



# DATA SCIENCE COMO INICIAR?



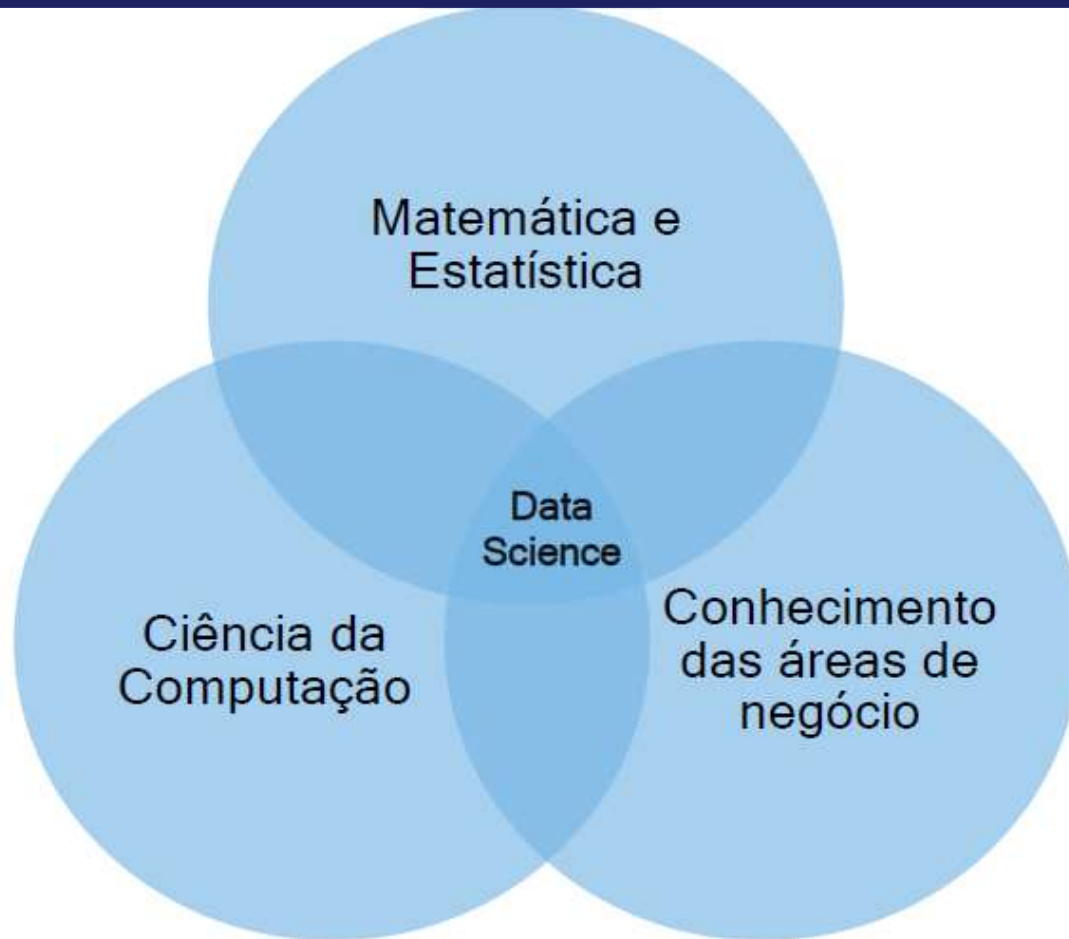
A LINUX DEVE SER A ÚNICA  
EMPRESA A DERRUBAR A  
PRÓPRIA LYNX

## POR QUE ESTUDAR DATA SCIENCE?

O diferencial competitivo  
das empresas HOJE  
devem ser os dados  
(estruturados ou não)



## ÁREAS DE CONHECIMENTO CIÊNCIA DE DADOS



- Arquiteto de Dados (Data Architect)
- Analista de Business Intelligence (BI)
- Analista de Dados (Data Analyst)
- Analista de Machine Learning
- Cientista de Dados (Data Scientist)
- Engenheiro de Dados (Data Engineer)

