

Data Engineering

TechGuide - Alura, FIAP e PM3

Nível 1

Engenharia de Dados - Fundamentos:

- Uma pessoa Engenheira de Dados desempenha um papel crucial ao projetar, implementar e manter as infraestruturas de dados que permitem que uma organização trabalhe de maneira eficiente com suas informações. A pessoa engenheira de dados cria pipelines de dados para integrar, limpar e transformar dados provenientes de várias fontes e formatos. Isso possibilita a geração de insights valiosos para o negócio e melhora a tomada de decisões estratégicas.

Conteúdos

- **Artigo** O que faz um Engenheiro de Dados? (<https://medium.com/data-hackers/o-que-faz-um-engenheiro-de-dados-fdcb0bca966b>)
- **YouTube** O que é Big Data? O que é Data Analytics? (<https://www.youtube.com/watch?v=4CNqiAgmqSc>)
- **YouTube** Millena Gená: O que faz um Engenheiro de Dados? (<https://www.youtube.com/watch?v=Zsos7iTylqo>)
- **Artigo** Os Tipos e Perfis de Engenheiros de Dados (<https://medium.com/data-hackers/os-tipos-de-engenheiros-de-dados-c3674b4d74d1>)
- **Artigo** O passado, presente e o futuro da Engenharia de Dados — Data Hackers Podcast (<https://medium.com/data-hackers/o-passado-presente-e-o-futuro-da-engenharia-de-dados-data-hackers-podcast-49-34be714b20f6>)

Conteúdos Alura:

- **Artigo** Engenharia de Dados: o que é, o que faz e um Guia completo (<https://www.alura.com.br/artigos/engenharia-dados>)
- **YouTube** Alura: O que é Engenharia de Dados? (https://www.youtube.com/watch?v=_rMb6S6kimM)
- **Artigo** O que faz uma pessoa Engenheira de Dados? (<https://www.alura.com.br/artigos/o-que-pessoa-engenheira-dados>)
- **Podcast** Hipsters.Tech: Data Engineering no Quinto Andar (<https://www.hipsters.tech/data-engineering-no-quinto-andar-hipsters-on-the-road-60/>)
- **Curso** Primeiros passos com Engenharia de Dados (<https://cursos.alura.com.br/formacao-engenharia-dados-primeiros-passos>)

Python - Fundamentos:

- Python é uma linguagem de programação de alto nível, de uso geral, amplamente utilizada em aplicações web, desenvolvimento de software, ciência de dados e Machine Learning. Sua filosofia de projeto enfatiza a legibilidade do código com o uso de indentação significativa. Python é dinamicamente tipada e tem um garbage collector.
- Conhecer os tipos primitivos
- Declarar variáveis, considerando os diferentes tipos
- Usar estruturas condicionais ('if', 'else')
- Conhecer os operadores de atribuição e comparação
- Usar estruturas de repetição e laços ('while', 'for')
- Usar funções, passando parâmetros e argumentos
- Manipular métodos
- Manipular arrays e listas
- Obter dados de uma API
- Criar construtores
- Funções anônimas

Conteúdos

- **Site** Documentação Python (<https://docs.python.org/pt-br/3/tutorial/>)
- **Artigo** Programação | Python — Parte 1 (<https://medium.com/turing-talks/turing-talks-4-python-parte-1-29b8d9efd0a5>)
- **Artigo** Introdução ao Python (<https://medium.com/@goularteduarda.a/introdu%C3%A7%C3%A3o-ao-python-d30c29eba0d6>)
- **YouTube** Eduardo Mendes: Curso introdutório de Python (<https://www.youtube.com/watch?v=yTQDbqmv8Ho>)

Conteúdos Alura:

- **Artigo** O que é Python? História, Sintaxe e um Guia para iniciar na Linguagem (<https://www.alura.com.br/artigos/python>)
- **Podcast** Hipsters.tech: Python – Hipsters 122 (<https://www.hipsters.tech/python-hipsters-122/>)
- **Podcast** Hipsters.tech: Python Fluente – Hipsters Ponto Tech 179 (<https://www.hipsters.tech/python-fluente-hipsters-ponto-tech-179/>)
- **YouTube** Alura: A linguagem Python - Alura Live 94 (<https://www.youtube.com/watch?v=Nbt0eQHChol>)
- **YouTube** Alura: O que é Python? (<https://www.youtube.com/watch?v=-LATVnPcvHI>)
- **Artigo** Python - Uma Introdução à Linguagem (<https://www.alura.com.br/artigos/python-uma-introducao-a-linguagem>)
- **Artigo** Python: Trabalhando com precisão em números decimais (<https://www.alura.com.br/artigos/precisao-numeros-decimais-python>)
- **Artigo** Listas em Python: operações básicas (<https://www.alura.com.br/artigos/listas-no-python>)
- **Artigo** Como comparar objetos no Python? (<https://www.alura.com.br/artigos/como-comparar-objetos-no-python>)
- **Artigo** Python datetime: Lidando com datas e horários (<https://www.alura.com.br/artigos/lidando-com-datas-e-horarios-no->

[python](#))

- **Curso** Curso Python: crie a sua primeira aplicação
(<https://cursos.alura.com.br/course/python-crie-sua-primeira-aplicacao>)
- **Curso** Formação Python e orientação a objetos
(<https://cursos.alura.com.br/formacao-linguagem-python>)

Java - Fundamentos:

- Java é uma linguagem de programação amplamente usada para codificar aplicações Web. Java é uma linguagem multiplataforma, orientada a objetos e centrada em rede que pode ser usada como uma plataforma em si. É uma linguagem de programação rápida, segura e confiável para codificar tudo, desde aplicações móveis e software empresarial até aplicações de big data e tecnologias do servidor.
- Conhecer os tipos primitivos
- Declarar variáveis, considerando os diferentes tipos
- Usar estruturas condicionais ('if', 'else')
- Conhecer os operadores de atribuição e comparação
- Usar estruturas de repetição e laços ('while', 'for')
- Usar funções, passando parâmetros e argumentos
- Manipular métodos
- Manipular arrays e listas
- Obter dados de uma API
- Criar construtores

Conteúdos

- **Site** O que é Java? (<https://aws.amazon.com/pt/what-is/java/>)
- **Artigo** Aprenda Comigo: Java — Parte 1
(<https://medium.com/clebertech/aprenda-comigo-java-parte-1-42ef8fddd8b6>)
- **YouTube** FernandaDev: Java (Aula 1 - Classes, Objetos, Atributos e Métodos)
(<https://www.youtube.com/watch?v=ohmHbdUhAGc>)

- **YouTube** Loiane Groner: Primeiro Programa em Java
(<https://youtu.be/mu2ti43cgwc>)

Conteúdos Alura:

- **Artigo** Java: o que é, linguagem e um Guia para iniciar na tecnologia
(<https://www.alura.com.br/artigos/java>)
- **Podcast** Hipster 313 - Ecossistema Java revisitado
(<https://www.hipsters.tech/ecossistema-java-revisitado-hipsters-ponto-tech-313/>)
- **Artigo** Como começar a desenvolver em Java?
(<https://www.alura.com.br/artigos/comecando-com-o-desenvolvimento-java>)
- **Artigo** Meu primeiro programa em Java
(<https://www.alura.com.br/artigos/meu-primeiro-programa-em-java>)
- **Artigo** Java: Conheça o método main
(<https://www.alura.com.br/artigos/metodo-main-em-java>)
- **Artigo** Desenvolvendo aplicações Java com o VS Code
(<https://www.alura.com.br/artigos/desenvolvendo-aplicacoes-java-vs-code>)
- **Artigo** Diferença entre int e Integer em Java
(<https://www.alura.com.br/artigos/diferenca-entre-int-e-integer-em-java>)
- **Artigo** Como converter String para Date em Java
(<https://www.alura.com.br/artigos/como-converter-string-para-date-em-java>)
- **Artigo** Iterando uma lista em Java
(<https://www.alura.com.br/artigos/iterando-uma-lista-em-java>)
- **Artigo** Importando classes no Java
(<https://www.alura.com.br/artigos/importando-classes-em-java>)
- **Site** Apostila: Java e Orientação a Objetos
(<https://www.alura.com.br/apostila-java-orientacao-objetos>)
- **Site** Apostila: Java para Desenvolvimento Web
(<https://www.alura.com.br/apostila-java-web>)

- **YouTube** Alura: O que é o Java? (<https://www.youtube.com/watch?v=90NcVNsKGik>)
- **Curso** Formação Java e Orientação a Objetos (<https://cursos.alura.com.br/formacao-java>)
- **Desafio** 7 Days of Code: Java (<https://7daysofcode.io/matricula/java>)

Bancos de Dados - Fundamentos:

- Os bancos de dados são sistemas de armazenamento de dados que permitem a organização, a manipulação e o acesso eficiente às informações. Existem dois principais tipos de bancos de dados, sendo eles SQL (Structured Query Language) e NoSQL (Not Only SQL).
- O conhecimento dos fundamentos de bancos de dados, tanto SQL quanto NoSQL, é essencial na engenharia de dados. Os bancos de dados são a base para armazenar e manipular os dados utilizados nas análises e nos processos de tomada de decisão. Compreender os diferentes tipos de bancos de dados permite aos engenheiros de dados escolher a solução mais adequada para cada caso, levando em consideração os requisitos de escalabilidade, desempenho, consistência e flexibilidade. Além disso, a habilidade de escrever consultas SQL e de utilizar as características específicas dos bancos de dados NoSQL é essencial para extrair, transformar e carregar dados de forma eficiente e eficaz.

Conteúdos

- **Site** DB-Engines Ranking (<https://db-engines.com/en/ranking>)
- **YouTube** Código Fonte TV: SQL // Dicionário do Programador (<https://www.youtube.com/watch?v=kMznyl7r2Tc>)
- **YouTube** Código Fonte TV: NoSQL // Dicionário do Programador (<https://www.youtube.com/watch?v=1B64oqE8PLs>)
- **Artigo** SQL — O que é e como funciona na prática? (<https://medium.com/pravaler-digital-team/sql-o-que-%C3%A9-e-como-funciona-na-pr%C3%A1tica-6ae7a322554a>)
- **Artigo** SQL: comandos básicos (<https://medium.com/nerdzao/sql-comandos-b%C3%A1sicos-483a8dc12909>)

- **Site** AWS: O que é banco de dados relacional? (<https://aws.amazon.com/pt/relational-database/>)
- **YouTube** 10 conceitos essenciais de SQL que todo dev deve saber (<https://www.youtube.com/watch?v=cIUTAyomVPQ>)
- **Site** AWS: O que é NoSQL? (<https://aws.amazon.com/pt/nosql/>)
- **Artigo** NoSQL — O que são? Onde habitam? Do que se alimentam? (<https://medium.com/permalink-univesp/nosql-o-que-s%C3%A3o-onde-habitam-do-que-se-alimentam-2ccd4cc39571>)

Conteúdos Alura:

- **YouTube** Alura: O que é SQL e NoSQL? (<https://www.youtube.com/watch?v=aure5d3B88g>)
- **Artigo** SQL e NoSQL: trabalhando com bancos relacionais e não relacionais (<https://www.alura.com.br/artigos/sql-nosql-bancos-relacionais-nao-relacionais>)
- **Artigo** Banco de Dados: o que é, principais tipos e um guia para iniciar (<https://www.alura.com.br/artigos/banco-de-dados>)
- **YouTube** Dicas para desenvolver um projeto de banco de dados relacional (<https://cursos.alura.com.br/extra/alura-mais/dicas-para-desenvolver-um-projeto-de-banco-de-dados-relacional-c9093>)
- **Artigo** JOIN e seus tipos (<https://www.alura.com.br/artigos/join-e-seus-tipos>)
- **Artigo** Trabalhando com funções de agregação (<https://www.alura.com.br/artigos/trabalhando-funcoes-de-agregacao>)
- **Artigo** SQL: Comandos básicos (<https://www.alura.com.br/artigos/sql-comandos-basicos>)
- **Artigo** SQL: consultas com SELECT (<https://www.alura.com.br/artigos/sql-consultas-com-select>)
- **Curso** Curso SQL com MySQL: manipule e consulte dados (<https://cursos.alura.com.br/course/mysql-manipule-dados-com-sql>)
- **Podcast** NoSQL: por quê? (<https://cursos.alura.com.br/extra/hipsterstech/nosql-por-que-hipsters-22->

[a564\)](#)

