



VINICIUS DUZAC CERUTTI · 3D AGO · 108 VIEWS



Edit My Copy

4



Cientista ou Analista de Dados? Qual é a diferença?

Python · [State of Data Brazil 2021](#), [Google Data Trend Brazil 2004-2022](#)

Notebook Data Logs Comments (5)

Run

26.8s

Version 4 of 4

Data Visualization Business Exploratory Data Analysis Data Analytics Jobs and Career +2

Cientista ou Analista de Dados? Qual é a diferença?



Por: [Vinicius Duzac Cerutti](#)



Sumário:

- [Introdução](#)
- [Reponsabilidade de cada Cargo](#)
- [Técnicas e Métodos](#)
- [Ferramentas e Serviços](#)
- [Nível de Ensino](#)
- [Faixa Salarial](#)
- [Conclusão](#)
- [Referências](#)



Fonte: [Tumblr](#)

Tradução: este é o sinal que você estava procurando

Show hidden code

Introdução

Com a popularidade da ciência de dados, várias pessoas estão buscando conhecimentos nesta área. Contudo, há uma quantidade enorme de informações disponibilizadas na internet referente a este assunto e isso pode gerar uma confusão desde termos técnicos até definição de papéis da área. Esta análise, tem como objetivo de mostrar a diferença entre uma posição de cientista e analista de dados utilizando dados da

pesquisa do [state of data 2021](#).

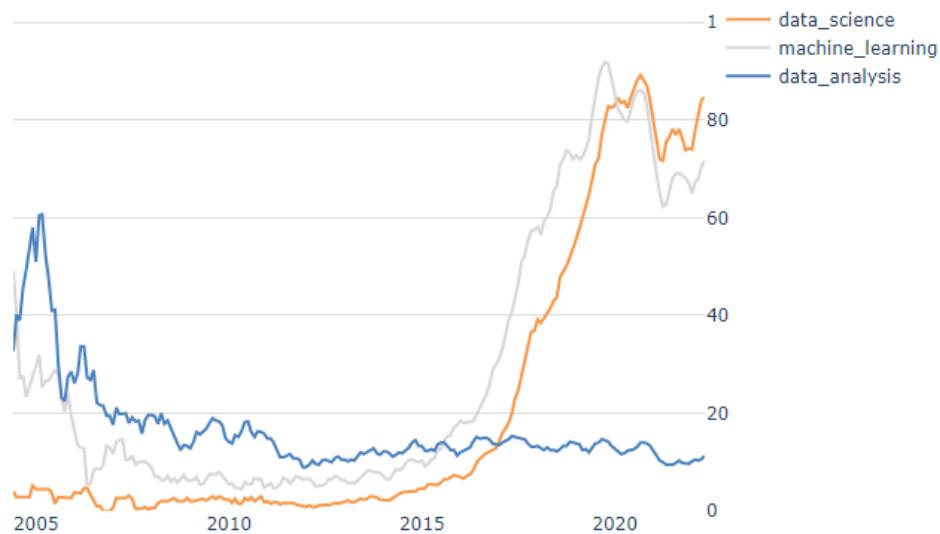
Aviso! Foi desconsiderado respondentes que se consideram desempregados, apenas estudantes, e que preferiram não informar a situação atual de trabalho.

▼ Show hidden code

Evolução de buscas por termos de áreas de Dados no Brasil

Aumento de procura dos termos machine learning e data science a partir do ano de 2015.

Fonte: [Google Trends](#)



▼ Show hidden code

Responsabilidade de cada Cargo

Analista de Dados:

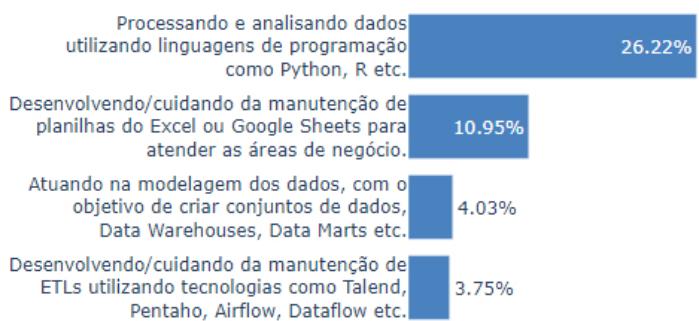
Um analista de dados tem o objetivo a ajudar a tomar decisões através de dados, respondendo perguntas da área de negócio e construindo indicadores para acompanhar a evolução. Para isso, realiza o processo de modelagem de dados em data warehouses até a construção de dashboards.

▼ Show hidden code

Tarefas que um analista de dados gasta a maior parte do seu tempo no trabalho

Realizando construções de dashboards em ferramentas de BI como PowerBI, Tableau, Looker, Olik etc.

