



escola
britânica de
artes criativas
& tecnologia

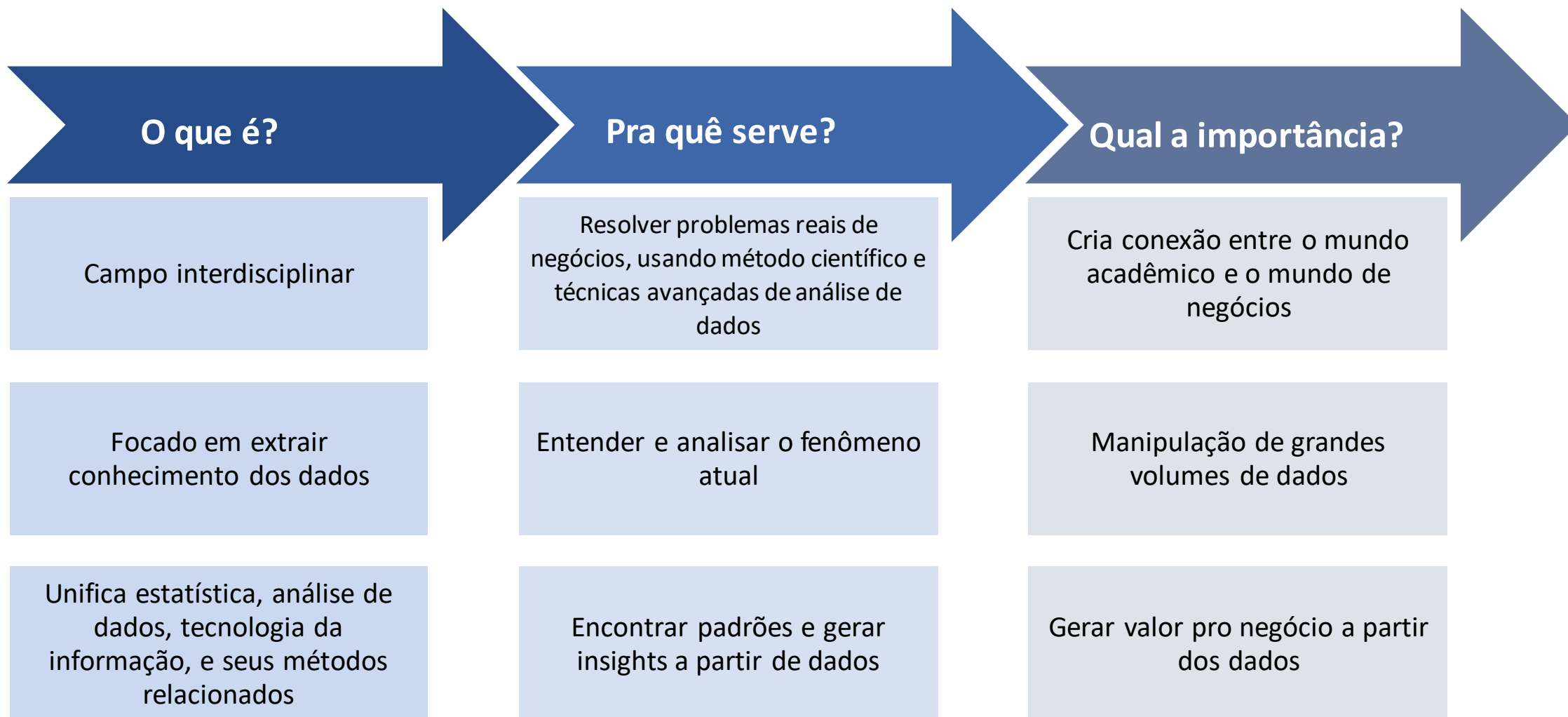
Profissão: Cientista de Dados

Módulo 1: Perspectivas de carreira

Introdução:

Ciência de Dados

O que é Ciência de Dados?



Glossário

Esses conceitos serão aprofundados ao longo do curso



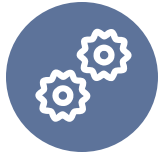
BIG DATA

Grande volume de dados, mais complexo e de novas fontes de dados. Tão volumoso que os tradicionais softwares de processamento de dados não gerenciam.



INTELIGÊNCIA ARTIFICIAL(IA)

Sistemas de computação capazes de realizar tarefas que normalmente necessitam da inteligência humana.



MACHINE LEARNING/APRENDIZADO DE MÁQUINA

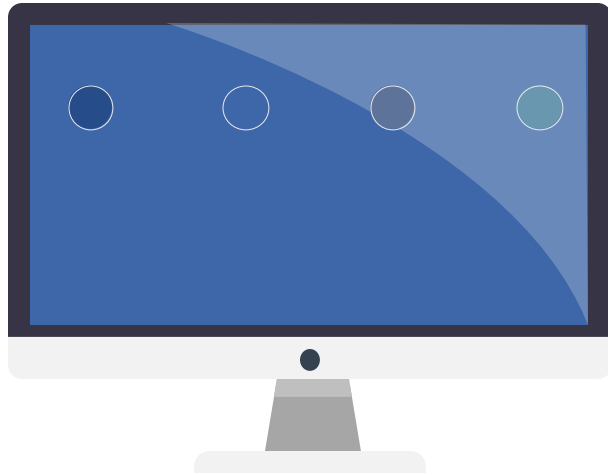
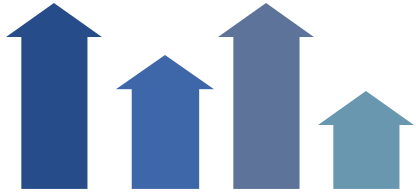
Um programa de computador que consegue aprender com os dados e se adaptar ou inferir coisas sobre dados novos sem a intervenção humana.