- **Curso** Curso MongoDB: conhecendo um banco de dados NoSQL (<https://cursos.alura.com.br/course/mongodb-banco-dados-nosql>)

Linux - Fundamentos:

- Linux é um termo popularmente empregado para se referir a sistemas operacionais que utilizam o Kernel Linux. As distribuições incluem o Kernel Linux, além de softwares de sistema e bibliotecas.
- Conhecer o sistema de diretórios do Linux
- Compactar e descompactar arquivos
- Editar arquivos no console com o VI
- Gerenciar os processos rodando na máquina
- Conhecer as variáveis de ambiente e o PATH
- Gerenciar pacotes
- Realizar comunicação remota com o SSH e SCP

Conteúdos

- **Artigo** Comandos linux (<https://pllara.medium.com/comandos-linux-4db7307c17b7>)
- **Artigo** Comandos básicos no Linux (<https://medium.com/linkapi-solutions/comandos-b%C3%A1sicos-no-linux-95079363ab0e>)
- **YouTube** Ninja do Linux: Comandos básicos da linha de comando do Linux (https://www.youtube.com/watch?v=rs_yshFGu8E)
- **Curso** MIT - The Missing Semester of Your CS Education (inglês) (<https://missing.csail.mit.edu/>)

Conteúdos Alura:

- **Artigo** Linux: download, instalação, distribuições e principais comandos (<https://www.alura.com.br/artigos/linux-download-instalacao-distribuicoes-e-principais-comandos>)

- **Artigo** Procurando arquivos no Linux
(<https://www.alura.com.br/artigos/procurando-arquivos-no-linux>)
- **Artigo** Criando comandos no Linux com Bash
(<https://www.alura.com.br/artigos/criando-comandos-no-linux-com-bash>)
- **Artigo** Linux - Visualizando tamanhos de Diretórios
(<https://www.alura.com.br/artigos/como-visualizar-o-tamanho-de-diretorios-no-linux>)
- **Artigo** Entendendo as permissões no Linux
(<https://www.alura.com.br/artigos/entendendo-as-permissoes-no-linux>)
- **Artigo** Linux: compactando e descompactando arquivos com o tar
(<https://www.alura.com.br/artigos/linux-compactando-e-descompactando-arquivos-com-o-tar>)
- **Artigo** SSH: o acesso remoto aos servidores
(<https://www.alura.com.br/artigos/como-acessar-servidores-remotamente-com-ssh>)
- **Curso** Linux Onboarding: usando a CLI de uma forma rápida e prática
(<https://cursos.alura.com.br/course/linux-onboarding-utilizar-cli-forma-rapida-pratica>)
- **Curso** Linux Onboarding: localizando arquivos e conteúdos
(<https://cursos.alura.com.br/course/linux-onboarding-arquivos-conteudos>)
- **Curso** Linux Onboarding: obtendo e tratando informações do sistema
(<https://cursos.alura.com.br/course/linux-onboarding-informacoes-sistema>)
- **Curso** Linux Onboarding: trabalhe com usuários, permissões e dispositivos
(<https://cursos.alura.com.br/course/linux-onboarding-usuarios-permissoes-dispositivos>)
- **Curso** DevOps: explorando conceitos, comandos e scripts no Linux CLI
(<https://www.alura.com.br/curso-online-devops-conceitos-comandos-scripts-linux-cli>)

Big Data - Fundamentos:

- Big Data refere-se a conjuntos de dados extremamente grandes e complexos, que não podem ser facilmente processados ou gerenciados por

métodos tradicionais. Esses conjuntos de dados são caracterizados por seu volume massivo, velocidade de geração e variedade de tipos e formatos. Com a explosão da quantidade de dados gerados diariamente, provenientes de diversas fontes como redes sociais, dispositivos IoT e transações comerciais, o Big Data apresenta desafios e oportunidades para extrair insights valiosos e tomar decisões informadas.

- A análise de Big Data envolve o uso de técnicas e tecnologias avançadas, como armazenamento distribuído, processamento em paralelo e aprendizado de máquina, para explorar e transformar esses dados em informações significativas para empresas, organizações e pesquisadores.

Conteúdos

- **Site** Oracle: O que é Big Data? (<https://www.oracle.com/br/big-data/what-is-big-data/>)
- **YouTube** Código Fonte TV: Big Data (<https://www.youtube.com/watch?v=lpfE8B9H9cl>)
- **YouTube** Nerdologia: BIG DATA (https://www.youtube.com/watch?v=hEFFCKxYbKM&pp=ygUIQkIHIERBVEE%3D&ab_channel=Nerdologia)
- **Artigo** A Era do Big Data (<https://medium.com/gabriel-luz/a-era-do-big-data-64ebad5859f2>)
- **YouTube** Curso de Big Data (<https://www.youtube.com/playlist?list=PLzKHEB8QCeI8UIA2JfUTYQaUgvJa6mUog>)

Conteúdos Alura:

- **Artigo** Big Data: Entenda mais sobre esse conceito (<https://www.alura.com.br/artigos/big-data>)
- **Podcast** Big Data e Apache Spark (<https://www.hipsters.tech/big-data-e-apache-spark-hipsters-ponto-tech-341/>)
- **Podcast** Hipsters.tech: Big Data e Data Science: pra quê afinal? (<https://www.hipsters.tech/big-data-e-data-science-pra-que-afinal-hipsters-73/>)
- **Podcast** Small Data, Big Data, Open Data (<https://www.hipsters.tech/small-data-big-data-open-data-hipsters-42/>)

- **YouTube** Alura: Big Data com Apache Spark
(https://www.youtube.com/watch?v=3aSfKsBhU5E&ab_channel=Alura)
- **Artigo** Iniciando um projeto Spark no Google Colab
(<https://www.alura.com.br/artigos/iniciando-projeto-spark-no-colab>)
- **Curso** Formação Apache Spark com Python
(<https://cursos.alura.com.br/formacao-apache-spark-python>)

Pipelines e ETL:

- Pipelines e ETL (Extract, Transform, Load) são conceitos essenciais no campo da Engenharia de Dados, desempenhando um papel fundamental na coleta, transformação e integração de dados. Compreender os fundamentos dessas áreas é essencial para aqueles que desejam trabalhar com a gestão e o processamento eficiente de grandes volumes de informações.

Conteúdos

- **Site** AWS: O que é um pipeline de dados?
(<https://aws.amazon.com/pt/what-is/data-pipeline/>)
- **Artigo** Glossário de dados: O que é Data Pipeline?
(<https://medium.com/datalakers-blog/gloss%C3%A1rio-de-dados-o-que-%C3%A9-data-pipeline-628e509e9cb5>)
- **Artigo** Pipeline de Dados (<https://www.linkedin.com/pulse/pipeline-de-dados-jose-r-f-junior/?originalSubdomain=pt>)
- **Artigo** Sete princípios para pipelines de dados confiáveis
(<https://medium.com/data-hackers/sete-princ%C3%ADpios-para-pipelines-de-dados-confi%C3%A1veis-de59d3c72cfc>)
- **YouTube** Supletivo DH - Apache NiFi (<https://www.youtube.com/watch?v=IVRG4ks83s0>)

Conteúdos Alura:

- **Curso** Curso Apache Beam: Data Pipeline com Python
(<https://cursos.alura.com.br/course/apache-beam-data-pipeline-python>)
- **Curso** Curso Engenharia de Dados: organizando dados na AWS
(<https://cursos.alura.com.br/course/engenharia-dados-organizando-dados->

[aws](#))

DevOps e Cloud - Conceitos:

- DevOps desempenha um papel importante na integração contínua e na entrega contínua de projetos, ao garantir a colaboração eficaz entre desenvolvedores, operadores de sistemas e outros stakeholders importantes em um projeto. Essa pessoa também é fundamental na implementação e gerenciamento de soluções de computação em nuvem, facilitando o acesso a recursos e serviços, bem como a escalabilidade e a segurança dos sistemas.
- DevOps é essencial para a engenharia de dados, pois permite a integração contínua e a entrega eficiente de projetos. A colaboração entre equipes de desenvolvimento, operações e engenharia de dados facilita a integração de fontes de dados, processamento em tempo real e entrega de insights. Além disso, práticas DevOps garantem testes automatizados, monitoramento e governança de dados, garantindo qualidade e segurança nas soluções analíticas. Essa abordagem agiliza a gestão de dados em escala, possibilitando tomadas de decisões informadas.

Conteúdos

- **Site** RedHat: Introdução ao DevOps (<https://www.redhat.com/pt-br/topics/devops>)
- **Artigo** Conceitos de DevOps e SRE (<https://medium.com/sysadminas/conceitos-de-devops-e-sre-bbb05bb3c9fe>)
- **YouTube** Código Fonte TV: DevOps // Dicionário do Programador (<https://www.youtube.com/watch?v=iwf6kcvxeD4>)
- **YouTube** Akita: Entendendo Devops para Iniciantes em Programação (Parte 1) (<https://www.youtube.com/watch?v=bwO8EZf0gLI>)
- **YouTube** Akita: Entendendo Devops para Iniciantes em Programação (Parte 2) (<https://www.youtube.com/watch?v=mcwnQVAn0pw>)

Conteúdos Alura:

- **YouTube** O que você precisa saber para começar em DevOps com Leonardo Sartorello (https://www.youtube.com/watch?v=IQ8-_khQATQ)
- **Curso** Formação Começando em DevOps (<https://cursos.alura.com.br/formacao-primeiros-passos-devops>)
- **Artigo** Cloud: o que é, História e Guia da computação em nuvem (<https://www.alura.com.br/artigos/cloud>)
- **YouTube** AWS, Google Cloud e Azure: Por onde começar? (<https://www.youtube.com/watch?v=z9k6rsdmWc0&t=300s>)
- **Curso** Formação Começando em Cloud Computing (<https://cursos.alura.com.br/formacao-cloud-computing>)

Git e GitHub - Fundamentos:

- Git é um sistema de controle de versão distribuído gratuito e de código aberto projetado para lidar com tudo, desde projetos pequenos a muito grandes com velocidade e eficiência.
- GitHub é um serviço de hospedagem para desenvolvimento de software e controle de versão usando Git.
- Criar um repositório
- Clonar um repositório
- Fazer commit, push e pull de e para o repositório
- Reverter um commit
- Criar branches e pull requests
- Lidar com merge e conflitos

Conteúdos

- **Site** GitHub Documentação (<https://docs.github.com/pt>)
- **Site** GitHub Pages Documentação (<https://docs.github.com/pt/pages/getting-started-with-github-pages/about-github-pages>)
- **Site** Git School - Visualizing Git (<https://git-school.github.io/visualizing-git/>)

- **Site** Dangit, Git!?! (<https://dangitgit.com/>)
- **YouTube** Rafaella Ballerini: O que é Git e GitHub? - definição e conceitos importantes 1/2 (<https://www.youtube.com/watch?v=DqTITcMq68k>)
- **YouTube** Rafaella Ballerini: Como usar Git e GitHub na prática! - desde o primeiro commit até o pull request! 2/2 (<https://www.youtube.com/watch?v=UBAX-13g8OM>)
- **YouTube** Mario Souto - Dev Soutinho: Git: Entendendo de vez como funciona do melhor e mais visual jeito possível (<https://www.youtube.com/watch?v=4-tfJ-ZyA0Q>)
- **YouTube** Mario Souto - Dev Soutinho: Como colocar seu projeto no ar DE GRAÇA via GitHub! | Hospedagem com GitHub Pages (https://www.youtube.com/watch?v=BU-w2_Aae54)
- **YouTube** CodandoTV(Rods) - 5 coisas que você precisa saber sobre Git (https://youtu.be/MqogPzjQyCY?si=VL_mT8EowuemptmU)
- **YouTube** CodandoTV(Rods) - README de Sucesso: Transforme seu Projeto ou seu Perfil em Destaque no GitHub (https://youtu.be/v9ZM2PVzctM?si=49ah_HRPIwpSEX4A)

Conteúdos Alura:

- **Artigo** Git e Github: O que são, Como Configurar e Primeiros Passos (<https://www.alura.com.br/artigos/o-que-e-git-github>)
- **Artigo** Mais git com o hub: a linha de comando do Github (<https://www.alura.com.br/artigos/github-na-linha-de-comando>)
- **Podcast** Hipsters 184: Guia do Iniciante em Github (<https://cursos.alura.com.br/extra/hipsterstech/guia-do-iniciante-em-github-hipsters-184-a378>)
- **Site** GitHub: diferentes maneiras de compartilhar seu projeto (<https://cursos.alura.com.br/extra/alura-mais/github-diferentes-maneiras-de-compartilhar-seu-projeto-c2002>)
- **Site** Websérie: Git e Github para Sobrevivência (<https://www.alura.com.br/webseries/git-e-github-para-sobrevivencia>)

- **Podcast** Hipsters 109: Git e Github
(<https://www.alura.com.br/podcast/hipsterstech-git-e-github-hipsters-109-a474>)
- **YouTube** Alura: Git e Github para Sobrevivência 01: Como o Git funciona?
(<https://www.youtube.com/watch?v=BAmvmaKQkIQ>)
- **Curso** Curso Git e GitHub: compartilhando e colaborando em projetos
(<https://cursos.alura.com.br/course/git-github-compartilhando-colaborando-projetos>)
- **Curso** Curso Git e GitHub: dominando controle de versão de código
(<https://cursos.alura.com.br/course/git-github-dominando-controle-versao-codigo>)
- **Desafio** 7 Days of Code: GitHub (<https://7daysofcode.io/matricula/github>)

Nível 2

Processamento distribuído de dados:

- Em processamento distribuído de dados usamos técnicas e ferramentas para realizar tarefas computacionais em um conjunto de máquinas conectadas, conhecido como cluster. Essas máquinas trabalham em conjunto para analisar e processar grandes volumes de dados de maneira mais eficiente do que seria possível em uma única máquina. O profissional pode usar várias ferramentas e frameworks para realizar esse trabalho, incluindo Apache Hadoop, Apache Spark e Databricks.

