

Data Science e o Mercado de Trabalho



07/12/2020

Data Science e o Mercado de Trabalho

Olá, sou Maria Marinho.

- **Cientista de Dados** na SulAmérica Saúde
- Técnica em **Processamento de Dados**
- Bacharela e licenciada em **Matemática** com ênfase em **Informática**
- Pós graduada em **Educação Matemática**
- Atuei na área de TI como **Desenvolvedora de Sistemas**
- Faço parte das **comunidades** PyLadies e R-Ladies São Paulo.



Data Science e o Mercado de Trabalho



```
def maria_marinho():
    profession = ['Data Scientist', 'System Analyst', 'Developer']
    hobbies = ['Music', 'Shows', 'Piano', 'Movies & Series']
    dogs = ['Iza', 'Lino', 'Luke']
    linkedIn = 'linkedin.com/in/maria-marinho-ds'
```

Data Science e o Mercado de Trabalho

- ✓ Trajetória no mundo de TI, Data Science e comunidades
- ✓ O que os dados me mostraram: projetos e workshops
- ✓ O que é Data Science?
- ✓ Perfil e Jornada do Cientista de Dados
- ✓ Mercado de Trabalho
- ✓ Data Science com Python e R
- ✓ O que são Soft Skills?
- ✓ Por onde começar? Portfólio e Networking
- ✓ Comunidades de Tecnologia
- ✓ Mulheres cientistas de dados
- ✓ Guia de bolso para começar a aventura no mundo dos dados

Minha História – anos 90

O começo: 1995

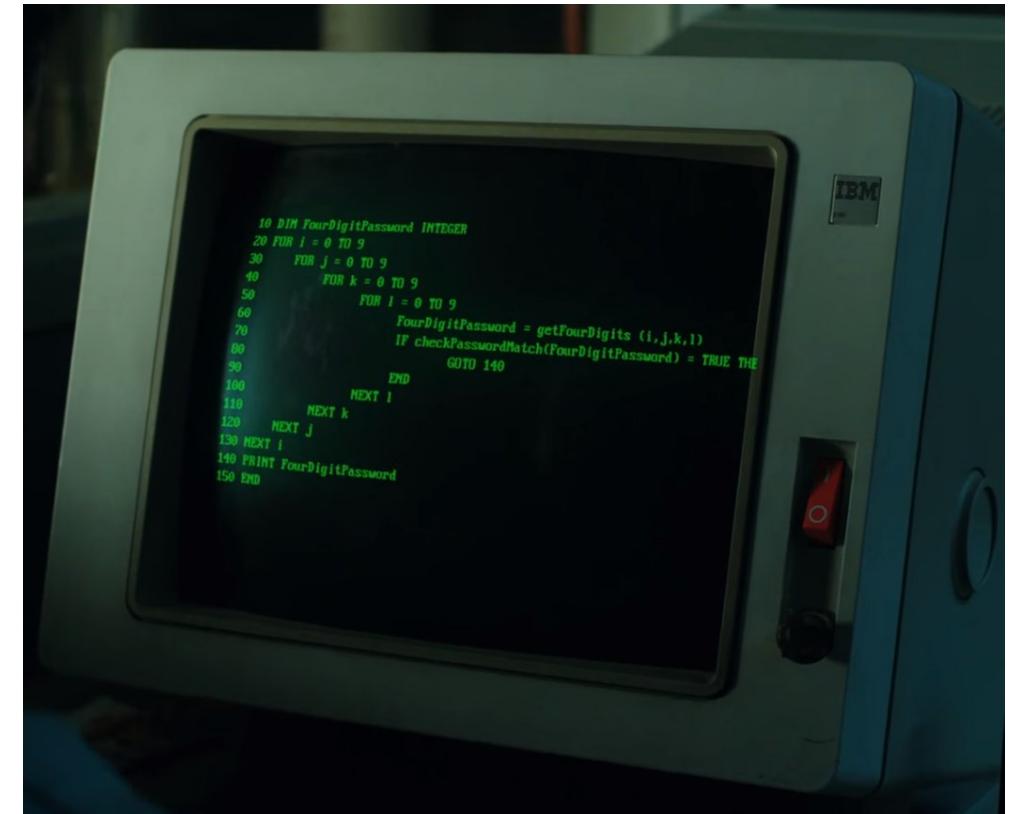
- Colegial Técnico em Processamento de Dados
- Qbasic, Pascal, Cobol, C
- Primeiro estágio: DOS, Windows 3.1, Clipper, dBase, Lotus 123



Minha História – anos 90

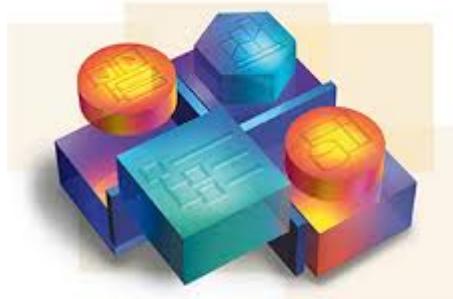
#Spoiler Stranger Things

- A foto ao lado mostra um código na linguagem **Basic** escrito pelo personagem Bob.

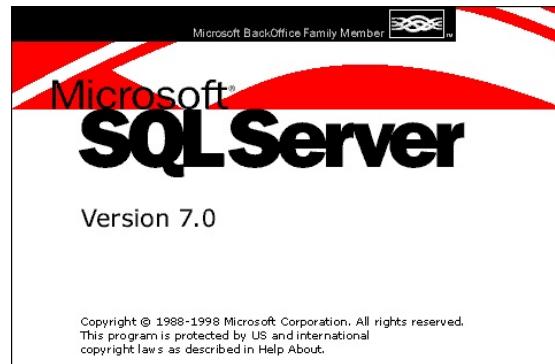


Minha História – anos 2000

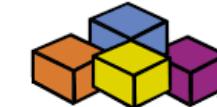
- Bacharelado em Matemática com Informática
- Pós graduação em Educação Matemática
- Analista desenvolvedora
- Professora de Matemática



Microsoft
Visual Basic 6.0



IBM Db2



Microsoft®
Visual Basic®
for Applications

Minha História – anos 2010

- Primeiro contato com um software estatístico em uma empresa de Pesquisa de Mercado: SPSS
- Intercâmbio
- Analista Desenvolvedora em empresas do campo bancário e farmacêutico: automatização de processos em VBA



Minha História – era da Ciência de Dados

- Descoberta da Ciência de Dados: um horizonte de possibilidades
- Cursos MOOC: Coursera
- Estudo sobre a Relação entre o Câncer de Mama e a Emissão de CO2
- Descoberta dos Meetups e Workshops das Comunidades de Tecnologia
- Bolsa de Estudos pela Udacity e Bertelsmann
- Monitora e Professora de cursos das PyLadies e R-Ladies São Paulo



UDACITY



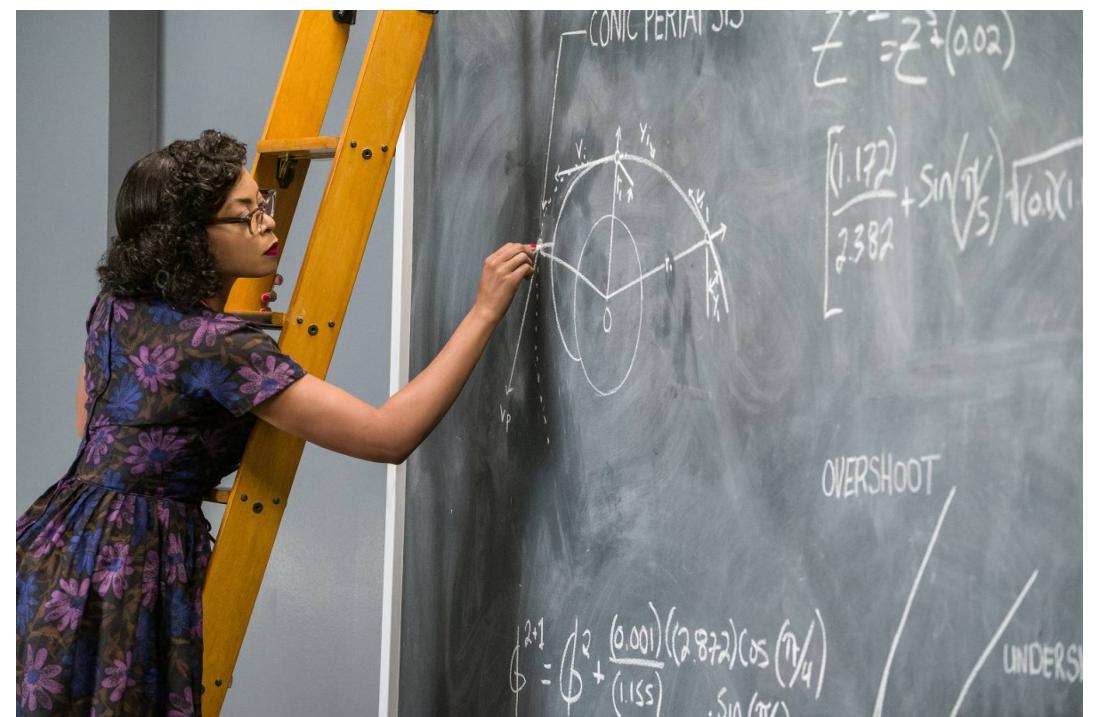
python™

coursera

Data Science e o Mercado de Trabalho

- O que os dados me mostraram?
- Projetos e Workshops

Cena do filme “Estrelas Além do Tempo”



