

Desenvolvimento de Habilidades em Ensino de Ciência de Dados: Relato de Experiência comparativa entre Metodologias

Elias Ytalo Silva Monteiro¹, César Henrique Araújo dos Santos¹, Fábio Henrique Hatem de Farias¹, Matheus Amorim de Andrade¹, Rodrigo Lins Rodrigues¹

¹Universidade Federal Rural de Pernambuco (UFRPE) – 52171-900 - Recife - PE - Brasil

{elias.ytalo, cesar.santos, fabio.hatem, matheus.andrade, rodrigo.linsrodrigues}@ufrpe.br

Resumo. Este artigo apresenta um relato de experiência sobre a realização do Estágio Supervisionado IV no curso de Licenciatura em Computação. O estágio teve como objetivo promover o ensino de computação por meio da linguagem R, aprimorando a interpretação de gráficos e o conhecimento estatístico dos alunos. Foram utilizadas duas abordagens metodológicas distintas: Aulas Expositivas Interativas e Peer Instruction, aplicadas em duas turmas diferentes. Os resultados indicam que ambas as metodologias são eficazes, mas apresentam impactos diferenciados no aprendizado. O estudo também destaca os desafios encontrados e as reflexões sobre o papel do docente na educação em computação.

Abstract. This article presents an experience report on the implementation of the Supervised Internship IV in the Bachelor's Degree in Computer Science program. The internship aimed to promote computer science education through the R programming language, enhancing students' ability to interpret graphs and their statistical knowledge. Two distinct methodological approaches were employed: Interactive Lecture Classes and Peer Instruction, applied in two different groups. The results indicate that both methodologies are effective but have distinct impacts on learning. The study also highlights the challenges encountered and reflections on the role of the teacher in computer science education.

1. Introdução

O estudo de dados tem se tornado cada vez mais essencial em diversas áreas do conhecimento, permitindo a análise e interpretação de informações para a tomada de decisões fundamentadas de acordo com Montgomery e Runger (2018). Com o crescimento exponencial do volume de dados disponíveis, impulsionado pelo avanço da tecnologia e da digitalização, torna-se indispensável compreender técnicas de manipulação, visualização e análise, bem como desenvolver habilidades estatísticas para interpretar padrões e tendências de forma precisa, segundo Freedman, Pisani e Purves (2017).

No contexto da educação, a Base Nacional Comum Curricular (BNCC) insere a ciência de dados no eixo de pensamento computacional, que abrange habilidades relacionadas à resolução de problemas, incluindo a descrição, análise e automação de processos. Especificamente, a habilidade EM13C012 estabelece que os estudantes devem ser capazes de produzir, analisar, gerir e compartilhar informações a partir de dados, utilizando princípios da ciência de dados Brasil (2018). Além disso, a BNCC de Matemática também enfatiza a importância da estatística no ensino médio, destacando

habilidades como EM13MAT305, que propõe a utilização de estatística descritiva para análise de dados, e EM13MAT403, que visa capacitar os estudantes a avaliar e interpretar informações estatísticas e probabilísticas. Dessa forma, compreender estatística e ciência de dados no ensino médio torna-se fundamental para a formação de cidadãos aptos a lidar com o grande volume de informações presente na sociedade atual.

Neste cenário a estatística desempenha um papel fundamental nesse cenário, sendo amplamente utilizada para sintetizar informações, identificar padrões e embasar conclusões em diferentes domínios, como saúde, economia, educação e ciências sociais em análise de Field (2022). Métodos como estatística descritiva e inferência estatística possibilitam a extração de insights a partir dos dados, contribuindo para a formulação de hipóteses e a tomada de decisões mais assertivas, como foi observado por Agresti e Franklin (2017). No contexto educacional, a BNCC enfatiza a necessidade de que os estudantes desenvolvam essas habilidades para a interpretação crítica de dados e informações, promovendo uma formação mais analítica e baseada em evidências.

Entretanto, a crescente disponibilidade de dados não assegura, por si só, a realização de análises eficazes. Estudos recentes, como o de Ben-Zvi e Garfield (2021), ressaltam que a ausência de uma formação estatística adequada pode resultar na reprodução de erros analíticos, comprometendo a confiabilidade dos resultados e afetando negativamente diversas áreas do conhecimento. Assim, surge a necessidade de capacitação em métodos estatísticos e ferramentas de análise para garantir a correta interpretação dos dados e a extração de informações relevantes. No ensino médio, esse desafio pode ser enfrentado com abordagens didáticas que integrem a ciência de dados aos componentes curriculares, alinhadas às diretrizes da BNCC.