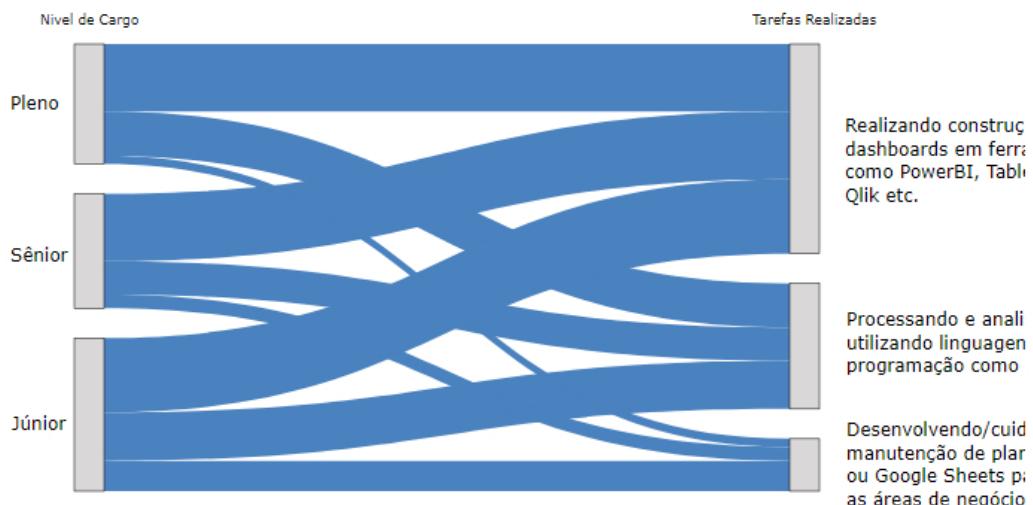
43.8%



Nota-se que poucos analistas de dados pleno e sênior realizam o desenvolvimento e manutenção de planilhas para atender as áreas de negócio. Além disso, a distribuição de tarefa pleno e sênior são parecidas, no entanto cada um tem suas responsabilidades específicas.

▼ Show hidden code

Tarefas que um analista de dados gasta a maior parte do seu tempo no trabalho por nível de cargo

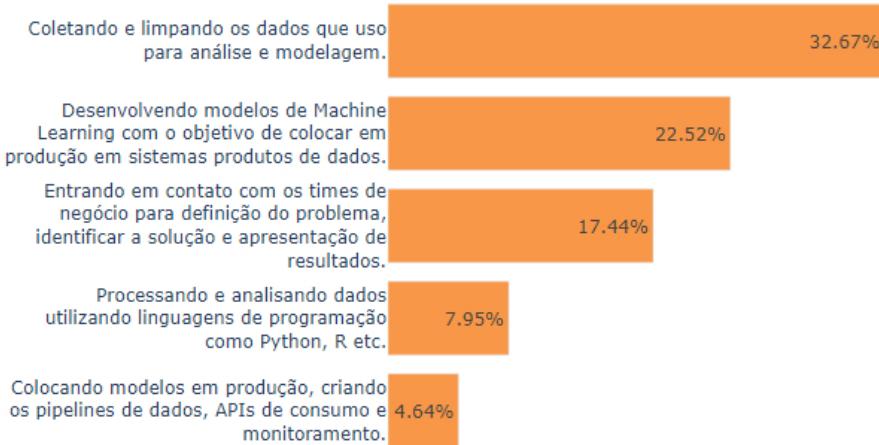


Cientista de Dados:

Tem como principal objetivo, entender o problema do negócio e identificar possíveis soluções para melhorar este processo, podendo ser desde de algoritmos de otimização até predição. Para isso, entra em contato com o negócio, buscando entender o problema e desenvolvendo uma solução desde o tratamento de dados até desenvolvimento de modelos e algoritmos.