Unidade 1:

Demanda por dados

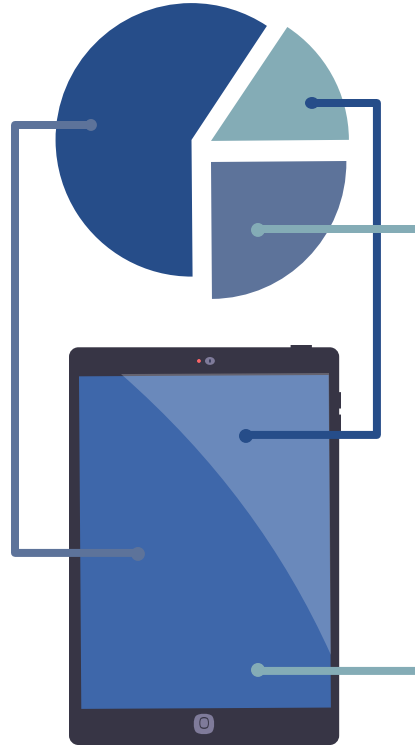
**Por que de repente
começamos a nos importar
com estatística e dados?**

Explosão na quantidade de dados disponíveis



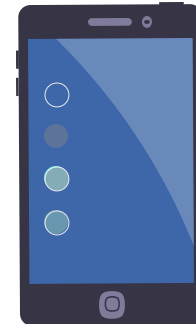
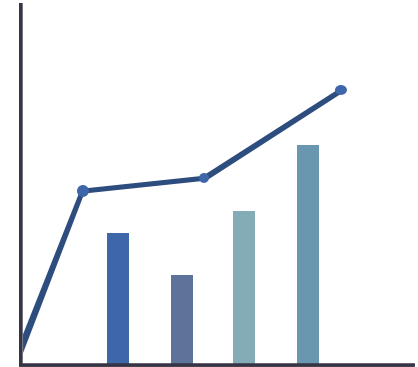
Nova realidade

Maior número de pessoas com celulares/computadores, maior número de sensores



Novas fontes de dados

Cartão de crédito, internet, celulares, sensores na aviação ou no chão de fábrica