Diante desse contexto, questiona-se: *quais metodologias podem ser utilizadas para melhorar o desempenho dos estudantes na análise estatística de dados?* Este estudo tem como objetivo comparar duas metodologias e discutir os desafios encontrados durante a implementação do ensino da estatística e ciência de dados, com um olhar especial para a implementação desses conceitos no ensino médio, conforme previsto na BNCC.

2. Referencial Teórico

2.1 A Metodologia Instrução por pares no Ensino de Programação

A integração de metodologias ativas no ensino de programação tem sido amplamente discutida na literatura, com destaque para a *Peer Instruction* ou em português Instrução por pares (IP), proposta por Eric Mazur na década de 1990. Estudos recentes, como os de Oliveira et al. (2017), Mariotto et al. (2018) e Ramos (2023), demonstram que a IP pode mitigar desafios comuns na aprendizagem de algoritmos e programação, como dificuldades de abstração e raciocínio lógico, fatores frequentemente associados à evasão em cursos de computação.

A eficácia da IP está ancorada em sua estrutura colaborativa, que substitui a passividade do modelo tradicional por ciclos de reflexão individual, discussão em pares e feedback imediato (Mariotto et al., 2018). Ramos (2023) destaca que a dinâmica de resolver questões conceituais, debater respostas divergentes e reavaliar escolhas após o diálogo estimula a construção ativa do conhecimento. Oliveira et al. (2017) corroboram

essa visão, apontando aumento de até 30% no desempenho de alunos em disciplinas introdutórias de programação após a adoção da IP.

Um diferencial dessa metodologia é sua adaptabilidade a ferramentas tecnológicas, como plataformas de votação em tempo real (ex.: Socrative), que agilizam a coleta de respostas e permitem ajustes didáticos imediatos (Mariotto et al., 2018; Ramos, 2023). Entretanto, Oliveira et al. (2017) alertam para a necessidade de calibrar questões conforme o perfil dos alunos (novatos vs. experientes), a fim de evitar desequilíbrios nas discussões.

No contexto deste estágio, a IP foi aplicada como complemento a aulas expositivas, visando superar limitações identificadas na literatura, como a superficialidade no entendimento de estruturas condicionais e loops, e pela preferência dos alunos, como visto em Mariotto et al., (2018). A escolha justifica-se pela ênfase da metodologia no desenvolvimento de argumentação lógica, habilidade crítica para a depuração de códigos.

2.2 Aulas expositivas

A aula expositiva permanece como uma das técnicas mais utilizadas no ensino, embora sua eficácia seja frequentemente questionada. Ribeiro (2007) reconhece sua utilidade na transmissão sintética de conteúdos complexos, especialmente em turmas heterogêneas, mas ressalta riscos como a passividade discente. Andreato (2019) ampliam essa crítica à luz da pedagogia freiriana, associando-a a uma concepção "bancária" de educação, onde o aluno é mero receptor.

Estudos apontam estratégias para revitalizar essa técnica. Ribeiro (2007) sugere a integração de organizadores avançados (mapas conceituais, analogias) e momentos de interação, como perguntas estratégicas ou atividades em duplas. Já Andreato (2019) defendem a incorporação de diálogos que conectem o conteúdo à realidade dos estudantes, princípio alinhado à abordagem de Paulo Freire.

3. Metodologia

3.1 Estrutura do Curso

O curso foi estruturado para ser desenvolvido ao longo de nove semanas, com dois encontros semanais, realizados às segundas e terças-feiras (para a turma 02), e, quarta-feira e sexta-feira (para a turma 01), totalizando uma carga horária média de seis horas por semana. As aulas aconteceram no período regular das atividades escolares, no turno da noite, considerando que os cursos da modalidade subsequente são ofertados nesse horário. O planejamento equilibrou teoria e prática, proporcionando aos alunos uma experiência abrangente no uso da linguagem R para análise de dados. Conforme segue o cronograma das duas turmas abaixo:

Tabela 1 - Cronograma das aulas da Turma 01

CRONOGRAMA PREVISTO DE AULAS - (Tuma 01 - QUA/SEX) - IA		
Módulo	Foco de Aprendizagem	Aulas Relacionadas
Módulo 1: Fundamentos e Ambiente	Introdução à Data Science, configuração do R/R Studio e lógica de programação básica (vetores e operações).	01, 02 e 03