Udacity Show Project: análise da base do ProUni

- Projeto em grupo para o Bertelsmann Data Science Challenge Scholarship Course

Universities' prices analysis based on Prouni database

- **Problem:** Is there any relevant difference between universities' prices grouped by UF (Federative Unit) for Computer Science graduation?
- **Hypothesis 1:** there isn't difference in monthly fees between UFs
- **Hypothesis 2:** there is difference in monthly fees between UFs



Datathon: Desafio de Dados sobre Saúde Pública no Brasil

- **Segurados de ambos sistemas: SUS e Planos de Saúde**

A análise tem como base os atendimentos de beneficiários de planos de saúde no Sistema Único de Saúde (SUS) e tem como objetivo responder a pergunta: Os planos de saúde resarcem o SUS de acordo com a lei nº 9.656/1998?



HackMobilidade 2018

- Análise de acidentes com pedestres (atropelamentos) biênio 2016-2017

Tema do HackMobilidade 2018: usar ciência de dados para propor soluções de segurança para a mobilidade ativa feminina.

Equipe de 4 mulheres:

- Alissa Mune
- Fabíola Canedo
- Maria Marinho
- Monica Craveiro



HackMobilidade 2018



HackMobilidade 2018



Workshops & Palestras

2018

- **Python para Casais** – PyLadies São Paulo
- **Data Science com Python para Todas** - PyLadies São Paulo
- **Os Primeiros Passos na Linguagem R** para estudantes da Fatec Zona Sul

2019

- **Workshops de R** (Básico e Intermediário) com as R-Ladies São Paulo
- **Introdução à Estatística com Pandas** com o Grupo de Estudos de Data Science das PyLadies São Paulo (GEDS)
- **Data Science e o Mercado de Trabalho** – Universidade Metodista
- **Do SQL para o R usando o pacote dplyr** no 1º satRday São Paulo

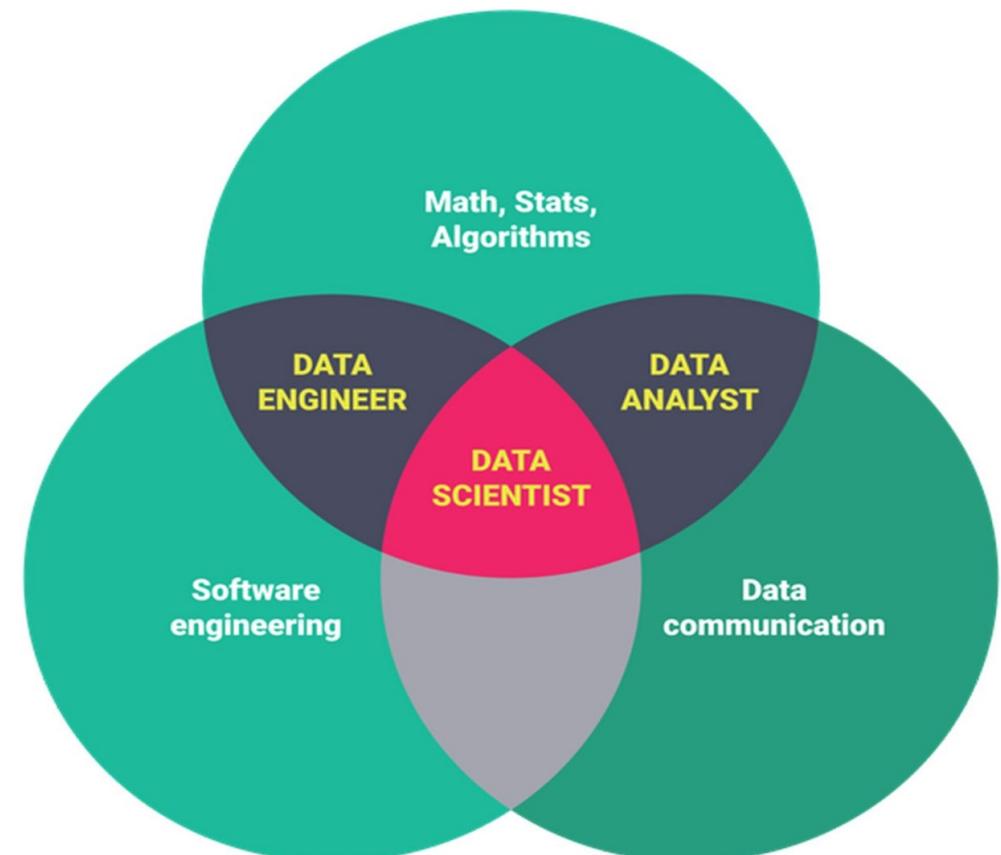
Workshops & Palestras

2020

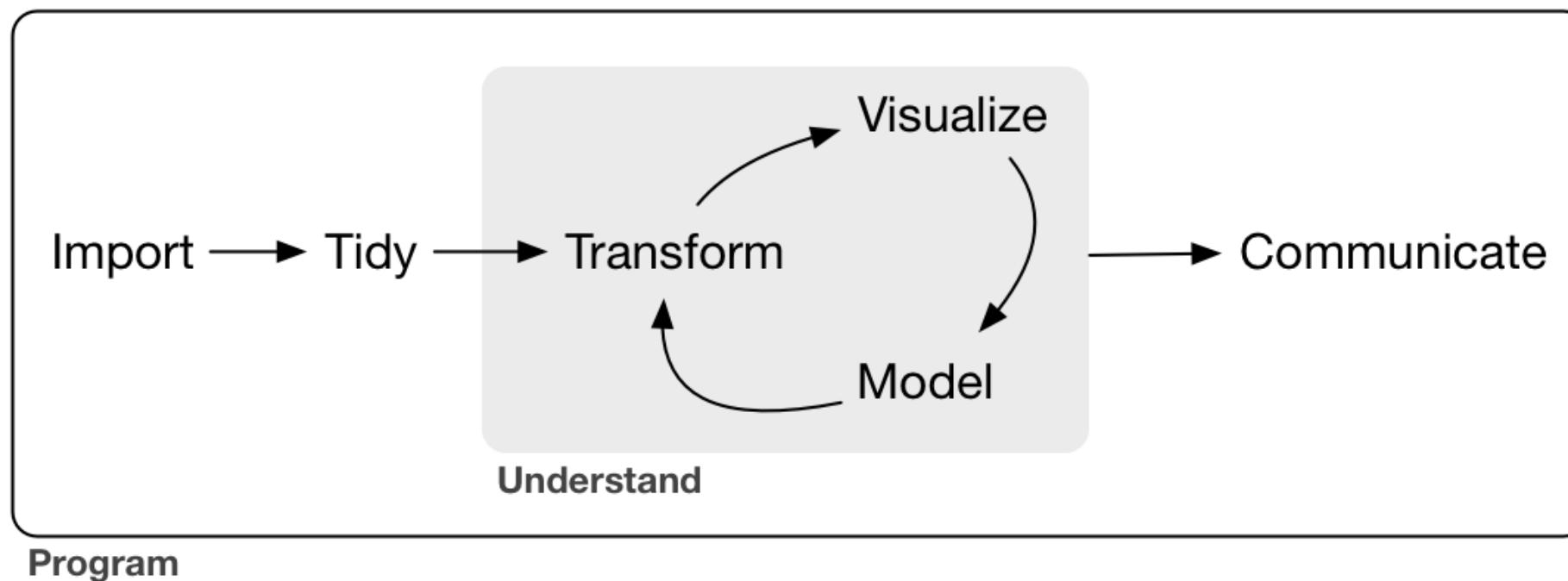
- **Do SQL para o R usando o pacote dplyr** – R-Ladies São Paulo e Nerdzão
- **Ciência de Dados** - SulAmérica
- **SQL e R: um ótimo inner join** – R-Ladies Gyn (Goiânia)
- **Gestão de Saúde Populacional e Ciência de Dados** – Tangerines
- **Estatística para Jornalistas** – Abraji
- **Data Science e o Mercado de Trabalho** – Universidade Metodista

Mas afinal, o que é Data Science?

