

**A ciência dos dados é um esporte coletivo.  
Você tem as habilidades necessárias  
para atuar em equipe?**

A large, stylized IBM logo consisting of the letters 'I', 'B', and 'M' in a bold, sans-serif font. The letters are filled with a pattern of dark blue and light blue squares, creating a checkerboard effect. The 'I' is dark blue on top and light blue on the bottom. The 'B' is light blue on top and dark blue on the bottom. The 'M' is dark blue on top and light blue on the bottom.

**IBM**

1

---

**Introdução**

2

---

**O cientista de dados**

3

---

**O engenheiro de dados**

4

---

**O desenvolvedor**

5

---

**O analista de negócios**

6

---

**Equipes de ciência de dados: os novos agentes da mudança**

7

---

**Recursos**

---

## Introdução

O maior desafio da revolução big data é entender todas as informações geradas pela vasta economia digital atual. É interessante que uma organização colete cada dado que ela tem o poder de alcançar, mas como ela extrai valor desse imenso volume de informações?

Quanto mais dados você tiver, melhor a qualidade dos seus relatórios e recomendações estratégicas, certo? Claro... se você for capaz de analisar esses dados de forma inteligente e rápida e conseguir agir a partir deles com insights de valor. Do contrário, mais dados podem significar mais problemas: dados desorganizados, problemas de armazenamento, riscos à segurança,



equipes corporativas frustradas e equipe de TI sobrecarregada.

**Então, como descobrir o que você quer de fato dos seus dados – e quais dados podem lhe dar essas respostas?**

Como transformar métricas e registros brutos em informações que tenham (e possam gerar) valor real para os negócios? Você tem as habilidades de que sua empresa precisa para gerar resultados?

*Organizações de sucesso montam equipes de ciência dos dados que incorporam diferentes conjuntos de habilidades e responsabilidades, em vez de depender de alguns indivíduos de elite.*

Na prática, várias pessoas trabalham em uma equipe para desenvolver produtos de dados. Suas análises terão a mesma qualidade da equipe responsável por coletar, desenvolver e analisar os dados da base.

**Qual é a contribuição da equipe de ciência dos dados?** O pessoal de programação mais profunda, como engenheiros de dados e desenvolvedores de front-end, contribuem com a consumibilidade para a ciência dos dados. Os matemáticos e cientistas de dados usam algoritmos estatísticos para encontrar padrões nos

dados. Ao longo do processo, tudo precisa ser alinhado com o resultado comercial pretendido, que é direcionado pelos olhos do analista de negócios. Trabalhando em conjunto, a equipe de ciência dos dados pode vencer os desafios e problemas de hoje para criar novas oportunidades e possibilidades para amanhã.

Que talentos e habilidades definem os integrantes de uma equipe de ciência dos dados, e como eles se complementam?

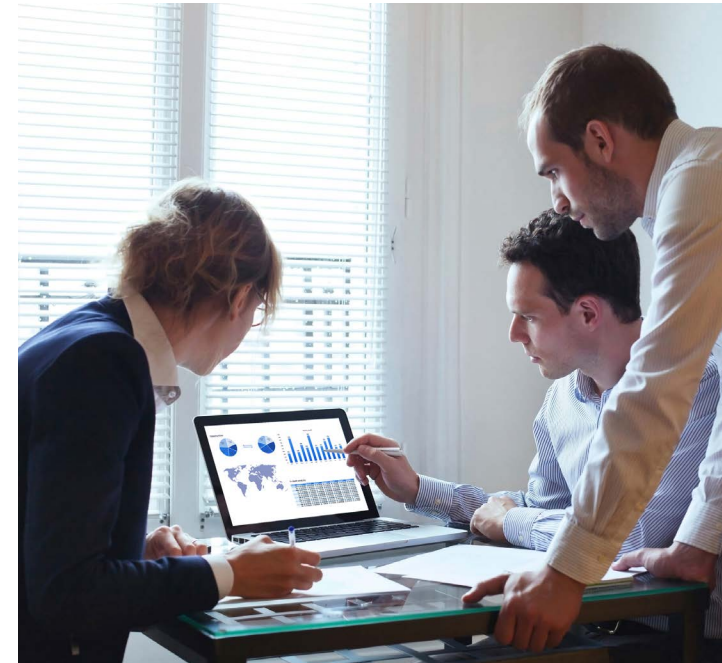
**Leia para saber.**

## O cientista de dados

Solucionar problemas e responder a perguntas usando análises de dados está se tornando rapidamente a norma no mundo atual, movido pelos dados. À medida que experimentos no mundo real se tornam onipresentes nos negócios modernos, o cientista de dados está evoluindo para a função que atiga, ajusta e alimenta esse motor operacional.

