

Título do plano	Exploração dos dados	Sequência: 9
Etapa/Ano	Ensino Médio: ☐ 1º Ano ☐ 2º Ano ☒ 3º Ano	
Nível de Maturidade [link]	Escola ☐ Emergente ☐ Básico ☒ Intermediário ☐ Avançado	Docente ☐ Básico ☒ Intermediário ☐ Avançado
Competências [link]	☐ C1 ☒ C2 ☒ C3 ☒ C4 ☒ C5	
Objetos de Conhecimento	• Exploração inicial de dados; • Estatística descritiva aplicada a dados escolares; • Frequência absoluta e frequência relativa; • Distribuição de respostas em dados categóricos e ordinais; • Medidas simples de síntese: contagem, proporção, percentual, moda e comparação entre grupos; • Visualização de dados: tabelas, gráficos de barras, gráficos de colunas, gráficos de setores, gráficos empilhados e tabelas dinâmicas; • Leitura e interpretação de gráficos; • Identificação de padrões iniciais nos dados; • Comparação entre perfil dos respondentes e prioridades indicadas; • Relação entre visualização, interpretação e tomada de decisão; • Diferença entre descrever dados e explicar causas; • Riscos de interpretação indevida de gráficos; • Uso de planilhas para análise exploratória; • Relação entre exploração de dados e etapas posteriores de IA.	
Recursos Educacionais	• Questionário padrão do projeto Vozes da Escola ou projeto regional; • Base sintética de respostas já estruturada, coletada, limpa e documentada nos planos anteriores; • Planilha eletrônica, como Google Planilhas, Excel ou LibreOffice Calc; • Projetor, quadro ou cartolina; • Base original preservada e base tratada; • Ficha de documentação do dataset elaborada no plano 308; • Dicionário de dados; • Lista oficial das iniciativas I01 a I12; • Modelo de tabela de frequência; • Modelo de tabela de leitura de gráficos; • Canetas, post-its ou documento colaborativo.	
Componentes Curriculares Relacionados	☐ Arte ☐ Língua Port. ☐ Ed. Física	☐ Geografia ☐ História ☒ Matemática
Palavras-chave	Exploração de dados, estatística descritiva, visualização de dados, distribuição, frequência, proporção, padrões, gráficos, análise exploratória, dataset.	
Perguntas Importantes	• O que os dados mostram em uma primeira leitura? • Quais iniciativas aparecem com maior frequência? • Como comparar respostas de diferentes grupos? • Qual é a diferença entre contar respostas e interpretar padrões?	

	• O que uma distribuição de dados revela? • Que tipo de gráfico é mais adequado para cada tipo de dado? • Quando um gráfico pode induzir a uma interpretação errada? • Uma iniciativa mais votada significa necessariamente que ela é a melhor? • Como identificar padrões sem exagerar nas conclusões? • O que podemos afirmar com segurança a partir dos dados? • O que ainda não podemos afirmar? • Como a exploração inicial prepara a etapa de modelagem ou uso de IA?		
Encontros	1	Total de Horas-aula	2
Objetivos	• Compreender a exploração de dados como etapa posterior à estruturação, coleta, limpeza e documentação do dataset; • Utilizar estatística descritiva simples para sintetizar respostas da comunidade escolar; • Calcular e interpretar frequências absolutas, frequências relativas e percentuais; • Construir tabelas e gráficos para visualizar distribuições de respostas; • Comparar padrões entre diferentes grupos de respondentes; • Identificar tendências iniciais sem transformar correlação ou frequência em explicação causal; • Reconhecer limites da interpretação de dados e gráficos; • Relacionar a análise exploratória com a preparação para etapas posteriores de IA; • Produzir uma síntese visual e interpretativa inicial sobre as prioridades do projeto Vozes da Escola ou projeto regional.		
Habilidades Relacionadas	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 consulta a dados em formato digital para pesquisas científicas. (EM13CO14) Avaliar a confiabilidade das informações encontradas em meio digital, investigando seus modos de construção e considerando autoria, estrutura e propósito da mensagem. (EM13CO18) Planejar e gerenciar projetos integrados às áreas de conhecimento de forma colaborativa, solucionando problemas, usando diversos artefatos computacionais. Referencial Curricular IA no EM: (EM01IA02) Criar com a IA realizando o reconhecimento de padrões, raciocínio simbólico e estatístico. (EM02IA06) Compreender o ciclo de projeto envolvido com o desenvolvimento de sistemas de IA. (EM03IA03) Saber treinar uma IA por meio da técnica de Machine Learning. (EM03IA04) Compreender a diferença entre Acurácia e Precisão e como as medidas de desempenho da IA são realizadas. (EM03IA05) Compreender que a IA resulta da qualidade e da diversidade das bases de dados com os quais interage.		

