Curso de Python: Fundamentos de Data Science I:

PROGRAMA NANODEGREE

TABLEAU

<https://br.udacity.com/course/python-fundamentos-data-science--nd110?gclid=EAIaIQobChMIqcDQ5IGP3gIVjw2RCh0SvwXUEAAYASAAEgLqEfD_BwE>

# Cientistas de Dados

# <https://www.sas.com/pt_br/insights/analytics/cientistas-de-dados.html>

# 

## **Quem são, o que fazem e por que você quer ser um?**

Os cientistas de dados são uma nova geração de especialistas analíticos que possuem as habilidades técnicas para resolver problemas complexos – e a curiosidade para explorar quais problemas precisam ser resolvidos.

Eles são parte matemáticos, parte cientistas da computação e parte observadores de tendências. E, como transitam entre os mundos dos negócios e da TI, eles são muito procurados e bem-remunerados. Quem não gostaria de ser um?  
  
Eles também são um sinal dos tempos. Os cientistas de dados não estavam sob muitos radares há uma década, mas sua popularidade repentina reflete como as empresas pensam hoje sobre [big data](https://www.sas.com/pt_br/insights/big-data/what-is-big-data.html). Essa massa desajeitada de informações não-estruturadas não pode mais ser ignorada e esquecida. É uma mina de ouro virtual que ajuda a aumentar a receita – contanto que exista alguém que a minere e descubra insights de negócios que ninguém pensou em procurar antes.

### **De onde eles vieram?**

Muitos cientistas de dados começaram suas carreiras como estatísticos ou analistas de dados. Mas, à medida que o big data (e as tecnologias de armazenamento e processamento de big data, como o [Hadoop](https://www.sas.com/pt_br/insights/big-data/hadoop.html)) começaram a crescer e evoluir, esses papéis também evoluíram. Os dados deixaram de ser uma reflexão tardia para a TI. Eles são uma informação fundamental que requer análise, curiosidade criativa e habilidade para traduzir ideias de alta tecnologia em novas maneiras de gerar lucro.

O papel do cientista de dados também tem origens acadêmicas. Alguns anos atrás, as universidades começaram a perceber que os empregadores queriam programadores que soubessem trabalhar em equipe. Então, os professores ajustaram suas aulas para acomodar isso – e alguns programas, como o Institute for Advanced Analytics da North Carolina State University, se prepararam para produzir a próxima geração de cientistas de dados. Hoje, existem mais de 60 programas semelhantes em universidades dos EUA.

“Meus dias podem ser muito parecidos, mas o trabalho pode variar com o tempo. Por algumas semanas eu poderia estar trabalhando em um projeto de mineração de texto, e depois disso eu poderia estar criando um modelo preditivo sobre o cliente. Entre essas atividades, estão as reuniões sobre analytics e como isso pode ajudar diferentes partes do negócio”.

Funções típicas dos cientistas de dados

Não há uma descrição definitiva quando se trata do trabalho de um cientista de dados. Mas aqui estão algumas tarefas que você provavelmente realizará:

Coletar grandes quantidades de dados não-estruturados e transformá-los em um formato mais utilizável;

Resolver problemas de negócios usando técnicas orientadas por dados;

Trabalhar com uma variedade de linguagens de programação, incluindo SAS, R e Python;

Ter uma sólida compreensão de estatística, incluindo testes e distribuições;

Estar sempre atualizado sobre técnicas analíticas, como machine learning, deep learning e análise de texto;

Comunicar e colaborar tanto com TI quanto com a gerência;

Procurar por ordem e padrões nos dados, além de identificar tendências que podem ajudar no resultado financeiro de uma empresa.

O que existe no cinto de utilidades de um cientista de dados?

Esses termos e tecnologias são comumente usados por cientistas de dados:

Visualização de dados: a apresentação de dados em um formato imagético ou gráfico, que possam ser facilmente analisados;

Machine learning: um ramo da inteligência artificial baseado em algoritmos matemáticos e automação;

Deep learning: uma área de pesquisa em machine learning que usa dados para modelar abstrações complexas;

Reconhecimento de padrões: uma tecnologia que reconhece padrões em dados (geralmente de maneira intercambiável com machine learning);

Data preparation: o processo de converter dados brutos em outro formato, para que eles possam ser consumidos mais facilmente;

Análise de texto: o processo de examinar dados não-estruturados para obter insights importantes sobre o negócio.

“Em um dia típico, faço um brainstorming e resolvo problemas ao responder perguntas com a minha equipe, reviso análises e recomendações concluídas por ela e participo de várias reuniões".

Kristin Carney

Cientista de Dados, World’s Foremost Bank

Read the story

abstract background

Como você pode se tornar um cientista de dados?

Se preparar para uma carreira em ciência de dados pode ser uma jogada inteligente. Você terá muitas oportunidades de emprego, além de trabalhar no campo da tecnologia com espaço para experimentação e criatividade. Então, qual é a sua estratégia?