*Um bom cientista de dados não apenas aborda problemas de negócios; ele se foca nos problemas que têm o maior valor para a organização.*

Os cientistas de dados costumam ser chamados de “unicórnios”, porque têm uma rara combinação de talentos: eles cuidam de diversas responsabilidades e conjuntos de habilidades, abrangendo matemática, estatística, expertise em domínios, comunicações e outros. Basicamente, o trabalho dos cientistas de dados é procurar padrões ocultos. Eles fazem isso aplicando técnicas de análise avançadas, como, por exemplo, aprendizado automatizado, modelagem, estatística e visualização. Com frequência, cientistas de dados constroem modelos para prever resultados ou descobrir padrões de base; o objetivo deles é gerar insights que possam ser postos em prática e usados para melhorar resultados futuros.



*Como os cientistas de dados estão envolvidos em cada etapa da jornada do desenvolvimento de produtos de dados, eles tendem a trazer uma visão holística para a solução de problemas com dados. No entanto, eles não podem ser especialistas em tudo; é aqui que a equipe deles pode ajudar.*

Eles experimentam continuamente implementando novos modelos preditivos, regras de negócios e lógica de orquestração nos aplicativos impulsionados pela segunda melhor ação possível.

Os cientistas de dados habilidosos exploram e examinam dados de diversas fontes separadas. Eles se debruçam sobre todos os dados que chegam com o objetivo de conectar novas informações a dados históricos para encontrar uma relação ou tendência que ofereça uma vantagem

competitiva crucial ou aborde um problema de negócios urgente. Eles não só coletam e elaboram relatórios sobre dados; eles os analisam de diversos ângulos, determinam o que significam e, depois, recomendam maneiras de aplicar as conclusões. Eles precisam garantir que suas consultas estejam corretas e ser capazes de sustentar suas conclusões com modelos consistentes e dados confiáveis, pois, com frequência, espera-se que o cientista de dados apresente recomendações às equipes gerenciais e executivas.





## Principais habilidades dos cientistas de dados

Cientistas de dados se distinguem por seu forte conhecimento dos negócios, além da capacidade de informar as conclusões tanto aos líderes corporativos quanto aos de TI de forma que possa influenciar a maneira como uma organização aborda um desafio comercial. Com frequência, o cientista de dados se torna o ponto de contato entre o departamento de TI e executivos de nível C. Portanto, ele precisa ser capaz de falar ambas as “línguas” e entender a hierarquia dos dados; não pode ser simplesmente um especialista em dados. Isso também significa que cientistas de dados precisam ter um entendimento consistente dos negócios, além da convicção para defender suas conclusões frente à oposição.

Cientistas de dados são inquisitivos e curiosos: exploram, perguntam, fazem análises de possibilidades e questionam princípios e processos existentes. Entre as habilidades técnicas de um cientista de dados, costumam estar diversas linguagens de programação, familiaridade com gerenciamento de big data e ferramentas de análise, como Apache Hadoop e Spark, e experiência com ferramentas que o ajudem a visualizar dados e insights.

## O engenheiro de dados

O cientista de dados pode ser responsável por descobrir padrões ocultos nos dados, mas de onde você acha que ele recebe os dados, e o que acontece quando as soluções dele precisam ser dimensionadas para milhares de usuários ou lidar com informações confidenciais? Há um limite para até onde os cientistas de dados podem ir sem o devido apoio para operacionalizar seu trabalho.

### É aí que entra o engenheiro de dados.

Em um alto nível, engenheiros de dados ajudam a coletar, organizar e limpar os dados que os cientistas de dados acabarão por usar para desenvolver sua análise. Se os cientistas de dados extraem valor de dados, engenheiros de dados garantem que os dados fluam sem problemas da origem ao destino, para que possam ser processados.

Engenheiros de dados são responsáveis por preparar sistemas e processos que outros trabalhadores de dados, como cientistas de dados, usam e do qual dependem para trabalhar com dados. Engenheiros de dados precisam entender como aprimorar o fluxo de dados para minimizar latências de movimentação e dar agilidade às análises. Além disso, eles trabalham com desenvolvedores de front-end ao colocar projetos de ciência dos dados em produção.