- **Data Science ou Ciência de Dados** é uma ciência interdisciplinar sobre o processamento de grandes conjuntos de dados usando métodos estatísticos para extrair insights sobre os dados brutos.

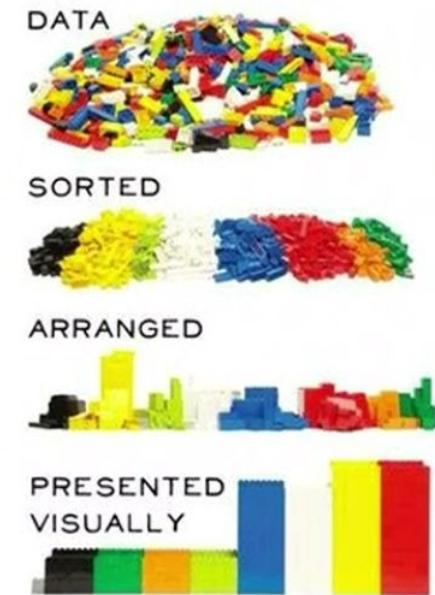


Ciclo da Ciência de Dados



Fonte: [R for Data Science](#)

Ciência de Dados... TADA!

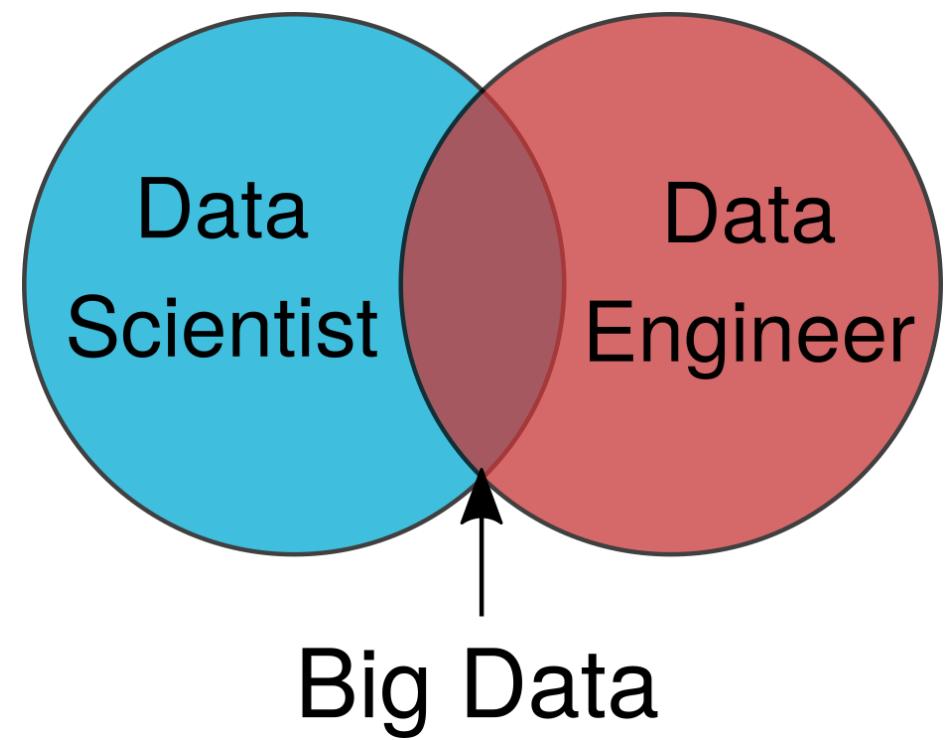


DATASCIENCE **GO**

Fonte: slide do Randy Lao na Conferência Data Science Go 2018

Data Science, Data Engineer: Big Data

- **Data Engineer ou Engenharia de Dados** é a área que se dedica a superar os “gargalos” de processamento de dados e problemas de manuseio de dados para aplicações que utilizam grandes volumes, variedades, e velocidades de dados.

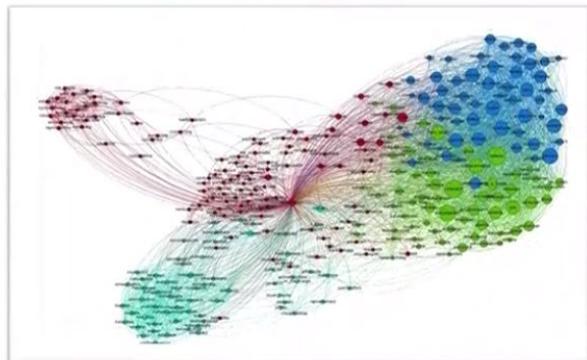


O que é Machine Learning?

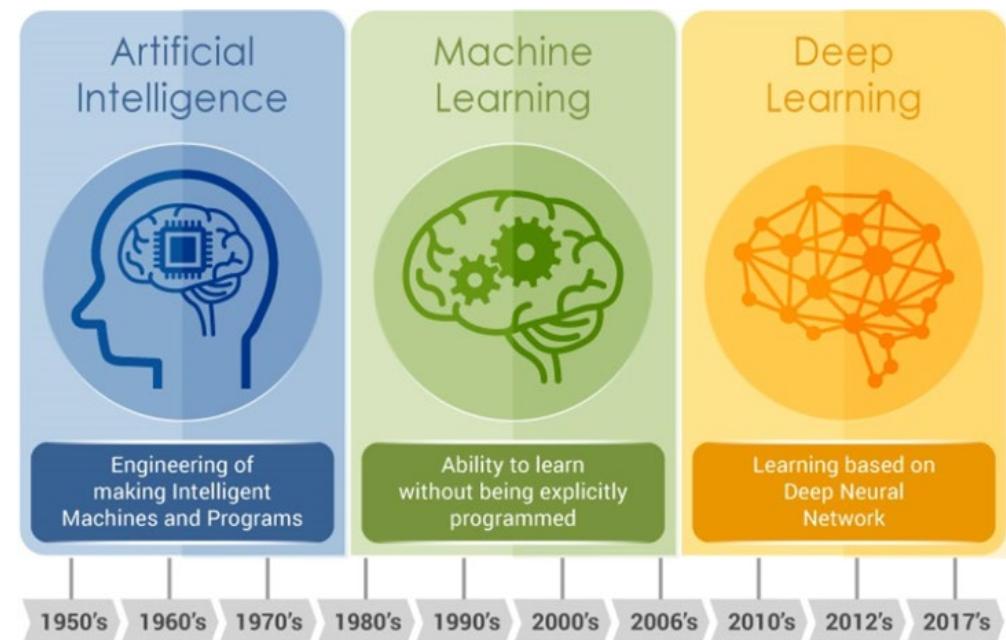
Machine Learning Explanation

Machine Learning is about:

- creating **algorithms** (set of rules)
- that **learns** from complex functions (patterns) from data
- to make **predictions** on it



**1. Algorithms
2. Learn
3. Prediction**



Fonte: slide do Randy Lao na Conferência Data Science Go 2018

Data Scientist

- **Data Scientist ou Cientista de Dados** é quem extrai insights de dados brutos (row data, messy data).