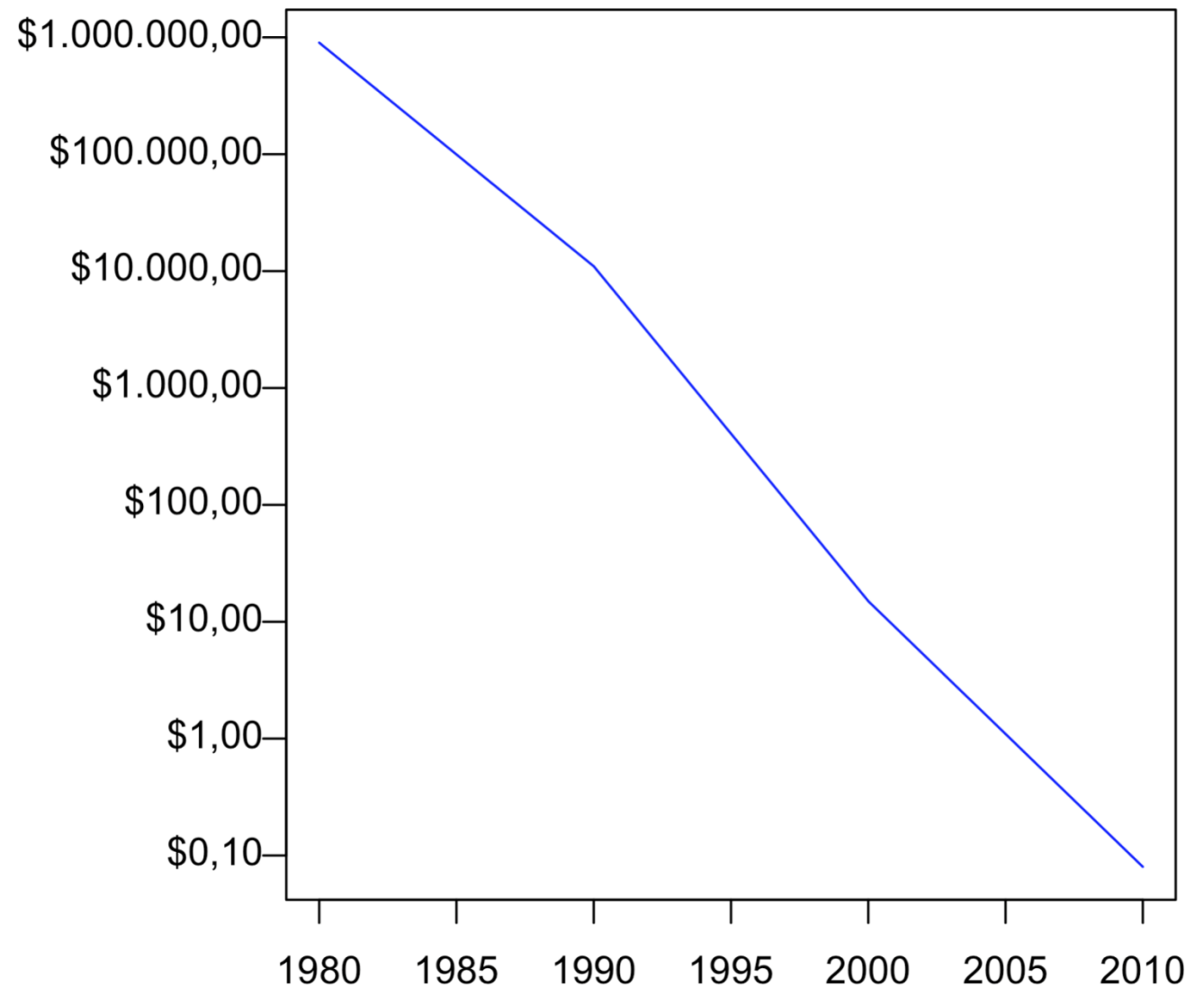


Capacidade de armazenamento

Influência da capacidade de armazenamento

Crescimento da demanda por dados

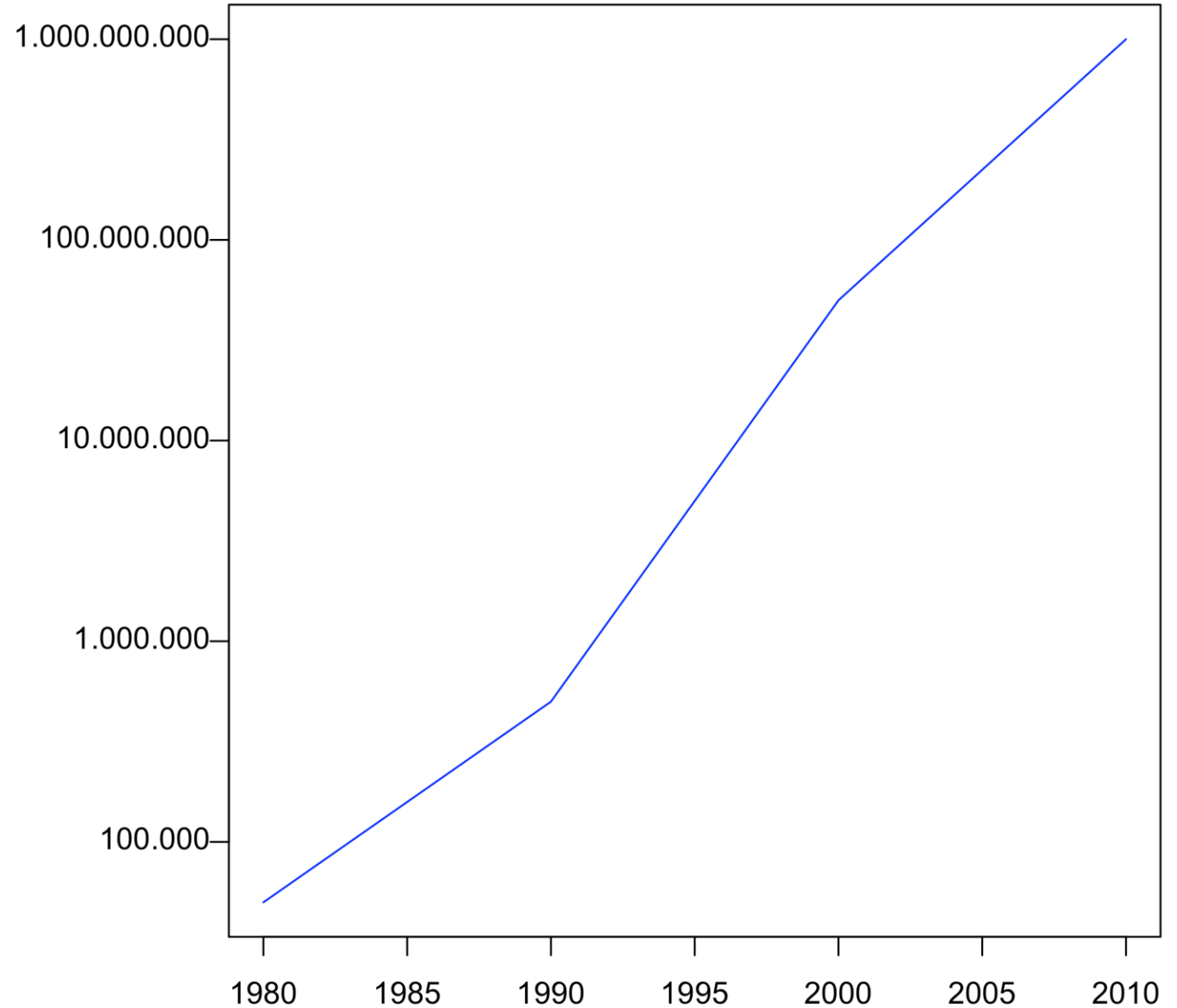
Custo do Gb em USD



Fonte: Steyn, J., Kirlidog, M.; A brief critique of the dominant approaches to ICT4D. A contextualist approach. Link: <https://bityli.com/0ONev>

Crescimento da demanda por dados

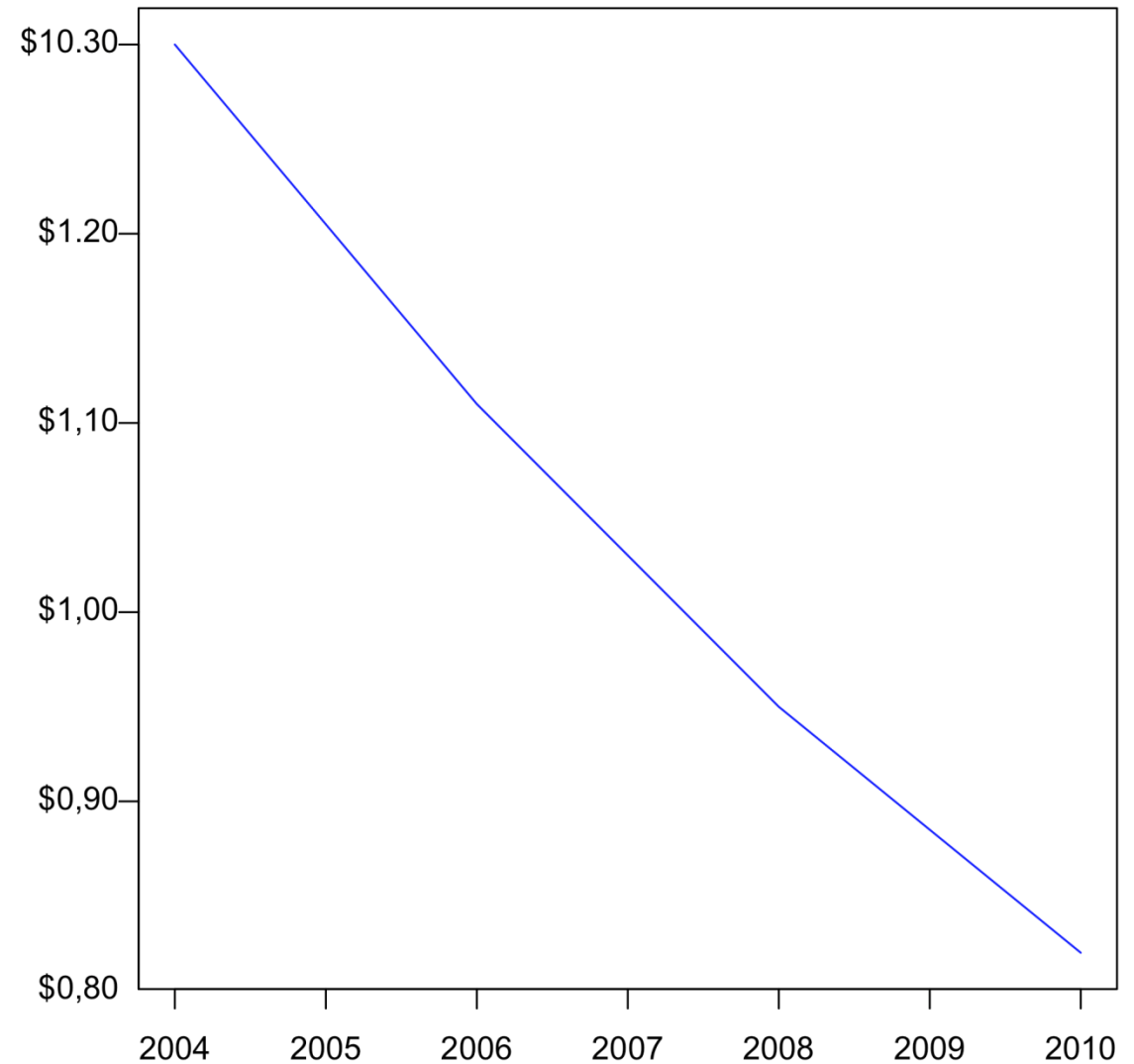
Evolução do número de transístores em microchips dobra a cada ano