Conteúdos

- **Site** O que é computação distribuída? (<https://aws.amazon.com/pt/what-is/distributed-computing/>)
- **Artigo** Apache Hadoop: O poder do processamento paralelo e distribuído
(<https://medium.com/data-hackers/apache-hadoop-o-poder-do-processamento-paralelo-e-distribu%C3%ADdo-6364bbc2ee8b>)
- **Artigo** Computação Distribuída: Processando Grandes Volumes de Dados em Larga Escala
(<https://www.linkedin.com/pulse/computa%C3%A7%C3%A3o->

distribuc3o.com.br/da-processando-grandes-volumes-de-gilmar-pupo/?originalSubdomain=pt)

- **YouTube** Código Fonte TV: Apache Spark (Data Analytics poderoso) // Dicionário do Programador (<https://www.youtube.com/watch?v=4TE6AGQ0IzI>)
- **YouTube** Código Fonte TV: Cluster // Dicionário do Programador (<https://www.youtube.com/watch?v=1cVUpoAR2Qc>)
- **Artigo** Arquitetura de um cluster Spark (<https://lorenadesouza.medium.com/bootcamp-de-dados-na-tw-spark-6633275480e4>)

Conteúdos Alura:

- **Artigo** Databricks: o que é e para que serve? (<https://www.alura.com.br/artigos/databricks-o-que-e-para-que-serve>)
- **Curso** Formação Apache Spark com Python (<https://www.alura.com.br/formacao-apache-spark-python>)
- **Curso** Formação Engenharia de dados com Databricks (<https://www.alura.com.br/formacao-databricks>)

Bancos de dados relacionais:

- Quando trabalhamos com bancos de dados relacionais somos responsáveis por projetar, implementar e gerenciar sistemas de bancos de dados que organizam dados em tabelas inter-relacionadas. Elas utilizam a linguagem SQL para manipular e consultar dados, além de técnicas de modelagem de dados para criar estruturas de banco de dados eficientes e eficazes. Os bancos de dados relacionais são usados em uma variedade de aplicações, incluindo sistemas de gestão de clientes, aplicações de e-commerce e sistemas de gerenciamento de recursos empresariais.

Conteúdos

- **Site** Oracle: O que é um banco de dados relacional (RDBMS)? (<https://www.oracle.com/br/database/what-is-a-relational-database/>)

- **Site** AWS: O que é banco de dados relacional?
(<https://aws.amazon.com/pt/relational-database/>)
- **YouTube** 10 conceitos essenciais de SQL que todo dev deve saber
(<https://www.youtube.com/watch?v=cIUTAyomVPQ>)
- **YouTube** Boson Treinamentos: Curso de Modelagem de Dados
(https://www.youtube.com/watch?v=Q_KTYFgvy1s&list=PLucm8g_ezqNoNHU8tjVeHmRGBFnjDlIxID)
- **YouTube** Otávio Miranda: Curso de SQL com MySQL (Completo)
(https://www.youtube.com/watch?v=IHYV_H1526Q&list=PLbIBj8vQhvm2WT-pjGS5x7zUz mh4VgvRk)

Conteúdos Alura:

- **Curso** Formação Modelagem de dados
(<https://cursos.alura.com.br/formacao-modelagem-dados>)
- **YouTube** O que é SQL e por que aprender SQL é essencial para você?
(<https://cursos.alura.com.br/extra/alura-mais/o-que-e-sql-e-por-que-aprender-sql-e-essencial-para-voce--c9092>)
- **Artigo** Saiba tudo sobre SQL - A linguagem padrão para trabalhar com banco de dados relacionais! (<https://www.alura.com.br/artigos/o-que-e-sql>)
- **Curso** Conhecendo SQL (<https://cursos.alura.com.br/formacao-conhecendo-sql>)
- **Artigo** SGBDs Relacionais (<https://www.alura.com.br/artigos/sgbds-relacionais>)
- **Artigo** SGBDs e suas linguagens internas
(<https://www.alura.com.br/artigos/sgbds-suas-linguagens-internas>)
- **Artigo** SQL: Comandos básicos (<https://www.alura.com.br/artigos/sql-comandos-basicos>)
- **Artigo** SQL: consultas com SELECT (<https://www.alura.com.br/artigos/sql-consultas-com-select>)
- **Artigo** SQL: cláusulas UPDATE e DELETE
(<https://www.alura.com.br/artigos/sql-clausulas-update-e-delete>)

- **Artigo** Trabalhando com funções de agregação (<https://www.alura.com.br/artigos/trabalhando-funcoes-de-agregacao>)
- **Artigo** JOIN e seus tipos (<https://www.alura.com.br/artigos/join-e-seus-tipos>)
- **Curso** SQL com MySQL Server da Oracle (<https://cursos.alura.com.br/formacao-oracle-mysql>)

Bancos de dados não relacionais:

- Os bancos de dados não relacionais, também conhecidos como bancos de dados NoSQL, são uma alternativa aos bancos de dados relacionais tradicionais. Eles são particularmente adequados para lidar com grandes volumes de dados estruturados e não estruturados e são capazes de escalar horizontalmente em vários servidores para suportar cargas de trabalho pesadas. Entre os bancos de dados NoSQL mais populares estão o MongoDB e o Datomic. Neste segmento, você aprenderá os fundamentos dos bancos de dados NoSQL, suas características e usos típicos.

Conteúdos

- **Site** AWS: O que é NoSQL? (<https://aws.amazon.com/pt/nosql/>)
- **Artigo** NoSQL — O que são? Onde habitam? Do que se alimentam? (<https://medium.com/permalink-univesp/nosql-o-que-s%C3%A3o-onde-habitam-do-que-se-alimentam-2ccd4cc39571>)
- **YouTube** Código Fonte TV: MongoDB (O Banco de Dados NoSQL mais Legal) (<https://www.youtube.com/watch?v=4dTI1mVLX3I>)
- **Artigo** Principais diferenças de um banco de dados tradicional e o MongoDB (<https://papito.medium.com/principais-diferen%C3%A7as-de-um-banco-de-dados-tradicional-e-o-mongodb-4fc1117453f8>)
- **YouTube** Ricardo Leme: Introdução ao NoSQL-MongoDB (<https://www.youtube.com/watch?v=kedLyo95fGU&list=PLyqIZW5s3wkoMhARQKp3s4YtaeKucb0Xj>)

Conteúdos Alura:

- **Artigo** MongoDB: o que é, quais suas características e benefícios e como trabalhar nessa ferramenta (<https://www.alura.com.br/artigos/mongodb>)
- **Artigo** Cassandra ou MongoDB, qual a melhor escolha para o meu projeto? (<https://www.alura.com.br/artigos/cassandra-ou-mongodb-qual-a-melhor-escolha-para-o-meu-projeto>)
- **Curso** MongoDB (<https://cursos.alura.com.br/formacao-mongodb>)
- **Curso** Datomic (<https://cursos.alura.com.br/formacao-datomic>)

Orquestração de Data Pipelines:

- A orquestração de data pipelines é um processo crucial na engenharia de dados que envolve o gerenciamento e a automação de fluxos de trabalho de dados. É através da orquestração que os pipelines de dados são configurados, programados e monitorados, garantindo a eficiência e a confiabilidade dos processos de extração, transformação e carga (ETL). Uma das ferramentas mais populares para orquestração de data pipelines é o Apache Airflow. Neste segmento, você aprenderá os fundamentos da orquestração de data pipelines, incluindo conceitos importantes, técnicas e melhores práticas.

Conteúdos

- **Artigo** Data Pipeline Orchestration (<https://towardsdatascience.com/data-pipeline-orchestration-9887e1b5eb7a>)
- **Artigo** Escolhendo uma tecnologia para orquestração de Dataworkflow (<https://johnosd.medium.com/escolhendo-uma-tecnologia-para-orquestra%C3%A7%C3%A3o-de-dataworkflow-b2ba99f6a497>)
- **Site** Apache Airflow (<https://airflow.apache.org/>)
- **Artigo** Apache Airflow: conceitos iniciais (<https://medium.com/datarisk-io/apache-airflow-conceitos-iniciais-e09c0dd18141>)
- **Artigo** Primeiros passos com o Apache Airflow: ETL fácil, robusto e de baixo custo (<https://medium.com/data-hackers/primeiros-passos-com-o-apache-airflow-etl-f%C3%A1cil-robusto-e-de-baixo-custo-f80db989edae>)

- **Artigo** Criando a primeira DAG no Airflow em 10 minutos (<https://estevestoni.medium.com/criando-a-primeira-dag-no-airflow-em-10-minutos-5b135618fe18>)
- **YouTube** Eduardo Mendes: Live de Python #123 - Apache Airflow com Juan Funez (<https://www.youtube.com/watch?v=LMk78Dmk0qU>)

Conteúdos Alura:

- **Artigo** Airflow - Entendendo os DAGs (<https://www.alura.com.br/artigos/airflow-entendendo-dags>)
- **Curso** Apache Airflow: orquestrando seu primeiro pipeline de dados (<https://cursos.alura.com.br/course/apache-airflow-primeiro-pipeline-dados>)
- **Curso** Formação Apache Airflow (<https://cursos.alura.com.br/formacao-apache-airflow>)

Contêineres:

- Os contêineres são pacotes de software que contêm todos os elementos necessários para serem executados em qualquer ambiente. Gerenciamento de contêineres é uma área crucial na computação em nuvem e DevOps, que envolve o uso de tecnologias para automatizar o processo de criação, implantação, escalonamento e monitoramento de contêineres. Contêineres são unidades de software padronizadas que permitem aos desenvolvedores empacotar todas as dependências de um aplicativo (código, bibliotecas, configurações, etc.) em um único pacote. Isso permite que o aplicativo seja executado de forma consistente em qualquer ambiente de infraestrutura.
- A tecnologia de contêineres, como exemplificada pelo Docker, fornece um ambiente consistente e portátil para desenvolvimento, teste e implantação de aplicativos, o que é vital para o trabalho eficiente de engenharia de dados. Além disso, o Kubernetes, um sistema de orquestração de contêineres, permite o gerenciamento, a automação e a escalabilidade de aplicações baseadas em contêineres em ambientes de produção. Dominar esses conceitos e tecnologias possibilita a engenheiros de dados construir e manter pipelines de dados eficientes e confiáveis.