▼ Show hidden code

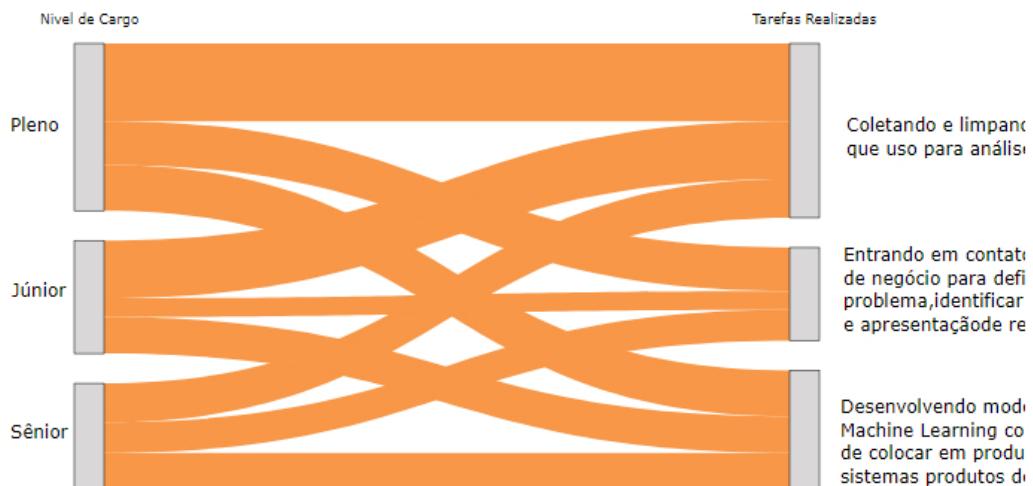
Tarefas que um cientista de dados gasta a maior parte do seu tempo no trabalho



Percebe-se que cientistas nível júnior ficam mais responsáveis pela coleta e tratamento de dados e pouco contato com times de negócios para definição de problema e solução, enquanto seniores exercem mais este papel e também.

▼ Show hidden code

Tarefas que um cientista de dados gasta a maior parte do seu tempo no trabalho por nível de cargo

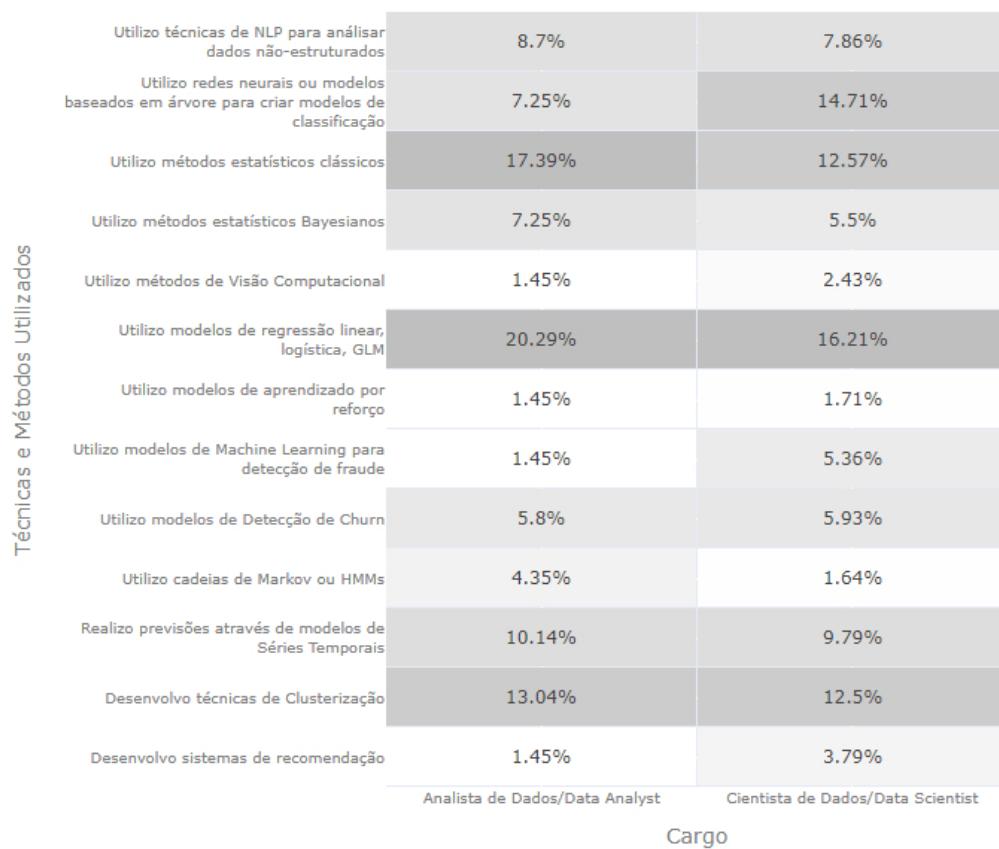


Técnicas e Métodos que cada Cargo utiliza no seu Trabalho

Os principais métodos utilizados para ambos os cargos são modelos de regressão linear, logística e GLM. Contudo, um analista de dados utiliza mais métodos estatísticos clássicos, como testes de hipóteses, análise multivariada etc. Já um cientista, utiliza mais métodos de redes neurais ou modelos baseados em árvore para classificação.