Fonte: https://en.wikipedia.org/wiki/Moore%27s_law

Crescimento da demanda por dados

**Custo do sensor industrial de internet das
coisas (IoT)**



Fonte: Link: <https://bit.ly/3eeS3Do>

**Crescimento da
demanda por
dados**



A importância dos dados

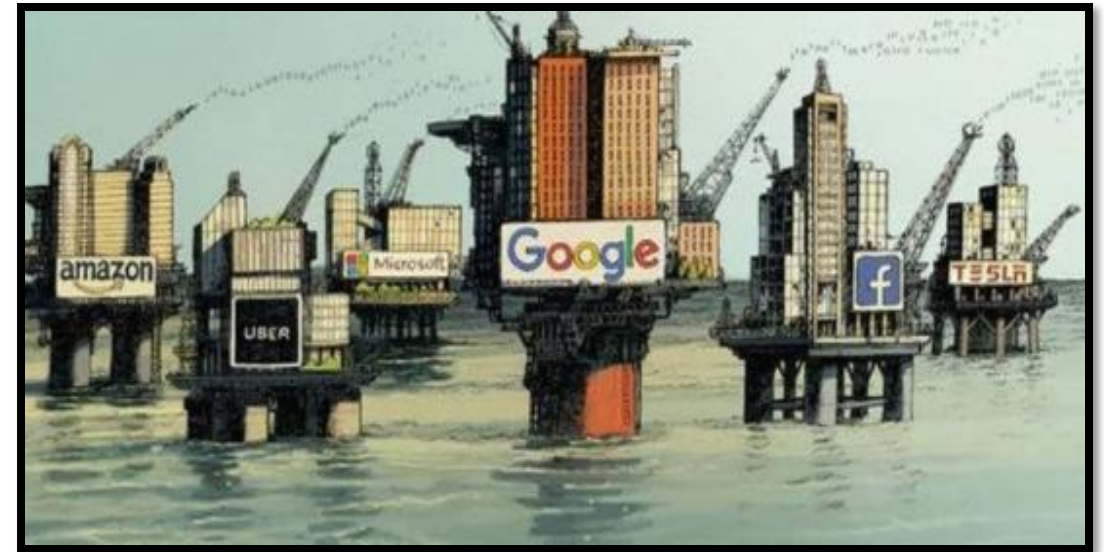
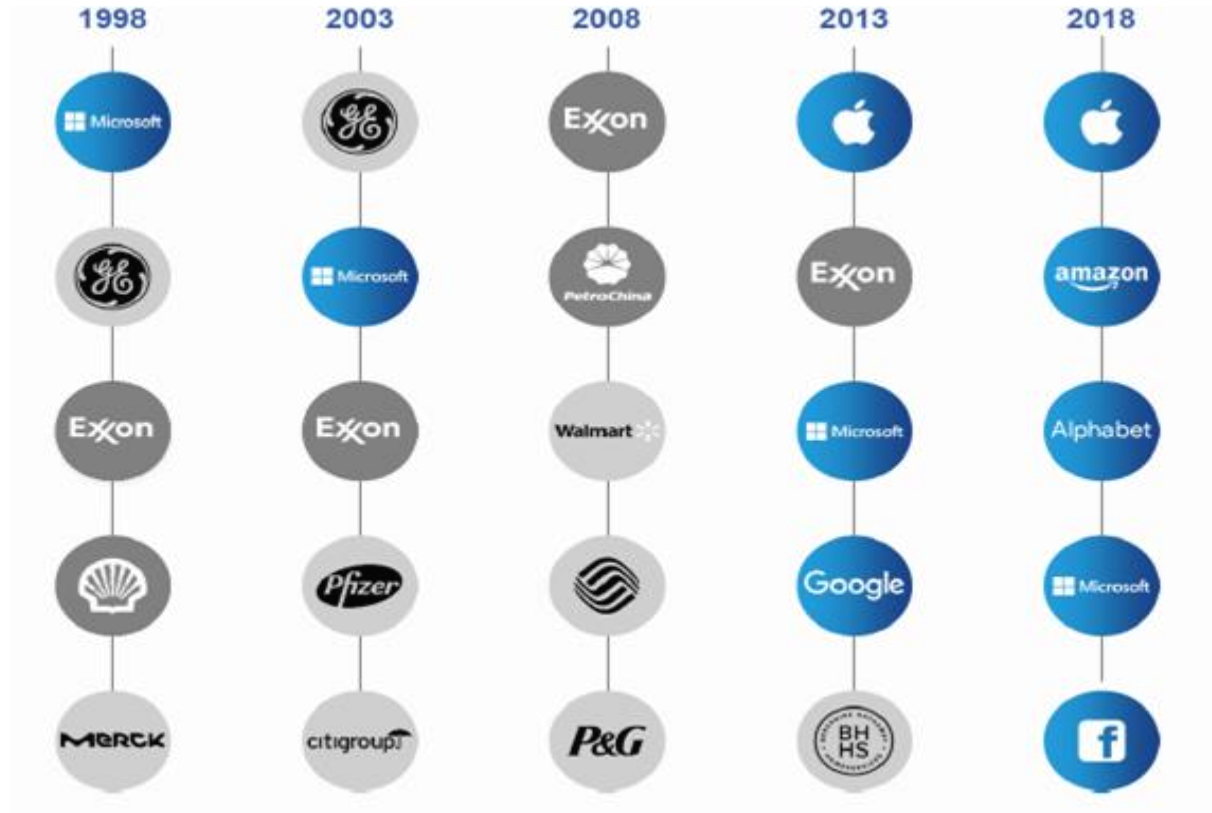
**“Dados são o novo petróleo”
Clive Humby, 2006**