- O Kubernetes (também conhecido como k8s ou kube) é uma plataforma de orquestração de containers open source que automatiza grande parte dos processos manuais necessários para implantar, gerenciar e escalar aplicações em containers.
- Isolar seu software para funcionar independentemente
- Implantar software em clusters
- Modularizar seu sistema em pacotes menores
- Conhecer a plataforma Docker
- Conhecer Kubernetes

Conteúdos

- **Site** IBM: Contêineres (<https://www.ibm.com/br-pt/cloud/learn/containers>)
- **Site** HPE: O que são contêineres? (<https://www.hpe.com/br/pt/what-is/containers.html>)
- **Site** Google Cloud: O que são contêineres? (<https://cloud.google.com/learn/what-are-containers?hl=pt-br>)
- **Site** Azure: O que é um contêiner? (<https://azure.microsoft.com/pt-br/resources/cloud-computing-dictionary/what-is-a-container/>)
- **Artigo** Red Hat: O que é Kubernetes (<https://www.redhat.com/pt-br/topics/containers/what-is-kubernetes>)
- **Artigo** Azure: O que é Kubernetes? (<https://azure.microsoft.com/pt-br/topic/what-is-kubernetes/>)
- **Artigo** Kubernetes - Guia de Referência Gratuito (<https://renatogroffe.medium.com/kubernetes-guia-de-refer%C3%A2ncia-gratuito-a2c1be8e7593>)
- **YouTube** Código Fonte TV: Containers // Dicionário do Programador (<https://www.youtube.com/watch?v=-pUZBovqRcU>)
- **YouTube** Código Fonte TV: Kubernetes // Dicionário do Programador (<https://www.youtube.com/watch?v=mVL0nOM3AGo>)
- **YouTube** Insight Lab: Revisão Docker - Curso de Introdução ao Kubernetes (<https://www.youtube.com/watch?v=bcRArpK00OU>)

- **YouTube** Pisani da Arch: Kubernetes - entenda e comece a arquitetar agora mesmo (<https://www.youtube.com/watch?v=XjaqTqm7g58>)

Conteúdos Alura:

- **Artigo** O que é Kubernetes (<https://www.alura.com.br/artigos/o-que-e-kubernetes>)
- **Artigo** Criando volumes com Docker (<https://www.alura.com.br/artigos/criando-volumes-com-docker>)
- **Artigo** Docker Compose para compor uma aplicação (<https://www.alura.com.br/artigos/compondo-uma-aplicacao-com-o-docker-compose>)
- **YouTube** Alura: O que são Containers? (https://www.youtube.com/watch?v=jv4_sLIGOS0)
- **YouTube** Alura: Containers, Docker e Kubernetes (com Giovanni Bassi) (<https://www.youtube.com/watch?v=wxLvMxzc1Q>)
- **Podcast** Hipsters.tech: Containers e Docker - Hipsters 75 (<https://www.alura.com.br/podcast/hipsterstech-containers-e-docker-hipsters-75-a511>)
- **Podcast** Hipsters.tech: Kubernetes - Hipsters 177 (<https://www.alura.com.br/podcast/kubernetes-hipsters-177-a389>)
- **Podcast** Hipsters.tech: Microsserviços e Kubernetes na Americanas.com - Hipsters On The Road 56 (<https://www.alura.com.br/podcast/microsservicos-e-kubernetes-na-americanas-com-hipsters-on-the-road-56-a857>)
- **Curso** Formação Orquestração de containers com Kubernetes (<https://www.alura.com.br/formacao-kubernetes>)
- **Curso** Curso Docker: criando e gerenciando containers (<https://www.alura.com.br/curso-online-docker-criando-gerenciando-containers>)
- **Curso** Curso Kubernetes: Pods, Services e ConfigMaps (<https://www.alura.com.br/curso-online-kubernetes-pods-services-configmap>)

- **Curso** Curso Kubernetes: Deployments, Volumes e Escalabilidade (<https://www.alura.com.br/curso-online-kubernetes-deployments-volumes-escalabilidade>)
- **Curso** Curso Google Kubernetes Engine (<https://www.alura.com.br/curso-online-google-kubernetes-engine-primeiros-passos>)
- **Curso** Curso Kubernetes na Digital Ocean: gerenciando aplicações containerizadas (<https://www.alura.com.br/curso-online-kubernetes-digital-ocean>)
- **Curso** Curso Amazon EKS: gerenciando aplicações containerizadas com Kubernetes (<https://www.alura.com.br/curso-online-amazon-eks-kubernetes>)
- **Curso** Curso Azure: gerencie Kubernetes com AKS e ACR (<https://www.alura.com.br/curso-online-azure-gerencie-kubernetes-com-aks-acr>)

Data Warehouses:

- Um Data Warehouse é um sistema de armazenamento de dados projetado para análise de negócios, suportando processamento de transações e consultas complexas. Ele permite a consolidação de dados de várias fontes, fornecendo um ambiente unificado para análise de dados. Com recursos de normalização e limpeza de dados, garante a qualidade e a consistência dos dados armazenados. Soluções de Data Warehouse populares incluem Google BigQuery, Amazon Redshift e Snowflake, cada uma com suas próprias características e benefícios únicos.
- Data Warehouses desempenham um papel crucial na engenharia de dados. Eles fornecem uma infraestrutura que facilita a gestão, o processamento e a análise de grandes volumes de dados. Sua capacidade de armazenar e processar grandes quantidades de dados de várias fontes permite análises avançadas e suporte à tomada de decisões, o que é fundamental no atual cenário empresarial orientado por dados. Os engenheiros de dados projetam, constroem e mantêm data warehouses para garantir que os dados estejam acessíveis, seguros e prontos para análise. Além disso, soluções modernas de data warehouse, como o BigQuery, o Amazon Redshift e o

Snowflake, oferecem capacidades de escalabilidade e flexibilidade para lidar com o crescente volume, velocidade e variedade de dados, tornando-se ainda mais valiosos para a engenharia de dados.

Conteúdos

- **Site** AWS: O que é um data warehouse? (<https://aws.amazon.com/pt/data-warehouse/>)
- **Site** Oracle: O que é um data warehouse? (<https://www.oracle.com/br/database/what-is-a-data-warehouse/>)
- **Artigo** Aprofundando em Data Warehouse... (<https://medium.com/@aasouzaconsult/aprofundando-em-data-warehouse-65ed2bca9a33>)
- **Site** O que é o BigQuery? (<https://cloud.google.com/bigquery/docs/introduction?hl=pt-br>)
- **Site** GCP: O que é o BigQuery? (<https://cloud.google.com/bigquery/docs/introduction?hl=pt-br>)
- **Site** AWS: O que é o Amazon Redshift? (https://docs.aws.amazon.com/pt_br/redshift/latest/mgmt/welcome.html)
- **Site** Snowflake: Conceitos-chave e arquitetura? (<https://docs.snowflake.com/pt/user-guide/intro-key-concepts>)
- **YouTube** Tutorial: Construindo seu primeiro Data Warehouse na Nuvem (<https://www.youtube.com/watch?v=Aeg8o5xJoN8>)

Conteúdos Alura:

- **Curso** Formação BigQuery (<https://cursos.alura.com.br/formacao-bigquery>)
- **Curso** Formação BI e Data Warehouse com SQL Server e Power BI (<https://cursos.alura.com.br/formacao-business-intelligence-data-warehouse>)
- **Curso** Data Warehouse com Data Marts e Power BI (<https://cursos.alura.com.br/formacao-data-warehouse-data-marts-power-bi>)

- **Curso** "Data Warehouse com Data Lake e SQL Server"
(<https://cursos.alura.com.br/formacao-business-intelligence-data-warehouse-data-lake-sql-server>)

Data Lakes:

- Um Data Lake é uma arquitetura de armazenamento de dados que permite o armazenamento de grandes volumes de dados brutos em seu formato original, provenientes de diversas fontes e em vários formatos, como estruturado, semiestruturado e não estruturado. Diferentemente do Data Warehouse, que exige que os dados sejam limpos e estruturados antes de serem armazenados, o Data Lake permite o armazenamento de dados "como estão", proporcionando flexibilidade para transformar e processar os dados conforme necessário para fins de análise.
- Data Lakes são fundamentais para a engenharia de dados devido à sua capacidade de lidar com a crescente variedade, volume e velocidade de dados gerados no mundo digital atual. Eles fornecem um meio econômico e escalável para armazenar e analisar grandes volumes de dados de diversas fontes, o que é fundamental para análises avançadas, como aprendizado de máquina e análise preditiva. Os engenheiros de dados são responsáveis pela construção e manutenção de Data Lakes, garantindo a ingestão de dados de maneira eficiente, o gerenciamento seguro dos dados e a implementação de processos de governança de dados. Além disso, os Data Lakes suportam a democratização dos dados, permitindo que diferentes partes interessadas, desde cientistas de dados até analistas de negócios, acessem os dados necessários para gerar insights valiosos.

Conteúdos

- **Site** AWS: O que é um data lake? (<https://aws.amazon.com/pt/big-data/datalakes-and-analytics/what-is-a-data-lake/>)
- **YouTube** Código Fonte TV: Data Warehouse / Data Lake / Data Lakehouse
(<https://www.youtube.com/watch?v=V07Pk4de-5M>)
- **Artigo** Data Lake: O componente essencial em uma Estratégia de Dados
(<https://www.linkedin.com/pulse/data-lake-o-componente-essencial-em-uma-estrat%C3%A9gia-de-brito-ruiz/?originalSubdomain=pt>)

- **Artigo** Conceito de Datalake e armazenamentos possíveis (<https://medium.com/rescuepoint/conceito-de-datalake-e-armazenamentos-poss%C3%ADveis-65f337924fc8>)
- **Artigo** Data Lake x Data Warehouse: Onde armazenar seus dados? (<https://medium.com/datalakers-blog/data-lake-x-data-warehouse-onde-armazenar-seus-dados-67534c9dcfad>)
- **YouTube** Seu primeiro passo no mundo analítico - Criando o primeiro datalake na AWS (<https://www.youtube.com/watch?v=sVFT-XK6rk4>)

Conteúdos Alura:

- **YouTube** Alura: O que são Data Lakes? (<https://www.youtube.com/watch?v=GkhuRcHsVMk>)
- **Artigo** Data Lake vs Data Warehouse (<https://www.alura.com.br/artigos/data-lake-vs-data-warehouse>)
- **Podcast** Hipsters.tech: Data Lakes (<https://www.hipsters.tech/data-lakes-hipsters-ponto-tech-269/>)
- **Curso** Formação Engenharia de dados com Databricks (<https://cursos.alura.com.br/formacao-databricks>)
- **Site** Hotmart e a Democratização de dados em um cenário de microsserviços (<https://cursos.alura.com.br/extra/cases/hotmart-e-a-democratizacao-de-dados-em-um-cenario-de-microservicos-z1066>)
- **Site** Banco Pan e a Jornada de construção de um Data Lake (<https://cursos.alura.com.br/extra/cases/banco-pan-e-a-jornada-de-construcao-de-um-data-lake-z1284>)
- **Curso** "Data Warehouse com Data Lake e SQL Server" (<https://cursos.alura.com.br/formacao-business-intelligence-data-warehouse-data-lake-sql-server>)

Cloud - Fundamentos:

- Cloud, ou computação em nuvem é a distribuição de serviços de computação pela Internet usando um modelo de preço pago conforme o uso. Uma nuvem é composta de vários recursos de computação, que abrangem desde os próprios computadores (ou instâncias, na terminologia

de nuvem) até redes, armazenamento, bancos de dados e o que estiver em torno deles. Ou seja, tudo o que normalmente é necessário para montar o equivalente a uma sala de servidores, ou mesmo um data center completo, estará pronto para ser utilizado, configurado e executado.