MODERN DATA SCIENTIST

Data Scientist, the sexiest job of the 21th century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ★ Machine learning
- ★ Statistical modeling
- ★ Experiment design
- ★ Bayesian inference
- ★ Supervised learning: decision trees, random forests, logistic regression
- ★ Unsupervised learning: clustering, dimensionality reduction
- ★ Optimization: gradient descent and variants

PROGRAMMING & DATABASE

- ★ Computer science fundamentals
- ★ Scripting language e.g. Python
- ★ Statistical computing packages, e.g., R
- ★ Databases: SQL and NoSQL
- ★ Relational algebra
- ★ Parallel databases and parallel query processing
- ★ MapReduce concepts
- ★ Hadoop and Hive/Pig
- ★ Custom reducers
- ★ Experience with xaaS like AWS

DOMAIN KNOWLEDGE & SOFT SKILLS

- ★ Passionate about the business
- ★ Curious about data
- ★ Influence without authority
- ★ Hacker mindset
- ★ Problem solver
- ★ Strategic, proactive, creative, innovative and collaborative

COMMUNICATION & VISUALIZATION

- ★ Able to engage with senior management
- ★ Story telling skills
- ★ Translate data-driven insights into decisions and actions
- ★ Visual art design
- ★ R packages like ggplot or lattice
- ★ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

Perfil do Cientista de Dados

- **Rachel Schutt**, uma das autoras do livro “**Doing Data Science**”, criou um gráfico para visualizar a si mesma como cientista de dados.
- ▶ **Reflexão:**
- Onde está o seu perfil de cientista de dados no momento?
- Onde você gostaria de estar daqui meses ou anos?

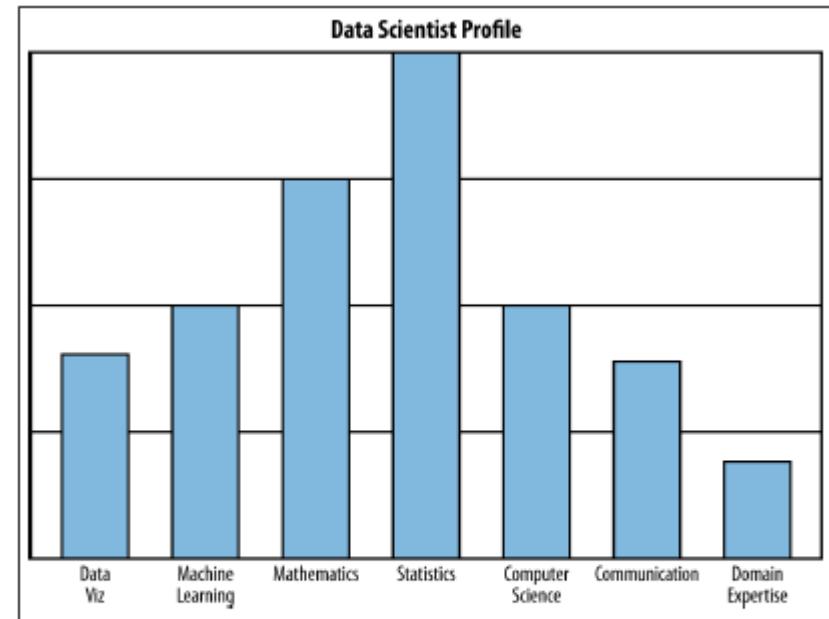
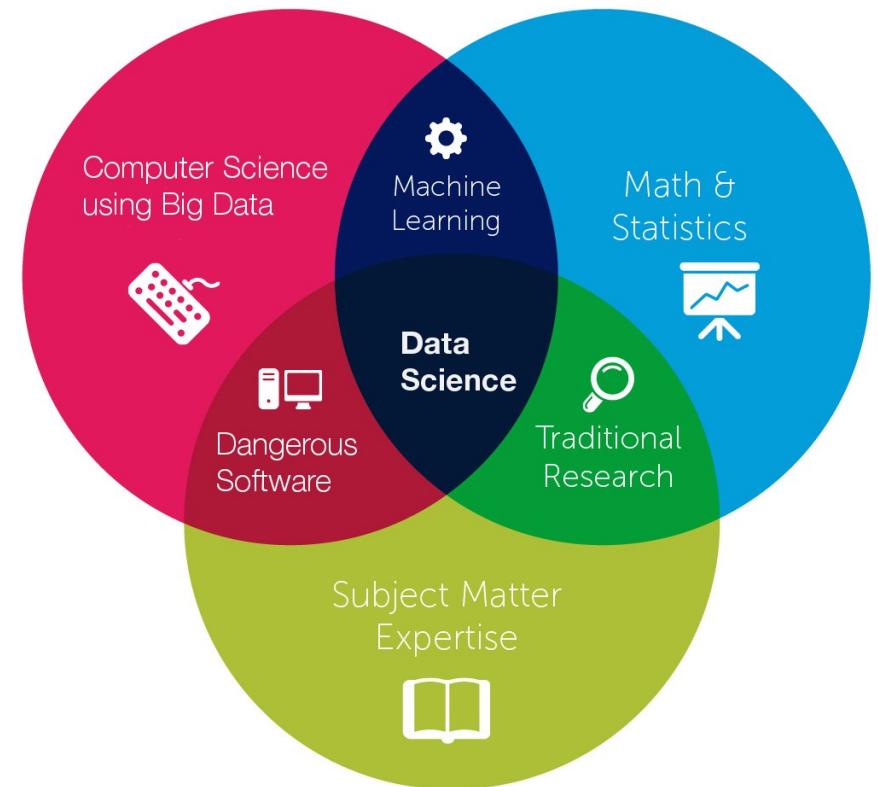


Figure 1-2. Rachel's data science profile, which she created to illustrate trying to visualize oneself as a data scientist; she wanted students and guest lecturers to “riff” on this—to add buckets or remove skills, use a different scale or visualization method, and think about the drawbacks of self-reporting

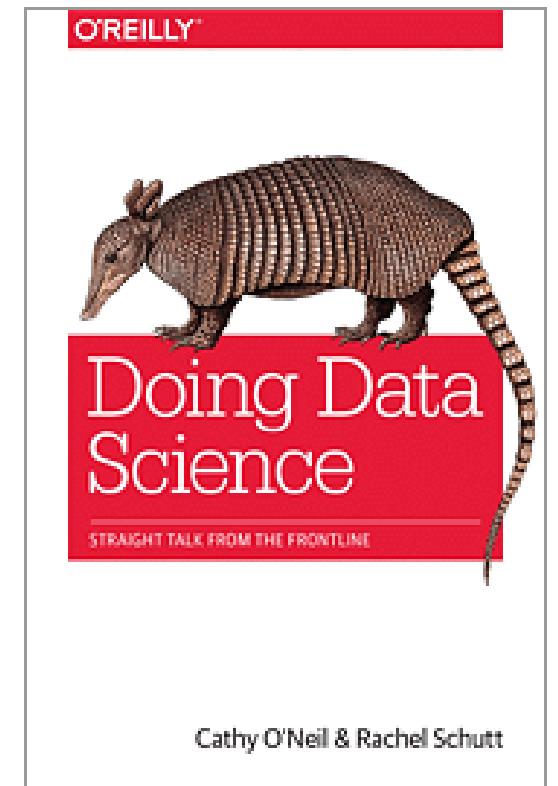
Perfil do Cientista de Dados

- Ciência da Computação
- Matemática
- Estatística
- Machine Learning (Aprendizagem de Máquina)
- Conhecimento da área que será “investigada”
- Habilidades de comunicação e apresentação
- Visualização de dados