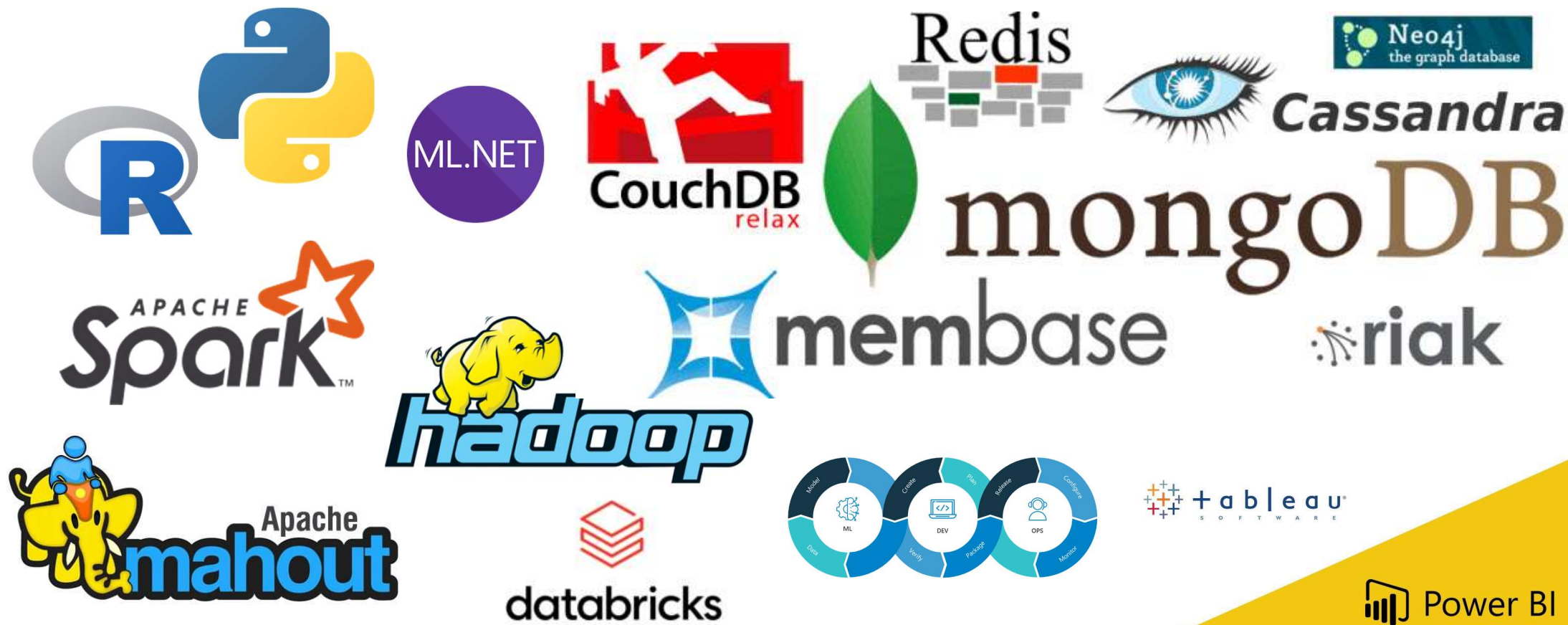
# HABILIDADES POR ÁREA

| Área                   | Habilidade   |
|------------------------|--|
| Business               | Design e Desenvolvimento de Produto                              |
|                        | Gestão de Projetos   |
|                        | Desenvolvimento de Negócios                                      |
|                        | Governança e Compliance  |
|                        | Finanças   |
| Tecnologia             | Gestão de dados estruturados (RDBMS, SQL, XML)                   |
|                        | Gestão de dados não-estruturados (Bancos de dados NoSQL)         |
|                        | Processamento de Linguagem Natural (NLP)                         |
|                        | Machine Learning (árvores de decisão, redes neurais, clustering) |
|                        | Big Data (Hadoop, MapReduce, Spark)                              |
| Matemática e Modelagem | Otimização   |
|                        | Matemática   |
|                        | Modelos gráficos   |
|                        | Algoritmos   |
|                        | Estatística Bayesiana  |

| Área                                    | Habilidade                      |
|---|---------------------------------|
| Programação e Administração de Sistemas | Administração de Sistemas       |
|   | Administração de Banco de Dados |
|   | Cloud                           |
|   | Programação Back-end            |
|   | Programação Front-end           |
| Estatística                             | Gestão de dados                 |
|   | Data Mining e Visualização      |
|   | Modelagem estatística           |
|   | Design de experimentos          |
|   | Comunicação                     |