- Conhecer a diferença entre IaaS, PaaS e SaaS
- Conhecer os maiores provedores de cloud
- Especializar-se em algum provedor

Conteúdos

- **Artigo** Computação em nuvem (<https://medium.com/sysadminas/computa%C3%A7%C3%A3o-em-nuvem-515930304cf9>)
- **Artigo** O que é cloud? (<https://gabriel-faraday.medium.com/o-que-%C3%A9-cloud-991109e708c6>)
- **YouTube** Gabs Ferreira: Por que investir e estudar cloud? (<https://www.youtube.com/watch?v=Z45BTNeZ1I0>)
- **YouTube** Andre Iacono: O que é MICROSOFT AZURE? Qual Certificação começar em 2022? (<https://www.youtube.com/watch?v=f-oVzkvMwnE>)
- **YouTube** AWS: O que é a AWS? (<https://www.youtube.com/watch?v=8JI9wQ8sUdQ>)
- **YouTube** O que é Google Cloud e por que aprender? (<https://www.youtube.com/shorts/Lzq3f1DHWcl>)
- **Artigo** AWS vs Google Cloud vs Azure: o que cada um tem de melhor? (<https://medium.com/data-hackers/aws-vs-google-cloud-vs-azure-o-que-cada-um-tem-de-melhor-52107174f7b7>)
- **YouTube** Código Fonte TV: Azure (A plataforma Cloud da Microsoft) (<https://www.youtube.com/watch?v=YgE-sZaCuJ0>)
- **YouTube** Mundo da Cloud: AWS do Zero ao Expert (<https://www.youtube.com/watch?v=HiBCv9DolxI&list=PLtL97Owd1gkQ0dfqGW8OtJ-155Gs67Ecz>)

Conteúdos Alura:

- **Podcast** Hipsters.tech: TechGuide - Fundamentos Cloud – Hipsters Ponto Tech #348 (<https://www.hipsters.tech/techguide-fundamentos-cloud-hipsters-ponto-tech-348/>)
- **Artigo** Cloud: o que é, História e Guia da computação em nuvem (<https://www.alura.com.br/artigos/cloud>)
- **Podcast** Hipsters.tech: Uma jornada Para o Cloud - Hipsters Deep Dive 005 (<https://www.alura.com.br/podcast/uma-jornada-para-o-cloud-hipsters-deep-dive-005-a1100>)
- **Podcast** Hipsters.tech: Histórias do Cloud - Hipsters 04 (<https://www.alura.com.br/podcast/historias-do-cloud-hipsters-04-a582>)
- **Artigo** Heroku, Vercel e outras opções de cloud como plataforma (<https://www.alura.com.br/artigos/heroku-vercel-outras-opcoes-cloud-plataforma>)
- **Artigo** AWS: Guia sobre o que é Amazon Web Services, seus Serviços e Certificações (<https://www.alura.com.br/artigos/aws>)
- **Artigo** Terraform: criando máquinas na Azure (<https://www.alura.com.br/artigos/terraform-maquinas-na-azure>)
- **YouTube** Alura: O que é cloud? (<https://www.YOUTUBE.com/watch?v=wev9fMrg-TU>)
- **YouTube** Alura: AWS, Google Cloud e Azure: Por onde começar? | Hipsters.Talks (<https://www.YOUTUBE.com/watch?v=z9k6rsdmWc0&t=300s>)
- **YouTube** Alura: Certificação em Cloud: Azure, AWS, Google | Hipsters.Talks (https://www.YOUTUBE.com/watch?v=W4K82n_WK5g&t=290s)
- **Curso** Formação Começando em Cloud Computing (<https://cursos.alura.com.br/formacao-cloud-computing>)
- **Curso** Formação Amazon Web Services (<https://cursos.alura.com.br/formacao-amazon-web-services>)
- **Curso** Formação Google Certified Associate Cloud Engineer (<https://cursos.alura.com.br/formacao-google-certified-associate-cloud->

[engineer\)](#)

- **Curso** Formação Certificação AWS Certified Cloud Practitioner (<https://cursos.alura.com.br/formacao-aws-certified-cloud-practitioner>)
- **Curso** Formação Containers com AWS ECS e EKS (<https://cursos.alura.com.br/formacao-containers-aws>)
- **Curso** Formação Google Cloud Platform (<https://cursos.alura.com.br/formacao-google-cloud>)
- **Curso** Formação Certificação Google Certified Associate Cloud Engineer (<https://cursos.alura.com.br/formacao-google-certified-associate-cloud-engineer>)
- **Curso** Formação Azure (<https://cursos.alura.com.br/formacao-conhecendo-azure>)
- **Curso** Formação Certificação AZ-900: Microsoft Azure Fundamentals (<https://cursos.alura.com.br/formacao-certificacao-az-900-microsoft-azure-fundamentals>)

Redes de Computadores - Fundamentos:

- Rede de computadores é uma malha que interliga milhares de sistemas computacionais para a transmissão de dados. Também conhecidos como nós, esses dispositivos interconectados enviam, recebem e trocam tráfego de dados, voz e vídeo, graças ao hardware e software que compõe o ambiente.
- Configurações de redes são essenciais para que seja possível acessar uma aplicação, principalmente se ela estiver na nuvem.
- Entenda melhor os componentes de rede e a suas atribuições
- Diferenciar os serviços disponibilizados a nível de rede
- Saber as diferenças entre as camadas de rede
- Saber o que é servidor web, proxy reverso e load balancer

Conteúdos

- **Site** IBM: Conceitos de rede e comunicação (<https://www.ibm.com/docs/pt-br/aix/7.3?topic=management-network-communication-concepts>)

- **Site** AWS: O que são redes de computadores?
(<https://aws.amazon.com/pt/what-is/computer-networking/>)
- **Site** Red Hat: O que é gerenciamento de rede? (<https://www.redhat.com/pt-br/topics/management/what-is-network-management>)
- **Artigo** Redes de Computadores — Sabendo o básico — Intro
(<https://medium.com/code-breaker/redes-de-computadores-sabendo-o-b%C3%A1sico-intro-1a7d5653e118>)
- **Artigo** O protocolo TCP/IP (<https://joaomarcuraa.medium.com/o-protocolo-tcp-ip-1dc2cdb88b07>)
- **YouTube** Código Fonte TV: IP (Entenda finalmente como funciona o IPv4 e o IPv6) | Dicionário do Programador (<https://www.youtube.com/watch?v=O8DmpmBMUSw>)
- **YouTube** Rede de Computadores - Animação em 3D
(<https://www.youtube.com/watch?v=lqcp3k8DgGw>)
- **Curso** Enterprise Networking, Security, and Automation
(<https://www.netacad.com/pt-br/courses/networking/ccna-enterprise-networking-security-automation>)
- **Curso** Switching, Routing, and Wireless Essentials
(<https://www.netacad.com/pt-br/courses/networking/ccna-switching-routing-wireless-essentials>)
- **Curso** Introduction to Networks (<https://www.netacad.com/pt-br/courses/networking/ccna-introduction-networks>)

Conteúdos Alura:

- **Artigo** Rede de computadores: um guia completo sobre o que são, como funcionam e quais cursos fazer na área
(<https://www.alura.com.br/artigos/rede-de-computadores>)
- **Artigo** Conhecendo algumas topologias de rede
(<https://www.alura.com.br/artigos/conhecendo-algumas-topologias-de-rede>)
- **Artigo** Como calcular máscaras de sub-rede de forma simples
(<https://www.alura.com.br/artigos/como-calcular-mascaras-de-sub-rede>)

- **Artigo** Entendendo os padrões de Wi-Fi
(<https://www.alura.com.br/artigos/entendendo-os-padroes-de-wi-fi>)
- **Artigo** Diferenças entre Hubs e Switches
(<https://www.alura.com.br/artigos/diferencas-entre-hubs-e-switches>)
- **Artigo** Entendendo os cabos de rede
(<https://www.alura.com.br/artigos/entendendo-os-cabos-de-rede>)
- **Artigo** SSH, Telnet e as diferenças para conectar em um servidor
(<https://www.alura.com.br/artigos/entendendo-as-diferencas-entre-telnet-e-ssh>)
- **Curso** Formação Redes de Computadores
(<https://cursos.alura.com.br/formacao-redes>)

Nível 3

Deploy de Data Pipelines:

- O deploy de data pipelines envolve a implantação e execução de fluxos automatizados para coletar, transformar e carregar dados de diferentes fontes para um destino específico. Ele garante a disponibilidade contínua dos dados para análise e outras aplicações, incluindo agendamento, monitoramento de desempenho e resolução de problemas durante o processo. O objetivo do deploy de data pipelines é assegurar a ingestão confiável de dados para suportar análises, relatórios e tomada de decisões informadas.
- Na engenharia de dados, o deploy de data pipelines desempenha um papel crucial ao garantir a disponibilidade, fluxo contínuo, escalabilidade, desempenho e monitoramento dos dados, proporcionando dados confiáveis e atualizados para impulsionar insights e ações.

Conteúdos

- **Artigo** How to bring your modern data pipeline to production
(<https://towardsdatascience.com/how-to-bring-your-modern-data-pipeline-to-production-2f14e42ac200>)
- **Site** AWS Data Pipeline (<https://aws.amazon.com/pt/datapipeline/>)

- **Site** Amazon MWAA (<https://aws.amazon.com/pt/managed-workflows-for-apache-airflow/>)
- **YouTube** O que é e quando usar o Azure Data Factory (<https://www.youtube.com/watch?v=Y Ez5jRGDtbk>)
- **YouTube** Transforme e catalogue dados com o AWS Glue (<https://www.youtube.com/watch?v=HFFiAy2J2OQ>)
- **YouTube** Best Practices for Building and Deploying Data Pipelines in Apache Spark - Vicky Avison (<https://youtu.be/1WUlua-xjJA>)
- **YouTube** Escritório de Dados: Aula 04 - Fazendo deploy de pipelines (visão geral) (<https://www.youtube.com/watch?v=g-Jl8tqQojY>)
- **YouTube** Escritório de Dados: Aula 05 - Fazendo deploy de pipelines (mais exemplos) (<https://www.youtube.com/watch?v=jNivNlva2sU>)
- **Curso** AWS Data Lake: construindo pipelines na AWS (<https://cursos.alura.com.br/formacao-aws-data-lake-construindo-pipelines-aws>)
- **Curso** Azure Data Lake: criando um pipeline de ingestão de dados (<https://cursos.alura.com.br/course/azure-data-lake-pipeline-ingestao-dados>)

Gerenciamento de logs:

- O gerenciamento de logs é o processo de coleta, armazenamento, monitoramento e análise de registros de eventos gerados por sistemas, aplicativos e dispositivos. Esses registros de eventos, conhecidos como logs, contêm informações importantes sobre o funcionamento e o desempenho do sistema, bem como eventos relevantes para fins de segurança e auditoria.
- O gerenciamento de logs desempenha um papel crucial na engenharia de dados, pois fornece insights valiosos para a monitorização, resolução de problemas, otimização de desempenho e análise de sistemas e aplicativos. Os logs podem ser usados para identificar e solucionar problemas, detectar atividades suspeitas ou não autorizadas, rastrear eventos relevantes e analisar o comportamento do sistema ao longo do tempo. Além disso, o gerenciamento de logs é essencial para fins de conformidade, auditoria e

conformidade regulatória, permitindo o acompanhamento e a documentação adequados das atividades relacionadas aos dados. Para isso, são utilizadas ferramentas e técnicas para coletar, centralizar, analisar e visualizar os logs, tornando-os mais eficientes e eficazes para a engenharia de dados.

Conteúdos

- **YouTube** Linhagem de dados e o poder da Alvin
(<https://www.youtube.com/watch?v=yKf1IR1DZzY>)
- **YouTube** Afinal o que é Observabilidade? Será que é igual a monitoração?
(<https://youtu.be/feZskFGTeRA>)
- **Artigo** What Is Data Monitoring?
(<https://www.egnyte.com/guides/governance/data-monitoring>)

Conteúdos Alura:

- **Curso** Observabilidade: coletando métricas de uma aplicação com Prometheus (<https://cursos.alura.com.br/course/observabilidade-prometheus>)
- **Curso** Monitoramento: Prometheus, Grafana e Alertmanager
(<https://cursos.alura.com.br/course/monitoramento-prometheus-grafana-alertmanager>)

Data Mesh:

- O Data Mesh promove uma abordagem descentralizada na arquitetura de dados, capacitando cada equipe ou domínio a ser responsável por seus próprios dados. Isso resulta em maior autonomia, colaboração e eficiência na gestão dos dados.
- A importância do Data Mesh na engenharia de dados está relacionada à descentralização da responsabilidade pelos dados. Essa abordagem promove a colaboração, agilidade e melhoria da qualidade e confiabilidade dos dados na organização.

Conteúdos

- **YouTube** Desafios na jornada ao Data Mesh - Data Hackers + Thoughtworks
(<https://www.youtube.com/watch?v=ONJWjO1H3F4>)

- **YouTube** Data Mesh - Thoughtworks no Agile Brazil 2021
(<https://www.youtube.com/watch?v=spS5qrDS8Ns>)
- **Artigo** Data Mesh Principles and Logical Architecture
(<https://martinfowler.com/articles/data-mesh-principles.html>)

Conteúdos Alura:

- **Podcast** Datamesh com a ThoughtWorks
(<https://cursos.alura.com.br/extra/hipsterstech/datamesh-com-a-thoughtworks-hipsters-ponto-tech-292-a1418>)
- **Curso** Formação Data Mesh (<https://cursos.alura.com.br/formacao-data-mesh>)

Streaming:

- Streaming é o processamento e transmissão contínua de dados em tempo real, à medida que são gerados. Diferentemente do armazenamento tradicional, onde os dados são coletados e processados posteriormente, o streaming permite a análise e o uso imediato dos dados à medida que são recebidos. Essa abordagem é ideal para lidar com grandes volumes de dados em alta velocidade, permitindo a detecção de padrões em tempo real, tomada de decisões instantâneas e resposta rápida a eventos em andamento.
- O streaming desempenha um papel essencial na engenharia de dados, capacitando as organizações a processar, analisar e tomar decisões com base em dados em tempo real. Isso resulta em maior agilidade, escalabilidade e reatividade, possibilitando uma tomada de decisões mais informada e orientada por insights atualizados.