>Show hidden code

Técnicas e Métodos que cada Cargo utiliza no seu trabalho



Tempo De Experiência em Dados e na TI/desenvolvimento de Software

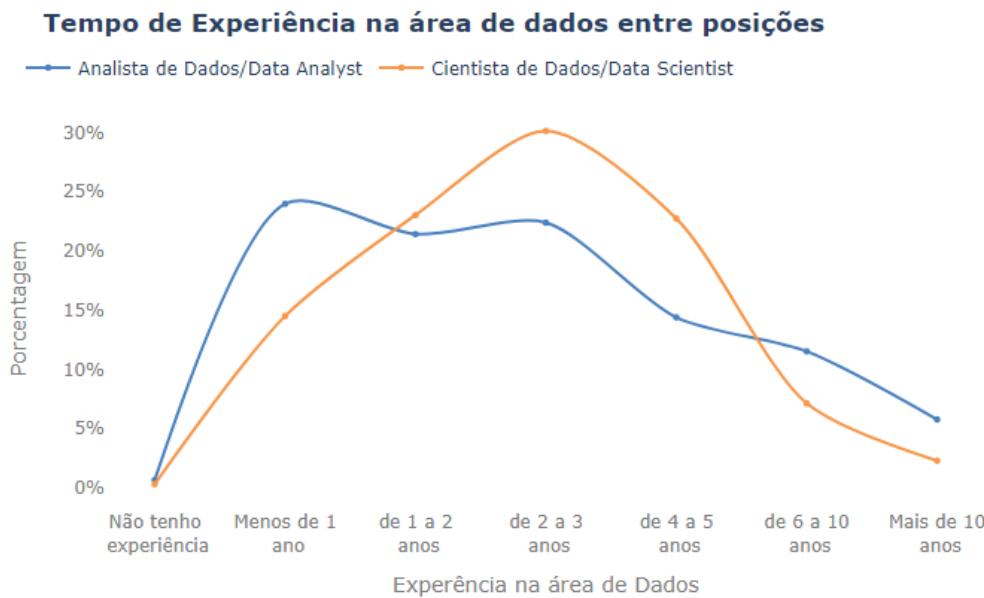
Tempo de Experiência em Dados:

Analistas de dados possuem uma menor experiência na área de dados, o que sugere que é um cargo de entrada para quem está começando nesta área. Além disso, este cargo apresenta uma porcentagem maior acima de 6 anos de experiência em relação ao de cientista de dados, isso pode estar relacionado com o fato que antigamente existia o cargo de analista de BI no qual desenvolvia dashboard e views para o negócio e hoje muitas empresas estão considerando como analistas de dados.

Para cientistas de dados, percebe-se que são pessoas que já possuem uma bagagem em dados e estão

querendo se aprofundar mais na parte de algoritmos e modelos. Vale lembrar que algumas vagas pedem tempo de experiência como requisito diferencial ou obrigatório.

>Show hidden code



Relação entre Tempo de Experiência em Dados e Idade:

Esperamos que quanto maior o tempo de experiência em dados, maior seria a idade, o que é verdade. Para pessoas que não possuem experiência na área, analistas de dados possuem a média de idade de 7 anos a mais do que um cientista, isto pode estar vinculado com a migração de pessoas para esta área.

Se analisarmos a diferença entre idade e tempo de experiência, percebe-se que estamos tendo uma maior idade ao passar dos anos, por exemplo: hoje as cientistas de dados iniciam na área com 27 anos. Se analisarmos cientistas que possuem de 2 a 3 anos de experiência em dados tem a média de 28 anos, se diminuirmos essa idade pelo tempo de experiência, eles começaram a trabalhar com dados com 26 anos, um ano mais novo do que hoje.