CRONOGRAMA PREVISTO DE AULAS - (Tuma 01 - QUA/SEX) - IA		
Módulo	Foco de Aprendizagem	Aulas Relacionadas
Módulo 2: Gestão de Estruturas	Domínio de Data Frames e técnicas de importação/exportação de dados em formatos CSV e Excel.	04 e 05
Módulo 3: Manipulação (Tidyverse)	Limpeza de dados (NAs, pivotagem) e transformação avançada com o pacote dplyr (filter, select, mutate, group_by).	06, 07, 08 e 09
Módulo 4: Visualização de Dados	Ciência da visualização com ggplot2 , abrangendo desde gráficos básicos até camadas avançadas e refinamento de temas.	10 e 11
Módulo 5: Estatística e Fechamento	Aplicação de Estatística Básica (médias, moda, desvio padrão) e conclusão do curso.	12 e 13

Tabela 2 - Cronograma das aulas da Turma 02

CRONOGRAMA PREVISTO DE AULAS - (Tuma 02 - SEG/TER) - BD		
Módulo	Foco de Aprendizagem	Aulas Relacionadas
Módulo 1: Fundamentos e Ambiente	Introdução à Data Science, configuração do R/R Studio e lógica de programação básica (vetores e operações).	01, 02 e 03
Módulo 2: Gestão de Estruturas	Domínio de Data Frames e técnicas de importação/exportação de dados em formatos CSV e Excel.	04 e 05
Módulo 3: Manipulação (Tidyverse)	Limpeza de dados (NAs, pivotagem) e transformação avançada com o pacote dplyr (filter, select, mutate, group_by).	06, 07, 08 e 09
Módulo 4: Visualização de Dados	Ciência da visualização com ggplot2 , abrangendo desde gráficos básicos até camadas avançadas e refinamento de temas.	10 e 11
Módulo 5: Estatística e Fechamento	Aplicação de Estatística Básica (médias, moda, desvio padrão) e conclusão do curso.	12 e 13

A organização das aulas seguiu uma abordagem progressiva, iniciando com conceitos fundamentais da programação em R e avançando para temas mais complexos. Os conteúdos abordados incluíram: Introdução à Data Science, Introdução à Linguagem R, Instalação do R e do RStudio, Introdução à Programação em R, Exploração de Data Frames, Importação e Exportação de Dados, Limpeza e Preparação de Dados, Manipulação Básica de Dados com dplyr e tidyr, Visualização de Dados com ggplot2, Visualização Avançada com ggplot2, Estatística Básica e Documentação de Análises com RMarkdown. Todo o percurso formativo esteve de acordo com os princípios da Base Nacional Comum Curricular (BNCC), promovendo o desenvolvimento de competências como a capacidade de produzir, analisar, gerenciar e compartilhar informações a partir de dados, com base nos princípios da ciência de dados, conforme a habilidade EM13C012. Além disso, o curso também contribuiu para o desenvolvimento das habilidades EM13MAT403 e EM13MAT305, que envolvem o uso de conceitos estatísticos e fundamentos da ciência de dados na resolução de problemas e na interpretação de informações.

O ambiente de aprendizagem foi planejado para ser dinâmico e interativo, buscando engajar os alunos por meio de atividades práticas e desafios progressivos. As aulas foram realizadas em dois laboratórios de informática equipados com projetores e computadores conectados à internet, permitindo que os alunos desenvolvessem suas atividades de forma prática e acompanhando as explicações dos instrutores em tempo real.

A infraestrutura também contemplou o uso de plataformas educacionais como o Google Classroom, utilizado para o compartilhamento de materiais, e o Google Forms, empregado para a realização de avaliações diagnósticas e formativas. Além disso, foi criado um grupo no WhatsApp para comunicação direta entre alunos e professores, o que facilitou a resolução de dúvidas e promoveu uma interação contínua ao longo do curso.

O curso foi ministrado na Escola Técnica Estadual Professor Lucilo Ávila Pessoa, instituição pública do estado de Pernambuco. As turmas participantes foram compostas por estudantes do curso técnico em Redes de Computadores, especificamente dos primeiros e segundos módulos da modalidade subsequente — voltada para alunos que já concluíram o ensino médio e retornam à escola para cursar a formação técnica.