Se você é um estudante

Escolher uma universidade que ofereça uma graduação em ciência de dados – ou ao menos um curso livre em ciência de dados e analytics – é um primeiro passo importante. A IESB, a FIA, o Centro Universitário Metodista Izabela Hendrix, a PUC Minas, a FGV e a USP são exemplos de universidades com aulas de ciência de dados.

Se você é um profissional que quer mudar de carreira

Embora a maioria dos cientistas de dados tenha experiência como estatísticos ou analistas de dados, outros vêm de campos não-técnicos, como administração ou economia. Como podem profissionais de origens tão diversas acabarem no mesmo campo? É importante analisar o que eles têm em comum: um jeito de resolver problemas, a capacidade de se comunicar bem e uma curiosidade insaciável sobre como as coisas funcionam. Saiba como o SAS Academy for Data Science fornece as ferramentas para se tornar um cientista de dados certificado.

Além dessas qualidades, você também precisará de um conhecimento sólido em:

Estatística e machine learning;

Linguagens de codificação como SAS, R ou Python;

Bancos de dados como MySQL e Postgres;

Tecnologias de visualização de dados e criação de relatórios;

Hadoop e MapReduce.

Se você não quiser aprender isso tudo por si só, faça um curso on-line ou inscreva-se em um bootcamp. Então, fale com mais pessoas. Conecte-se a outros cientistas de dados da sua empresa ou encontre uma comunidade on-line. Eles fornecem informações privilegiadas sobre o que os cientistas de dados fazem – e onde você encontrará os melhores empregos.

Quando uma empresa está pronta para contratar um cientista de dados?

Antes de aceitar uma posição de cientista de dados, há algumas coisas sobre a organização que você deve avaliar:

Ela lida com grandes quantidades de dados e tem problemas complexos que precisam de solução? As organizações que realmente precisam de cientistas de dados têm duas coisas em comum: elas gerenciam grandes quantidades de dados e enfrentam problemas sérios no dia-a-dia. Elas normalmente são de indústrias como financeira, governo e farmacêutica;

Ela valoriza seus dados? A cultura de uma empresa impacta a decisão sobre se ela deve ou não contratar um cientista de dados. Existe nela um ambiente que favoreça análises? Há buy-in executivo? Se não, investir em um cientista de dados seria dinheiro indo pelo ralo;

Ela está pronta para se transformar? Como cientista de dados, você espera ser levado a sério e parte disso implica em ver seu trabalho se concretizar. Você dedica seu tempo em descobrir maneiras pelas quais sua empresa pode funcionar melhor. Em resposta, uma empresa precisa estar pronta – e disposta – a seguir com os resultados de suas descobertas.

Contratar um cientista de dados para orientar as decisões de negócios com base em dados pode ser um salto de fé para algumas organizações. Certifique-se de que o negócio para o qual você está trabalhando tenha a mentalidade certa – e esteja pronto para fazer algumas mudanças.

“Eu trabalho para uma empresa ágil, o que requer que eu seja flexível e me adapte às circunstâncias. Na semana passada, por exemplo, eu estava fazendo várias tarefas, como melhorando as pontuações de recomendação, ajustando a integração com o sistema de gerenciamento de conteúdo operacional, criando novas variáveis transformadas com base no comportamento do consumidor a serem usadas para modelos de afinidade e fazendo algumas refatorações de relatórios de desempenho e de painéis analíticos existentes".

Manuel-David Garcia

Cientista de Dados para uma empresa de médio porte em Heidelberg, Alemanha

Leia a história

Tecnologias para cientistas de dados

High-Performance Analytics High-Performance Analytics

Take full advantage of powerful SAS Analytics to get fast, accurate insights from all your data.

Contextual Analysis SAS® Contextual Analysis

Combine machine learning with subject-matter expertise so you can make sense of all your text data.

STAT SAS/STAT®

Take advantage of extensive statistical capabilities to meet the data analysis needs of your entire organization.

Data Loader Hadoop SAS® Data Loader for Hadoop

Manage big data on your own terms – and avoid burdening IT – with self-service data integration and data quality.

Enterprise Miner SAS® Enterprise Miner™

Streamline the data mining process to create highly accurate predictive and descriptive models based on large volumes of data.

Icon Visual Statistics SAS® Visual Statistics

Create and modify predictive models faster than ever using a visual interface and in-memory processing.

MBA ANALYTICS EM BIG DATA

O LABDATA da FIA é pioneiro no Brasil no lançamento de cursos de Big Data e Data Science. Este curso foi desenvolvimento para apresentar, de forma prática, as principais tecnologias para coleta e processamento de dados e diversas técnicas de modelagem com foco em Analytics e Inteligência Artificial recomendadas para um Data Science. Apesenta também um módulo executivo com disciplinas de Gestão. O programa contempla aplicações em Hadoop, Spark, Cassandra, Cloud Computing, Machine Learning, Inteligência Artificial e muito mais. As aulas são realizadas em laboratório com ambiente Cloudera e ferramentas como SAS, R, IBM WATSON, SPSS, SPARK e PYTHON.