Valioso / Cobiçado

Precisa ser refinado

**Diferentes níveis de
usabilidade**

A importância dos dados



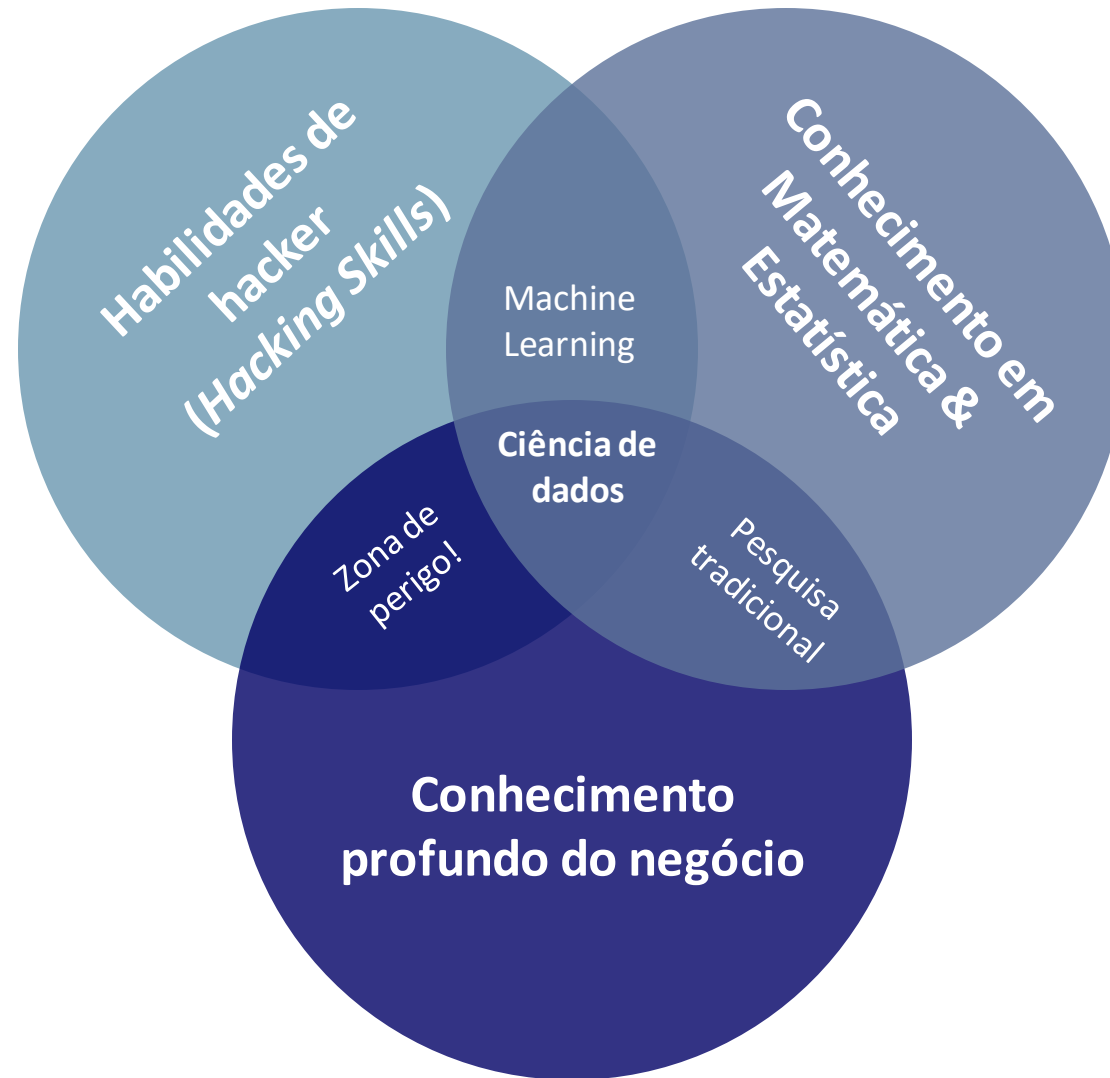
Maiores empresas americanas por valor de mercado