	Habilidades articuladoras da progressão anterior: (EF69IA04) Entender a relação entre Inteligência Artificial e estatística. A aula retoma essa habilidade como fundamento para mostrar que estatística e visualização são etapas importantes antes de qualquer uso de IA. (EF69IA05) Avaliar dados e sua qualidade, aplicando métodos de tratamento inicial. A exploração dos dados permite verificar, depois da limpeza, se ainda existem problemas de distribuição, categorias pouco representadas ou padrões que exigem cautela interpretativa.
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:

“Depois que limpamos e documentamos a base, o que os dados começam a nos contar?”

Agora, os estudantes passam para uma nova etapa: **explorar os dados para identificar distribuições, padrões iniciais e perguntas para análises futuras.**

Sugestões de apontamentos:

- Explorar os dados é olhar para a base antes de tirar conclusões fortes.
- A primeira análise deve ser descritiva: contar, comparar, visualizar e perguntar.
- Gráficos ajudam a enxergar padrões, mas também podem induzir interpretações equivocadas.
- Antes de treinar qualquer IA, precisamos entender como os dados estão distribuídos.
- Uma base pode estar limpa e documentada, mas ainda apresentar desequilíbrios.

Perguntas para iniciar a discussão:

- Quais iniciativas vocês imaginam que aparecerão mais?
- Será que estudantes, professores e familiares priorizam as mesmas iniciativas?
- O que significa uma iniciativa aparecer muitas vezes?
- O que significa uma categoria aparecer poucas vezes?
- Um gráfico pode enganar? Como?
- Que cuidados precisamos ter antes de afirmar que encontramos um padrão?

2. O que é explorar dados?

Nesta etapa, o professor apresenta a ideia de análise exploratória de dados em linguagem acessível.

Sugestões de explicações:

- Explorar dados é observar a base para entender sua forma geral.
- É uma etapa de perguntas, não de respostas definitivas.
- Na exploração, contamos respostas, calculamos percentuais, construímos gráficos e procuramos padrões.
- A exploração ajuda a decidir quais análises fazem sentido depois.
- Em IA, essa etapa é importante porque mostra se os dados têm equilíbrio, diversidade e coerência.

Exemplo simples:

Se a base possui um conjunto de respostas, a exploração inicial pode ajudar a responder:

- quantas pessoas responderam;
- quantos estudantes, professores, familiares e funcionários participaram;
- quais iniciativas foram mais escolhidas como prioridade 1;
- se há diferença entre prioridades de estudantes e professores;
- se determinadas iniciativas aparecem mais em certos grupos;
- se há grupos pouco representados;
- quais gráficos ajudam a comunicar melhor esses resultados.

Perguntas para discussão:

- Qual é a primeira coisa que devemos contar em uma base?
- Por que o total de respostas é importante?
- Por que não basta olhar apenas para a iniciativa mais votada?
- O que devemos observar antes de comparar grupos?

3. Frequência absoluta, frequência relativa e percentual

Esta etapa introduz os conceitos estatísticos básicos que serão usados na aula.

Sugestões de explicações:

- Frequência absoluta é o número de vezes que uma resposta aparece.
- Frequência relativa é a proporção dessa resposta em relação ao total.
- Percentual é essa proporção expressa em porcentagem.
- Contar respostas é importante, mas comparar percentuais ajuda quando os grupos têm tamanhos diferentes.

Exemplo:

Iniciativa	Frequência absoluta	Frequência relativa	Percentual
------------	---------------------	---------------------	------------

I03 Oficina de leitura e escrita	45	45/150	30%
I06 Roda de convivência e mediação	30	30/150	20%
I11 Sustentabilidade	15	15/150	10%

Sugestões de perguntas:

- Qual iniciativa apareceu mais vezes?
- O que o percentual acrescenta à contagem?
- Se uma escola tem 50 respostas e outra tem 300, podemos comparar apenas os números absolutos?
- Quando o percentual é mais adequado do que a contagem?

Mensagem importante:

A frequência absoluta mostra o volume de respostas. A frequência relativa e o percentual ajudam a comparar grupos de tamanhos diferentes.

