

Educação Profissional Paulista

Técnico em
**Ciência de
Dados**

Lógica de programação e algoritmos

Decomposição de problemas

Aula 2

Código da aula: [DADOS]ANO1C3B3S19A2

Lógica de
programação e
algoritmos

Mapa da unidade 1 Componente 4

semana
21

Recursividade

semana
23

Busca e
ordenação

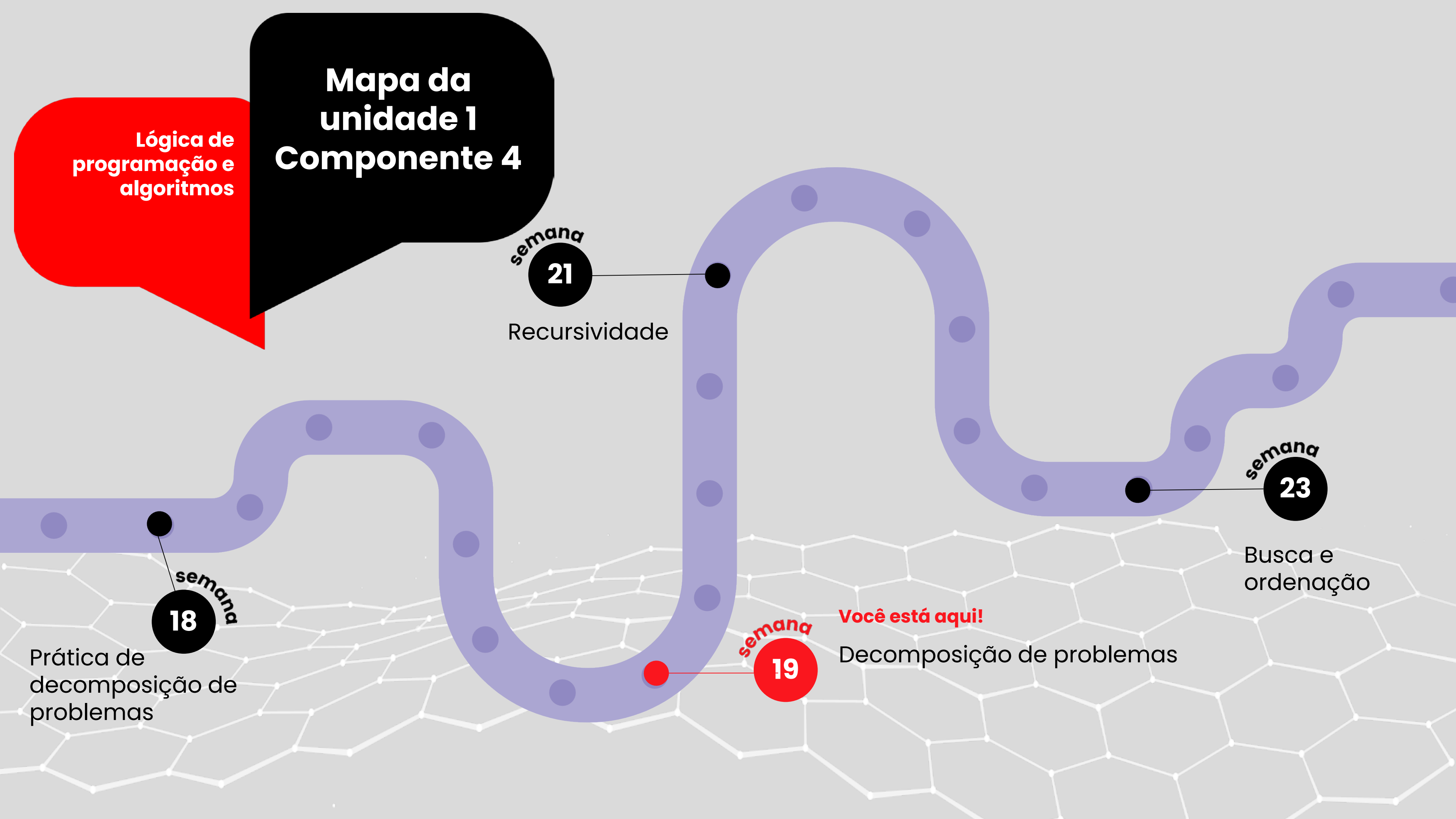
semana
18

Prática de
decomposição de
problemas

semana
19

Você está aqui!

Decomposição de problemas



Lógica de
programação e
algoritmos

Mapa da unidade 1 Componente 4

Você está aqui!