Conteúdos

- **Site** O que são fluxos de dados dados em tempo real?
(<https://aws.amazon.com/pt/what-is/real-time-data-streaming/>)
- **YouTube** Kafka (Plataforma de Mensageria e Streaming)
(https://www.youtube.com/watch?v=qOqXz5Qv_-8)

- **YouTube** Processando 1bi de mensagens com open source (<https://www.youtube.com/watch?v=SEBv4N595eY>)
- **YouTube** Streaming de Dados (<https://www.youtube.com/watch?v=PaUeW5a-PTs>)

Conteúdos Alura:

- **Artigo** Spark Streaming (<https://www.alura.com.br/artigos/spark-streaming>)
- **Curso** Kafka: Produtores, Consumidores e streams (<https://cursos.alura.com.br/course/kafka-introducao-a-streams-em-microservicos>)
- **Curso** Formação Mensageria com Apache Kafka (<https://cursos.alura.com.br/formacao-kafka>)

Proteção de dados:

- A proteção de dados é o processo de proteger informações importantes de forma que garanta a confidencialidade, integridade e a disponibilidade destes dados.

Conteúdos

- **YouTube** ENTENDA O QUE É LGPD (<https://www.youtube.com/watch?v=dPipsp5arAc>)
- **YouTube** Tudo que você precisa saber sobre a LGPD (<https://www.youtube.com/watch?v=hu6Xlc7QVnE>)
- **YouTube** O que você precisa saber sobre o GDPR (<https://www.youtube.com/watch?v=egGvTgk3HH8>)
- **YouTube** DLP – Data Loss Prevention (<https://www.youtube.com/watch?v=VHeX3D2WOUc>)
- **Artigo** Como estudei para certificação Privacy and Data Protection Foundation (PDPF) (<https://caiovicentinni.medium.com/como-estudei-para-certifica%C3%A7%C3%A3o-privacy-and-data-protection-foundation-pdpf-81a824aa9f5e>)

- **Artigo** Com a Lei em vigor, quais as diferenças entre a LGPD e GDPR? (<https://medium.com/@rogosantos/com-a-lei-em-vigor-quais-as-diferen%C3%A7as-entre-a-lgpd-e-gdpr-8e6d008c898d>)
- **Artigo** Como a LGPD pode impactar na Experiência do Usuário (UX) (<https://brasil.uxdesign.cc/como-a-lei-geral-de-prote%C3%A7%C3%A3o-de-dados-pode-afetar-a-experi%C3%Aancia-do-usu%C3%A1rio-e514270bf59c>)

Infraestrutura como código (IaC):

- Infraestrutura como código (IaC) é uma abordagem em que a infraestrutura de TI, incluindo servidores, redes e recursos relacionados, é gerenciada e provisionada por meio de código. Em vez de configurar manualmente a infraestrutura, o IaC utiliza arquivos de configuração ou scripts que descrevem a infraestrutura desejada de forma automatizada e reproduzível.
- A infraestrutura como código é de extrema importância na engenharia de dados, pois permite uma gestão eficiente e escalável dos recursos de infraestrutura necessários para processamento e armazenamento de dados. Com o IaC, é possível definir, provisionar e configurar ambientes de maneira consistente, rápida e controlada. Isso facilita a implantação e manutenção de pipelines de dados, ambientes de desenvolvimento e testes, clusters de processamento distribuído e outros componentes necessários para a engenharia de dados. Além disso, o IaC permite a automação e o versionamento da infraestrutura, proporcionando maior agilidade, rastreabilidade e garantia de qualidade no gerenciamento dos recursos de TI.

Conteúdos

- **Site** Red Hat: O que é infraestrutura como código (IaC)? (<https://www.redhat.com/pt-br/topics/automation/what-is-infrastructure-as-code-iac>)
- **YouTube** O que é Infrastructure as code? (IaC) Infraestrutura como código (https://youtu.be/_iho2OK_LDg)
- **Artigo** IaC — Infraestrutura como código (<https://medium.com/sysvale/iac-infraestrutura-como-c%C3%B3digo-c514a869b88d>)

- **YouTube** Código Fonte TV: Terraform em 10 Minutos // Dicionário do Programador (<https://www.youtube.com/watch?v=0EAjJe8aPkc>)
- **YouTube** Terraform: infraestrutura como código (<https://www.youtube.com/watch?v=tE1WZg9ib8k>)

Conteúdos Alura:

- **Curso** Curso Infraestrutura como código: preparando máquinas na AWS com Ansible e Terraform (<https://cursos.alura.com.br/course/infraestrutura-codigo-maquinas-aws-ansible-terraform>)
- **Curso** Formação DevOps (<https://cursos.alura.com.br/formacao-devops>)
- **Curso** Formação Infraestrutura como código (<https://cursos.alura.com.br/formacao-infraestrutura-codigo>)

Governança de Dados:

- Governança de dados é um conjunto de políticas, processos e práticas que estabelecem a responsabilidade, a integridade, a qualidade e o uso adequado dos dados em uma organização. Ela envolve a definição de regras e padrões para a coleta, armazenamento, gerenciamento, compartilhamento e uso dos dados, garantindo que sejam confiáveis, consistentes e seguros ao longo de sua vida útil.
- A governança de dados desempenha um papel fundamental na engenharia de dados, pois garante que os dados sejam gerenciados de forma eficiente, confiável e em conformidade com as regulamentações e políticas internas. Ela estabelece diretrizes para a gestão dos dados, incluindo a definição de metadados, a identificação de proprietários de dados, a documentação de políticas de acesso e privacidade, a implementação de medidas de segurança e a garantia da qualidade dos dados.
- A governança de dados também promove a colaboração entre as equipes, estabelecendo processos de tomada de decisão baseados em dados confiáveis e padronizados. Isso resulta em uma melhor qualidade das análises, uma base sólida para tomada de decisões estratégicas e uma maior confiança nos dados utilizados para impulsionar as iniciativas de engenharia de dados.

Conteúdos

- **Site** Azure: O que é governança de dados? (<https://azure.microsoft.com/pt-br/resources/cloud-computing-dictionary/what-is-a-data-governance/>)
- **YouTube** Cultura de qualidade e governança de dados no Grupo Boticário (<https://www.youtube.com/watch?v=LTxHPUykq6k>)
- **YouTube** A política de governança de dados no Brasil (<https://www.youtube.com/watch?v=Ac9J19C5x30>)
- **Artigo** Cartilha orienta sobre governança de dados (<https://www.gov.br/economia/pt-br/assuntos/noticias/2022/dezembro/cartilha-orienta-sobre-governanca-de-dados>)

Conteúdos Alura:

- **Artigo** O que é Governança de Dados? (<https://www.alura.com.br/artigos/o-que-e-governanca-de-dados>)
- **Curso** "LGPD, conhecendo e entendendo seus impactos" (<https://cursos.alura.com.br/course/lgpd-visao-impactos>)
- **Curso** "Governança de TI, gestão de demandas de serviços" (<https://cursos.alura.com.br/course/governanca-ti-demandas-servicos>)

Entrega e integração contínuas (CI/CD):

- CI/CD é a abreviação de Continuous Integration/Continuous Delivery, traduzindo para o português "entrega e integração contínuas". Trata-se de uma prática de desenvolvimento de software que visa tornar a integração de código mais eficiente por meio de builds e testes automatizados.
- Automatizar a integração de código entre varias partes da equipe se tornou cada vez mais importante, ja que assim é possível acelerar o desenvolvimento e diminuir o tempo de entrega de software.
- Executar testes automatizados da aplicação para verificar seu funcionamento.
- Realizar a entrega de atualizações de forma automática e com segurança.

- Realizar testes de conexão e testes de carga para evitar que a aplicação apresente problemas ao ser atualizada.

Conteúdos

- **Artigo** O que é CI/CD? (<https://www.redhat.com/pt-br/topics/devops/what-is-ci-cd>)
- **Artigo** Sobre integração contínua (<https://docs.github.com/pt/actions/automating-builds-and-tests/about-continuous-integration>)
- **YouTube** Código Fonte TV: Pipeline (CD/CI) | Dicionário do Programador (<https://www.youtube.com/watch?v=AZtTd3pFVTY>)

Conteúdos Alura:

- **Artigo** Entrega e integração contínua de aplicações Vue (<https://www.alura.com.br/artigos/entrega-e-integracao-continua-de-aplicacoes-vue>)
- **Artigo** GitOps: o que é? (<https://www.alura.com.br/artigos/o-que-e-git-ops>)
- **Artigo** Tipos de testes: quais os principais e por que utilizá-los? (<https://www.alura.com.br/artigos/tipos-de-testes-principais-por-que-utiliza-los>)
- **Podcast** Integração Contínua, Deploy Contínuo e Github Actions (<https://cursos.alura.com.br/extra/hipsterstech/integracao-continua-deploy-continuo-e-github-actions-hipsters-213-a335>)
- **Podcast** Testes Automatizados (<https://cursos.alura.com.br/extra/hipsterstech/testes-automatizados-hipsters-51-a535>)
- **Curso** Curso Integração Contínua: mais qualidade e menos risco no desenvolvimento (<https://cursos.alura.com.br/course/desenvolvimento-software-integracao-continua>)
- **Curso** Curso Entrega Contínua: confiabilidade e qualidade na implantação de software (<https://cursos.alura.com.br/course/entrega-continua-confiabilidade-qualidade>)

- **Curso** Formação: Integração Contínua e Entrega Contínua (<https://cursos.alura.com.br/formacao-integracao-continua-entrega-continua>)

Habilidade Auxiliar: Arquitetura de software

Criação e manipulação de APIs:

- Criação e manipulação de APIs envolvem o desenvolvimento e interação com Interfaces de Programação de Aplicativos. APIs são conjuntos de regras que permitem que diferentes sistemas se comuniquem e compartilhem dados. A criação de APIs implica em projetar endpoints que permitem o acesso e interação com recursos específicos, enquanto a manipulação envolve utilizar APIs existentes para realizar tarefas específicas, como autenticação, envio de solicitações e processamento de dados retornados. Essas práticas são cruciais para a integração eficiente de sistemas e aplicativos em diferentes áreas de desenvolvimento.
- A criação e manipulação de APIs são fundamentais na engenharia de dados, permitindo a troca eficiente de dados entre sistemas, a integração com serviços externos e a colaboração entre equipes. As APIs simplificam a coleta, processamento e distribuição de informações, ajudando a construir pipelines de dados robustos e escaláveis.