<https://www.cienciaedados.com/as-10-habilidades-de-um-cientista-de-dados/>

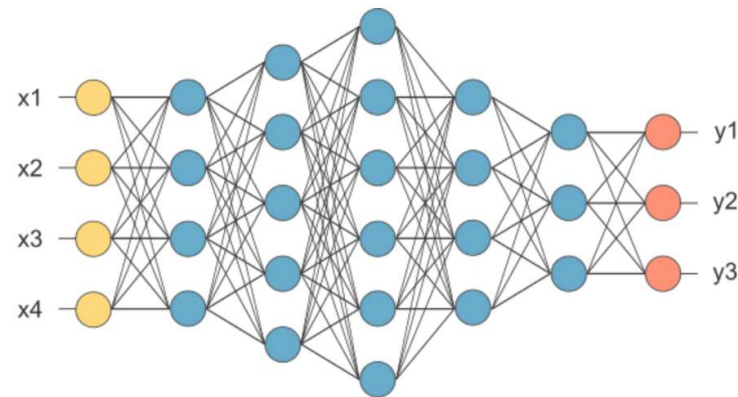
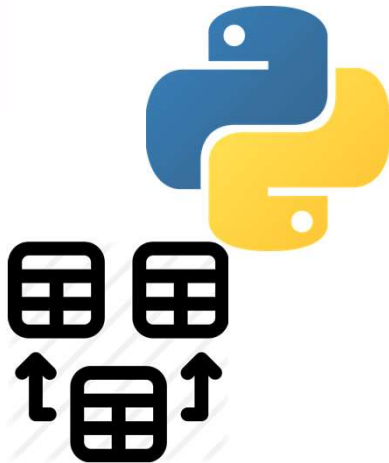
CIÊNCIAS DE DADOS E BIG DATA



COMO INICIAR?

Quantos meses?

Quantas horas por dia?