4. Distribuições: como as respostas se espalham?

Nesta etapa, o professor apresenta a ideia de distribuição dos dados.

Sugestões de explicações:

- Distribuição é a forma como as respostas se espalham entre as categorias.
- Uma distribuição pode ser concentrada, equilibrada ou muito desigual.
- Observar a distribuição ajuda a perceber se uma opção domina a base ou se há diversidade de escolhas.
- Em IA, distribuições muito desequilibradas podem influenciar análises e modelos.

Exemplo de interpretação:

Situação observada	Possível interpretação inicial
Uma iniciativa concentra 60% das escolhas	Pode indicar forte prioridade, mas exige verificar quem respondeu
Todas as iniciativas aparecem com percentuais próximos	Pode indicar diversidade de interesses
Uma iniciativa aparece muito em um grupo e pouco em outro	Pode indicar diferença de percepção entre grupos
Um grupo tem poucas respostas	As comparações com esse grupo devem ser feitas com cautela

Perguntas para discussão:

- A distribuição está concentrada ou espalhada?
- Alguma iniciativa aparece muito mais do que as outras?
- Algum grupo quase não aparece na base?
- O que pode ser um padrão real?
- O que pode ser efeito da amostra?

Mensagem importante:

Uma distribuição não explica sozinha por que algo aconteceu. Ela mostra uma forma inicial dos dados. A explicação exige contexto, documentação, discussão com a comunidade e, quando necessário, novas coletas.

5. Tipos de gráficos e escolhas de visualização

Nesta etapa, os estudantes relacionam tipos de dados a tipos de gráficos.

Sugestões de explicações:

- Nem todo gráfico serve para todo tipo de dado.
- Gráfico de barras ou colunas é bom para comparar categorias.
- Gráfico de setores pode mostrar partes de um todo, mas fica ruim quando há muitas categorias.
- Gráfico empilhado ajuda a comparar grupos dentro de categorias, mas pode ser mais difícil de construir.
- Tabela é melhor quando precisamos ver valores exatos.
- Visualizar bem é escolher uma forma de mostrar os dados sem distorcer a interpretação.

Orientação didática:

Para esta aula, priorizar gráficos de barras ou colunas. Gráficos de setores e gráficos empilhados podem ser usados apenas se a turma já tiver familiaridade com planilhas.

Quadro de apoio:

Tipo de informação	Visualização adequada	Exemplo no projeto
Contagem por iniciativa	Gráfico de barras ou colunas	Quantas vezes cada iniciativa foi escolhida
Percentual por vínculo	Gráfico de colunas	Distribuição dos respondentes por vínculo
Comparação entre grupos	Barras agrupadas ou tabela	Prioridades por estudantes, professores e familiares
Escala baixo/médio/alto	Barras empilhadas ou tabela	Impacto ou esforço percebido para cada iniciativa
Valores exatos	Tabela	Frequência e percentual por iniciativa

Muitas categorias	Tabela ou barras horizontais	Iniciativas I01 a I12
-------------------	------------------------------	-----------------------

Perguntas para discussão:

- Que gráfico é melhor para comparar as 12 iniciativas?
- O gráfico de pizza é adequado quando há muitas categorias?
- Como um gráfico pode destacar demais uma diferença pequena?
- Que informações precisam aparecer no título e nos rótulos?
- O gráfico deve mostrar números absolutos, percentuais ou ambos?

6. Atividade prática: construindo a primeira visão dos dados

A turma será organizada em grupos. Cada grupo receberá a base tratada e documentada do projeto Vozes da Escola ou de um projeto regional equivalente.

A atividade deve produzir uma primeira exploração com apoio de planilha eletrônica.

Para garantir a execução em 2 horas-aula, recomenda-se que:

- **Tarefa 1** e **Tarefa 2** sejam obrigatórias;
- **Tarefa 3** seja realizada se houver tempo ou por grupos selecionados;
- **Tarefa 4** seja obrigatória como fechamento interpretativo.

Tarefa 1 — Perfil dos respondentes

Os estudantes deverão construir uma tabela e um gráfico para responder:

- Quantas pessoas responderam?
- Qual a distribuição por vínculo com a escola?
- Qual a distribuição por etapa de ensino?
- Há algum grupo pouco representado?