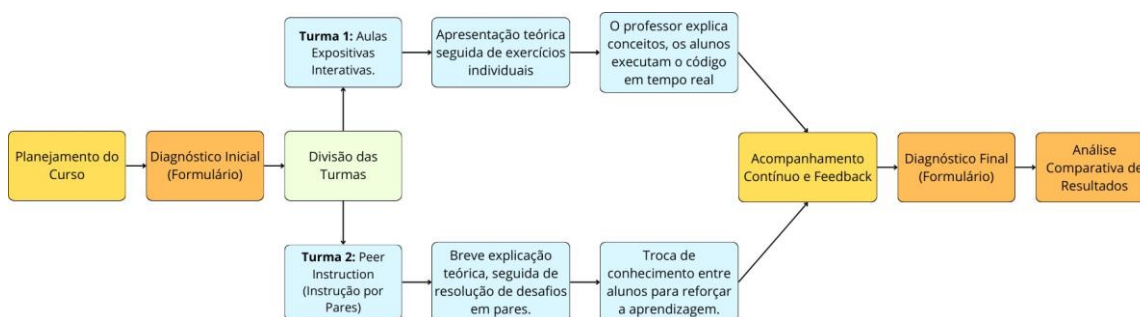


Figura 1. Fluxograma da Metodologia do Curso

O fluxograma apresentado ilustra a organização metodológica adotada ao longo do curso, evidenciando o planejamento estruturado desde a concepção inicial até a avaliação final dos resultados. A sequência de etapas demonstra um ciclo de ensino-aprendizagem que equilibra diferentes abordagens didáticas, garantindo um acompanhamento contínuo e uma análise comparativa do desempenho dos alunos.

Inicialmente, a aplicação de um diagnóstico inicial permitiu identificar o nível de conhecimento prévio dos participantes, possibilitando a divisão das turmas em dois

grupos distintos: um seguindo o modelo de Aulas Expositivas Interativas e outro adotando a estratégia de Instrução por pares. Essa divisão teve como objetivo avaliar os impactos de cada metodologia no aprendizado dos alunos.

Acompanhando todo o processo, a etapa de feedback contínuo possibilitou ajustes e intervenções pedagógicas conforme necessário, garantindo que as dificuldades enfrentadas pelos alunos fossem identificadas e tratadas de maneira eficaz. O diagnóstico final, realizado ao término do curso, permitiu comparar a evolução dos estudantes e analisar os impactos das diferentes metodologias de ensino. A partir dessa análise, foi possível realizar uma avaliação comparativa, identificando pontos fortes e desafios de cada abordagem, contribuindo para futuras melhorias na estruturação de cursos similares.

A pesquisa foi desenvolvida no âmbito da disciplina de Estágio Supervisionado IV, tendo como locus a Escola Técnica Estadual (ETE) Professor Lucilo Ávila Pessoa, situada na Avenida Caxangá, em Recife-PE. A escolha desta instituição e das turmas envolvidas ocorreu de forma intencional, fundamentada no vínculo prévio de um dos pesquisadores, Elias Ytalo Silva Monteiro, que é ex-aluno da referida escola e foi aluno do professor regente das turmas onde o curso foi ministrado, o Professor Renan Beserra Pereira da Silva. Essa proximidade facilitou o diálogo pedagógico e a viabilização da aplicação das metodologias propostas, garantindo um ambiente de colaboração entre o estagiário e o professor supervisor durante todas as etapas da intervenção.

3.2 Abordagens Metodológicas

Foram adotadas duas metodologias de ensino distintas para analisar seus impactos no aprendizado dos alunos: Aulas Expositivas Interativas e Instrução por pares.

3.2.1 Aulas Expositivas Interativas

Na Turma 1, a metodologia utilizada foi a de Aulas Expositivas Interativas, caracterizada por um modelo no qual o professor apresenta o conteúdo de forma estruturada, intercalando explicações teóricas com exercícios práticos individuais. O aprendizado foi pautado na autonomia dos alunos, incentivando-os a resolverem os desafios por conta própria após a exposição do conteúdo. Os estudantes tiveram a oportunidade de executar comandos em R logo após a explicação, consolidando imediatamente os conceitos abordados. Apesar do alto nível de organização, observou-se que alguns alunos apresentaram dificuldades para acompanhar o ritmo da exposição, o que exigiu um acompanhamento mais individualizado em determinados momentos.

3.2.2 Instrução por pares

Na Turma 2, a abordagem escolhida foi a de Instrução por pares, um método baseado na aprendizagem colaborativa, no qual os alunos discutem e resolvem problemas em pares ou pequenos grupos. Após uma breve explicação introdutória, os estudantes eram desafiados a resolver questões e apresentar suas soluções aos colegas, promovendo um ambiente de troca de conhecimento e fortalecimento do aprendizado. Esse formato incentivou a participação ativa dos alunos, desenvolvendo habilidades como argumentação, raciocínio crítico e comunicação técnica. No entanto, desafios surgiram no gerenciamento do tempo e na necessidade de equilibrar as contribuições individuais dentro dos grupos. Alguns alunos com menor familiaridade com a linguagem R enfrentaram dificuldades para acompanhar os colegas mais experientes, o que exigiu estratégias para garantir que todos participassem ativamente da construção do conhecimento.

