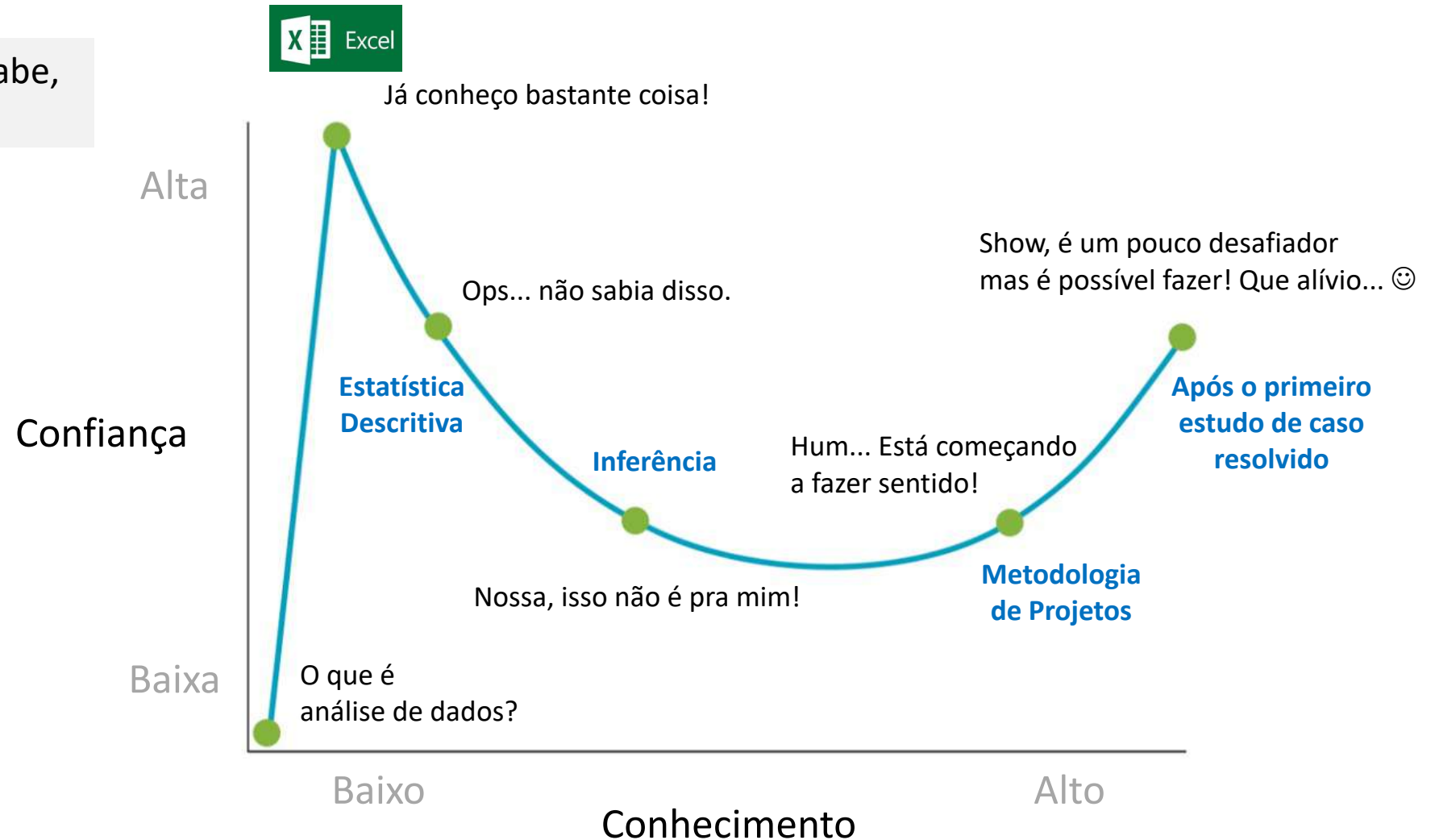
Você sabe que lei é essa? Basicamente 80% das oportunidades e resultados que um analista e/ou cientista de dados pode entregar nas empresas são entregues usando apenas 20% das técnicas e ferramentas existentes. Ou seja, só 20% desse mapa realmente vale a pena ser dominado.