Modelo de tabela:

Vínculo com a escola	Frequência	Percentual
Aluno(a)		
Professor(a)		
Familiar/Responsável		
Gestão		
Funcionário(a)/Colaborador(a)		
Comunidade		
Outro		

Perguntas para interpretação:

- Quem participou mais?
- Quem participou menos?
- Isso afeta a leitura das prioridades?
- O perfil dos respondentes parece equilibrado?

Tarefa 2 — Prioridade 1 das iniciativas

Os estudantes deverão construir uma tabela e um gráfico com as iniciativas escolhidas como Prioridade 1.

Modelo de tabela:

Código	Iniciativa	Frequência	Percentual
I01	Clube de estudo orientado		
I02	Tutoria entre pares		
I03	Oficina de leitura e escrita		
I04	Mural/boletim da escola		
I05	Feira/mostra de projetos		
I06	Roda de convivência e mediação		
I07	Ação cultural/esportiva		
I08	Uso pedagógico do celular		
I09	Organização de espaços de estudo		
I10	Cuidado do ambiente escolar		
I11	Sustentabilidade		
I12	Parceria local		

Perguntas para interpretação:

- Qual iniciativa aparece mais?
- As três primeiras estão próximas ou há uma diferença grande?
- Há iniciativas quase não escolhidas?
- Uma iniciativa pouco escolhida deve ser descartada automaticamente?
- Como a documentação do dataset ajuda a interpretar esses resultados?

Tarefa 3 — Comparação entre grupos

Se houver tempo, os estudantes deverão escolher uma comparação simples:

- prioridade por vínculo;
- prioridade por etapa de ensino;
- impacto percebido por iniciativa;
- esforço percebido por iniciativa;
- tempo estimado para implementação por iniciativa.

Exemplo:

Vínculo	I03 Oficina de leitura	I06 Roda de convivência	I11 Sustentabilidade
Aluno(a)			
Professor(a)			

Familiar/Responsável

Perguntas para interpretação:

- Os grupos indicam prioridades parecidas?
- Algum grupo escolhe mais uma iniciativa específica?
- O grupo comparado tem respostas suficientes?
- A diferença observada é grande ou pequena?
- O que ainda precisaríamos investigar?

Tarefa 4 — Primeira síntese interpretativa

Cada grupo deverá escrever uma síntese curta com a seguinte estrutura:

- **O que observamos:** descrição dos principais resultados;
- **O que parece ser um padrão:** tendência inicial identificada;
- **O que exige cautela:** limitação ou risco de interpretação;
- **O que precisamos investigar melhor:** pergunta para a próxima etapa.

Modelo:

Observamos que a iniciativa ____ apareceu entre as mais escolhidas como prioridade 1. Esse padrão pode indicar _____. No entanto, é necessário ter cuidado porque _____. Para compreender melhor esse padrão, seria importante investigar _____.

Exemplo:

Observamos que a iniciativa I03 — Oficina de leitura e escrita apareceu entre as mais escolhidas como prioridade 1. Esse padrão pode indicar interesse da comunidade por ações de leitura e escrita. No entanto, a maioria dos respondentes da base é formada por estudantes, então ainda é necessário verificar se professores, familiares e funcionários apresentam a mesma tendência. Antes de

concluir que esta é a principal prioridade da escola, precisamos analisar a distribuição por vínculo e etapa de ensino.

7. Leitura crítica dos gráficos

Nesta etapa, os grupos trocam os gráficos entre si. Cada grupo analisa a visualização produzida por outro grupo.

Modelo de análise:

Critério	Pergunta
Título	O gráfico tem título claro?
Eixos	Os eixos estão nomeados corretamente?
Escala	A escala facilita ou distorce a comparação?
Categorias	As categorias estão legíveis?
Percentuais	O gráfico mostra contagem, percentual ou ambos?
Interpretação	A conclusão está apoiada nos dados?
Limitações	O grupo indicou algum cuidado na leitura?

Perguntas para discussão:

- O gráfico ajuda ou confunde?
- A visualização permite comparar bem as categorias?
- O título induz uma conclusão antes da análise?
- A interpretação feita pelo grupo está apoiada nos dados?
- Há alguma conclusão exagerada?

Mensagem importante:

Um gráfico não é neutro. Escolhas como escala, título, ordem das categorias e tipo de visualização influenciam a leitura.

8. Relação com IA: por que explorar dados antes de modelar?

Esta etapa explicita a relação entre exploração de dados e sistemas de IA.