Em muitas organizações, um engenheiro de dados será encarregado de integrar dados, inclusive projetar, desenvolver e mensurar as vias de ingestão de dados e integração para grandes volumes de dados temporais vindos de diferentes fontes. Entre os exemplos estão extrações de banco de dados,

*Engenheiros de dados costumam ser incumbidos da tarefa de preparar a fundação para que um analista de dados ou cientista de dados recupere facilmente os dados necessários para suas avaliações e experimentos.*

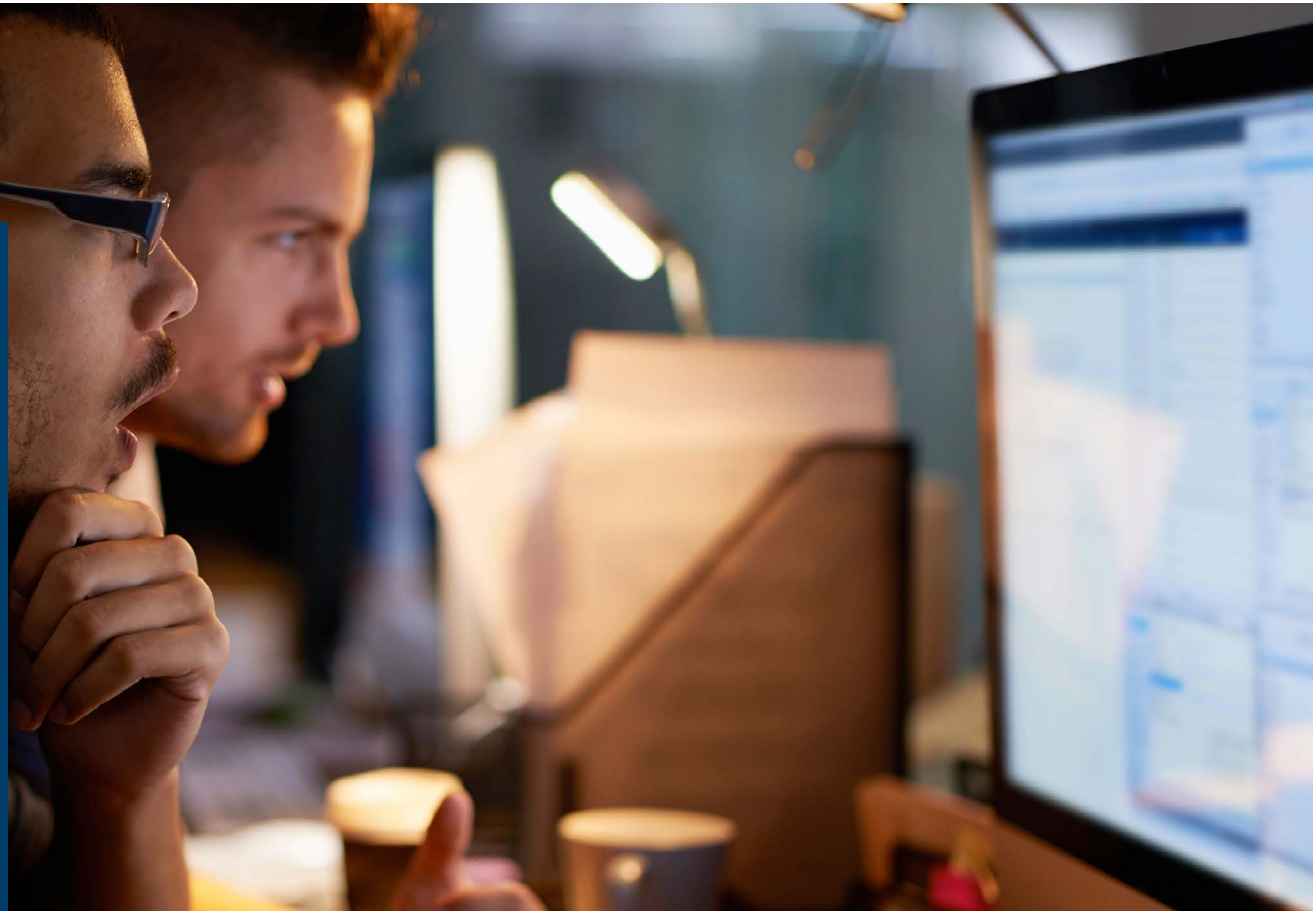
registros de servidor do aplicativo, imagens digitalizadas, gravações de voz, streams de Twitter, sites e dados de sensores de integridade. Quando vias contínuas estiverem instaladas, indo e vindo desses imensos “pools” de informações filtradas, os cientistas de dados podem retirar conjuntos de dados relevantes para suas análises.