O processo de aprendizado

O efeito Dunning-Kruger



“Quanto **menos** uma pessoa sabe, mais ela acha que sabe”



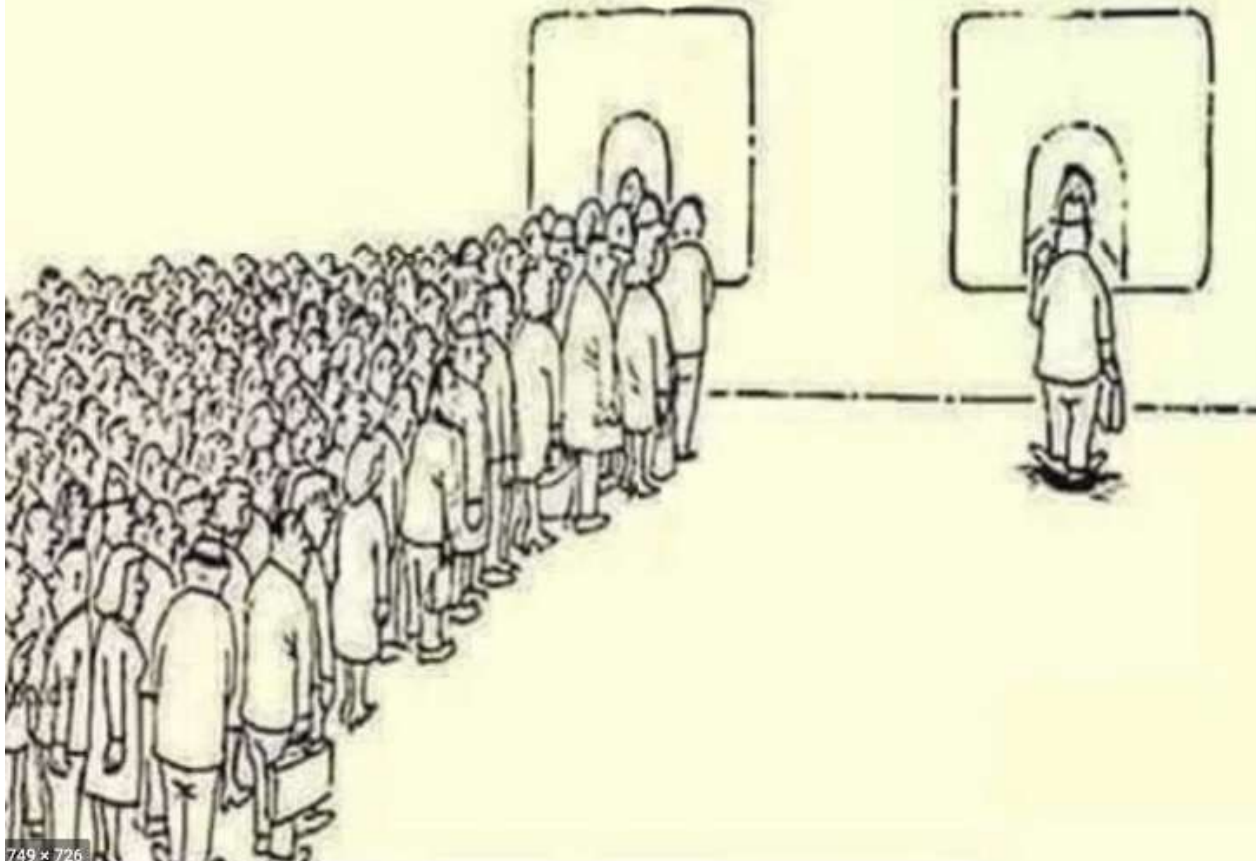
O processo de aprendizado

Não caia nessa



Curso de Python e Power BI
torne-se um especialista em
Análise de Dados

Aprenda Técnicas Estatísticas, Ferramentas
e Metodologia de Projetos de Dados,
pratique e melhore suas habilidades
analíticas nos próximos meses
(Lifelong Learning)



Não acredite em milagre.

Desenvolver sua habilidade em dados é o melhor investimento hoje e para os próximos anos, mas é preciso dedicação e resiliência para ser um bom profissional.

Ferramenta, embora bacana, é somente parte do processo!

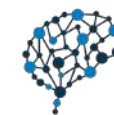
Não seja um “apertador de botão”. O mercado já tem vários e você precisa fazer mais para se destacar.