Show hidden code

Média de Idade por Tempo de Experiência em Dados

Analista de Dados/Data Analyst Cientista de Dados/Data Scientist



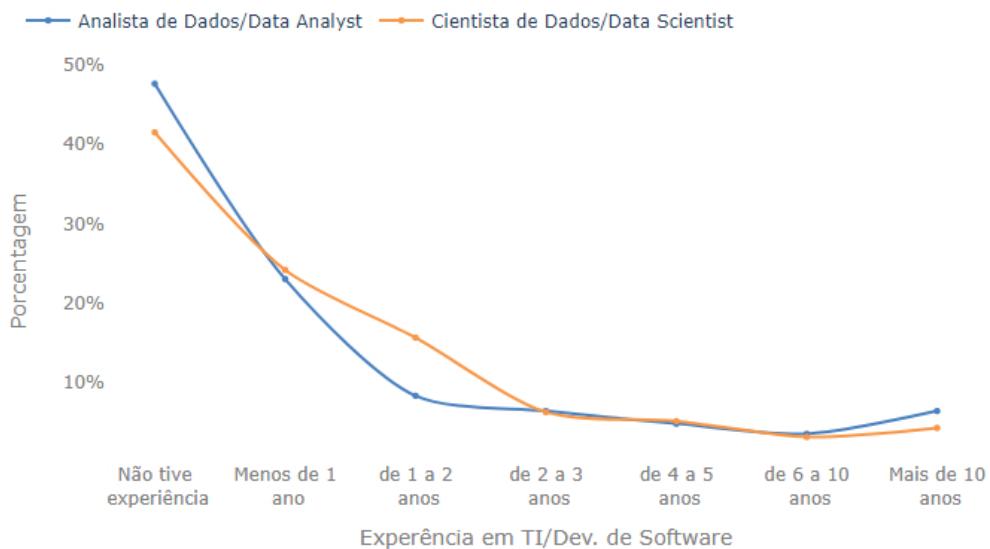


Tempo de Experiência em TI:

Percebe-se que para ambas as posições a maioria dos respondentes não tiveram experiência em TI antes de se tornar um analista ou cientista de dados. Contudo, para cientista de dados 15,62% já tiveram experiência em TI ou desenvolvimento de software.

>Show hidden code

Tempo de Experiência em TI entre posições



Relação entre Experiência em TI vs Dados

Analista de Dados:

Poucos analistas de dados tiveram experiência em TI ou desenvolvimento de software. Desta forma, para pessoas que não são da área ou que não possuem muito conhecimento em programação, talvez analista de dados, seja uma boa opção para dar início na carreira em dados.

>Show hidden code

Relação entre Experiência em TI e Dados para Analistas de Dados

Experiência em TI	Mais de 10 anos	de 6 a 10 anos	de 4 a 5 anos	de 2 a 3 anos	de 1 a 2 anos	Menos de 1 ano	Não tive experiência	Experiência em Dados
	1	0	0	0	0	1	0	
Mais de 10 anos	1	3	2	5	3	3	3	
de 6 a 10 anos	0	2	1	5	1	1	1	
de 4 a 5 anos	0	5	1	2	2	3	2	
de 2 a 3 anos	0	3	4	4	3	3	3	
de 1 a 2 anos	0	7	5	2	4	6	2	
Menos de 1 ano	1	19	18	13	12	6	3	
Não tenho experiência	0	36	36	39	20	14	4	
	Não tive experiência	Menos de 1 ano	de 1 a 2 anos	de 2 a 3 anos	de 4 a 5 anos	de 6 a 10 anos	Mais de 10 anos	Experiência em Dados

Cientista de Dados:

Para cientista de dados, ao mesmo tempo que possuem experiência com dados, possuem experiência em TI ou desenvolvimento de software. Isto sugere, que pelo menos uma vez foi necessário realizar algum processo de TI ou desenvolvimento de software ao longo da sua jornada como cientista, por exemplo: configuração de banco de dados, criação de APIs e entre outros.

▼ Show hidden code

Relação entre Experiência em TI e Dados para Cientista de Dados

Experiência em TI	Mais de 10 anos	de 6 a 10 anos	de 4 a 5 anos	de 2 a 3 anos	de 1 a 2 anos	Menos de 1 ano	Não tive experiência	Experiência em Dados
	0	0	0	0	0	22	1	
Mais de 10 anos	0	1	3	4	2	3	2	
de 6 a 10 anos	0	1	2	6	2	0	0	
de 4 a 5 anos	0	1	4	5	3	5	0	
de 2 a 3 anos	0	3	5	6	6	1	1	
de 1 a 2 anos	0	4	14	22	12	2	1	
Menos de 1 ano	0	22	16	25	19	3	0	
Não tenho experiência	1	19	37	38	36	11	4	
	Não tive experiência	Menos de 1 ano	de 1 a 2 anos	de 2 a 3 anos	de 4 a 5 anos	de 6 a 10 anos	Mais de 10 anos	Experiência em Dados

Ferramentas e Serviços

Linguagem de Programação e Banco de Dados:

Um analista de dados utiliza muito SQL para realizar as extrações de dados e construções de dashboard e