Decomposição de problemas

19

Aula 2

Código da aula: [DADOS]ANO1C3B3S19A2



Objetivo da aula

- Conhecer a importância de quebrar problemas complexos em partes menores e mais gerenciáveis, facilitando, assim, a análise e a solução.



Recursos didáticos

- Recurso audiovisual para exibição de vídeos e imagens;
- Acesso ao laboratório de informática e/ou internet.



Duração da aula

50 minutos.



Competências técnicas

- Usar técnicas para explorar e analisar dados, aplicar modelos estatísticos, identificar padrões, realizar inferências e tomar decisões baseadas em evidências;
- Compreender e dominar técnicas de manipulação de dados; extrair, transformar e carregar conjuntos de dados de diferentes fontes, garantindo a qualidade e a integridade dos dados; criar e compreender visualizações gráficas.



Competências socioemocionais

- Colaborar efetivamente com outros profissionais, como cientistas de dados e engenheiros de dados;
- Trabalhar em equipes multifuncionais colaborando com colegas, gestores e clientes.

Construindo o conceito

Passos básicos para aplicar a decomposição de problema em Ciência de Dados

Passos básicos para aplicar a decomposição de problema em Ciência de Dados

1. Identificação do problema

2. Decomposição em subproblemas

3. Coleta de dados

4. Análise exploratória

5. Pré-processamento de dados

6. Modelagem

7. Validação do modelo

8. Interpretação dos resultados

9. Comunicação dos resultados

Nas aulas anteriores, foram abordadas as estratégias existentes para a decomposição de problemas: **Dividir e conquistar**, **Abordagem top-down**, **Abordagem bottom-up** e **Abstração**.

Em Ciência de Dados, independentemente da estratégia escolhida de decomposição de problema, os tópicos 1, 3, 4, 5, 7, 8, e 9 serão frequentemente aplicados, seja qual for o tipo de problema. Em relação aos tópicos 2 e 6, esses poderão ser remodelados de acordo com a estratégia.

Elaborado especialmente para o curso.

Construindo
o **conceito**

Exemplo prático - *E-Commerce Boost*

Imagine que você é um técnico em Ciência de Dados em uma empresa de comércio eletrônico e sua equipe está trabalhando para aumentar as vendas.

Você recebeu um conjunto de dados com informações sobre os clientes, como idade, gênero, histórico de compras e valor gasto.

Seu objetivo é criar um algoritmo (modelo) que preveja quais clientes têm maior probabilidade de realizar compras novamente.

Situação fictícia elaborada especialmente para o curso.

Construindo
o conceito

Aplicando os passos de decomposição de problemas em Ciência de Dados para o estudo de caso do E-Commerce Boost

Observação: Independentemente do tipo de problema em Ciências de Dados, os tópicos 1, 3, 4, 5, 7, 8, e 9 são aplicados de forma comum.

Tópico	Dividir e conquistar	Abordagem top-down	Abordagem bottom-up	Abstração
1 Identificação do problema	O problema é criar um modelo que preveja quais clientes têm maior probabilidade de realizar compras novamente, com base em suas características ou seu histórico de compras.	O problema é criar um modelo que preveja quais clientes têm maior probabilidade de realizar compras novamente, com base em suas características ou seu histórico de compras.	O problema é criar um modelo que preveja quais clientes têm maior probabilidade de realizar compras novamente, com base em suas características ou seu histórico de compras.	O problema é criar um modelo que preveja quais clientes têm maior probabilidade de realizar compras novamente, com base em suas características ou seu histórico de compras.
3 Coleta de dados	Coletar e organizar os dados dos clientes, inclusive idade, gênero, histórico de compras e valor gasto.	Coletar e organizar os dados dos clientes, inclusive idade, gênero, histórico de compras e valor gasto.	Coletar e organizar os dados dos clientes, inclusive idade, gênero, histórico de compras e valor gasto.	Coletar e organizar os dados dos clientes, inclusive idade, gênero, histórico de compras e valor gasto.
4 Análise exploratória	Analisar as variáveis mais relevantes para identificar padrões e tendências, como a relação entre idade e valor gasto.	Analisar as variáveis mais relevantes para identificar padrões e tendências, como a relação entre idade e valor gasto.	Analisar as variáveis mais relevantes para identificar padrões e tendências, como a relação entre idade e valor gasto.	Analisar as variáveis mais relevantes para identificar padrões e tendências, como a relação entre idade e valor gasto.
5 Pré-processamento de dados	Limpar dados, preencher valores faltantes e transformar as variáveis necessárias para o modelo.	Limpar dados, preencher valores faltantes e transformar as variáveis necessárias para o modelo.	Limpar dados, preencher valores faltantes e transformar as variáveis necessárias para o modelo.	Limpar dados, preencher valores faltantes e transformar as variáveis necessárias para o modelo.
7 Validação do modelo	Avaliar o desempenho do modelo usando métricas como precisão, recall e F1-Score.	Avaliar o desempenho do modelo usando métricas como precisão, recall e F1-Score.	Avaliar o desempenho do modelo usando métricas como precisão, recall e F1-Score.	Avaliar o desempenho do modelo usando métricas como precisão, recall e F1-Score.
8 Interpretação dos resultados	Interpretar os resultados do modelo para entender quais características dos clientes tem maior influência nas previsões.	Interpretar os resultados do modelo para entender quais características dos clientes tem maior influência nas previsões.	Interpretar os resultados do modelo para entender quais características dos clientes tem maior influência nas previsões.	Interpretar os resultados do modelo para entender quais características dos clientes tem maior influência nas previsões.
9 Comunicação dos resultados	Apresentar os resultados do modelo para a equipe de marketing, destacando as principais conclusões e recomendações para aumentar as vendas.	Apresentar os resultados do modelo para a equipe de marketing, destacando as principais conclusões e recomendações para aumentar as vendas.	Apresentar os resultados do modelo para a equipe de marketing, destacando as principais conclusões e recomendações para aumentar as vendas.	Apresentar os resultados do modelo para a equipe de marketing, destacando as principais conclusões e recomendações para aumentar as vendas.