Cultura data driven

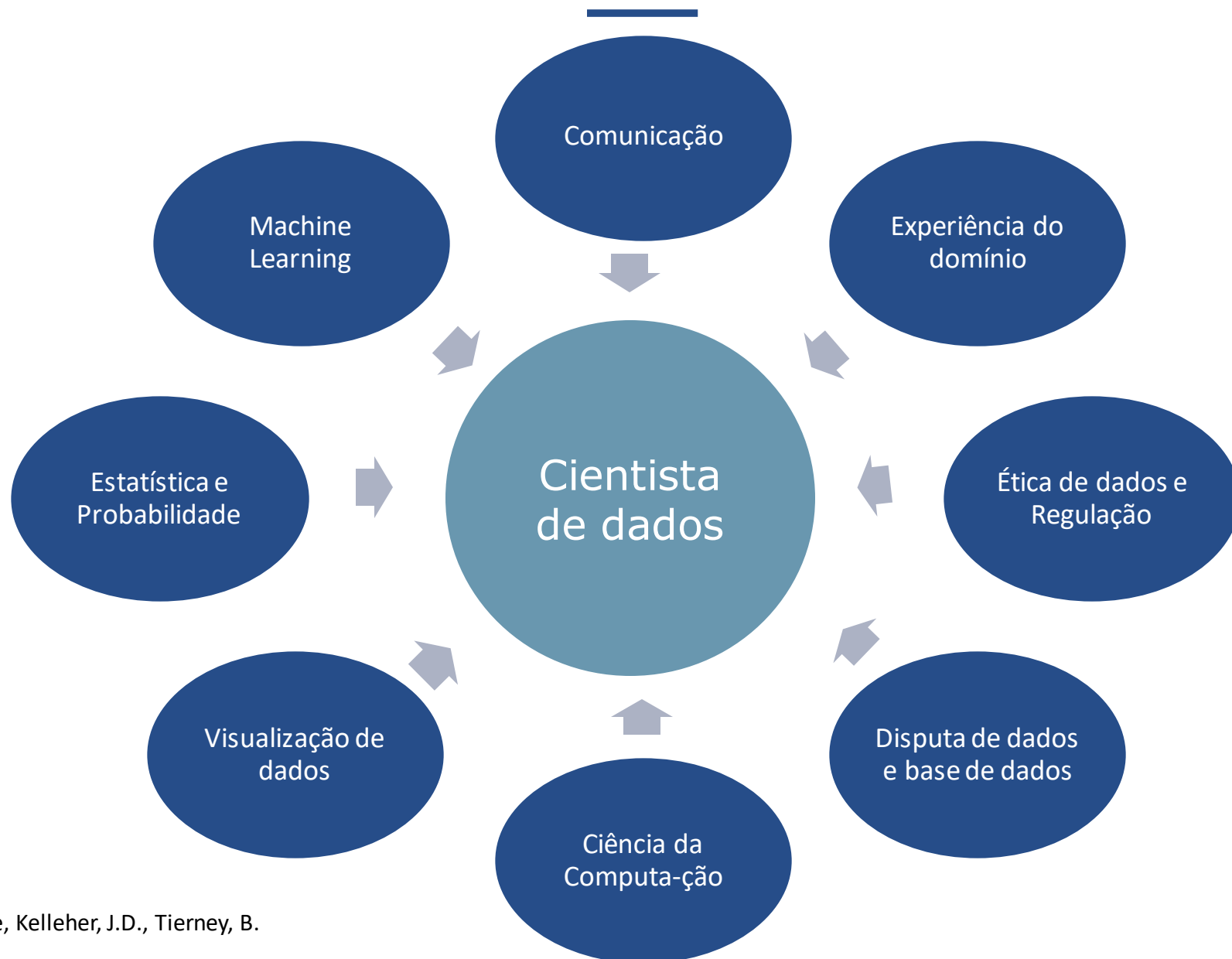


Unidade 2: O profissional

O que é um cientista de dados?

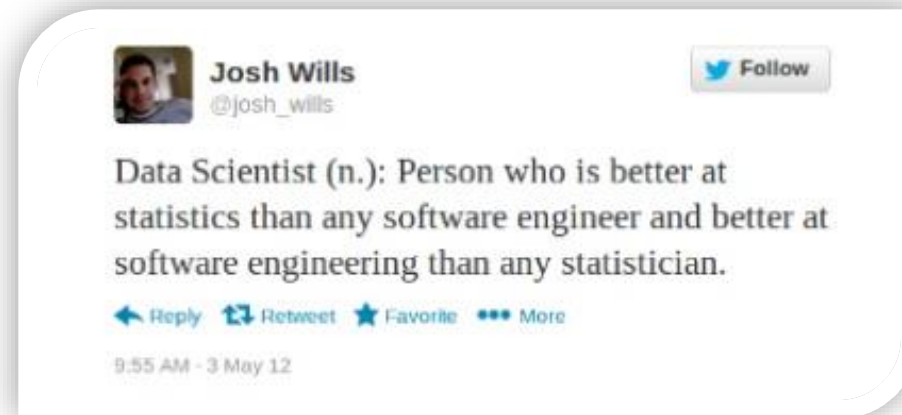


O que é um cientista de dados?



O que é um cientista de dados?

Cientista de Dados (S.): Pessoa que é melhor em estatística que um engenheiro de software e melhor em engenharia de software que um estatístico.