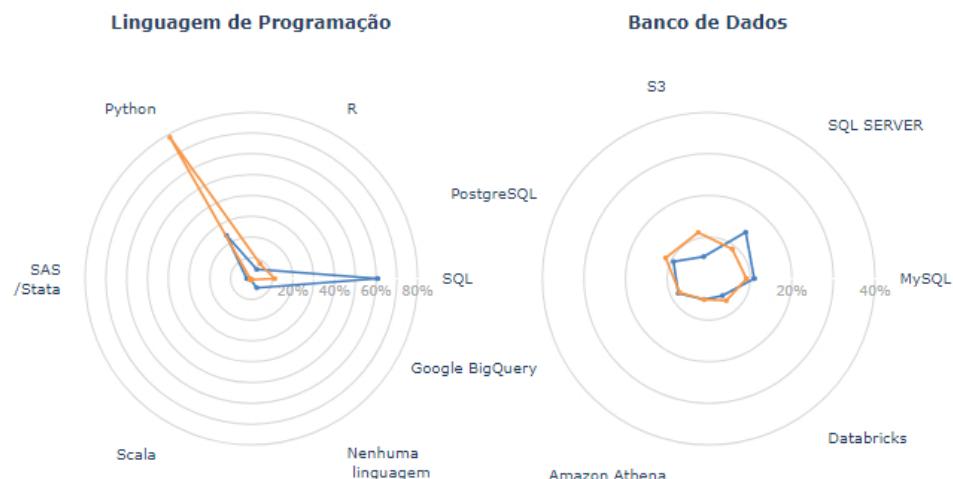
Python para auxiliar no processo de extração e transformação de dados e análise exploratória. Já para cientistas de dados, a principal linguagem de programação é Python, isto pode estar vinculado com a versatilidade que a ferramenta possui em vários processos de ciência de dados. Outro fato, é que 5% e 1% dos analistas e cientistas de dados respectivamente, não utilizam nenhuma linguagem de programação, talvez estas pessoas utilizem outras plataformas.

Em banco de dados, percebe-se que cientistas de dados trabalham com mais de um banco de dados, enquanto analistas a maioria utiliza SQL server. Contudo, vale lembrar que estes bancos de dados, o que altera de fato para o usuário além da performance é a sintaxe. Então, se tem conhecimento em alguns deles a transição de conhecimento para outro é facilitada.

▽ Show hidden code

Comparação entre Linguagens de Programação e Banco de Dados por Cargo

Analista de Dados/Data Analyst Cientista de Dados/Data Scientist



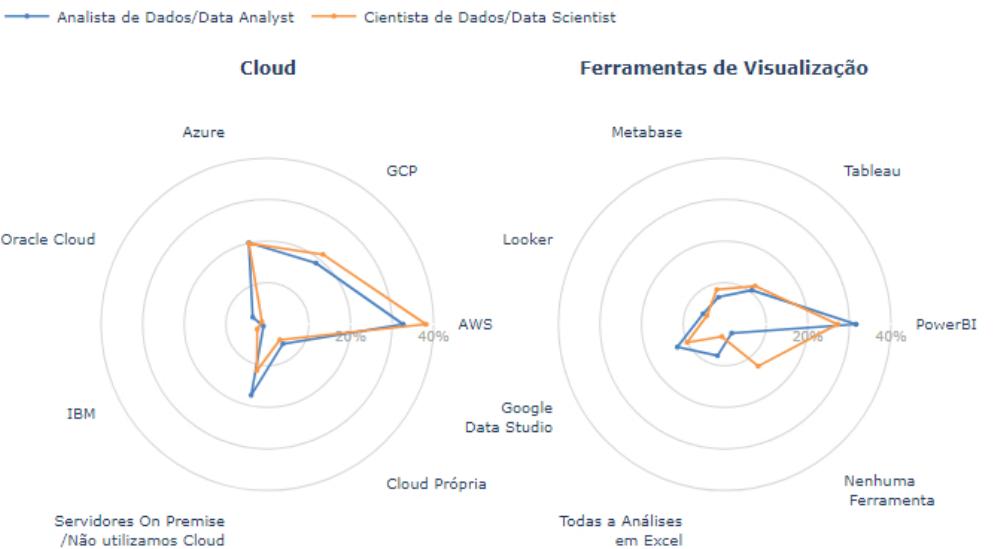
Cloud e Visualização de dados:

Já para serviços de cloud a maioria dos respondentes, utilizam AWS independente do cargo e logo em seguida google cloud (GCP) ou não utilizam cloud/servidores on premise. Nota-se que não há uma diferença entre os cargos para o uso de cloud, então se está começando na carreira ou migrando de área independente de qual profissão queira seguir vale a pena estudar cloud.

Para ferramentas, de visualização, temos o domínio de power BI tanto para cientistas e analistas e logo em seguida o google data studio que é uma plataforma que consegue fazer o link facilmente com google analytics. Além disso, percebe-se que 13% dos cientistas de dados, não utilizam nenhuma ferramenta de BI, certamente são os cientistas que só trabalham com algoritmos e análises. Por fim, ainda temos o tradicional Excel que 8% dos analistas utilizam como ferramenta de visualização e análise.