Elaborado especialmente para o curso.

Construindo o conceito

Aplicando os passos de decomposição de problemas em Ciência de Dados para o estudo de caso do *E-Commerce Boost*

Quando aplicamos os **passos básicos para a decomposição de problema em Ciência de Dados** para o estudo de caso do *E-Commerce Boost*, em cada uma das estratégias percebemos que:

- Independentemente da estratégia escolhida, os tópicos 1, 3, 4, 5, 7, 8 e 9, indicados na tabela, mantiveram-se iguais.

Você notou também, na tabela, que o que muda em relação às estratégias são as etapas de **Decomposição em subproblemas** e **Modelagem**? Elas serão abordadas na próxima aula e, dessa forma, você entenderá as diferenças entre as estratégias dentro do tópico de Ciência de Dados



Colocando
em **prática**

Coleta de dados em Ciência de Dados



Documento texto



Durante a aula



Em grupo



ORGANIZEM-SE EM GRUPOS

Para entender como a **coleta de dados** é um dos passos básicos na aplicação da decomposição de problemas em Ciência de Dados, vocês irão realizar uma pesquisa na internet.



PESQUISEM SOBRE O TEMA

Abram o navegador de internet e acessem o site de busca (Google, Bing, etc.). Na barra de pesquisa, digite "**Como fazer coleta de dados para Ciência de Dados**".