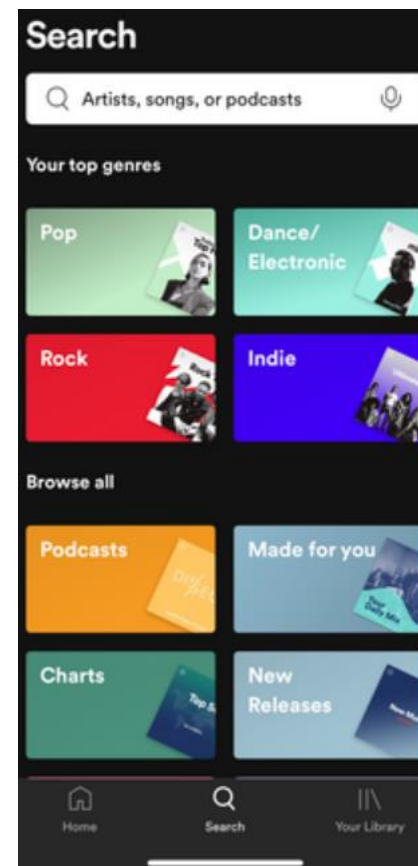
Descrição de uma vaga de cientista de dados



É um profissional “unicórnio”

- Doutorado em uma disciplina quantitativa ou experiência equivalente (por exemplo, Estatística, Bioinformática, Ciência da Computação, Matemática, Física ...)
- Profundo conhecimento em Estatística e Matemática
- Proficiência em tecnologias ETL (Talend, Pentaho DI, SQL Server Integration Services).
- Forte experiência com ferramentas de visualização (Tableau, Qlik, Power BI, ...)
- Proficiência em ecossistemas de Big Data (Spark, Hadoop, Mahout, Kafka, ...)
- Proficiente em linguagens de programação: Python, R, Java, C ++.
- Experiência em Frameworks de Machine Learning e Deep Learning: TensorFlow, PyTorch, Keras, Caffe,...
- 3+ anos de experiência em construção de APIs

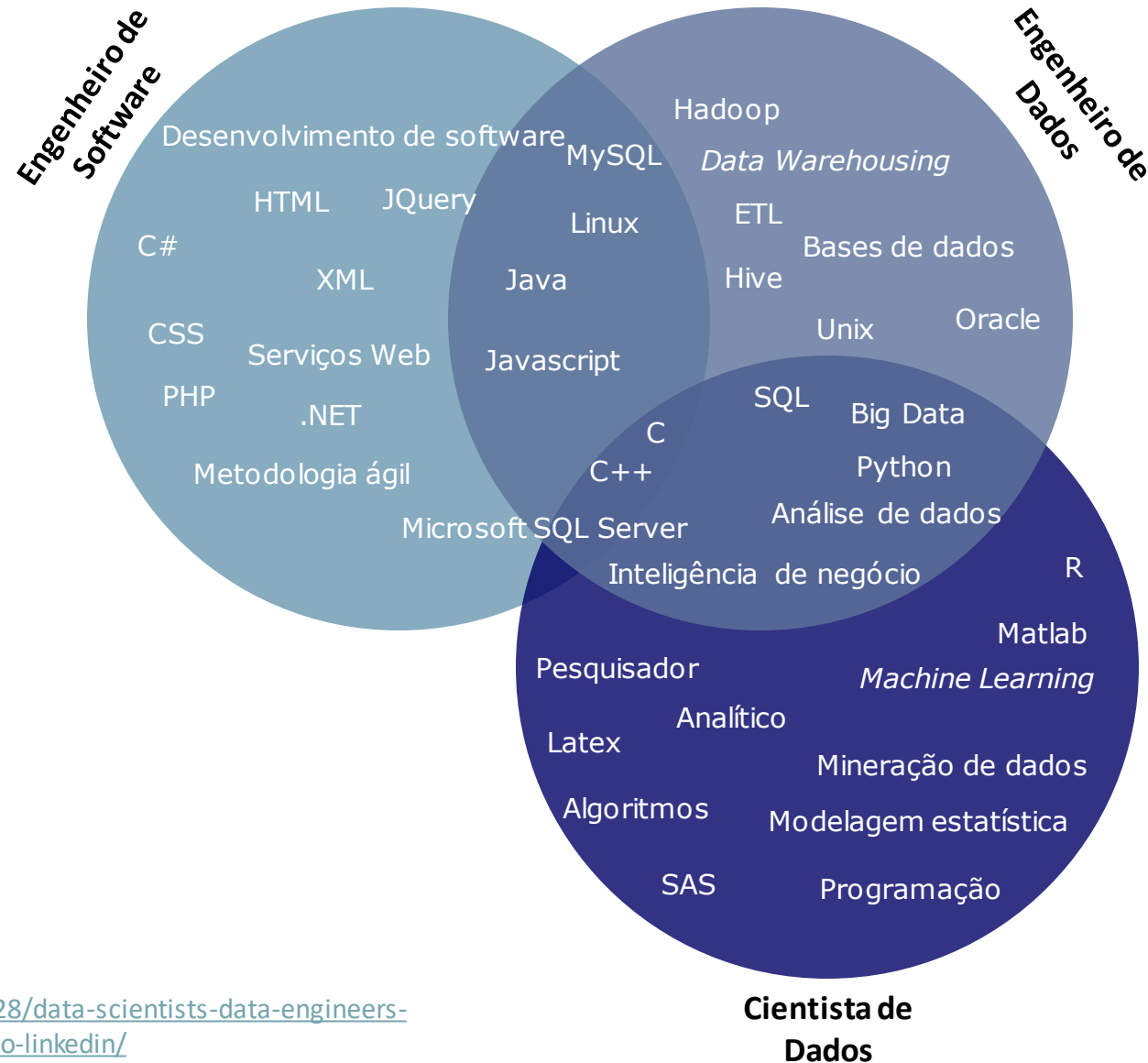
O que faz um cientista de dados?



Unidade 3:

Linhas de desenvolvimento

Diferença entre os profissionais de dados



Diferenças entre os profissionais de dados



Passado		Futuro		
Engenheiro de dados	Analista de dados	Cientista de dados		
Dado bruto	O quê	Como	Porquê	Onde
Números	Descrição	Experiência	Causa e Efeito	Predições
Letras	Contexto	Testado	Provado	O que é melhor
Símbolos	Relacionamentos	Instrução	O que não se sabe	
O que se sabe			Modelos	
Sinais	Relatórios	Programas		

FONTE: Carlos Somohano
Founder Data Science
London

Linhas de desenvolvimento



Full Stack Data Scientist



Especialistas em técnicas

NLP; Computer vision; Regressão logística; Séries temporais



Especialistas em áreas

CRM/Marketing: Segmentação; CLTV; Churn



Riscos

Credit score; Behavior score; Perda esperada; Precificação x risco; Fraude



Unidade 4:

Mercado do cientista de dados

Perspectivas de carreira

“A profissão mais atraente nos próximos 10 anos será a estatística”

Hal Varian, Chief Economist do Google

**“Cientista de dados: A
profissão mais atraente do
século 21”**

D. J. Patil, Harvard Business Review

Salários no Brasil

Valores em R\$		Salário Mensal
Salário Nominal (Bruto Mensal)	Piso Salarial	6.123,35
	Média Salarial	6.709,13
	1º Quartil	3.586,31
	Salário Mediana	6.000,00
	3º Quartil	10.560,32
	Teto Salarial	13.847,01

UF	Jornada	Piso Salarial	Média Salarial
São Paulo	41	6.808,75	7.460,10

Quem contrata um cientista de dados?