3.3 Avaliação e Análise dos Resultados

A avaliação do curso foi composta por três elementos principais: exercícios práticos semanais, participação ativa nas discussões e um projeto final. Os exercícios semanais foram estruturados para reforçar os conceitos trabalhados em aula, permitindo que os alunos praticassem a aplicação dos comandos em R de forma independente. A participação nas atividades foi monitorada e registrada para analisar o nível de envolvimento dos estudantes e identificar possíveis dificuldades. O projeto final consistiu na aplicação dos conhecimentos adquiridos ao longo do curso para resolver um problema real de análise de dados, incentivando os alunos a consolidarem seu aprendizado e desenvolverem uma abordagem mais autônoma para a resolução de problemas computacionais.

Com base nessas estratégias, foi possível comparar as abordagens e seus impactos no aprendizado dos alunos. Enquanto as Aulas Expositivas Interativas proporcionaram um ambiente mais estruturado e individualizado, o Instrução por pares demonstrou um potencial significativo para desenvolver habilidades interpessoais e reforçar a aprendizagem por meio da colaboração. Essa análise foi fundamental para compreender os desafios e benefícios de cada metodologia, fornecendo insights valiosos para futuras aplicações no ensino de computação.

4. Resultados e Discussão

4.1 Prova de Nivelamento

No primeiro dia, os alunos realizaram uma prova de nivelamento com questões do Enem sobre análise de dados e interpretação de gráficos. Foram aplicadas provas distintas, porém equivalentes, para cada turma. O objetivo foi mensurar a proficiência inicial antes da introdução das metodologias ativas.

A Turma 1 (19 alunos) teve média inicial de 2,56, com 7 alunos acima da média, 9 abaixo e 3 ausentes. O boxplot indica que a mediana passou de 2,0 para 6,0 e o terceiro quartil, de 3,25 para 8,5, demonstrando significativa evolução, especialmente entre os alunos de menor desempenho inicial. Esses resultados orientaram práticas pedagógicas personalizadas, alinhadas à abordagem de Bloom (1984), que enfatiza a adaptação do ensino às necessidades individuais.

A Turma 2 (27 alunos) obteve média inicial de 3,68, com 11 alunos acima da média, 7 abaixo e 9 ausentes. O boxplot revela aumento da mediana de 4,0 para 6,0 e do terceiro quartil de 5,0 para 8,75, evidenciando um aprendizado mais equilibrado. A melhoria geral reforça a eficácia da metodologia aplicada e a importância de estratégias pedagógicas adaptadas. Os dados confirmam que a personalização do ensino contribuiu para a progressão dos alunos, permitindo que desenvolvessem habilidades essenciais para o curso.

4.2 Avaliação Final

No final do curso, realizamos o teste pós-curso com os alunos, aplicando as mesmas provas utilizadas no nivelamento inicial, porém desta vez invertendo as turmas onde foram aplicadas. Dessa forma, os alunos da Turma 1 responderam à prova originalmente aplicada à Turma 2, e vice-versa. O objetivo dessa avaliação foi medir o progresso dos alunos após semanas de aprendizado baseado em metodologias ativas. A comparação

entre os resultados do nivelamento inicial e do pós-teste permitiu identificar avanços individuais e coletivos.

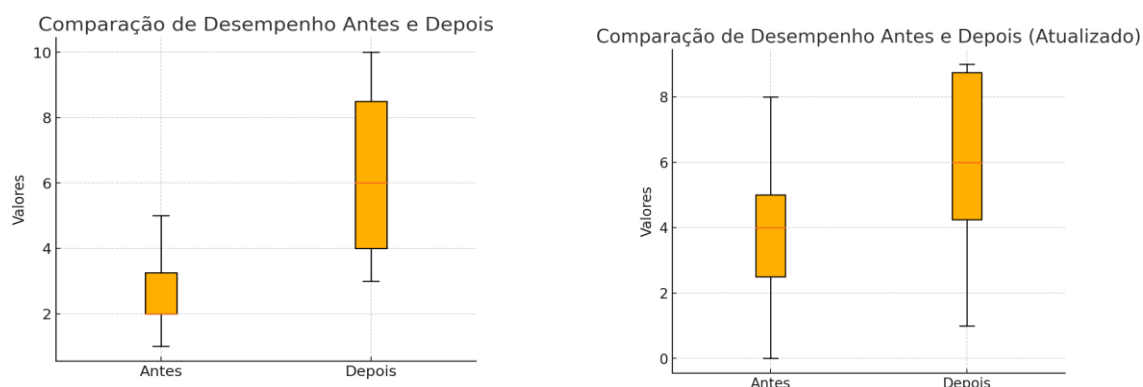


Figura 2. Comparação de desempenho da Turma 01 e Turma 02

Os dados analisados evidenciaram uma evolução significativa, demonstrando a eficácia das metodologias ativas adotadas. Os resultados da reaplicação das provas no pós-teste indicam uma clara progressão de proficiência dos alunos, reforçando o impacto positivo da abordagem pedagógica utilizada.

