



## EDUCAÇÃO ESTATÍSTICA MEDIADO POR LINGUAGEM COMPUTACIONAL NA FORMAÇÃO DE PROFESSORES

### STATISTICAL EDUCATION MEDIATED BY COMPUTATIONAL LANGUAGE IN TEACHER TRAINING

**Profa. Dra. Kátia Milani Lara Bossi<sup>1</sup>**

**Prof. Dr. Juliano Schimiguel<sup>2</sup>**

**Prof. Dr. Marcio Eugen Klingenschmid Lopes dos Santos<sup>3</sup>**

**Prof. Dr. Marcelo Eloy Fernandes<sup>4</sup>**

**DOI: 10.5281/zenodo.21864695**

#### RESUMO

O presente artigo apresenta um recorte da tese intitulada “Educação Estatística mediada por Linguagem Computacional na Formação de Professores”, com o objetivo de divulgar os resultados da pesquisa que articulou três áreas do conhecimento: Linguagem de Programação, Formação Docente e Educação Estatística. A investigação contou com a participação de estudantes de pós-graduação (mestrado e doutorado) em Ensino de Ciências e Matemática de uma universidade localizada na cidade de São Paulo. Trata-se de um estudo exploratório que adotou uma abordagem metodológica mista, com coleta de dados quantitativos e qualitativos. As técnicas utilizadas foram observação e aplicação de questionários, sendo os dados analisados por meio da Análise de Conteúdo. Os resultados indicaram que a integração entre essas competências na formação de professores é tanto viável quanto pertinente. Ademais, a introdução da programação no ensino de estatística demonstrou potencial para fomentar uma integração futura mais consistente e eficaz dessas ferramentas no ambiente educacional, contribuindo significativamente para o fortalecimento do ensino e da aprendizagem.

**Palavras-chave:** Ensino de Estatística. Formação Docente. Tecnologia Digital. Linguagem de Programação.

#### ABSTRACT

This article presents an excerpt from the dissertation entitled “Statistical Education Mediated by Computational Language in Teacher Education”, aiming to share the results of a study that integrated three areas of knowledge: Programming Language, Teacher Education, and Statistical Education. The participants were postgraduate students

<sup>1</sup> Universidade Cruzeiro do Sul – ORCID: <https://orcid.org/0000-0001-9583-8743>

<sup>2</sup> Universidade Cruzeiro do Sul – ORCID: <https://orcid.org/0000-0001-8552-7984>

<sup>3</sup> Universidade Cruzeiro do Sul – ORCID: <https://orcid.org/0000-0002-9812-5981>

<sup>4</sup> Centro Paula Souza – ORCID: <https://orcid.org/0000-0001-9646-6646>

(Master's and Doctorate) in Science and Mathematics Education at a university located in São Paulo, Brazil. The research was exploratory in nature and adopted a mixed-methods approach, combining both quantitative and qualitative data. The techniques used included observation and questionnaires, and the data were analyzed through Content Analysis. The findings revealed that integrating these skills into teacher education is both feasible and relevant. Moreover, the introduction of programming into statistical education laid the groundwork for a more robust and effective future integration of these tools within educational contexts, significantly enhancing teaching and learning processes.

**Keywords:** Teaching Statistics. Teacher Training. Digital Technology. Programming Language.

## 1 INTRODUÇÃO

O presente artigo traz um recorte da Tese sob o título: “Educação Estatística mediado por Linguagem Computacional na Formação de Professores” vinculado ao Programa de Pós-graduação em Ensino de Matemática da Universidade Cruzeiro do Sul (BOSSI, 2024). A finalidade é apresentar os resultados da pesquisa que integrou três áreas Linguagem de Programação, Formação Docente e Educação Estatística.

A sociedade cada vez mais necessita de cidadãos bem-informados que devem avaliar informações vindas de diversas fontes: mídia, governo e entre outros. Como afirma Toledo e Lopes (2021) que o conhecimento

estatístico propicia o desenvolvimento do pensamento científico, necessário para o exercício de várias profissões, levando os alunos a concluir que o acesso ao saber matemático e estatístico é essencial na vida das pessoas, tanto no campo profissional quanto no pessoal (Toledo; Lopes, 2021, p. 17).

Coletar dados para obter *insights* e prever o comportamento a partir de descobertas de padrões utilizando técnicas e metodologias analíticas é o objetivo da Ciência de Dados. Conhecer os dados que serão analisados é extremamente importante e perguntas como essas surgirão: Como esses dados estão distribuídos? Quais as médias e desvios padrões? Como os dados estão relacionados? Eis que, surge a necessidade da Análise Exploratória de Dados, também conhecida pela sigla AED, que tem como finalidade verificar e conhecer os dados antes de tirar conclusões, sendo a primeira etapa no processo de análise

de dados a fim de identificar padrões, tendências, dados discrepantes e entre outros para fornecer uma visão geral do conjunto de dados (AMARAL, 2016).

