<table>
<tr><td>Título do plano</td><td colspan="3">Coleta e Amostragem</td><td>Sequência:
6</td></tr>
<tr><td>Etapa/Ano</td><td colspan="4">Ensino Médio: ☐ 1º Ano | ☐ 2º Ano | ☑ 3º Ano</td></tr>
<tr><td>Nível de Maturidade
[link]</td><td>Escola</td><td>☐ Emergente
☐ Básico
☑ Intermediário
☐ Avançado</td><td>Docente</td><td>☐ Básico
☑ Intermediário
☐ Avançado</td></tr>
<tr><td>Competências
[link]</td><td colspan="4">☐ C1 ☑ C2 ☑ C3 ☐ C4 ☐ C5</td></tr>
<tr><td>Objetos de Conhecimento</td><td colspan="4">• Processo de coleta de dados em contextos reais
• Diferença entre população e amostra
• Tipos de amostragem (aleatória, por conveniência, estratificada)
• Viés de amostragem e representatividade
• Janela de coleta e contexto dos dados
• Registro de metadados (tempo, local, perfil)
• Relação entre coleta de dados e qualidade do dataset
• Impacto da amostragem no funcionamento de sistemas de IA</td></tr>
<tr><td>Recursos Educacionais</td><td colspan="4">• Questionário padrão já utilizado anteriormente;
• Quadro, projetor ou cartolina;
• Conjunto de respostas (real ou simulado);
• Fichas ou post-its para simulação de coleta;</td></tr>
<tr><td>Componentes Curriculares Relacionados</td><td>☐ Arte
☐ Língua Port.
☐ Ed. Física</td><td>☐ Geografia
☐ História
☐ Matemática</td><td colspan="2">☐ Língua Inglesa
☐ Ciências
☑ Computação</td></tr>
<tr><td>Palavras-chave</td><td colspan="4">Coleta de dados, amostragem, viés, representatividade</td></tr>
<tr><td>Perguntas Importantes</td><td colspan="4">• Quem está representado nos dados coletados?
• Quem ficou de fora da coleta?
• A amostra representa a realidade que queremos estudar?
• Como o momento da coleta influencia os dados?
• O que acontece quando os dados são coletados apenas de um grupo específico?
• Como evitar vieses na coleta de dados?
• Qual a diferença entre coletar todos os dados e trabalhar com uma amostra?</td></tr>
<tr><td>Encontros</td><td colspan="2">1</td><td>Total de Horas-aula</td><td>2</td></tr>
<tr><td>Objetivos</td><td colspan="4">• Compreender que a coleta de dados é uma etapa crítica na construção de um dataset;
• Diferenciar população e amostra, reconhecendo limitações e implicações;
• Identificar possíveis vieses na coleta de dados;
• Entender como as decisões na coleta impactam diretamente os resultados de análises e sistemas de IA.</td></tr>
<tr><td>Habilidades Relacionadas</td><td colspan="4">BNCC Computação:
(EM13CO12) Produzir, analisar, gerir e compartilhar informações a partir de dados, utilizando princípios de ciência de dados.
(EM13CO13) Analisar e utilizar as diferentes formas de representação e</td></tr>
</table>

CC BY-NC-SA 4.0 VICARI, Rosa Maria; GALAFASSI, Cristiano; BRACKMANN, Christian; MIZUSAKI, Lucas.
IA@Escola. 2024. Disponível em: www.computacional.com.br/ia.

	consulta a dados em formato digital para pesquisas científicas. **Referencial Curricular IA no EM:** EM02IA06 — Compreender o ciclo de projeto envolvido com o desenvolvimento de sistemas de IA. Essa habilidade se relaciona ao plano porque a coleta de dados e a definição da amostra fazem parte das etapas iniciais do ciclo de desenvolvimento de IA . EM03IA05 — Reconhecer que os resultados da IA dependem tanto da qualidade e diversidade dos dados quanto dos algoritmos utilizados. É a habilidade mais diretamente relacionada ao plano, pois a aula discute representatividade, diversidade, viés e qualidade dos dados .
Práticas Pedagógicas Inovadoras	☑ Aula Enriquecida com Tecnologia ☐ Ensino Híbrido: Sala de aula invertida ☐ Ensino Híbrido: rotação por estação ☐ Ensino Híbrido: rotação individual ☑ Aulas Mão na massa ☑ Aprendizagem baseada em projetos ☐ IA Desconectada

Sequência Didática

1. Introdução ao Tema

A aula pode iniciar com a pergunta:
"Os dados que coletamos representam realmente a realidade?"

A intenção é fazer a turma perceber que os dados não são neutros e que a forma como são coletados influencia diretamente os resultados.

Sugestões de apontamentos:
- "Antes de organizar ou analisar dados, precisamos entender de onde eles vêm."
- "Nem todo conjunto de respostas representa bem a realidade."
- "A forma como coletamos os dados pode gerar distorções."