Sugestões de explicações:

- Antes de treinar um modelo, precisamos entender a base.
- Se uma categoria aparece muito mais do que outra, o modelo pode aprender mais sobre ela.
- Se um grupo quase não aparece nos dados, o modelo pode ter dificuldade para representar esse grupo.

- Se uma visualização mostra desequilíbrio, isso é um alerta para a etapa de IA.
- A exploração dos dados ajuda a decidir se a base está pronta ou se precisa de nova coleta, nova limpeza ou nova documentação.

Exemplos:

Achado exploratório	Implicação para IA
Poucos familiares responderam	O modelo pode representar pouco a visão das famílias
Uma iniciativa domina a base	O modelo pode priorizar essa categoria em excesso
Muitos campos textuais vazios	Análises de texto podem ficar limitadas
Algumas categorias têm poucos casos	Comparações e classificações podem ser instáveis
Grupos diferentes têm prioridades distintas	Pode ser necessário analisar por segmento, não apenas no total geral

Perguntas para discussão:

- Esta base está equilibrada para análise?
- Ela poderia ser usada para treinar um modelo?
- Que riscos existem se treinarmos uma IA sem observar as distribuições?
- O que a exploração mostra sobre a qualidade e a diversidade dos dados?
- Que decisão humana ainda precisa orientar o processo?

9. Fechamento

Ao final, a turma constrói coletivamente um checklist de exploração inicial dos dados.

Sugestão de checklist:

Item	Pergunta de verificação
Total de respostas	Sabemos quantas respostas existem na base?
Perfil dos respondentes	Sabemos quem respondeu?
Distribuição por grupo	Há grupos muito representados ou pouco representados?
Prioridades	Sabemos quais iniciativas aparecem com maior frequência?
Percentuais	Calculamos percentuais, não apenas contagens?
Comparações	Comparamos grupos com cautela?

Gráficos	Escolhemos gráficos adequados ao tipo de dado?
Escalas	As escalas dos gráficos não distorcem a leitura?
Interpretação	Diferenciamos descrição de explicação?
Limitações	Registramos cuidados e limites da análise?
IA	Relacionamos os padrões encontrados com possíveis impactos em sistemas de IA?

Sugestões para encerramento:

- Hoje vimos que explorar dados é observar antes de concluir.
- A estatística descritiva nos ajuda a enxergar distribuições e padrões iniciais.
- A visualização torna os dados mais compreensíveis, mas precisa ser construída com cuidado.
- Em projetos de IA, entender a distribuição dos dados é parte da responsabilidade técnica e ética.
- O próximo passo será usar essas primeiras leituras para formular perguntas mais específicas e preparar análises mais avançadas.

Produto da aula

Ao final da aula, cada grupo deverá entregar:

- uma tabela com o perfil dos respondentes;
- uma tabela com a distribuição das iniciativas escolhidas como prioridade 1;
- pelo menos dois gráficos produzidos a partir da base;
- uma comparação entre grupos ou categorias, se houver tempo;
- uma síntese interpretativa curta contendo:
 - o que os dados mostram;
 - qual padrão inicial foi observado;
 - qual cuidado interpretativo é necessário;
 - que pergunta deve orientar a próxima análise;
- uma breve resposta à pergunta:

“O que a exploração inicial dos dados revela e o que ela ainda não permite concluir?”

Avaliação	A avaliação pode ser processual e baseada na participação dos estudantes, na construção das tabelas e gráficos, na leitura crítica das visualizações e na qualidade da síntese interpretativa. O professor pode observar se os estudantes conseguem calcular frequências e percentuais, escolher gráficos adequados, interpretar distribuições, comparar grupos com cautela e diferenciar descrição de explicação causal. O produto avaliativo pode incluir: tabela com perfil dos respondentes, tabela com distribuição das iniciativas escolhidas como prioridade 1, pelo menos dois gráficos, uma comparação entre grupos ou categorias, quando possível, e uma síntese interpretativa
------------------	---

	curta indicando o que os dados mostram, qual padrão inicial foi observado, qual cuidado interpretativo é necessário e que pergunta deve orientar a próxima análise.
Material Complementar	• Material Didático do Professor — Tema 309: Exploração dos Dados; • Questionário padrão do projeto Vozes da Escola ou projeto regional; • Base sintética do projeto Vozes da Escola ou projeto regional; • Base limpa produzida no Plano 307; • Ficha de documentação do dataset produzida no Plano 308; • Dicionário de dados; • Lista oficial das iniciativas; • Site do projeto IA@Escola.