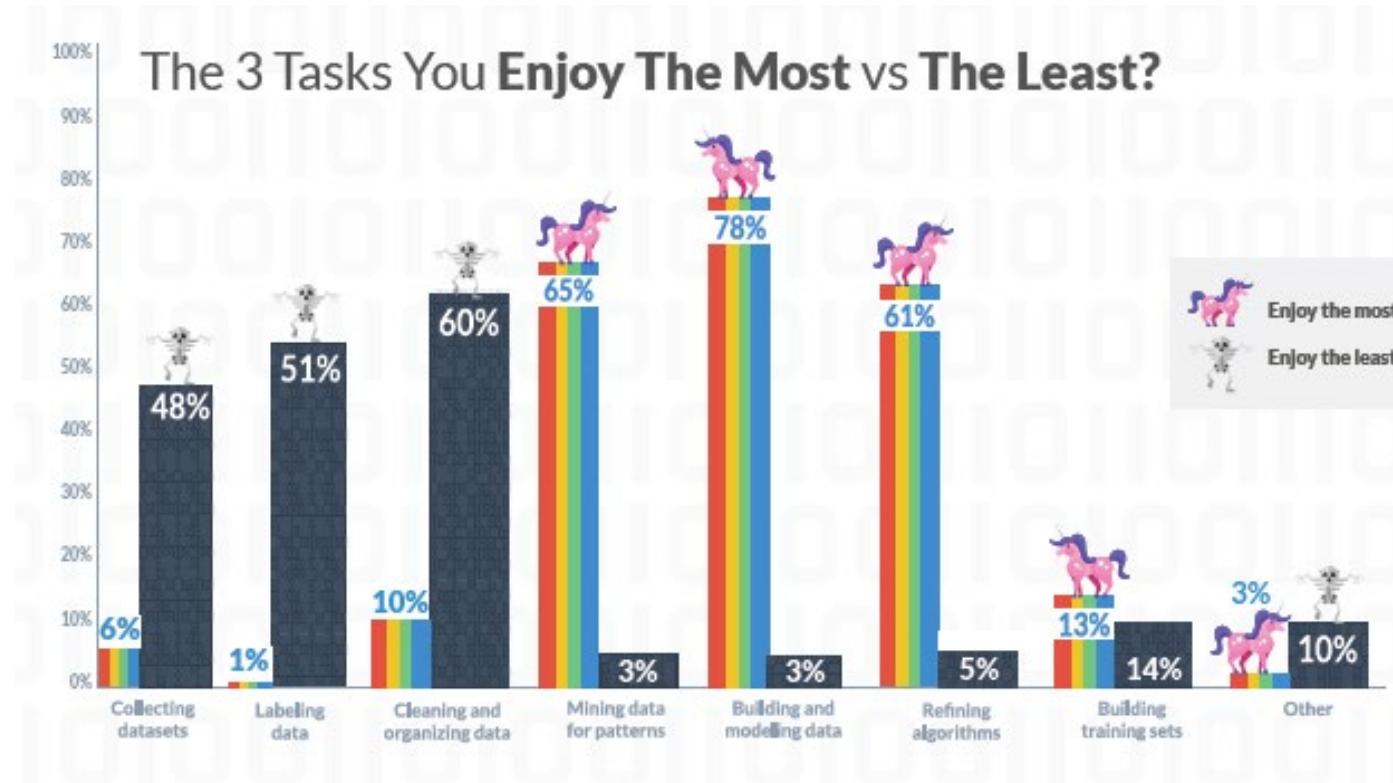


Perfil do Cientista de Dados

- **Opinião das autoras do livro “Doing Data Science”:**
“Uma equipe de ciência de dados funciona melhor quando diferente habilidades (perfis) são representadas em diferentes pessoas, porque ninguém é bom em tudo. Isso nos faz pensar se poderia ser mais vale a pena definir uma “equipe de ciência de dados”, do que definir um cientista de dados”.



Perfil do Cientista de Dados



Data Scientist – O Mercado de Trabalho

inteligência artificial

Oportunidades aumentam para quem estuda ciência de dados

Fonte: <https://www1.folha.uol.com.br/sobretudo/carreiras/2018/06/1971998-oportunidades-aumentam-para-quem-estuda-ciencia-de-dados.shtml>

Por que a ciência de dados está provocando uma corrida de empresas à USP São Carlos?

Elas buscam profissionais e parcerias que possam trazer novas perspectivas para seus negócios



Texto: Denise Casatti / Assessoria de Comunicação do ICMC
Arte: Beatriz Abdalla

25/04/2019

Fonte: <https://jornal.usp.br/universidade/por-que-a-ciencia-de-dados-esta-provocando-uma-corrida-de-empresas-a-usp-sao-carlos/>

USP abre inscrições para mestrado com foco em ciência de dados

Programa é voltado para profissionais com formação em matemática, estatística ou computação; inscrições ficam abertas até 30 de abril

Da Redação
08/04/2019 às 11h30

Fonte: <https://cio.com.br/usp-abre-inscricoes-para-mestrado-com-foco-em-ciencia-de-dados/>

USP oferece novo curso de Ciência de Dados a partir do próximo ano

O Bacharelado em Ciência de Dados tem currículo interdisciplinar e integra conhecimentos da computação, da estatística e da matemática; curso será oferecido no campus de São Carlos

25/06/2020

Fonte: <https://jornal.usp.br/universidade/usp-oferece-novo-curso-de-ciencia-de-dados-a-partir-do-proximo-ano/>

Data Scientist – O Mercado de Trabalho

Área de Dados deve gerar 6,1 milhões de empregos com salários de até R\$ 34 mil

Segundo o Guia Salarial 2020, cientista de Dados podem receber até R\$ 13.100,00 e o analista de BI pode ganhar até R\$ 20.450, no nível de especialista.

Por Lidianne Porto — Last updated 16 mar, 2020

Fonte: <https://escolaeducacao.com.br/area-de-dados-deve-gerar-61-milhoes-de-empregos-com-salarios-de-ate-r-34-mil/>

NOTÍCIAS CARREIRA E PROFISSÃO

OPORTUNIDADE

Carreira de cientista de dados oferece salários de até R\$ 12 mil no Brasil

Especialistas explicam os motivos da maior procura por profissionais, além de habilidades e perfil ideal para quem deseja ingressar na área

EM Estado de Minas

postado em 12/03/2020 14:26 / atualizado em 12/03/2020 14:27

Fonte: https://www.em.com.br/app/noticia/emprego/2020/03/12/interna_emprego,1127924/carreira-de-cientista-de-dados-oferece-salarios-de-ate-r-12-mil.shtml

CARREIRA - VOCÊ S/A

Por que o Nubank sempre busca cientistas de dados e paga até R\$ 25 mil

O Nubank não exige background de programação para contratar. Confira o que a fintech valoriza e como funciona o trabalho

Por Udacity
© 30 jun 2018, 09h00

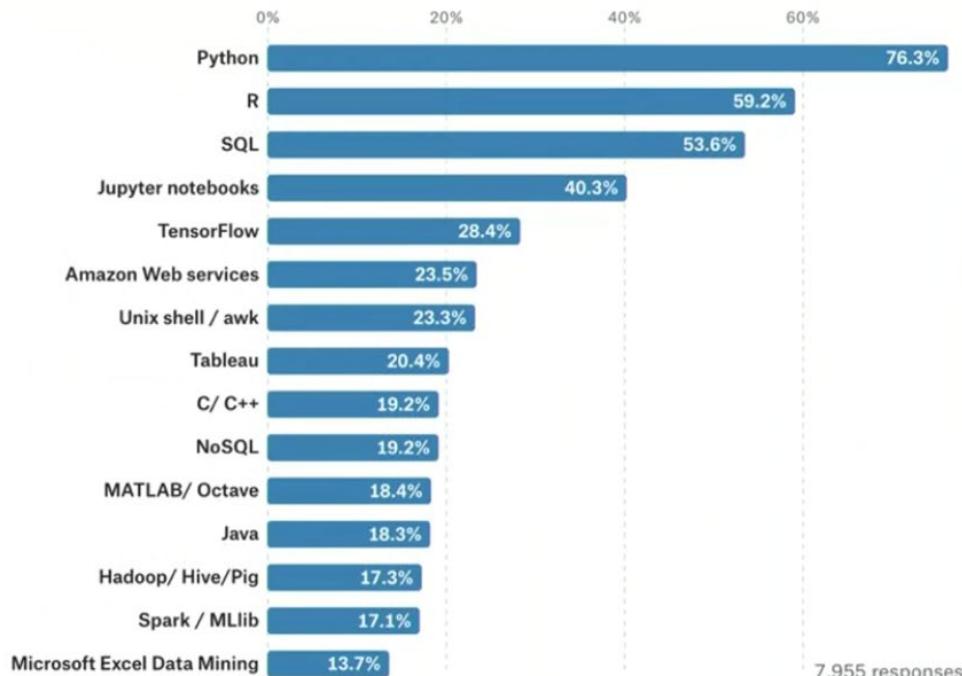
Fonte: <https://exame.com/carreira/por-que-o-nubank-sempre-busca-cientistas-de-dados-e-paga-ate-r-25-mil/>

A Ciéncia de Dados é um processo e se tornar um Cientista de Dados é uma longa jornada.



Fonte da figura: [Becoming a Data Scientist – Curriculum via Metromap](#)

As ferramentas da(o) Cientista de Dados



Source: <https://www.kaggle.com/surveys/2017>

Data Science com Python, por quê?