▽ Show hidden code

Comparação entre Cloud e BIs por Cargo

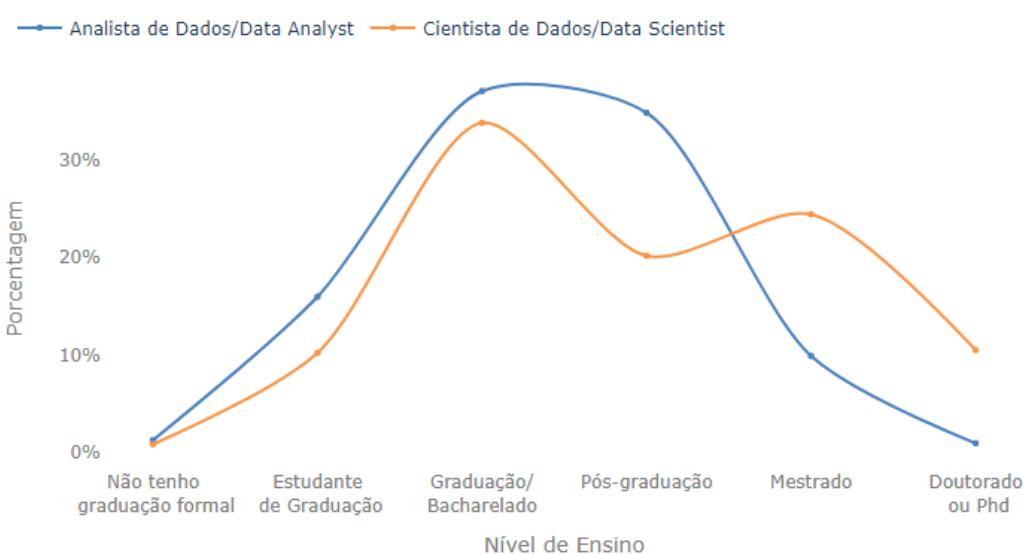


Nível de Ensino

Para nível de ensino, percebe-se que poucos respondentes não possuem uma graduação formal. Seguindo a lógica dos respondentes, para um analista de dados é necessário terminar a graduação e realizar uma pós-graduação. Geralmente, cursos de faculdade não abordam todas as metodologias e assuntos iniciais de dados. Já para cientista de dados há uma divisão entre graduação e mestrado, talvez, alguns cientistas respondentes não realizam pós-graduação e vão direto para o mestrado. Além disso, algumas vagas para cientistas de dados consideram um diferencial ter mestrado.

▼ Show hidden code

Nível de Ensino entre as posições

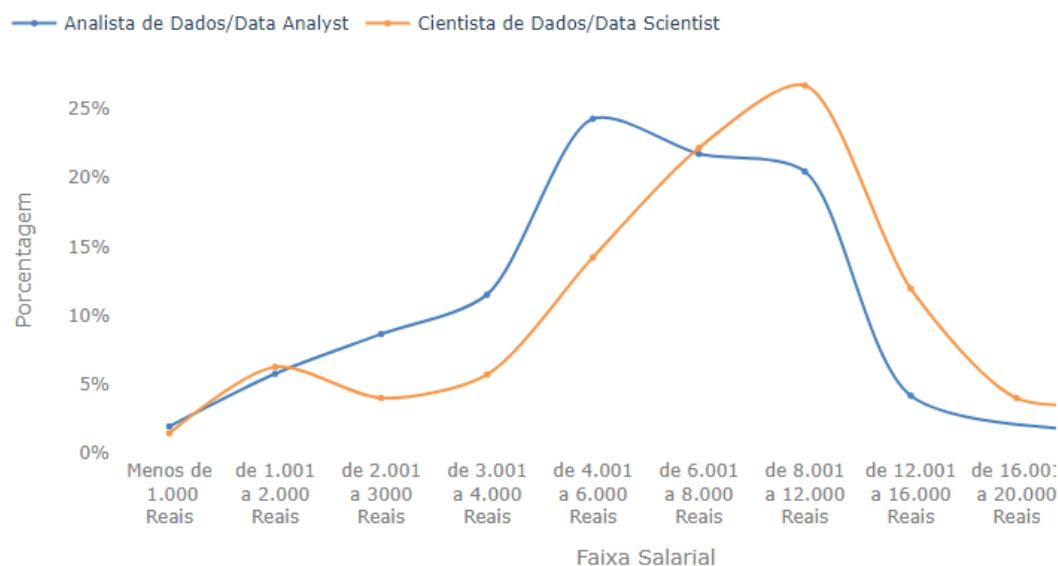


Faixa Salarial

Nota-se que analistas de dados ganham menos em relação a cientistas. Outro fato, é que salários abaixo de 4.001 reais podem ser considerados abaixo do mercado para ambos os cargos. Além disso, ao contrário que algumas notícias e cursos online comentam que esta profissão ganha mais de 20 mil reais por mês, poucos respondentes ganham isso de verdade.