## Principais habilidades dos engenheiros de dados

Engenheiros de dados são engenheiros de alto nível que entendem o funcionamento interno de um software de banco de dados. Eles compilam e instalam sistemas de bancos de dados, compõem consultas complexas, dimensionam-nas para várias máquinas, gerenciam backups e implementam sistemas de recuperação de desastre. Eles desenvolvem, constroem, testam e mantêm arquiteturas como banco de dados e sistemas de processamento de dados em larga escala.

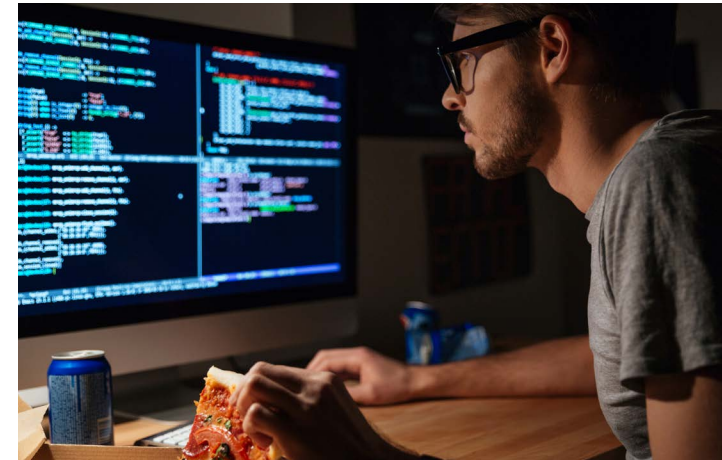
Bons engenheiros de dados estão sempre aprendendo e pensando em quais novas tecnologias os ajudarão a gerar negócios. Isso os leva a desenvolver um profundo pano de fundo de programação e também a cultivar familiaridade com tecnologias baseadas em Hadoop, como MapReduce, Hive e Pig. Engenheiros de dados costumam ter experiência significativa com tecnologias baseadas em SQL e tecnologias NoSQL, além de metodologias e soluções de data warehousing como extrair, transformar, carregar (ETL).



## O desenvolvedor

Embora desenvolvedores nem sempre tenham uma posição totalmente dedicada dentro de uma equipe de ciência dos dados, eles são uma peça essencial do quebra-cabeça. Um projeto de sucesso costuma transformar o trabalho da ciência dos dados em produto, para que ele atenda a um público interno ou a clientes externos. É transformando ciência dos dados em processos de negócios que as empresas desenvolvem uma vantagem competitiva usando os dados.

Com frequência, os desenvolvedores entram no fim do fluxo de trabalho da ciência dos dados. Eles são responsáveis por desenvolver os aplicativos nos quais ficarão os modelos, aplicativos que utilizam os insights e os dados coletados do resto da equipe de ciência dos dados. Isso exige uma quantidade considerável de tempo de programação e pode ser um trabalho desafiador. A falta de tecnologias integradas pode prejudicar o progresso dos desenvolvedores, dificultando a integração da análise da equipe de ciência dos dados.



*Desenvolvedores trazem consumibilidade para a ciência dos dados, permitindo que públicos internos e externos aproveitem os dados e as análises em seu trabalho cotidiano.*

Outra exigência: os desenvolvedores precisam ter um conjunto completo de habilidades de programação a postos. Isso inclui expertise no desenvolvimento de serviços web, habilidades de desenvolvimento de front-end e um forte conhecimento de funcionalidade e recursos de interface de usuário. Além disso, eles devem ter familiaridade com interfaces de programação de aplicativos (APIs) e usá-las para integrar diversos produtos e fontes de dados a aplicativos. Quando feito da maneira correta, a habilidade do desenvolvedor cria um aplicativo tão consistente que você sequer perceber que existe um produto de dados por trás de tudo.

Para desenvolver aplicativos móveis, web e corporativos que mudam todo o panorama e balançam mercados, os desenvolvedores precisam dos dados – e das ferramentas – para transformar sua visão em realidade. O desenvolvedor precisa colaborar com o cientista de dados, o engenheiro de dados e o analista de negócios para garantir o alinhamento entre os objetivos de negócios e a parte de análises.