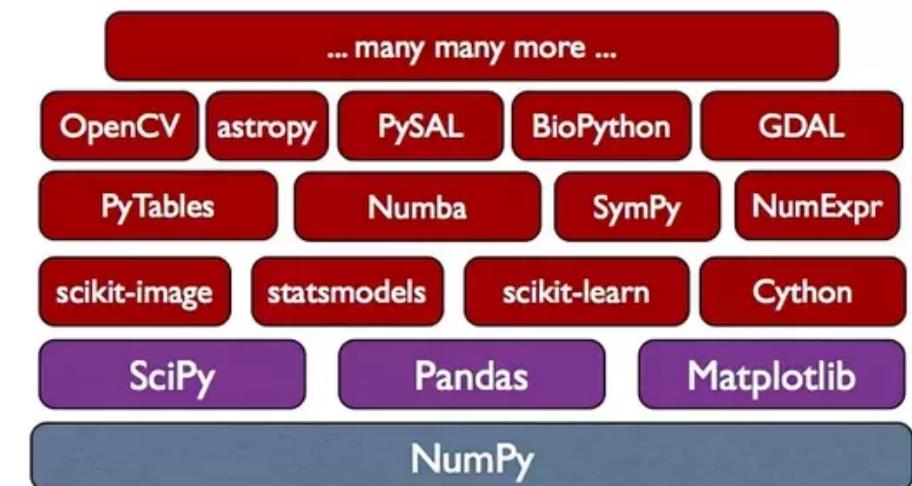
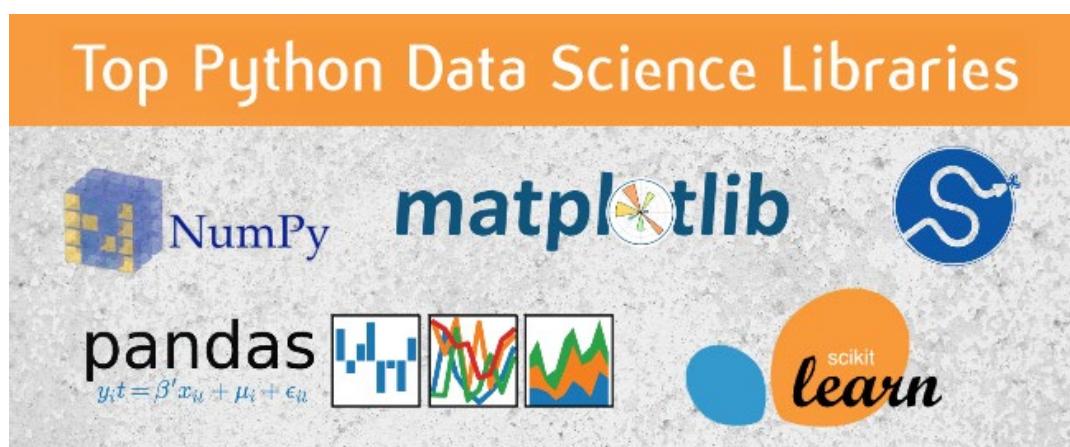
O **Python** tem vários **recursos** que o tornam adequado para aprender (e fazer) ciência de dados:

- É gráatis ☺
- É relativamente simples codificar
- Tem grande intensidade computacional e **poderosas bibliotecas** de análise de dados
- Se integra bem com outras bases de dados e ferramentas usadas pelo Engenheiro de Dados (como o Hadoop e Spark, por exemplo).



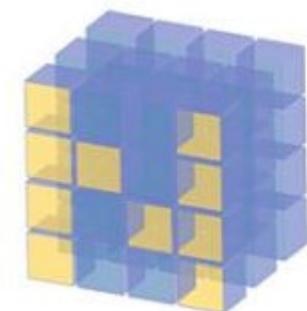
Data Science com Python

Há muitas bibliotecas de dados, frameworks, módulos e toolkits que implementam eficientemente algoritmos e técnicas de Data Science.



Bibliotecas do Python para Data Science

- ▶ **NumPy:** suporte para **Python numérico**. A característica mais poderosa de NumPy é o **array n-dimensional**. Esta biblioteca também contém funções básicas de **álgebra linear, transformações de Fourier**, capacidades avançadas de números aleatórios e ferramentas para integração com outras linguagens de baixo nível, como Fortran, C e C ++.



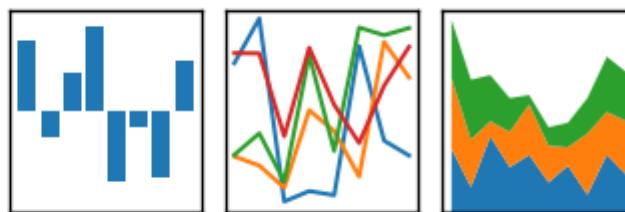
NumPy

Bibliotecas do Python para Data Science

- ▶ **Pandas:** para operação e **manipulação de dados estruturados**. É amplamente utilizado para **preparação de dados**. A biblioteca Pandas foi adicionada há relativamente pouco tempo no Python e têm sido fundamentais para impulsionar o uso do Python na comunidade de cientistas de dados.

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



Bibliotecas do Python para Data Science

- ▶ **SciPy:** suporte para **Scientific Python**. SciPy é construída sobre NumPy e é uma das bibliotecas mais úteis para uma variedade de ciências de alto nível e engenharia como transformação de Fourier discreta, Álgebra Linear e otimização e matrizes esparsas.



Bibliotecas do Python para Data Science

- ▶ **Matplotlib**: para traçar **grande variedade de gráficos**, desde histogramas até gráficos de calor.
- ▶ **Scikit Learn**: para a **Machine Learning ou Aprendizagem de Máquina**. Construído sobre NumPy, SciPy e matplotlib, esta biblioteca contém uma grande quantidade de ferramentas eficientes para aprendizado de máquina e **modelagem estatística**, incluindo classificação, regressão, clustering e redução de dimensionalidade.



Ferramentas do Python para Data Science

► O que é Anaconda?

A Anaconda é uma **distribuição Python** (ou R) que possui uma série de **ferramentas para Ciência de Dados**, Análise Preditiva, Computação Científica e Machine Learning. Inclui o núcleo da linguagem Python (no caso da versão Python da distribuição) e, ainda, mais de **100 bibliotecas Python**, um editor de código chamado **Spyder**, o **Jupyter notebook** e o **Conda** que é um gerenciador de pacotes do Anaconda.



Data Science com R, por quê?

- É uma linguagem de programação para análise de dados
- É open source
- Possui uma comunidade ativa de desenvolvedores
- É flexível, permitindo desenvolver funções e pacotes para facilitar o trabalho
- Está disponível em diferentes plataformas: Windows, Linux e Mac