Casos de uso real de ciência de dados

FINANCEIRO

- Detecção de fraudes
- Segmentação de clientes
- Score de crédito
- Preço x Risco

VAREJO

- Análise de carrinho
- Previsão de demanda
- Pricing
- Logística

MANUFATURA

- Manutenção preditiva
- Detecção de anomalias
- Previsão de demanda
- Sustentabilidade

Exemplos de empresa: Setor Financeiro

FINANCEIRO

- Detecção de fraudes
- Segmentação de clientes
- Score de crédito
- Preço x Risco



Exemplos de empresa: Varejo

VAREJO

- Análise de carrinho
- Previsão de demanda
- Pricing
- Logística

RENNER

RaiaDrogasil

Droga Raia

DROGASIL

via**varejo**
BALIA pontofrio

Grupo Pão de Açúcar

FLAVIA extra ASIA pontofrio
partiu boniteiro conviva

Exemplos de empresa: Manufatura

MANUFATURA

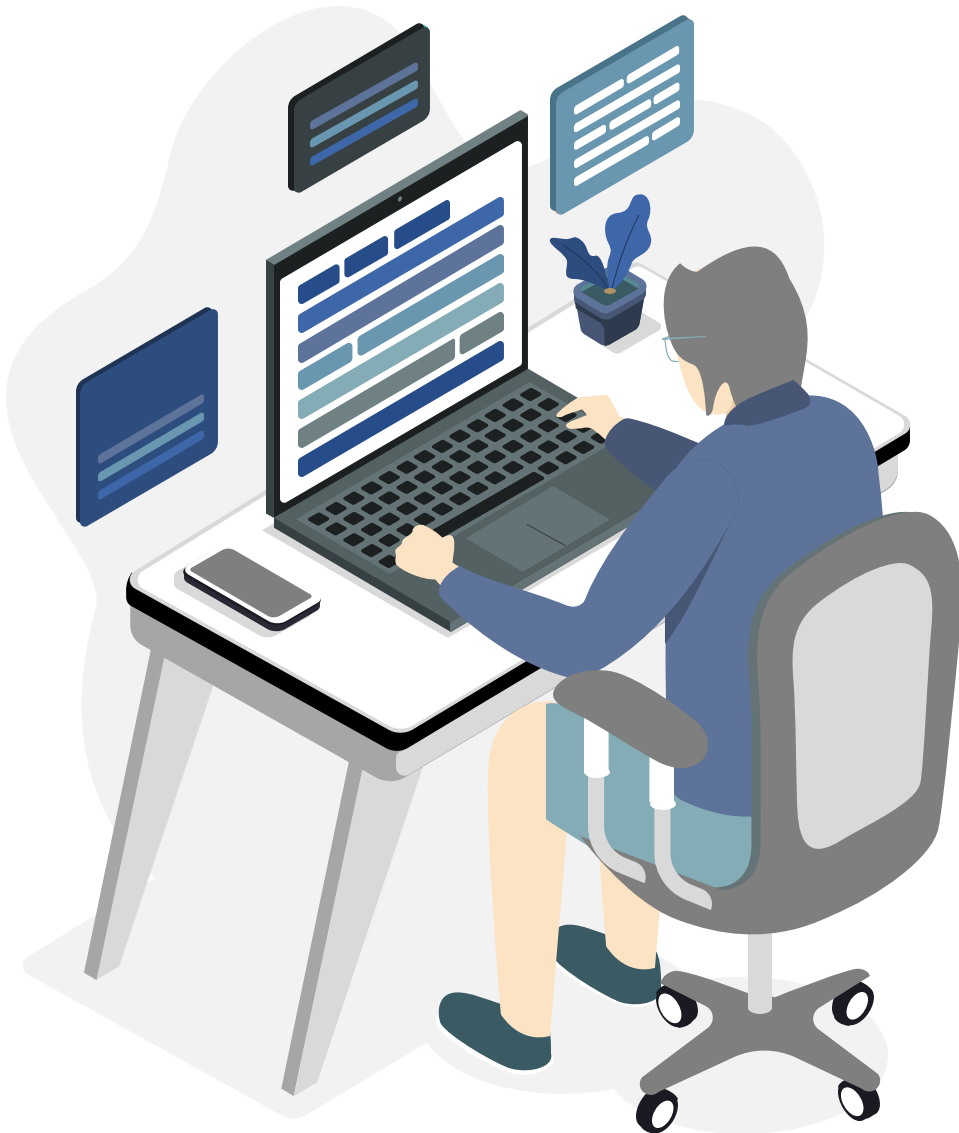
- Manutenção preditiva
- Detecção de anomalias
- Previsão de demanda
- Sustentabilidade



Unidade 4:

Próximos passos

Como eu faço para me tornar um cientista de dados?



O primeiro passo foi dado: Você está aqui!

- Estudar sem um plano pode ser difícil e frustrante.
- Seguir a cadência de um curso vai facilitar o processo de aprendizagem
- Ter o apoio de uma comunidade de estudantes também será importante.

Ciência de dados necessita de diversas habilidades diferentes e ao longo desse curso você vai entender melhor sobre elas e como aperfeiçoá-las.