ANALISEM AS INFORMAÇÕES

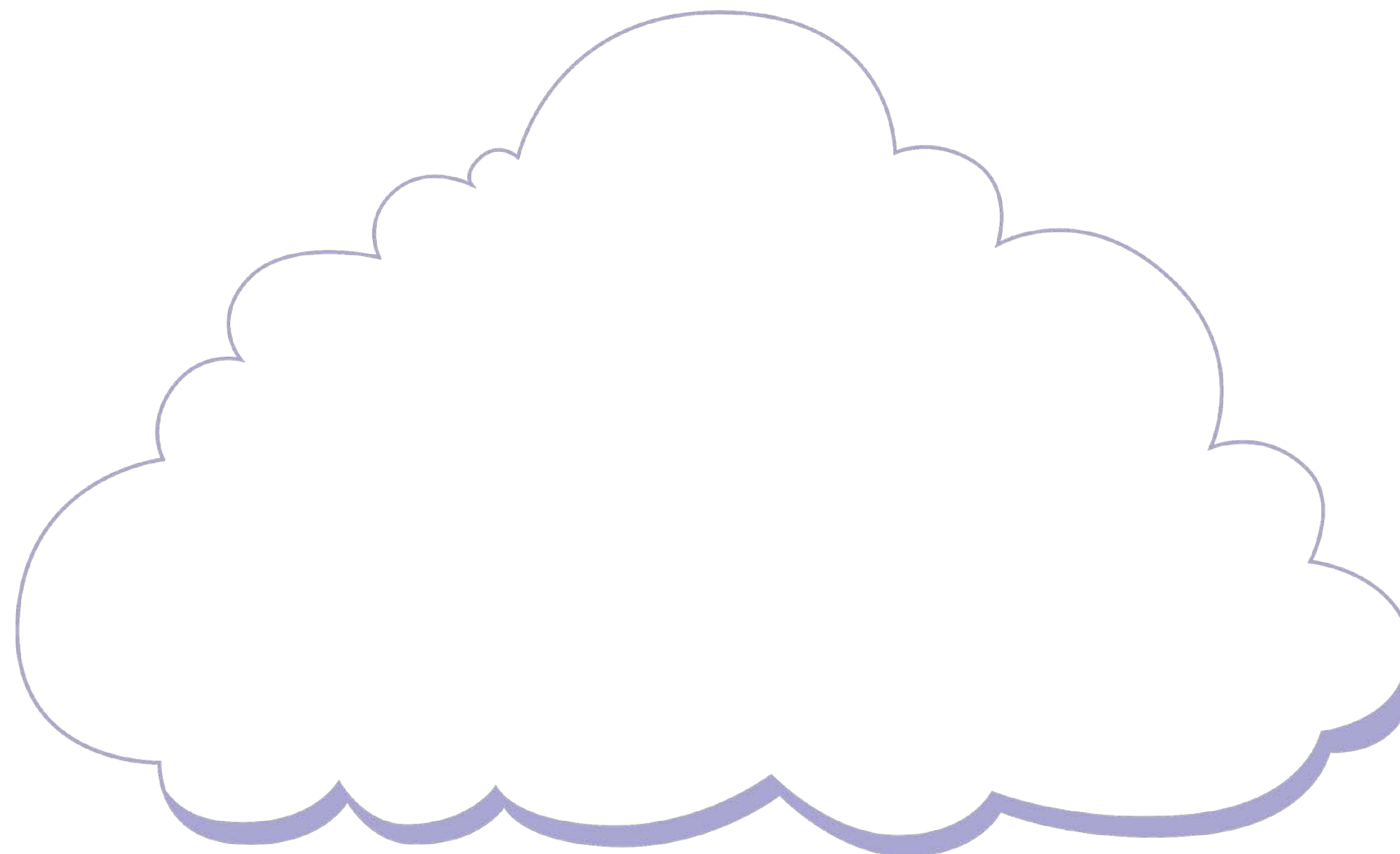
- Quais são as principais etapas da coleta de dados?
- Quais são as fontes de dados mais utilizadas em projetos de Ciência de Dados?
- Quais são as melhores práticas para garantir a qualidade dos dados coletados?



ELABOREM UM RESUMO DA PESQUISA

Faça um resumo da pesquisa e envie pelo AVA (Ambiente Virtual de Aprendizagem).

Nuvem de palavras



© Getty Images

O que nós
aprendemos
hoje?



© Getty Images

O que nós
**aprendemos
hoje?**

Então ficamos assim...

- 1** Na decomposição de problemas em Ciência de Dados, começa-se identificando claramente o problema e definindo objetivos. Depois, coletam-se dados de fontes confiáveis para fundamentar a análise;
- 2** O segundo passo na decomposição de problemas em Ciência de Dados é a análise exploratória dos dados para encontrar padrões e tendências, seguido pelo desenvolvimento e teste de modelos para prever ou classificar dados;
- 3** Os modelos em Ciência de Dados são continuamente aprimorados e testados para resolver o problema proposto. Esse ciclo de análise, modelagem e validação é crucial para o sucesso do projeto. Ao final, os resultados são claramente explicados para auxiliar na tomada de decisões.

Saiba mais

Que tal conhecer os primeiros passos sobre a análise exploratória de dados (ou Exploratory Data Analysis – EDA, em inglês)?

Por meio dessa abordagem, cientistas de dados podem ter uma visão panorâmica sobre os dados. E mais do que isso: conseguem extrair algum conhecimento a partir deles.

GONZALEZ, V. Análise Exploratória: primeiros passos | *Alura*, 24 set. 2020. Disponível em: <https://www.alura.com.br/artigos/analise-exploratoria>. Acesso: 10 maio 2024.

Referências da aula

CAMARGO, R; RIBAS, T. *Gestão ágil de projetos*: as melhores soluções para suas necessidades. São Paulo: Saraiva, 2019.

CHIAVENATO, I. *Administração*: teoria, processo e prática. São Paulo: Manole, 2014.

Identidade visual: Imagens © Getty Images

**Educação
Profissional
Paulista**

Técnico em
**Ciência de
Dados**