## O analista de negócios

O analista de negócios, por vezes, chamado de analista de cidadãos, fornece expertise e orientação no nível dos negócios às equipes de ciência dos dados. Sua responsabilidade é aplicar conhecimento do domínio e tomar decisões informadas por dados.

O analista de negócios entende do que a empresa precisa, mas não tem o conhecimento técnico para desenvolver uma análise detalhada. Melhorias na tecnologia e uma compreensão mais profunda do perfil

do usuário os transformaram em “consumidores analíticos”, permitindo que tomem decisões baseadas em insights de dados. Tecnologias flexíveis e de fácil utilização permitem que eles desenvolvam análises de nível de negócios sem codificar e sem precisar pedir ao cientista de dados. Isso dá aos analistas de dados mais independência, aumentando a chance de que usem as conclusões da equipe de ciência dos dados em suas análises cotidianas.



## Equipes de ciência dos dados: os novos agentes da mudança



Habilidosos cientistas de dados, engenheiros de dados, desenvolvedores e analistas de negócios são figuras de transformação nos negócios modernos. São o coração pulsante da economia de big data. Eles não apenas projetam novos sistemas; serão responsáveis por novas fontes de dados e novas maneiras de usar esses dados. Claro, a TI ainda precisará desenvolver o sistema, mas serão os profissionais de ciência dos dados que ajudarão os departamentos a colaborar para solucionar problemas e acelerar a inovação.

Os melhores produtos de dados são aqueles que o usuário final sequer nota. Isso exige um cuidadoso pensamento sobre a interface, os dados que são capturados e a automação permitida pelo back-end de análises. A tecnologia para coletar e analisar volumes de dados de negócios gigantescos já está disponível, e você pode explorá-la para beneficiar sua empresa. Você está pronto para atuar em equipe no jogo de alto risco da ciência dos dados?



## A ciência dos dados é um esporte coletivo. Você tem as habilidades necessárias para atuar em equipe?

### Treinando os profissionais de ciência de dados do futuro

Para continuar preparado e capacitado para ser um bom integrante da equipe, você precisa melhorar seus atributos desenvolvendo conhecimento e aprimorando suas habilidades. A IBM é uma dentre muitas empresas a ajudar os profissionais de ciência de dados a ficarem melhores e mais rápidos em seu trabalho.

Todo ramo de crescimento rápido precisa de profissionais com novas habilidades e expertise. A IBM tem sido uma apoiadora

ativa da comunidade de ciência de dados e planeja continuar dando esse apoio em fóruns educativos online e presenciais. Por meio de eventos, encontros, cursos, conteúdo, contribuições com a comunidade de código aberto e outros, a IBM apoia os profissionais de dados de hoje e do futuro próximo, ajudando-os a se preparar para atender às demandas de alto volume e alta velocidade de dados do futuro.





# Recursos

Pronto para aprimorar suas habilidades de ciência de dados? Aqui estão alguns recursos para começar:

1. Desenvolva suas habilidades de ciência de dados com a [Big Data University](#)
2. Comece a usar as ferramentas de que você precisa com o [IBM Data Science Experience](#)



---

© Copyright IBM Corporation 2016

IBM Analytics  
Route 100  
Somers, NY 10589

Produzido nos Estados Unidos da América  
Junho de 2016

IBM, o logotipo IBM e **ibm.com** são marcas comerciais da International Business Machines Corp., registradas em muitas jurisdições em todo o mundo. Outros nomes de produtos e serviços podem ser marcas comerciais da IBM ou de outras empresas. Uma lista atual das marcas comerciais da IBM está disponível na Internet sob o título "Copyright and trademark information" em [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

Este documento encontra-se atualizado na data inicial de sua publicação e pode ser alterado pela IBM a qualquer momento. Nem todas as ofertas estão disponíveis em todos os países em que a IBM opera.

AS INFORMAÇÕES CONTIDAS NESTE DOCUMENTO SÃO FORNECIDAS "NA FORMA EM QUE SE ENCONTRAM" SEM QUALQUER GARANTIA, EXPRESSA OU IMPLÍCITA, INCLUINDO NENHUMA GARANTIA DE COMERCIALIZAÇÃO, ADEQUAÇÃO A UMA DETERMINADA FINALIDADE E NENHUMA GARANTIA OU CONDIÇÃO DE NÃO-VIOLAÇÃO. Os produtos da IBM são garantidos de acordo com os termos e condições dos acordos sob os quais eles são fornecidos.



Recycle