## DATASET

<https://www.kaggle.com/datasets>

<https://data.fivethirtyeight.com/>

<https://github.com/BuzzFeedNews/everything>

<https://www.data.gov/>

<https://brasil.io/datasets/>

<http://dados.gov.br/>

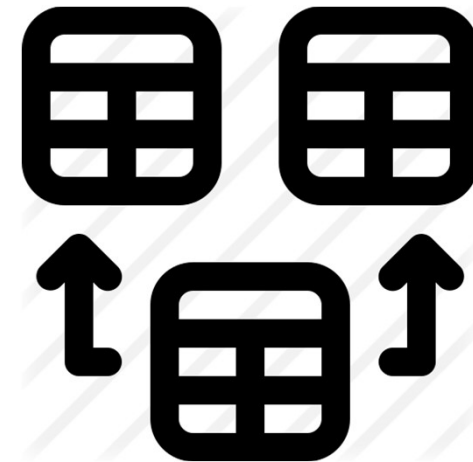
<https://www.reddit.com/r/datasets/>

<https://academictorrents.com/browse.php?cat=6>

<https://dumps.wikimedia.org/>

<https://www.quandl.com/search>

<http://www.portaldatransparencia.gov.br/>





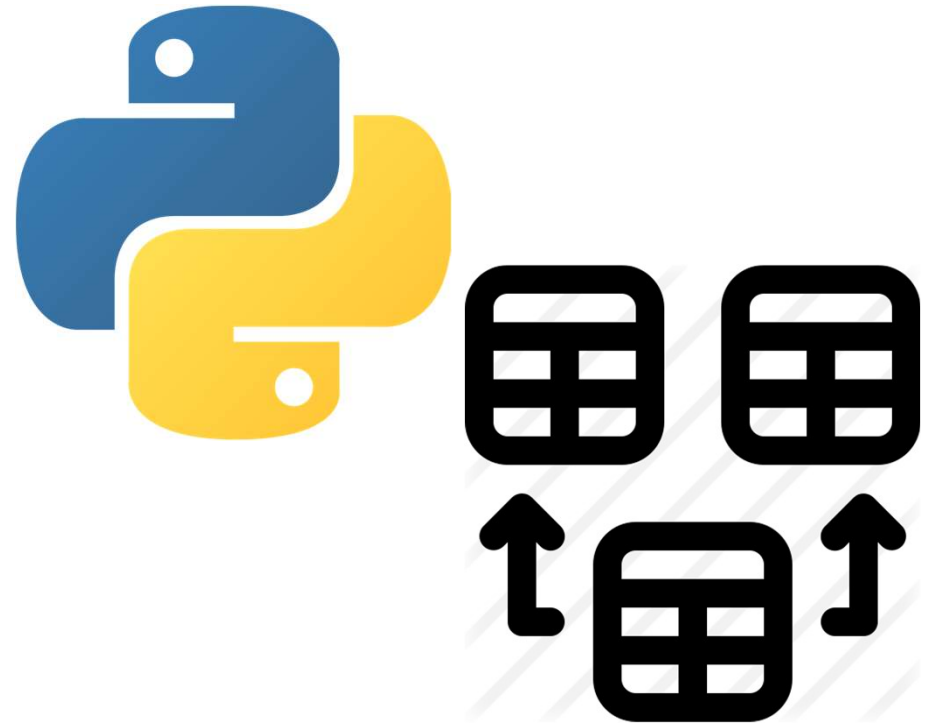
# PYTHON COM DATASET

## MANIPULAÇÃO DE DADOS

- Biblioteca Pandas
  - Selecionar colunas
  - Filtrar linhas
  - Agregação (Group by)
  - Funções Apply Lambda
  - União de dados (Merge, Concat)

## ANÁLISE DE DADOS

- Biblioteca Seaborn e Matplotlib
  - Gráficos
  - Correlações
  - Validação de hipóteses



## CURSOS PYTHON PARA DATA SCIENCE

<https://www.datascienceacademy.com.br/course?courseid=python-fundamentos>

<https://anandology.com/python-practice-book/index.html>

<https://www.youtube.com/watch?v=vmEHCJofslg>

[https://youtu.be/nIHq1MtJaKs?list=PLHz\\_AreHm4dm6wYOIW20Nyg12TAjmMGT-](https://youtu.be/nIHq1MtJaKs?list=PLHz_AreHm4dm6wYOIW20Nyg12TAjmMGT-)