Perguntas para iniciar a discussão:
- Se apenas metade da turma responder, o resultado representa todos?
- Quem costuma responder mais facilmente um questionário?
- Existe diferença entre opinião e representatividade?

2. Coleta de dados: como os dados são obtidos

Retomar o questionário padrão já estudado e discutir como ele é aplicado na prática.

CC BY-NC-SA 4.0 VICARI, Rosa Maria; GALAFASSI, Cristiano; BRACKMANN, Christian; MIZUSAKI, Lucas.
IA@Escola. 2024. Disponível em: www.computacional.com.br/ia.

Sugestões de explicações:
- “Coletar dados é decidir quem vai responder, quando e como.”
- “Mesmo com um bom questionário, a coleta pode gerar problemas.”

Apontar aspectos importantes:
- local da coleta;
- momento da aplicação;
- perfil dos respondentes;
- taxa de resposta.

Perguntas para discussão:
- O que muda se aplicarmos o questionário em horários diferentes?
- Quem pode não responder?
- O ambiente influencia as respostas?

3. Amostragem: trabalhar com parte da realidade

Introduzir a ideia de população e amostra.

Sugestões de explicações:
- “População é o conjunto completo que queremos estudar.”
- “A amostra é apenas uma parte desse conjunto.”
- “Nem sempre conseguimos coletar dados de todos.”

Apresentar tipos de amostragem:
- aleatória: escolha sem critério específico;
- por conveniência: quem está disponível;
- estratificada: grupos definidos (ex: séries, turmas).

Atividade prática:
Dividir a turma e coletar respostas de apenas um grupo.
Comparar com respostas da turma inteira.

Perguntas para discussão:
- Essa amostra representa todos?
- O que mudou entre os grupos?
- Qual tipo de amostragem foi usado?

4. Viés e representatividade

Esta é a parte central da aula.

Sugestões de explicações:
- “Viés acontece quando os dados não representam corretamente a realidade.”
- “Isso pode acontecer sem percebermos.”

CC BY-NC-SA 4.0 VICARI, Rosa Maria; GALAFASSI, Cristiano; BRACKMANN, Christian; MIZUSAKI, Lucas.
IA@Escola. 2024. Disponível em: www.computacional.com.br/ia.

Exemplos para explorar:
- coletar apenas de uma turma;
- coletar apenas de quem quis responder;
- coletar em um único horário.

Perguntas para discussão:
- Quem está sendo excluído dos dados?
- Esse dataset pode gerar conclusões injustas?
- Como evitar esse problema?

Importante destacar:
- O viés pode comprometer análises e decisões;
- Sistemas de IA aprendem com esses dados e reproduzem esses padrões.

5. Registro de contexto (metadados)

Apresentar a importância de registrar informações sobre a coleta.

Sugestões de explicações:
- "Além das respostas, precisamos registrar o contexto."
- "Essas informações ajudam a interpretar os dados."

Exemplos de metadados:
- data da coleta;
- local;
- perfil dos respondentes;
- responsável pela coleta.

Perguntas para discussão:
- Por que é importante saber quando o dado foi coletado?
- O contexto pode explicar diferenças nos resultados?

6. Fechamento

Construir coletivamente um quadro com:

Elemento	Decisão tomada	Impacto
Quem respondeu		
Quando coletou		
Tipo de amostra		
Possíveis vieses		

Sugestões para encerramento:
- "Hoje vimos que dados não surgem prontos."
- "A coleta influencia diretamente a

CC BY-NC-SA 4.0 VICARI, Rosa Maria; GALAFASSI, Cristiano; BRACKMANN, Christian; MIZUSAKI, Lucas.
IA@Escola. 2024. Disponível em: www.computacional.com.br/ia.

qualidade do dataset.” ● “Antes de analisar dados, precisamos confiar em como eles foram obtidos.”	
Avaliação	A avaliação pode ser processual, considerando a participação dos estudantes nas discussões, a capacidade de diferenciar população e amostra, a identificação de tipos de amostragem e a análise crítica de possíveis vieses. Como produto avaliativo, os estudantes podem construir um quadro com os seguintes campos: quem respondeu, quem ficou de fora, quando a coleta ocorreu, tipo de amostra, possíveis vieses, metadados registrados e impacto esperado na análise. O professor pode observar se os estudantes conseguem explicar como decisões aparentemente simples de coleta podem comprometer ou fortalecer a qualidade do dataset e os resultados de sistemas de IA.
Material Complementar	Material Didático do Professor — Tema 306 questionário padrão utilizado no projeto https://www.computacional.com.br/

CC BY-NC-SA 4.0 VICARI, Rosa Maria; GALAFASSI, Cristiano; BRACKMANN, Christian; MIZUSAKI, Lucas.
IA@Escola. 2024. Disponível em: www.computacional.com.br/ia.