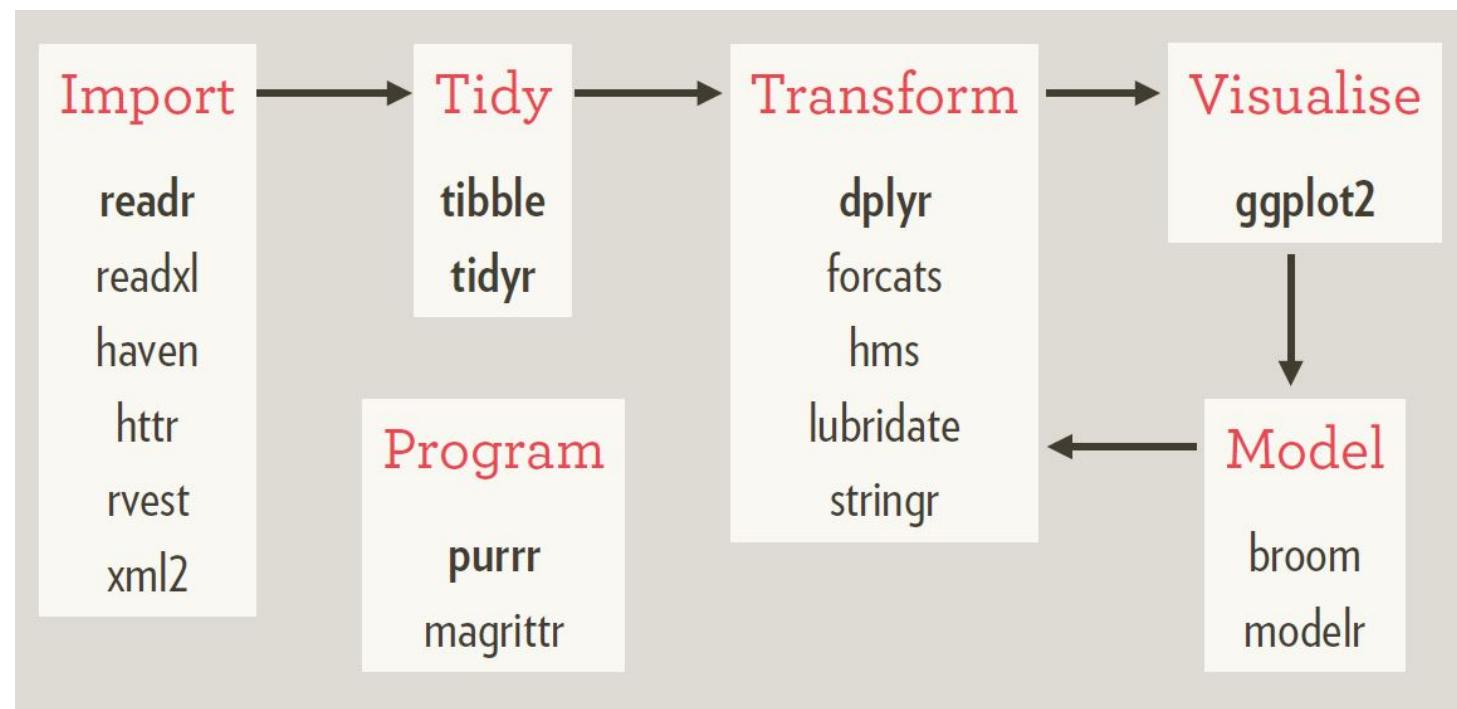


Data Science com R

Tidyverse é uma coleção de pacotes R projetados para a ciência de dados. Todos os pacotes compartilham uma mesma filosofia de desenvolvimento, sintaxe e estruturas de dados.



Data Science com R



Fonte: [Data Science with R by Hadley Wickham, rstudio-conf 2017](#)

O que são Soft Skills?

Soft Skills são **competências comportamentais** e estão relacionadas à forma de se relacionar e interagir com as pessoas:

- Comunicação
- Storytelling
- Criatividade
- Intuição
- Saber entrevistar
- Trabalho em equipe

Por onde começar? Portfólio e Networking

1. Faça um plano de estudos que inclua no mínimo:

- Python ou R (ou os dois!) e as suas bibliotecas para Data Science
- SQL
- Estatística
- Machine Learning
- E não menos importante: inglês!!!!

2. Leia blogs sobre o assunto, **ouça** podcasts, assim você ficará informada(o) das novidades

3. Frequento os meetups e faça parte das comunidades, desta forma você aprenderá com outras pessoas e terá a oportunidade de fazer networking.

4. Faça projetos! Pense em um problema e tente resolvê-lo. Escolha um dataset do mundo real (viva os dados abertos!). Desta maneira poderá criar o seu portfólio.

5. Prepare-se para as entrevistas e os testes técnicos.

Comunidades de Tecnologia

Você já foi em um **meetup**?

Meetups são eventos onde as pessoas se encontram para falar de um tema de interesse em comum. Esses eventos podem ser palestras, workshops, cursos, etc. Muitos deles são gratuitos!

Há varias comunidades “tech” que promovem esses encontros, principalmente na cidade de São Paulo e no ABC. Encontre eventos nessas plataformas: [meetup](#) e [sympla](#)

O Facebook e o Telegram também possuem grupos sobre Data Science, Big Data, Python, R e muito mais.

Comunidades de Tecnologia

Estas são algumas comunidades que valem a pena você acompanhar no [Meetup](#), [Facebook](#), [Telegram](#) ou outras redes sociais:

- **Python:** [Grupy-SP](#), [Grupy-ABC](#), [PyData São Paulo](#), [PyLadies São Paulo](#), [Pt-BR Data Science & Python](#), [Scipy Sampa](#), [Python Brasil Programadores](#)
- **R:** [R Brasil](#), [R-Ladies São Paulo](#), [R Brasil - Programadores](#)
- **Data Science, ML & AI:** [Nubank Data Science & Machine Learning](#), [Data Science in Action Brasil](#), [Women in Data Science](#), [WIMLDS São Paulo](#), [Estatística ^Data Science](#), [Dados Abertos.BR](#), [Brasil Big Data](#), [AI Girls](#), [AI Brasil](#), [Tangerines](#)
- **Dev:** [Developers-BR](#), [Nerdzão](#), [NerdGirlz](#), [Elas Programam](#), [Cloud Girls São Paulo](#), [Garoa Hacker Clube](#), [Blue Talks](#), [GDG-SP](#), [GDG-ABC](#), [Codamos Club](#)

GEDS - Grupo de Estudos de Data Science das



- ▶ O grupo iniciou as atividades em 07/2018.
- ▶ Estudamos juntas tópicos de Estatística e Python. Analisamos datasets para colocar em prática os estudos e assim, obter insights. Recentemente compartilhamos nosso aprendizado em um workshop de Introdução à Estatística e Pandas.
- ▶ **Teleconferências** uma vez por semana.
- ▶ **Encontros presenciais** uma vez por mês aos sábados.
- ▶ “Agora é melhor que nunca” – Zen do Python



Mulheres Brasileiras Cientistas de Dados



► **Gabriela de Queiroz**

Cientista de Dados, Fundadora das R-Ladies, Gerente de Machine Learning, Palestrante, Advocate na IBM (USA)

linkedin.com/in/gabrieladequeiroz

► **Isa Oliveira**

Head de Dados no Grupo SBF, ex-Head de Ciência de Dados na SulAmérica, PhD em Astrofísica

linkedin.com/in/isa-oliveira-phd



Mulheres Brasileiras Cientistas de Dados



► **Angélica Custódio**

Cientista de Dados na Recovery, Estatista, Mestranda em Ciência da Computação, PyLady e R-Lady

linkedin.com/in/angelicacustodio

► **Patricia Novais**

Head de Ciência de Dados na SulAmérica, Física, PhD em Astrofísica, Professora, PyLady

linkedin.com/in/patricianovais



Mulheres Cientistas de Dados

► **Lillian Pierson**

Cientista de Dados, autora do livro Data Science for Dummies

[linkedin.com/in/lillianpierson](https://www.linkedin.com/in/lillianpierson)



► **Rachel Schutt**

Diretora na BlackRock (AI Labs) e autora do livro Doing Data Science

[linkedin.com/in/rachelschutt](https://www.linkedin.com/in/rachelschutt)

Guia de Bolso: “Ciência de Dados para Todos”

► **Estatística Básica**

- Curso Udacity: [Introdução à Estatística Descritiva](#)
- Curso Udacity: [Introdução à Estatística Inferencial](#)
- Livro: Guia Mangá de Estatística ❤️
- Material Workshop PyLadies São Paulo (GEDS): [Introdução à Estatística com Pandas](#)
- Workshop Online PyLadies São Paulo: [Estatística Descritiva com Python](#)
- Blog: [O Estatístico](#)



► **Estatística Avançada e Machine Learning**

- Livro: [An Introduction to Estatistical Learning with Applications in R](#)
- Aulas em vídeo WiMLDS São Paulo: [Capítulos: An Introduction to Estatistical Learning](#)
- Livro: [Aprendizado de máquina: uma abordagem estatística](#)

Guia de Bolso: “Ciência de Dados para Todos”

► Python Básico

- Curso Coursera/USP: [Introdução à Ciência da Computação com Python Parte 1](#)
- Curso Coursera/USP: [Introdução à Ciência da Computação com Python Parte 2](#)
- Curso SoloLearn: [Python](#)
- Curso Kaggle: [Python](#)
- Curso em vídeo: [Python para Zumbis](#)



**KEEP
CALM
and
CODE IN
PYTHON**

Guia de Bolso: “Ciência de Dados para Todos”

► Data Science

- Curso Data Science Academy: [Introdução à Ciência de Dados 2.0](#)
- Podcast: [Pizza de Dados](#)
- Blog: [Cientista de Dados com GIFs](#)