<https://www.cursoemvideo.com/course/python-3-mundo-1/>

<https://www.igti.com.br/custom/desenvolvedor-python/>



CRIAR E RODAR O NOTEBOOK



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

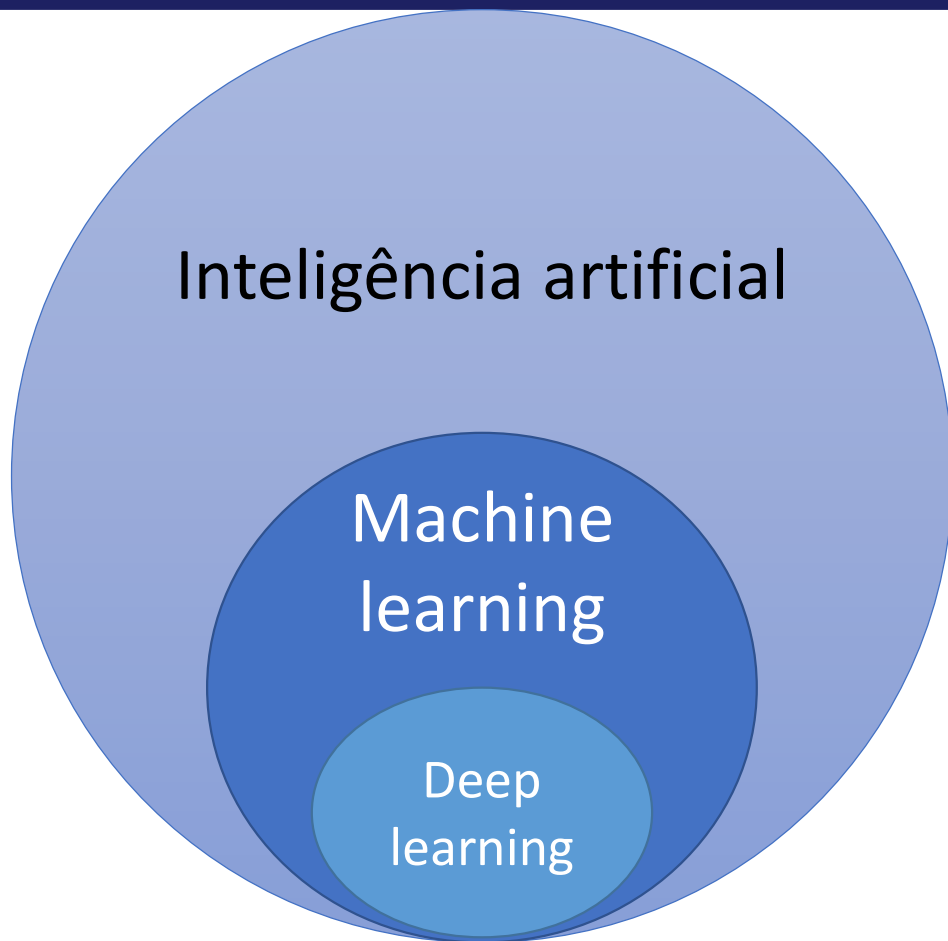


<https://www.kaggle.com>



<https://jupyter.org>

# INTELIGÊNCIA ARTIFICIAL e MODELOS

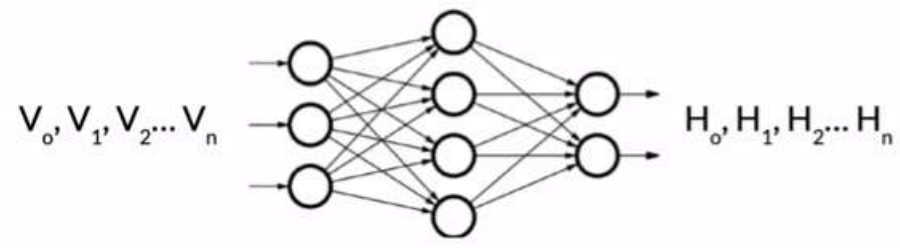


$$v = \sqrt{2gH}$$

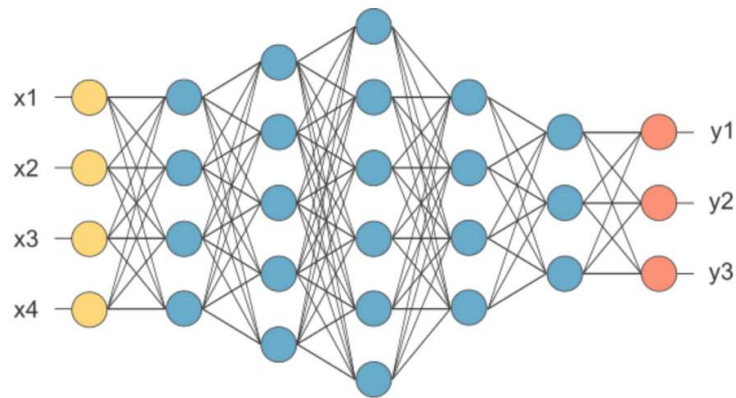
Fórmula Analítica



Aproximação Numérica



# MACHINE LEARNING



REGRESSÃO

CLASSIFICAÇÃO

CLUSTERIZAÇÃO

## MODELOS DE MACHINE LEARNING

- Linear Regression
- Decision Tree
- Random Forest
- Extra trees
- K-Means
- Cluster Hierárquico
- ARIMA

## ENTENDER

- O que é cada modelo?
- Como uso (treino e teste)?
- Qual métrica mostra a performance?