▼ Show hidden code

Comparação entre Faixa Salarial por Cargo



Relação entre Faixa Salarial e Tarefas

Analistas de Dados:

Independente da faixa salarial a principal tarefa é a construção de dashboards e o processamento e análise de dados. Contudo, salários abaixo de 3.001 reais realizam mais desenvolvimento de manutenção de planilhas de Excel do que o processamento e análise de dados, isto pode estar vinculado com o conhecimento que esta pessoa possui em dados. Além disso, salários acima de 6.001 reais até 12.000 reais exercem mais a tarefa de modelar dados com objetivo de criar conjunto de dados, data warehouses. Para salários acima 12.001 reais as tarefas são construção de dashboards e análise de dados.

▼ Show hidden code

Relação entre Faixa Salarial e Tarefas para Analistas de Dados

Faixa Salarial



Cientistas de Dados:

Para cientistas de dados, nota-se que a distribuição de tarefas é similar independente da faixa salarial. Contudo, pessoas com faixa salarial abaixo de 4.001 reais tendem a realizar mais tarefas de processamento e análise de dados do que entrar em contato com times de negócios, ao contrário de respondentes com faixa salarial entre 12.001 à 16.000 reais que exercem bastante este papel e desenvolvimento de modelos de machine learning.

  Show hidden code

Relação entre Faixa Salarial e Tarefas para Cientistas de Dados

Processando e analisando dados utilizando linguagens de programação como Python, R etc.	1	4	3	6	4	6	10	0	1
Entrando em contato com os times de negócio para definição do problema, identificar a solução e apresentação de resultados.	0	1	1	3	8	22	27	13	0
Desenvolvendo modelos de Machine Learning com o objetivo de colocar em produção em sistemas produtos de dados.	1	7	3	7	12	25	27	9	6
Colocando modelos em produção, criando os pipelines de dados, APIs de consumo e monitoramento.	0	2	1	0	1	7	3	4	3
Coletando e limpando os dados que uso para análise e modelagem.	2	10	3	10	27	35	38	18	2
	Menos de 1.000	de 1.001 a 2.000	de 2.001 a 3.000	de 3.001 a 4.000	de 4.001 a 6.000	de 6.001 a 8.000	de 8.001 a 12.000	de 12.001 a 16.000	de 16.001 a 20.000

Conclusão

Imaginamos que todo mundo que trabalha com dados é nomeado nas chamadas de vídeos como "o pessoal de dados". Contudo, nota-se que existe uma diferença entre cientistas e analistas de dados:

Analista de Dados:

Analistas possuem o papel de auxiliar o negócio a tomar decisões através de dados desenvolvendo análises, dashboards e views em banco de dados. Para isso, utiliza como principais ferramentas SQL e Python e para banco de dados SQL Server. Como métodos e conhecimentos técnicos são modelos de regressão linear, logística e estatística clássica: teste de hipóteses, análise multivariada entre outros.

Para iniciantes neste cargo, as principais tarefas são desenvolvimento de dashboards e manutenção de planilhas enquanto para pleno/seniores a principal tarefa é construção de dashboards e indicadores. Além disso, é um cargo que a maioria dos respondentes apresentam ter graduação ou pós-graduação e seu salário é inferior ao de cientista, também é um cargo que não exige muita experiência com dados e nem com TI/desenvolvimento de software.

Cientista de Dados:

Cientistas possuem o papel de entender o problema da área de negócio e propor soluções através de dados utilizando algoritmos de otimização ou predição. Como ferramental utiliza principalmente Python e métodos de redes neurais ou modelos de árvores para classificação. Independentemente do nível de cargo, todos realizam as mesmas tarefas como desenvolvimento de modelos, tratamento de dados entre outros.

Cargos com nível mais pleno/sênior realizam mais a tarefa de entrar em contato com área de negócio para entender o problema. Além disso, a maioria dos respondentes, apresentam ter graduação ou mestrado como nível de educação. Seu salário é superior ao de um analista, no entanto, poucos ganham acima de 20 mil reais, também é um cargo que exige já experiência em dados.

E é isso! Foi um prazer fazer esse kernel, aprendi muito! Obrigado por ler e se você gostou, por favor, dê um upvote!



Fonte: [quotefancy](#)

Tradução: sonhe grande, mantenha-se positivo, trabalhe duro e aproveite a jornada.

Referências

Schubert de Abreu (2021) [Data Scientists & Analysts: What's the difference?](#)

Andre Sionek (2019) [Is there any job out there? Kaggle vs Glassdoor](#)

Andre Sionek (2019) [How to Create Award Winning Data Visualizations](#)

License

This Notebook has been released under the [Apache 2.0](#) open source license.

Continue exploring



Data

2 input and 0 output



Logs

26.8 second run - successful



Comments

5 comments