Esse avanço pode ser compreendido à luz do conceito de Zona de Desenvolvimento Proximal (ZDP), proposto por Vygotsky (1978), que destaca que a aprendizagem ocorre de maneira mais eficaz quando os alunos são desafiados além de suas capacidades atuais, mas recebem suporte adequado para superar essas dificuldades. Durante o curso, esse suporte foi oferecido de forma contínua, permitindo que os alunos ampliassem suas habilidades.

4.3 Análise do Desempenho por Conteúdo

A análise dos dados obtidos a partir do nivelamento e da avaliação final por turma revela uma evolução significativa nos diferentes conteúdos abordados ao longo do curso. Conforme se evidencia nos gráficos, nas figuras 3 e 4:

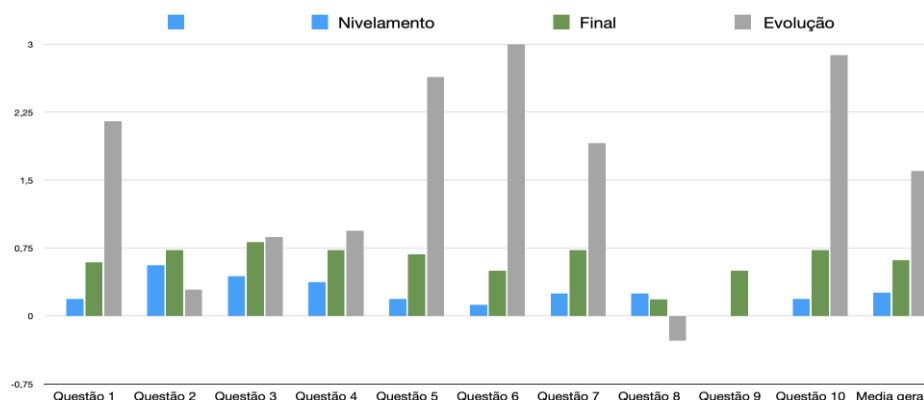


Figura 3. Comparação do desempenho da Turma 01 por área de conhecimento (ou questões) no pré-teste (barras azuis) e no pós-teste (barras verdes), seguidos da evolução em questão (barras cinzas).

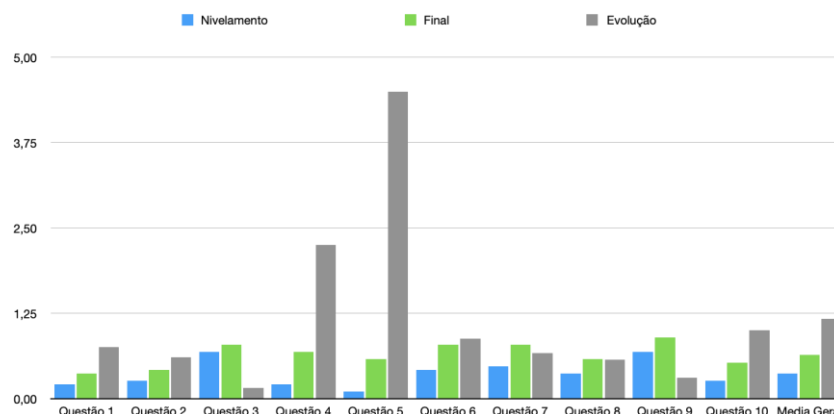


Figura 4. Comparação do desempenho da Turma 02 por área de conhecimento (ou questões) no pré-teste (barras azuis) e no pós-teste (barras verdes), seguidos da evolução em questão (barras cinzas).

A reorganização dos resultados por áreas de conhecimento permitiu uma compreensão mais clara do progresso dos alunos, destacando tanto os avanços quanto os desafios persistentes. Os Fundamentos de Programação e Estruturas de Controle apresentaram um crescimento de 19% para 59,09%, demonstrando que as estratégias adotadas para reforçar a lógica de programação e o uso de estruturas condicionais foram eficazes. Da mesma forma, em Manipulação de Dados e Operações Matemáticas, a taxa de acertos passou de 56% para 72,73%, indicando que a prática contínua e os exercícios aplicados consolidaram os conceitos essenciais dessa área.