## CURSOS DE MACHINE LEARNING

<https://pt.coursera.org/learn/ai-for-everyone>

<https://work.caltech.edu/telecourse>

<https://www.udemy.com/course/machine-learning-e-data-science-com-python-y/>

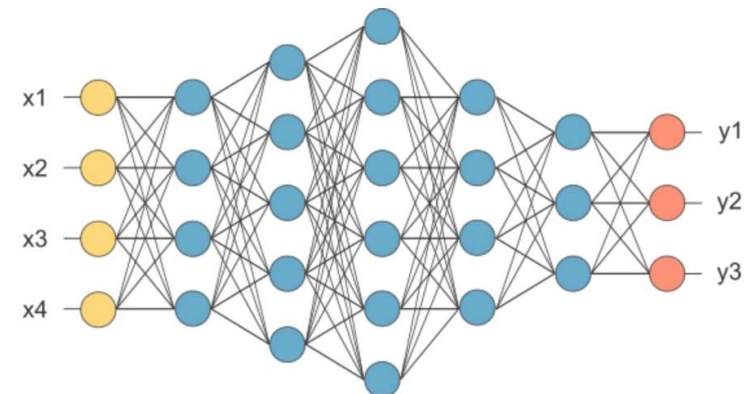
<https://www.igti.com.br/custom/bootcamp-analista-de-machine-learning/>

<https://www.coursera.org/learn/machine-learning>

<https://www.udemy.com/course/inteligencia-artificial-e-m>

<https://www.udacity.com/>

<https://www.edx.org/course/subject/computer-science>



# SQL e Storytelling de Resultados

<https://www.sqlcourse.com>

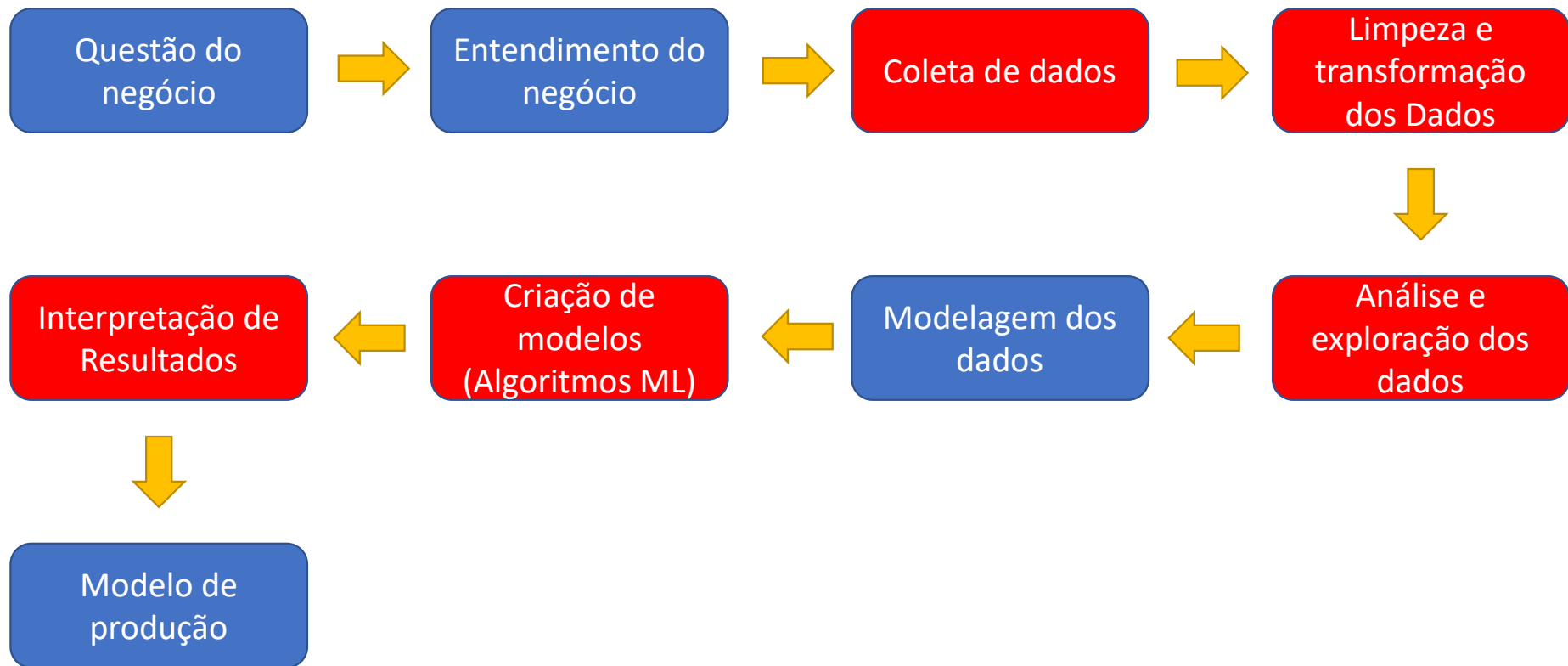
## SQL

- Tabelas e Schemas
- Relações entre tabelas
- Modelos de dados
- Sintaxe de SQL
- SubQueries
- Agregações
- Funções especiais