► SQL

- Curso Solo Learn: [SQL Fundamentals](#)
- Curso Kaggle: [Intro to SQL](#)
- Tutorial W3schools: [SQL Tutorial](#)



Guia de Bolso: “Ciência de Dados para Todos”

► Data Science com Python

- Materiais Grupo de Estudos de Data Science (GEDS) PyLadies São Paulo: [Oficinas e Trilha de Estudos \(Estatística e Python\)](#)
- Curso Data Science Academy: [Python Fundamentos para Análise de Dados](#)
- Cursos Kaggle: [Pandas](#), [Intro to Machine Learning](#), [Intermediate Machine Learning](#) e [Data Visualization](#)
- Trilha Cognitive Class (3 cursos): [Applied Data Science with Python](#)
- Disciplina de Python do curso de Engenharia Civil da UFPR(4 cursos):
[Introdução à Computação Científica com Python](#)
- Curso em vídeo LabHacker (para leigos):
[Análise de Dados em Python: Aula 01](#) e [Análise de Dados em Python: Aula 02](#)
- Tutorial: [Para aprender Data Science com Python do zero](#)

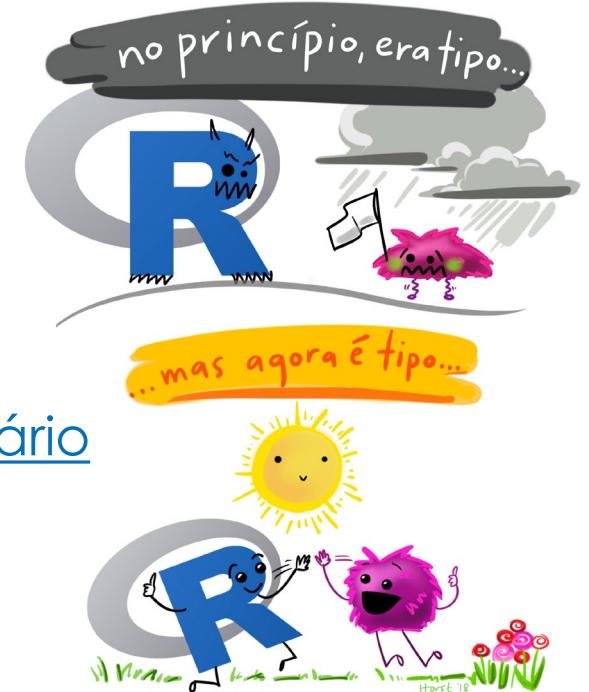


Guia de Bolso: “Ciência de Dados para Todos”

► Linguagem R

- Material Workshops R-Ladies São Paulo: [R Básico e Intermediário](#)
- Workshops Online R-Ladies São Paulo: [R Básico e Intermediário](#)
- Vídeos R-Ladies Belo Horizonte: [R Básico](#)
- Material Workshops R-Ladies GYN (Goiânia): [R Básico e Intermediário](#)
- Livro Curso-R: [Zen do R](#)

Um livro de programação em R
para não-programadores



Guia de Bolso: “Ciência de Dados para Todos”

► Data Science com R

- Livros:
 - IBPAD: [Ciência de Dados com R – Introdução](#)
 - Curso-R: [Ciência de Dados em R](#)
 - [R for Data Science](#)
- Material de Cursos: [Curso-R](#)
- Curso Udacity: [Data Analysis with R](#)
- Site: [Tidyverse](#)
- Blog: [R-Bloggers](#)



Guia de Bolso: “Ciência de Dados para Todos”

- Uma dúvida comum entre quem quer começar a estudar Ciência de Dados é: **o que estudar?**
- Agora **você tem um guia** que pode ser uma referência para o **seu roteiro de estudos**.
- Mas atenção: **não aprendemos por osmose, mágica ou milagre.**
- A **dica** é este método milenar... →



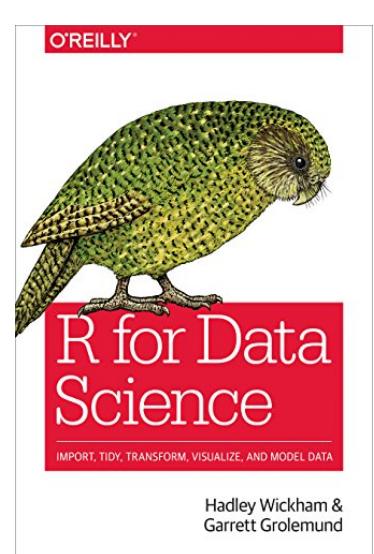
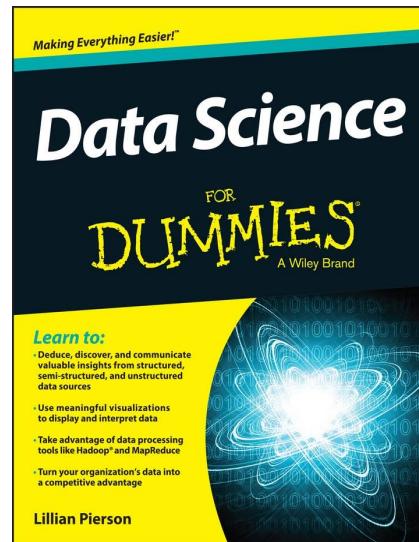
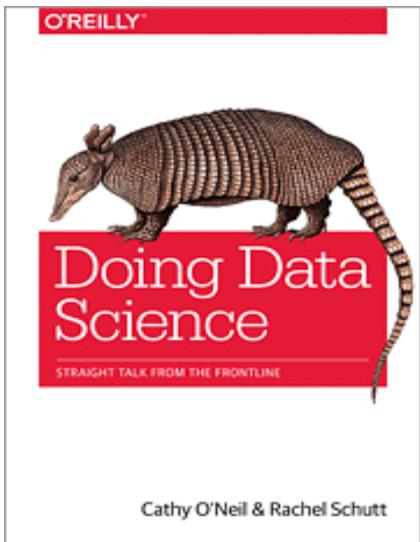
Referências

► Livros

Doing Data Science, autoras: Cathy O'Neil & Rachel Schutt

Data Science For Dummies, autora: Lillian Pierson

R for Data Science, autores: Hadley Wickham (o pai do Tidyverse!) & Garrett Grolemund



Referências

► Sites

- <https://www.vooo.pro/insights/um-tutorial-completo-para-aprender-data-science-com-python-do-zero/>
- <http://felipegalvao.com.br/blog/2016/02/29/manipulacao-de-dados-com-python-pandas/c>
- <https://www.quora.com/What-is-the-relationship-among-NumPy-SciPy-Pandas-and-Scikit-learn-and-when-should-I-use-each-one-of-them>
- <https://dadosedecisoes.com.br/anaconda/>
- https://github.com/rstudio/rstudio-conf/blob/master/2017/The_Tidyverse-Hadley_Wickham/tidyverse.pdf
- <https://medium.com/yottabytes/i-wanna-be-a-data-scientist-but-how-779219ce6311>
- <https://towardsdatascience.com/the-economics-of-getting-hired-as-a-data-scientist-e3882933b43c>

Referências

► Figuras

Data Scientist Report 2017 CrowdFlower

<https://www.kisspng.com/png-data-science-data-analysis-analytics-big-data-data-5415739/>

<https://www.oreilly.com/ideas/data-engineers-vs-data-scientists>

<http://datadriven.tv/blog/modern-data-scientist-infographic/>

<http://www.discoversdk.com/blog/top-python-data-science-libraries>

<https://medium.com/@sunnerli/get-start-with-pandas-822db89705c9>

<https://www.fullstackpython.com/scipy-numpy.html>

<https://softwareengineeringdaily.com/2016/02/01/matplotlib-with-ben-root/>

<http://wasduk.com/old-floppy-disk/>

<https://medium.com/cutshort/how-to-become-a-data-scientist-a-detailed-step-by-step-guide-635b079937e2>

<https://i2.wp.com/thedatascientist.com/wp-content/uploads/2017/02/data-science-wordcloud.jpeg?ssl=1>

<https://towardsdatascience.com/notes-on-artificial-intelligence-ai-machine-learning-ml-and-deep-learning-dl-for-56e51a2071c2>

Just do it!

If you want to go fast,

go alone.

If you want to go far,

go together.

Muito obrigada!

► Maria Marinho

 linkedin.com/in/maria-marinho-ds

 github.com/MaryMS