Preditiva.ai
Especialista em Analytics

Introdução ao Mundo dos Dados

Perguntas mais frequentes



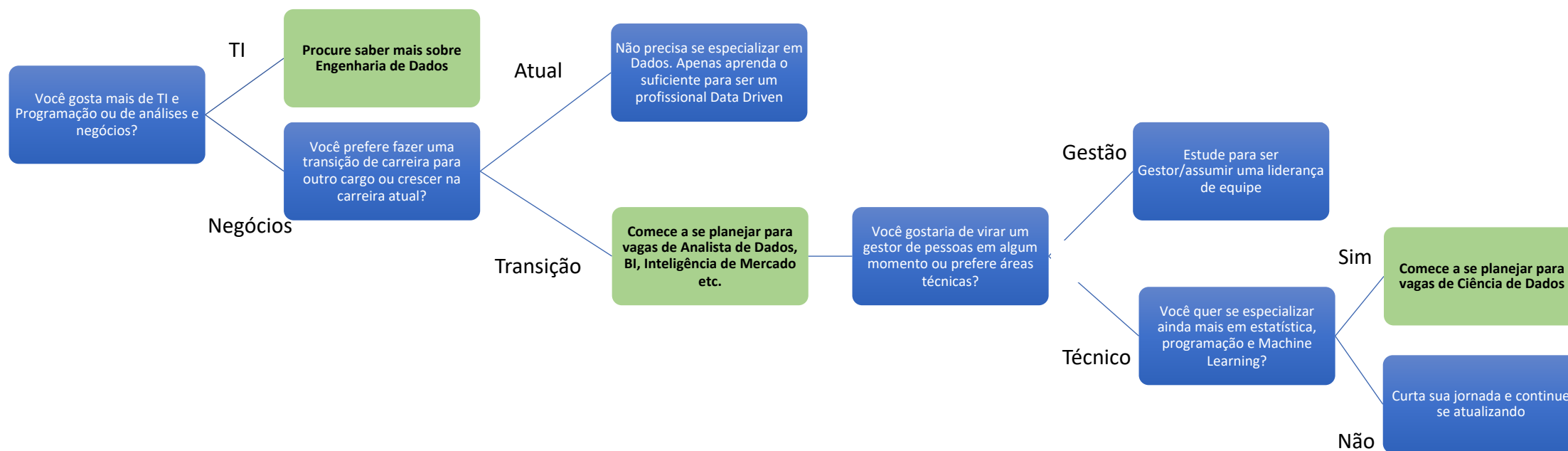
**"Dados" serve pra minha
área? Essa habilidade é
para mim?**



"Não sei se devo virar Engenheiro, Analista ou Cientista de Dados: o que fazer?"



Uma forma simples de planejar seus próximos passos





Quero migrar para áreas de dados logo...

Dá para fazer isso com rapidez?

O passo a passo para conquistar sua vaga

Cronograma



	Mês 1	Mês 2	Mês 3	Próximos meses
Estudo	Inscrição na Preditiva :-)	Estudo de Probabilidade	Estudo de Power BI	Aperfeicoe os maiores gaps
	Excel	Segundo estudo de caso resolvido	Estudo de Regressão	Continue resolvendo estudos de caso e analisando bases
	Estudo de Estatística	Estudo de Inferência		
	Primeiro estudo de caso resolvido	Metodologia de Projetos	Terceiro estudo de caso resolvido	
Linkedin	Criação do Perfil	Início da criação do Portfólio (use o estudo de caso resolvido)	Finalização do Portfólio - Versão inicial	Se inscreva para cada vaga e mande mensagem para o recrutador
	Título	Inscrição para 30 vagas (após primeiro trabalho do portfólio)	Inscrição para 30 vagas	Localize pessoas do seu LinkedIn que trabalham na empresa e peça para avaliar seu Portfólio e se coloque a disposição para trabalhar
	Descrição	Mensurar sua taxa de empregabilidade (seu funil)	Atualização do funil e mudanças com aprendizados e feedbacks	Continue crescendo a rede
	Cargos	Adicionar recrutadores e enviar mensagens	Adicionar 50 empresas brasileiras que recrutam	Continue adicionando recrutadores
	Cursos e Certificados	Postar algum conteúdo relevante (post próprio) a cada 2 dias	Postar algum conteúdo relevante (post próprio) a cada 2 dias	Continue postando conteúdo
			Entrar em contato com profissionais da área	
	500 conexões (sempre analistas de dados, cientistas etc)	1000 conexões/seguidores	2000 conexões/seguidores	