# TAREFA DE UM DATA SCIENTIST



<https://semanadatascience.com>

- O que é Data Science
- Desenvolver um projeto completo
- Como se tornar um Cientista de Dados
- O que NÃO fazer para aprender
- Data Science
- Como começar da forma certa
- Qual o melhor método

## SEMINÁRIO ONLINE GRATUITO

### **Data Universe**

O universo dos dados em um evento prático e interativo com quem está na linha de frente. Mais de 20 horas com sessões ao vivo sobre business intelligence, ciência de dados, engenharia de dados e data analytics.

<https://vrlps.co/Q63mW31/cp>

## BLOG E COMUNIDADES

<https://www.notion.so/IGTI-Cientista-de-Dados-73457fa8817d4578a8867824f4ac753a>

<https://ml-ops.org/>

<https://sejaumdatascientist.com>

<https://medium.com/@rehaddad/primeiros-passos-com-machine-learning-ml-net-81e428bdc8f>

<https://medium.com/@lauradamaceno/quem-quer-ser-um-cientista-de-dados-7f9bcfd27a57>

<https://cio.com.br/os-7-melhores-cursos-gratuitos-online-de-ciencia-de-dados-para-iniciantes/>

<http://www3.decom.ufop.br/toffolo/pt-br/publicacoes>

<https://community.cloud.databricks.com>

<https://www.meetup.com/pt-BR/Sao-Paulo-Women-in-Machine-Learning-Data-Science>

<https://www.youtube.com/playlist?list=PLqPPg1-6FvSukt2YbVDrOLdMJHo08ADi5>

Fundamentos de Big Data

<https://www.cienciaedados.com/7-e-books-gratuitos-sobre-big-data-ia-machine-learning-e-deep-learning>

<http://deeplearningbook.com.br>

[https://drive.google.com/drive/folders/1\\_KnOIJeqsPdshfC-RR8GehjLmEQ-Ld2W](https://drive.google.com/drive/folders/1_KnOIJeqsPdshfC-RR8GehjLmEQ-Ld2W)

<https://drive.google.com/drive/folders/1JeA2wksnzw11SDrz8NiZegnTmfNYUAgc>

# CURSOS DE QUALIFICAÇÃO

[Produtos](#) ▾[Serviços](#) ▾[Indústrias](#) ▾[Desenvolvedores](#) ▾[Suporte](#)[Procurar](#)

## Prepare-se para o exame da IBM Cloud

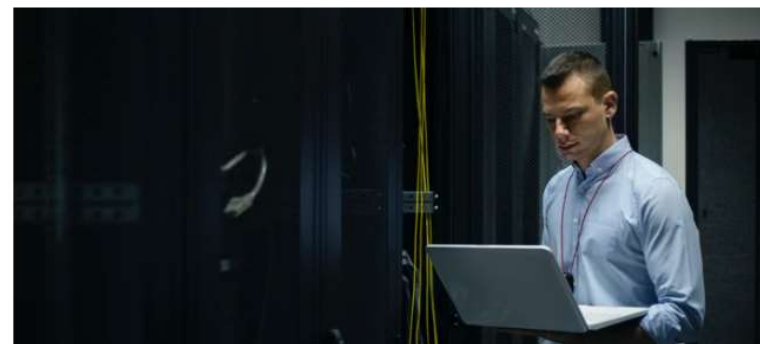
Participe do curso online e gratuito que desenvolvemos para ajudar você a se preparar para o exame da certificação IBM Cloud Foundations!