Conteúdos

- **Site** Red Hat: O que é API? (<https://www.redhat.com/pt-br/topics/api/what-are-application-programming-interfaces>)
- **YouTube** Código Fonte TV: API // Dicionário do Programador (<https://www.youtube.com/watch?v=vGuqKIRWosk>)
- **Artigo** Como CRIAR a sua PRIMEIRA API em python com o FastAPI (<https://medium.com/data-hackers/como-criar-a-sua-primeira-api-em-python-com-o-fastapi-50b1d7f5bb6d>)
- **Artigo** Criação de uma API Rest com Python (<https://nataniel-paiva.medium.com/cria%C3%A7%C3%A3o-de-uma-api-rest-com-python-76696d17bfb9>)

- **YouTube** Eduardo Mendes: API REST - Live de Python
(https://www.youtube.com/watch?v=1_nQ5A2HcgU)

Conteúdos Alura:

- **Curso** Formação Django REST APIs: crie aplicações REST em Python
(<https://cursos.alura.com.br/formacao-django-rest>)
- **Curso** Formação Começando com Flask: framework web de Python
(<https://cursos.alura.com.br/formacao-flask>)
- **Artigo** Django ou Flask: características, semelhanças e diferenças
(<https://www.alura.com.br/artigos/django-ou-flask>)
- **Curso** Curso Java: consumindo API, gravando arquivos e lidando com erros
(<https://cursos.alura.com.br/course/java-consumindo-api-gravando-arquivos-lidando-erros>)
- **Curso** Curso Swagger: documentando suas APIs
(<https://cursos.alura.com.br/course/swagger-documentando-apis>)

HTTP - Fundamentos:

- HTTP significa Hyper Text Transfer Protocol. A comunicação entre computadores cliente e servidores web é feita enviando solicitações HTTP e recebendo respostas HTTP.
- Entender a diferença dos verbos HTTP
- Testar os requests e ver os status codes no navegador
- Saber fazer uma requisição HTTP na linha de comando com WGET
- Baixar uma imagem com WGET
- Fazer um post

Conteúdos

- **Site** MDN Web Docs: Uma visão geral do HTTP
(<https://developer.mozilla.org/pt-BR/docs/Web/HTTP/Overview>)
- **Site** MDN Web Docs: Métodos de requisição HTTP
(<https://developer.mozilla.org/pt-BR/docs/Web/HTTP/Methods>)
- **Site** HTTP Cats (<https://http.cat/>)

- **Site** HTTP Dogs (<https://http.dog/>)
- **YouTube** Fabiano Gabardo Lemos: Requisições HTTP - GET, POST, PUT, PATCH DELETE (<https://www.youtube.com/watch?v=kncOJZrnkTg>)
- **YouTube** Programador a Bordo: Protocolo HTTP e TCP/IP - Introdução (<https://www.youtube.com/watch?v=V4XZ81vRGtM>)

Conteúdos Alura:

- **Artigo** HTTP: Desmistificando o protocolo da Web (<https://www.alura.com.br/artigos/desmistificando-o-protocolo-http-parte-1>)
- **Artigo** Métodos de requisição do HTTP (<https://www.alura.com.br/artigos/metodos-de-requisicao-do-http>)
- **Artigo** Qual é a diferença entre HTTP e HTTPS? (<https://www.alura.com.br/artigos/qual-e-diferenca-entre-http-e-https>)
- **Podcast** Hipsters.tech: HTTP/2: magia com o novo protocolo - Hipsters 13 (<https://www.alura.com.br/podcast/http-2-magia-com-o-novo-protocolo-hipsters-13-a573>)
- **Curso** Curso HTTP: Entendendo a web por baixo dos panos (<https://www.alura.com.br/curso-online-http-entendendo-web-por-baixo-dos-panos>)

Testes de software:

- Testes de software referem-se ao processo de avaliar e validar um software para garantir que ele esteja funcionando corretamente e atenda aos requisitos estabelecidos. Esses testes envolvem a execução de diferentes cenários e casos de uso, identificando bugs, erros e garantindo a qualidade do software.
- Na engenharia de dados, os testes de software são essenciais para garantir a precisão e confiabilidade dos pipelines de dados e dos sistemas de armazenamento, processamento e distribuição de dados. Ao realizar testes rigorosos, os engenheiros de dados podem identificar e corrigir erros ou inconsistências nos dados, garantindo a integridade e a consistência dos fluxos de dados. Além disso, os testes de software ajudam a validar as

transformações e os processos de manipulação de dados, certificando-se de que os dados sejam processados corretamente e estejam em conformidade com as regras de negócio. Isso resulta em um ambiente de engenharia de dados mais confiável, seguro e de alta qualidade, garantindo que os dados sejam consistentes e precisos para análise, tomada de decisões e outras atividades relacionadas aos dados.

Conteúdos

- **Artigo** Você sabe o que é teste de software e por que ele é necessário? (<https://medium.com/cwi-software/voc%C3%AA-sabe-o-que-%C3%A9-teste-de-software-e-por-que-ele-%C3%A9-necess%C3%A1rio-23158c439267>)
- **Artigo** O que são testes unitários e como executá-los em Python (<https://medium.com/itautech/o-que-s%C3%A3o-testes-unit%C3%A1rios-e-como-execut%C3%A1-los-em-python-4d4a1b780fd6>)
- **Artigo** Como testar seus ETLs e tornar seu sistema de Dados mais confiável (<https://medium.com/@bernardo.costa/como-testar-seus-etls-e-tornar-seu-sistema-de-dados-mais-confi%C3%A1vel-20811d4d4a6>)
- **YouTube** Mauro de Boni: Curso de Teste de Software para iniciantes (<https://www.youtube.com/watch?v=dc3DBWgt8gw&list=PLx6gdu4s3nkf4OcLZ--TgMCHr22H7TDCs>)
- **Artigo** Tutorial: Testes Unitários e Python — Parte I (<https://medium.com/mercos-engineering/tutorial-testes-unit%C3%A1rios-e-python-parte-i-bb77182db93f>)
- **Artigo** Testes Unitários e Python — Parte II: Mocks, APIs & Star Wars (<https://medium.com/mercos-engineering/testes-unit%C3%A1rios-e-python-parte-ii-mocks-apis-star-wars-ou-como-consumir-uma-api-rest-3a24d153db59>)

Conteúdos Alura:

- **Curso** Curso Python e TDD: explorando testes unitários (<https://cursos.alura.com.br/course/python-tdd-explorando-testes-unitarios>)

- **Curso** Java e refatoração: melhorando códigos com boas práticas (<https://cursos.alura.com.br/course/java-refatoracao-melhorando-codigos-boas-praticas>)

Design Patterns:

- Na engenharia de software, um "padrão de projeto" (Design Pattern em inglês) é uma solução geral e reutilizável para um problema que ocorre normalmente dentro de um determinado contexto de projeto de software.
- Conhecer e aplicar os principais Design Patterns

Conteúdos

- **Artigo** Design Patterns — O que são e quais os benefícios? (<https://djesusnet.medium.com/design-patterns-gof-o-que-s%C3%A3o-e-quais-os-benef%C3%ADcios-9cd0cfdd6ebf>)
- **Artigo** Design Patterns (<https://medium.com/xp-inc/design-patterns-727494af001d>)
- **YouTube** Matheus Castiglioni: O que são design patterns (padrões de projetos) e para que servem? (<https://www.YOUTUBE.com/watch?v=KNWd1aWtsMw>)
- **YouTube** RinaldoDev: Design Patterns - o que são? Por que aprender a utilizar? (<https://www.YOUTUBE.com/watch?v=AWOf6Wo6gtg>)

Conteúdos Alura:

- **Podcast** Hipsters.tech: TechGuide - Design Patterns – Hipsters Ponto Tech #347 (<https://www.hipsters.tech/techguide-design-patterns-hipsters-ponto-tech-347/>)
- **Artigo** Design patterns: Breve introdução aos padrões de projeto (<https://www.alura.com.br/artigos/design-patterns-introducao-padroes-projeto>)
- **Podcast** Hipsters.tech: Design Patterns - Hipsters 206 (<https://www.alura.com.br/podcast/design-patterns-hipsters-206-a345>)
- **Curso** Curso Design Patterns em Java I: boas práticas de programação (<https://www.alura.com.br/curso-online-introducao-design-patterns-java>)

- **Curso** Curso Design Patterns C# I: boas práticas de programação (<https://www.alura.com.br/curso-online-design-patterns-dotnet>)
- **Curso** Curso Design Patterns em PHP: padrões comportamental (<https://www.alura.com.br/curso-online-php-design-pattern-comportamental>)
- **Curso** Curso Design Patterns em PHP: padrões estruturais (<https://www.alura.com.br/curso-online-php-design-pattern-estrutural>)
- **Curso** Curso Design Patterns em PHP: padrões criacionais (<https://www.alura.com.br/curso-online-php-design-pattern-criacional>)

Estruturas de Dados:

- No contexto dos computadores, uma estrutura de dados é uma forma específica de armazenar e organizar os dados na memória do computador para que esses dados possam ser facilmente recuperados e utilizados de forma eficiente quando necessário posteriormente.
- Conhecer as principais estruturas de dados
- Implementar as principais estruturas de dados

Conteúdos

- **Artigo** Estrutura de dados (<https://medium.com/php-word/estrutura-de-dados-6afb386a6fc1>)
- **Artigo** Estrutura de Dados e Algoritmo - parte 1 (<https://medium.com/@PaoloProdossimoLopes/01-estrutura-de-dados-e-algoritmo-9cb6a9b14aa>)
- **Artigo** Estrutura de Dados e Algoritmo - parte 2 (<https://medium.com/@PaoloProdossimoLopes/02-estrutura-de-dados-e-algoritmo-deba0b384a0f>)
- **YouTube** Mario Souto - Dev Soutinho: Estruturas de dados com JavaScript (<https://www.youtube.com/watch?v=MweeZn1rR8s>)
- **YouTube** Fabio Akita: Árvores (<https://www.youtube.com/watch?v=9GdesxWtOgs>)

- **YouTube** Loiane Groner: Estrutura de Dados com Java | Lista Encadeada
(https://www.youtube.com/watch?v=RW0oD2L_tSg)
- **YouTube** Loiane Groner: Estrutura de Dados com Java | Filas
(https://www.youtube.com/watch?v=RdoLTgJSV_c)
- **YouTube** Loiane Groner: Estrutura de Dados com Java | Pilhas
(<https://www.youtube.com/watch?v=ZdU4wMyiTSs>)

Conteúdos Alura:

- **Site** Apostila Python: Estruturas de dados
(<https://www.alura.com.br/apostila-python-orientacao-a-objetos/estrutura-de-dados>)
- **Artigo** Estruturas de dados: uma introdução
(<https://www.alura.com.br/artigos/estruturas-de-dados-introducao>)
- **Artigo** Curso Estrutura de Dados: computação na prática com Java
(<https://www.alura.com.br/artigos/estrutura-dados-computacao-na-pratica-com-java>)
- **Podcast** Hipsters.tech: Algoritmos e estrutura de dados - Hipsters 186
(<https://www.alura.com.br/podcast/hipsterstech-algoritmos-e-estrutura-de-dados-hipsters-186-a375>)
- **Curso** Curso PHP: dominando as Collections
(<https://cursos.alura.com.br/course/php-dominando-collections>)
- **Curso** Curso Python: crie a sua primeira aplicação
(<https://cursos.alura.com.br/course/python-crie-sua-primeira-aplicacao>)
- **Desafio** 7 Days of Code: Estruturas de Dados
(<https://7daysofcode.io/matricula/estruturas-de-dados>)

Analytics by Design:

- Analytics By Design é uma metodologia para melhorar a prontidão analítica dos projetos. Um projeto pronto para análise permitirá que cientistas de dados, analistas e até mesmo equipes de negócios comecem a analisar e obtenham insights com mais rapidez.

Conteúdos

- **Artigo** Analytics by design: como o iti revolucionou sua forma de trabalhar com o auxílio de dados (<https://medium.com/ititech/analytics-by-design-como-o-iti-revolucionou-sua-forma-de-trabalhar-com-o-aux%C3%ADlio-de-dados-b4acae469839>)
- **Podcast** Data Transformers Podcast: Data Analytics By Design with Dr. Kirk Borne (inglês) (<https://datatransformerspodcast.com/podcast/data-analytics-by-design-with-dr-kirk-borne/>)

Habilidade Auxiliar: Dados e Machine Learning

Business Intelligence (BI) - Fundamentos:

- Business Intelligence é um conjunto de teorias, metodologias, processos e tecnologias que possibilitam a transformação dos dados "crus" em informações extremamente relevantes para tomada de decisão de uma empresa.
- Conhecer o processo de ETL
- Realizar a modelagem e estruturação de tabelas em um Data Warehouse
- Criar visualizações que façam sentido
- Conhecer o PowerBI

Conteúdos

- **Artigo** Business Intelligence em Bibliotecas com Power BI (<https://franciscofoz.medium.com/business-intelligence-em-bibliotecas-com-power-bi-4065df7b232c>)
- **YouTube** Rafael Piton: O que é BI - Business Intelligence (Guia Definitivo) (https://www.youtube.com/watch?v=PoCLfN6sF_8)
- **Artigo** O que é o Microsoft Power BI? (<https://medium.com/data-are-lakes/o-que-%C3%A9-o-microsoft-power-bi-f98476ef5c48>)
- **YouTube** Karine Lago: Introdução ao Power BI (<https://www.youtube.com/watch?v=IMfh2c7GaWI>)

Conteúdos Alura:

- **Artigo** Business Intelligence: O que é? (<https://www.alura.com.br/artigos/business-intelligence>)
- **Artigo** Power BI ou Excel para Business Intelligence? (<https://www.alura.com.br/artigos/power-bi-excel-business-intelligence>)
- **Curso** Formação Business Intelligence com Excel (<https://www.alura.com.br/formacao-business-intelligence-excel>)
- **Curso** Formação Power BI (<https://www.alura.com.br/formacao-power-bi>)
- **Curso** Formação BI e Data Warehouse com SQL Server e Power BI (<https://www.alura.com.br/formacao-business-intelligence-data-warehouse>)
- **Curso** Formação Qlik Sense (<https://www.alura.com.br/formacao-qlik-sense>)