O passo a passo para conquistar sua vaga

Cronograma

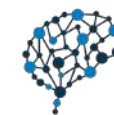


	Mês 1	Mês 2	Mês 3	Próximos meses
Estudo	Inscrição na Preditiva :-)	Estudo de Probabilidade	Estudo de Power BI	Aperfeicoe os maiores gaps
	Mas calma, não existe milagre.			
	Estudo de Estatística	Estudo de Inferência		
	Primeiro estudo de caso resolvido	Metodologia de Projetos	Terceiro estudo de caso resolvido	
Linkedin	Criação do Perfil	Início da criação do Portfólio (use o estudo de caso resolvido)	Finalização do Porfólio - Versão inicial	Se inscreva para cada vaga e mande mensagem para o recrutador
	Título	Inscrição para 30 vagas (apos primeiro trabalho do portfólio)	Inscrição para 30 vagas	Localize pessoas do seu Linkedin que trabalham na empresa e peça para avaliar seu Porfolio e se coloque a disposição para trabalhar
	Descrição	Mensura sua taxa de empregabilidade (seu funil)	Avalie o funil de vagas com aprendizados e feedbacks	Continue crescendo a rede
	Cargos	Adicione recrutadores e enviar mensagens	Adicionar 50 empresas brasileiras que possam te ajudar	Continue adicionando recrutadores
	Cursos e Certificados	Postar algum conteúdo relevante (post próprio) a cada 2 dias	Postar algum conteúdo relevante (post próprio) a cada 2 dias	Continue postando conteúdo
			Entrar em contato com profissionais da área	
	500 conexões (sempre analistas de dados, cientistas etc)	1000 conexões/seguidores	2000 conexões/seguidores	

Não caia na conversa daquele mentor que te promete 5k-10k em 3 meses...



**Estou desempregado... como
conciliar o curso com a
necessidade de pagar as contas?**



**Tenho mais do que 40 anos.
Será que consigo fazer a
transição após essa idade?**

FAQ

Perguntas mais frequentes

Trabalho, transição e promoção

Bom dia a todos da Preditiva Analytics.


Eu gostaria de compartilhar que, há 3 meses eu consegui um ótimo trabalho como Analista de Negócios e que esta semana já recebi uma promoção, antes mesmo de acabar o período de experiência. A partir de 01 de setembro serei Supervisor Comercial. Estou me sentindo muito feliz e realizado, pois eu estava desejando muito esta transição de carreira há muito tempo. Eu sempre trabalhei na área comercial, com vendas externas, e fui muito bem assim por muitos anos; porém, eu desejava agora, utilizar essa minha experiência comercial, mas para um trabalho mais analítico e de gestão. E sinceramente estava perdido, não sabia por onde começar. Meu irmão me sugeriu que eu estudasse Análise de Dados e então, pesquisando, encontrei o curso gratuito de Excel para Análise de Dados, fiz e adorei. Comprei o curso completo e estou cursando. Consegui o trabalho e meu desempenho está tão eficiente que fui promovido. Eu estou fazendo a minha parte, me dedicando, mas quero fazer um agradecimento especial à Preditiva Analytics, pois estes conhecimentos adquiridos estão sendo fundamentais para o sucesso que estou tendo no trabalho. É realmente incrível como tudo mudou e finalmente encontrei o meu caminho. Além do curso, as lives e mentorias, que atualmente quase nem consigo assistir ao vivo, só as gravações, tem conteúdos muito enriquecedores e dicas preciosas. Enfim, tudo me ajudou e ajuda muito. Estou cursando agora o módulo de SQL e Power BI, irei até o final na formação e céu é o limite. PREDITIVA ANALYTICS, vocês cumprem a missão de vocês na Terra, entregam o que prometem e além, sem enrolação. Quero parabenizar Vini e toda a equipe Preditiva pela excelência no que fazem. Vocês são sensacionais. Eu super recomendo a Preditiva Analytics. Muito obrigado. Um grande abraço a todos.



Liked by Matheus and 17 others



Preditiva.ai



Comentários


13

Avaliações

228

Pontuações

17.225

 Login atual em 26 de agosto, 10:27

Acompanhamento

1) Testes A/B (Testes de Hipóteses)	100%
5) Como trabalhar com amostras de dados	100%
Não pule este módulo!	100%
Excel para Análise de Dados	100%
4) Prevendo o futuro	100%
2) Regressão Linear Múltipla	100%
1) Microsoft Excel para Analytics	100%
4) Prevendo o futuro - Parte 2	100%
6) Método CRISP-DM	100%
7) Calculando o Valor de um Projeto de Dados e Apresentando os Resultados	100%
Case) Precificação de Garantias de Crédito por Safras de Café	100%
3) Análise de Correlação e Associação	90%
2) Como resumir e interpretar dados?	86%
Certificados	50%
Materiais complementares	40%
3) Fundamentos de BI, Extração e Visualização de Dados usando SQL e Power BI	20%
Mentorias em grupo	1%