A Análise Exploratória de Dados (EDA) registrou um dos avanços mais expressivos, evoluindo de 44% para 81,82%. Isso evidencia que, inicialmente, os alunos tinham dificuldades em interpretar e explorar dados, mas desenvolveram essas habilidades de forma significativa. Já em Visualização de Dados, o aumento de 38% para 72,73% demonstra que a introdução de ferramentas gráficas foi essencial para facilitar a compreensão dos dados e aprimorar a interpretação de gráficos. Em Modelagem Estatística e Probabilidade, a taxa de acertos subiu de 19% para 68,18%, mostrando que, apesar dos desafios iniciais, as atividades práticas e a abordagem contextualizada foram determinantes para a assimilação dos conteúdos. A Introdução ao Machine Learning também apresentou um crescimento notável, de 13% para 50%, embora ainda represente um desafio para os alunos. Esse resultado sugere que mais tempo e prática são necessários para consolidar os conceitos básicos dessa área.

A Aplicação de Modelos Estatísticos mostrou um avanço significativo, indo de 25% para 72,73%, o que demonstra que a abordagem utilizada para aplicar esses modelos foi eficiente, permitindo uma melhor compreensão da sua utilização na análise de dados. No entanto, a Interpretação de Resultados foi a única área em que houve uma leve queda no desempenho, passando de 25% para 18,18%. Isso sugere que os alunos encontraram dificuldades na análise crítica dos resultados gerados pelos modelos estatísticos e gráficos, apontando para a necessidade de reforço nesse aspecto.

Esses resultados demonstram que, de maneira geral, os alunos evoluíram significativamente ao longo do curso. No entanto, algumas áreas, como a interpretação

crítica de resultados e a introdução ao aprendizado de máquina, ainda exigem maior atenção para garantir uma assimilação mais sólida e aprofundada dos conteúdos.

A análise por conteúdo evidencia os avanços e desafios dos alunos ao longo do curso. Os maiores ganhos ocorreram em Análise Exploratória de Dados, Aplicação de Modelos Estatísticos e Visualização de Dados, indicando eficácia da metodologia adotada. Já Interpretação de Resultados e Introdução ao Machine Learning ainda requerem reforço para maior aprofundamento.

Na Prova 2, houve melhora significativa no desempenho, com reorganização dos resultados por conteúdo para facilitar a análise do progresso. Interpretação de alertas e monitoramento de dados cresceram de 21% para 37%, refletindo avanços na compreensão dos critérios de emissão, embora a interpretação de relatórios e gráficos ainda precise de melhorias. Análise de tendências evoluiu de 26% para 42%, demonstrando maior entendimento de previsões de mercado, mas exigindo reforços para análises mais críticas. Em Estatística Descritiva e Média Aritmética, o desempenho subiu de 68% para 79%, consolidando uma base prévia dos alunos. Análise de ruído e interpretação de gráficos avançaram de 21% para 68%, indicando melhora na leitura de gráficos ambientais. Cálculo de lucro e análise de custos subiram de 11% para 58%, mostrando eficiência da abordagem utilizada.

A codificação e frequência de dados tiveram progresso expressivo, de 42% para 79%, assim como modelagem estatística e medidas de tendência central, que passaram de 47% para 79%. Previsão de lucro e projeções econômicas melhoraram de 37% para 58%, enquanto mediana e estatística aplicada cresceram de 68% para 89%, reforçando a importância da estatística em fenômenos naturais. Por outro lado, cálculo percentual e análise de amostras, apesar do avanço de 26% para 53%, ainda representam desafios, exigindo reforço na interpretação de percentuais e cálculos amostrais. No geral, os temas com maior progresso foram Análise de Ruído, Estatística Descritiva e Modelagem Estatística, enquanto Cálculo Percentual e Interpretação de Alertas ainda precisam de intervenções pedagógicas adicionais.

Os resultados apontam um crescimento expressivo, especialmente entre alunos com desempenho inicial mais baixo, o que se alinha à teoria da Zona de Desenvolvimento Proximal (ZDP). O uso de metodologias ativas, conforme Brennan e Resnick (2012), contribuiu para o engajamento e a autonomia dos alunos no aprendizado de análise de dados. Para edições futuras do curso, recomenda-se reforçar os conteúdos com menor desempenho e expandir as estratégias eficazes nos tópicos de maior progresso, garantindo uma evolução equilibrada das habilidades dos alunos.

4.4 Desafios Encontrados

Os desafios encontrados durante o curso evidenciaram algumas dificuldades específicas dos alunos. A interpretação de resultados se mostrou um obstáculo, pois muitos tiveram dificuldades em compreender criticamente os dados apresentados e extrair conclusões relevantes, o que indica a necessidade de um reforço nesse aspecto. Da mesma forma, a introdução ao Machine Learning revelou-se desafiadora, sugerindo que os conceitos fundamentais dessa área podem demandar uma abordagem mais gradual e prática para facilitar o aprendizado. Além disso, a interpretação e aplicação de percentuais ainda não estão totalmente consolidadas, apontando para a necessidade de mais exercícios e

aplicações contextualizadas. Outro ponto observado foi a dificuldade na compreensão correta da emissão de alertas e de seus critérios, o que indica a importância de um treinamento mais aprofundado na análise de dados categorizados e na tomada de decisão baseada em indicadores.