Extração e Tratamento de Dados:

- A extração de dados é o processo de coleta ou recuperação de tipos diferentes de dados de uma variedade de fontes, muitos dos quais podem estar mal organizados ou completamente desestruturados.
- Obter os dados que serão analisados
- Tratar os dados obtidos, transformando-os, alterando sua estrutura e valores a fim de deixar a base de dados mais coerente e garantir que os dados que serão trabalhados estejam nas melhores condições para serem analisados
- Utilizar as bibliotecas Pandas e Scikit-learn para tratar os dados

Conteúdos

- **Artigo** Tratamento e Transformação de Dados NaN: Uma visão geral e prática (<https://medium.com/data-hackers/tratamento-e-transforma%C3%A7%C3%A3o-de-dados-nan-uma-vis%C3%A3o-geral-e-pr%C3%A1tica-54efa9fc7a98>)
- **YouTube** Awari: Importação e Tratamento de Dados em Data Science (<https://www.youtube.com/watch?v=i1TK8pWRQas>)

- **YouTube** Celso Cestaro: Iniciando em coleta e tratamento de dados com Python (<https://www.youtube.com/watch?v=3DF54MnH9KM>)
- **YouTube** Celso Cestaro: Python para data science - limpando dados (<https://www.youtube.com/watch?v=t18Ogc8rNyo>)
- **Artigo** Como fazer uma limpeza de dados completa em Python (<https://medium.com/turing-talks/como-fazer-uma-limpeza-de-dados-completa-em-python-7abc9dfc19b8>)
- **YouTube** xavercoding: Manipulação de Dados em Python/Pandas (<https://www.youtube.com/watch?v=qKeM93NTKQM>)
- **YouTube** Bix Tecnologia: Tratamento de Dados com Python (Webinar) (<https://www.youtube.com/watch?v=slv1f2YpMJE>)

Conteúdos Alura:

- **Artigo** Pandas: o que é, para que serve e como instalar (<https://www.alura.com.br/artigos/pandas-o-que-e-para-que-serve-como-instalar>)
- **Curso** Curso Pandas: conhecendo a biblioteca (<https://cursos.alura.com.br/course/pandas-conhecendo-biblioteca>)
- **Curso** Curso Pandas I/O: trabalhando com diferentes formatos de arquivos (<https://cursos.alura.com.br/course/pandas-io-trabalhando-diferentes-formatos-arquivos>)
- **Curso** Curso Pandas: seleção e agrupamento de dados (<https://cursos.alura.com.br/course/pandas-selecao-agrupamento-dados>)
- **Curso** Curso Pandas: transformação e manipulação de dados (<https://cursos.alura.com.br/course/pandas-transformacao-manipulacao-dados>)
- **Desafio** 7 Days of Code: Python Pandas (<https://7daysofcode.io/matricula/pandas>)

Machine Learning - Fundamentos:

- O Aprendizado de Máquina ou Machine Learning é um subcampo da Engenharia e da Ciência da Computação que evoluiu do estudo de

reconhecimento de padrões e da teoria do aprendizado computacional em inteligência artificial.

- Aprendizado supervisionado
- Utilizar algoritmos de classificação
- Utilizar algoritmos de regressão
- Utilizar o Scikit-learn para criar modelos de machine learning

Conteúdos

- **Site** Amazon AWS: O que é machine learning?
(<https://aws.amazon.com/pt/what-is/machine-learning/>)
- **Artigo** A explicação mais simples sobre aprendizado de máquina que você lerá (<https://kozyrk.medium.com/a-explica%C3%A7%C3%A3o-mais-simples-sobre-aprendizado-de-m%C3%A1quina-que-voc%C3%AA-ler%C3%A1-6d4e61002401>)
- **Artigo** Introdução ao Aprendizado de Máquina
(<https://medium.com/@avinicius.adorno/introdu%C3%A7%C3%A3o-a-aprendizado-de-m%C3%A1quina-e39ec5ef459b>)
- **Artigo** Árvores de Decisão (<https://medium.com/machine-learning-beyond-deep-learning/%C3%A1rvores-de-decis%C3%A3o-3f52f6420b69>)
- **YouTube** Nerdologia: Machine Learning - como ensinar uma máquina a aprender (<https://www.youtube.com/watch?v=mhe5e2B9bL8>)
- **YouTube** Filipe Deschamps: Qual a diferença entre Inteligência Artificial, Machine Learning, Data Science, Deep Learning, etc?
(<https://www.youtube.com/watch?v=ccZ2pyr3YDw>)
- **YouTube** Filipe Deschamps: Machine Learning - Tutorial prático usando apenas o navegador (<https://www.youtube.com/watch?v=JyGGMMyR3x5I>)

Conteúdos Alura:

- **Artigo** Direto ao ponto: o que é Machine Learning com exemplos reais
(<https://www.alura.com.br/artigos/machine-learning>)
- **Artigo** Desmistificando termos em Machine Learning
(<https://www.alura.com.br/artigos/desmistificando-termos-machine->

[learning](#))

- **Artigo** Quais são os 4 tipos de aprendizagem na IA, algoritmos e usos no dia a dia (<https://www.alura.com.br/artigos/desmistificando-termos-machine-learning-tipos-aprendizado>)
- **Artigo** Machine Learning: o que é aprendizado semi-supervisionado (<https://alura.com.br/artigos/machine-learning-aprendizado-semi-supervisionado>)
- **YouTube** Alura: O que é Machine Learning? (https://www.youtube.com/watch?v=luz_jc96bQk)
- **Podcast** Hipsters.tech: Machine Learning – Hipsters 89 (<https://www.hipsters.tech/machine-learning-hipsters-88/>)
- **YouTube** Alura: Como funciona o Machine Learning? (<https://www.youtube.com/watch?v=CAu97npy7zQ>)
- **Podcast** Hipsters.tech: Desafios em Machine Learning – Hipsters 137 (<https://www.hipsters.tech/desafios-em-machine-learning-hipsters-137/>)
- **Podcast** Hipsters.tech: Engenharia de machine learning – Hipsters Ponto Tech 248 (<https://www.hipsters.tech/engenharia-de-machine-learning-hipsters-ponto-tech-248/>)
- **YouTube** Alura: O que faz uma pessoa engenheira de Machine Learning? (<https://www.youtube.com/watch?v=BRhz6v-jfMM>)
- **Curso** Formação Machine Learning (<https://cursos.alura.com.br/formacao-machine-learning>)
- **Curso** Machine Learning com Python: Classificação (<https://cursos.alura.com.br/formacao-machine-learning-python-classificacao>)
- **Desafio** 7 Days of Code: Machine Learning (<https://7daysofcode.io/matricula/machine-learning>)
- **Livro** Machine Learning: Introdução à classificação (<https://www.casadocodigo.com.br/products/livro-machine-learning>)
- **Livro** Inteligência Artificial e ChatGPT: Da revolução dos modelos de IA generativa à Engenharia de Prompt

(<https://www.casadocodigo.com.br/products/livro-inteligencia-artificial-chatgpt>)

MLOps:

- O MLOps permite que os modelos de Machine Learning sejam implantados de maneira rápida e confiável, o que é especialmente importante em empresas que lidam com grandes volumes de dados.
- Ajuda a garantir a qualidade e confiabilidade dos modelos de Machine Learning em produção, além de facilitar a manutenção e atualização desses modelos.
- Permite que as pessoas engenheiras de dados trabalhem em conjunto com cientistas de dados e desenvolvedores de software para implementar soluções de Machine Learning em larga escala.
- Ajuda a garantir a governança e a conformidade dos modelos de Machine Learning com as políticas e regulamentações da empresa.

Conteúdos

- **Artigo** ML Ops: Machine Learning como Disciplina de Engenharia (<https://medium.com/@cbreuel/ml-ops-machine-learning-como-disciplina-de-engenharia-a058770b93dc>)
- **Artigo** MLOps: por onde começar (<https://medium.com/datalab-log/mlops-por-onde-começar-a7679db9b10e>)
- **YouTube** O que é MLOps? (https://youtu.be/Z42UL_4AQvI)
- **YouTube** MLOps com Airflow: Pipeline automatizado de aprendizado de máquina (<https://www.youtube.com/live/oO3ToUlo3cg?feature=share>)
- **Artigo** O que é MLOps e como funcionam operações de Machine Learning (<https://blog.somostera.com/data-science/mlops>)
- **YouTube** Introdução a MLOps - Conceitos e Prática (https://www.youtube.com/watch?v=JC_Hxs1ZIk)

Conteúdos Alura:

- **Artigo** ML Engineer, o Dev em (<https://www.alura.com.br/artigos/ml-engineer-o-dev-em-t>)
- **Curso** MLOps: deploy de modelos (<https://cursos.alura.com.br/course/mlops-deploy-modelos>)
- **Curso** MLOps: Machine Learning e APIs (<https://cursos.alura.com.br/course/mlops-machine-learning-e-apis>)
- **Curso** MLflow: gestão do ciclo de vidas de modelos de ML (<https://cursos.alura.com.br/course/mlflow-gestao-ciclo-vidas-modelos-ml>)
- **YouTube** O que faz uma pessoa engenheira de Machine Learning? (https://www.youtube.com/watch?v=BRhz6v-jfMM&t=1s&ab_channel=Alura)
- **Podcast** Engenharia de machine learning (<https://www.alura.com.br/podcast/hipsterstech-engenharia-de-machine-learning-hipsters-ponto-tech-248-a853>)
- **Podcast** Machine Learning e o MLOps (<https://www.alura.com.br/podcast/hipsterstech-machine-learning-e-o-mlops-hipsters-171-a398>)
- **Podcast** Ops em Machine Learning Revisitado (<https://www.alura.com.br/podcast/hipsterstech-ops-em-machine-learning-revisitado-hipsters-ponto-tech-333-a1902>)

Estatística e Matemática - Fundamentos:

- Equações, funções e limites
- Logaritmos
- Matrizes, determinantes, vetores e espaço vetorial
- Derivadas e integrais
- Diferença entre média, mediana e moda
- Distribuição de frequência
- Variância e desvio padrão
- Distribuição binomial, poisson e normal

- Nível e intervalo de confiança
- Técnicas de amostragem
- Introdução à Regressão linear
- Séries temporais

Conteúdos

- **Artigo** Preciso saber matemática para aprender Ciência de Dados?
(<https://beatrizyumi.medium.com/preciso-saber-matem%C3%A1tica-para-aprender-ci%C3%A2ncia-de-dados-9a9b7c7846a9>)
- **YouTube** Mario Filho: As 3 Áreas da Matemática Usadas em Data Science
(https://www.youtube.com/watch?v=niF_T21vdZ4)
- **YouTube** PrograMaria: Estatística básica para Ciência de Dados
(<https://www.youtube.com/watch?v=xiZwte8D1Xs>)
- **YouTube** Data Viking: Estatística básica para ciência de dados
(https://www.youtube.com/watch?v=K_q6uCgV0Mk)
- **YouTube** FC Nuvem: Introdução a Estatística para Ciência de Dados
(<https://www.youtube.com/watch?v=NdRKlIFa63c>)

Conteúdos Alura:

- **Artigo** Média ou mediana? Entendendo cada uma
(<https://www.alura.com.br/artigos/media-ou-mediana-entendendo-cada-uma>)
- **Artigo** Modelos matemáticos utilizados em séries temporais
(<https://www.alura.com.br/artigos/modelos-matematicos-utilizados-em-series-temporais>)
- **YouTube** Alura: Preciso saber matemática para programar?
(<https://www.youtube.com/watch?v=WZo8s-zsfuY>)
- **Curso** Formação Estatística com Python
(<https://cursos.alura.com.br/formacao-estatistica-python>)
- **Curso** Data Science: testando hipóteses
(<https://cursos.alura.com.br/course/data-science-testando-hipoteses>)

- **Livro** Introdução à Estatística para Ciência de Dados: Da exploração dos dados à experimentação contínua com exemplos de código em Python e R (<https://www.casadocodigo.com.br/products/livro-estatistica-datascience>)

TechGuide - Alura
Alura, PM3 e FIAP
O Techguide.sh é um projeto open source