[Inscriva-se](#)

<https://forms.integrate-events.com/#/events/ibm/la/318515?background=f4f4f4&font=286b94&page=ffffff>

Conforme empresas avançam na adoção de cloud, a procura por profissionais certificados aumenta. Segundo a IDC, em 2025, mais de 90% das novas aplicações serão nativas em nuvem.

As certificações de **IBM Cloud** fornecem credenciais reconhecidas pelo mercado, que ajudam a demonstrar o conhecimento e as habilidades necessárias para aplicar a tecnologia IBM em soluções de negócios.

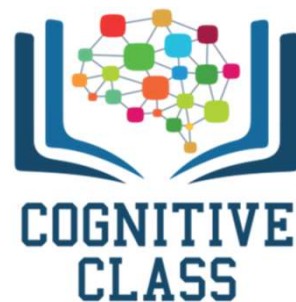




## CURSOS DE QUALIFICAÇÃO

The Kaggle logo is displayed in a light blue, lowercase, sans-serif font.

<https://www.kaggle.com/learn/overview>



<https://cognitiveclass.ai>

The 4Linux logo consists of the word "4LINUX" in a large, bold, black, sans-serif font. Below it, the phrase "OPEN SOFTWARE SPECIALISTS" is written in a smaller, black, sans-serif font.

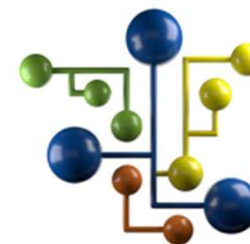
<https://4linux.com.br/cursos/cursos-gratuitos>

The IBM Academic Initiative logo features the IBM logo (a series of horizontal lines) above the text "IBM Academic Initiative" in a black, sans-serif font.

<https://www.research.ibm.com/university/br-pt>

The CódigoFluente logo features the text "CÓDIGOFLUENTE" in a black, sans-serif font. Below it, the words "CÓDIGOFLUENTE" are written in a large, bold, black, sans-serif font, with the letters slightly tilted and overlapping.

<https://www.codigofluente.com.br>



**Data Science  
Academy**

<https://www.datascienceacademy.com.br>

CURSOS DE QUALIFICAÇÃO



<https://somostera.com>



<https://www.udemy.com>



<https://www.edx.org>



<https://www.edx.org>



<https://iaexpert.academy>



<https://educacao-executiva.fgv.br/cursos/online/curta-media-duracao-online/introducao-ciencia-de-dados>



UDACITY

<https://www.udacity.com>



<https://www.startse.com>



<https://fullture.com/data-science>



<https://www.igti.com.br/>

- AULAS INTERATIVAS
- TREINAMENTO PRÁTICO
- CARGA HORÁRIA DE 148 H
- 5 MÓDULOS - 2 MESES



<https://databootcamp.com.br>



**PUC Minas**  
**Virtual**  
educação sem distância

[https://www.pucminas.br/PucVirtual/  
Pos-Graduacao/](https://www.pucminas.br/PucVirtual/Pos-Graduacao/)

- CIÊNCIA DE DADOS E BIG DATA
- INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA

<https://www.igti.com.br/>

- CIÊNCIA DE DADOS
- ENGENHARIA DE DADOS
  - DEEP LEARNING
- APRENDIZADO DE MÁQUINA
- VISÃO COMPUTACIONAL
  - DATA ANALYTICS





[https://mbauspesalq.com/  
cursos/mba-em-data-  
science-e-analytics/](https://mbauspesalq.com/cursos/mba-em-data-science-e-analytics/)

MBA em Data Science e  
Analytics

[http://www.cce.puc-rio.br/sitecce/website/website.dll/folder?nCuro=bi-master-\(traco\)-business-intelligence-master-\(traco\)-sistemas-inteligentes-de-apoio-a-decisao-em-negocios&nInst=ccpg](http://www.cce.puc-rio.br/sitecce/website/website.dll/folder?nCuro=bi-master-(traco)-business-intelligence-master-(traco)-sistemas-inteligentes-de-apoio-a-decisao-em-negocios&nInst=ccpg)

BI MASTER - Business Intelligence Master -  
Sistemas Inteligentes de Apoio à Decisão em  
Negócios







[https://www.fiap.com.br/  
mba/](https://www.fiap.com.br/mba/)