Esses desafios evidenciam a necessidade de intervenções pedagógicas adicionais, com estratégias mais interativas e contextualizadas, para fortalecer a assimilação desses conceitos e aprimorar o aprendizado dos alunos.

5. Reflexões e Aprendizados como Futuros Docentes

No contexto educacional, a adaptação às necessidades dos alunos é um fator determinante para a eficácia do ensino, uma vez que cada turma apresenta um perfil de aprendizado distinto. Dessa forma, torna-se fundamental que o professor adote uma postura flexível, ajustando suas abordagens pedagógicas e o ritmo das aulas de acordo com as demandas específicas dos estudantes. Além disso, a avaliação contínua desempenha um papel essencial nesse processo, possibilitando o monitoramento do progresso dos alunos e permitindo a implementação de ajustes metodológicos e intervenções pedagógicas mais assertivas ao longo do curso.

A combinação de diferentes estratégias metodológicas contribui significativamente para a maximização do aprendizado, equilibrando o desenvolvimento da autonomia individual com a construção coletiva do conhecimento. Nesse sentido, o papel do docente transcende a mera transmissão de conteúdos, exigindo uma atuação como mediador do processo de ensino-aprendizagem. Ao assumir essa posição, o professor favorece um ambiente dinâmico e participativo, estimulando a interação, o pensamento crítico e a colaboração entre os alunos, aspectos fundamentais para a consolidação do conhecimento e o desenvolvimento de competências essenciais.

6. Considerações Finais

O estudo da estatística e da ciência de dados, conforme discutido neste trabalho, revela-se essencial não apenas para a análise e interpretação de informações, mas também para a tomada de decisões fundamentadas em diversas áreas do conhecimento. No contexto educacional, a BNCC reforça a necessidade de desenvolver habilidades estatísticas e de análise de dados no ensino médio, preparando os estudantes para uma sociedade cada vez mais orientada por informações quantitativas.

A introdução desses conceitos na educação básica é crucial para capacitar os alunos a interpretar dados de forma crítica e precisa, evitando equívocos que podem comprometer a confiabilidade das inferências realizadas. Além disso, a integração da estatística ao ensino médio, conforme proposto pela BNCC, possibilita uma formação mais alinhada às demandas contemporâneas, incentivando o pensamento analítico e o uso consciente das informações.

Portanto, compreender os desafios da análise estatística e aprimorar as estratégias de ensino nessa área é fundamental para garantir que os estudantes adquiram as competências necessárias para lidar com a complexidade dos dados na atualidade. O fortalecimento do ensino de estatística e ciência de dados contribuirá significativamente

para uma sociedade mais informada e preparada para enfrentar os desafios impostos pelo crescente volume de informações disponíveis.

7. Referência Bibliográfica

- Agresti, A. & Franklin, C. (2017). Estatística: a arte e a ciência de aprender com os dados (3ª ed.). Pearson.
- Ben-Zvi, D. & Garfield, J. (2021). Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice. Springer.
- Brasil, Ministério da Educação. (2018). Base Nacional Comum Curricular. Brasília, DF: MEC.
- Field, A. (2022). Descobrindo a estatística usando o SPSS (5ª ed.). SAGE Publications.
- Freedman, D., Pisani, R. & Purves, R. (2017). Estatística (4ª ed.). W. W. Norton & Company.
- Mariotto, P. et al. (2018). Aplicação do Método de ensino Peer Instruction para o Ensino de Lógica de Programação com acadêmicos do Curso de Ciência da Computação.
- Montgomery, D. C. & Runger, G. C. (2018). Estatística aplicada e probabilidade para engenheiros (6ª ed.). John Wiley & Sons.
- Oliveira, M. A. F. et al. (2017). Aplicação do método Peer Instruction no ensino de Algoritmos e programação de computadores. RENOTE, 15(1).
- Ramos, T. D. F. (2023). Peer Instruction (Instrução entre Pares): Uma proposta metodológica para o ensino e aprendizagem na educação básica. CAPES.
- Ribeiro, C. (2007). A aula magistral ou simplesmente aula expositiva. Máthesis, 16(16), 189–201.
- Tukey, J. W. (1977). Exploratory Data Analysis. Addison-Wesley.
- Andreata, M. A. (2019). Aula expositiva e Paulo Freire. Ensino em Re-Vista, 26(3), 700–724.