As tecnologias digitais podem ser importantes ferramentas para o processo de ensino-aprendizagem modificam a prática pedagógica com um novo pensar e novas formas de aprender, pois permitem a criação de representações visuais e interativas que tornam certos conceitos mais concretos e acessíveis (BOSSI e SCHIMIGUEL, 2023). Na Análise de Dados o uso de *softwares* e linguagens de programação como *Python* permite a criação de representações visuais e interativas que tornam certos conceitos mais concretos e acessíveis e ainda, facilitando a manipulação das grandes bases de dados.

Sendo assim, consideramos fundamental capacitar o docente que atua no Ensino de Estatística para propor atividades de aprendizagem nas quais os alunos possam atribuir significados e desenvolver o pensamento estatístico por meio da integração de tecnologias.

A formação em novas tecnologias permite ao docente que perceba na sua própria realidade, interesses e expectativas, como as tecnologias podem ser úteis a ele (LIMA, 2020). É imprescindível que os docentes tenham a possibilidade de participarem de formações contínuas que possibilitam o contato com novas opções tecnológicas e como utilizá-las como ferramentas pedagógicas.

A inclusão da "Linguagem de Programação" na formação docente pode permitir que os docentes adquiram habilidades que são essenciais para a análise de dados e enriqueçam suas práticas pedagógicas.

Sendo assim, o objetivo geral da pesquisa foi investigar e aplicar linguagem computacional como estratégia pedagógica no Ensino da Estatística envolvendo alunos do Programa de Pós-graduação para proporcionar aos alunos uma experiência enriquecedora no uso da linguagem de programação Python para a Análise de Dados, com um foco específico na Análise Exploratória de Dados (AED).

O presente artigo está dividido em Referencial Teórico, Metodologia e por fim Resultados e Discussões.

## **2 REFERENCIAL TEÓRICO**

Nessa seção traremos o referencial teórico sobre as temáticas Educação Estatística, a Formação Docente da Estatística e o uso da tecnologia no ensino da Estatística.

### **2.1 EDUCAÇÃO ESTATÍSTICA**

A Estatística visa desenvolver métodos para coletar, organizar e analisar dados e a Educação Estatística é uma área de pesquisa que nasceu muito tempo depois e surgiu da necessidade de investigar meios para sanar as dificuldades encontradas por docentes, ao ensinar os conceitos e procedimentos estatísticos aos alunos. Com isso, a Educação Estatística pretende estudar como as pessoas ensinam e aprendem Estatística (CAZORLA, KATAOKA e SILVA, 2010).

É importante promover uma educação que estimule a criatividade e a reflexão do aluno. Promovendo a inserção crítica dele na realidade em que vive para tornar-se um ator que não só assiste ao mundo, mas que dele participe. Campos (2007) afirma que a Estatística deve ser trabalhada por meio de projetos, que utilize exemplos reais, sempre contextualizados dentro de uma realidade condizente a fim de preparar o aluno para interpretar o mundo praticando o discurso da responsabilidade social, mas utilizando bases tecnológicas a fim de valorizar e desenvolver competências para um aluno que vive numa sociedade eminentemente tecnológica.

Os objetivos da Educação Estatística congregam com os da Educação Crítica a fim de produzir uma pedagogia democrática, reflexiva com responsabilidade social. Lopes (2008) afirma que uma Educação Estatística Crítica

requer do professor uma atitude de respeito aos saberes que o estudante traz à escola, que foram adquiridos por sua vida em sociedade. Em nosso modo de entender, seria necessária a discussão de temas, como a poluição dos rios e mares, os baixos níveis do bem-estar das populações, o abandono da saúde pública, entre outros; questões que estão em manchetes de jornais diários e revistas e em reportagens de televisão. Trabalhando a análise dessas questões que estão sempre envolvidas em índices, tabelas, gráficos etc., podemos estar viabilizando a formação de cidadãos críticos, éticos e reflexivos (Lopes, 2008, p. 62).

A estatística desempenha um papel crucial nas empresas, ajudando na tomada de decisões estratégicas e operacionais permitindo que tomem decisões com base em análises objetivas e, ainda analisar o comportamento do consumidor e entender melhor as necessidades e preferências do seu público-alvo. Na área financeira, as previsões ajudam a projetar receitas, despesas e fluxos de caixa.

A AED é uma abordagem na estatística que determina na exploração a obtenção de insights a partir de um conjunto de dados a fim de resumir suas principais características.

Ela “olha” para os dados de forma cuidadosa e profunda para resumir a quantidade de informação presentes neles. Esta exploração inicial do conjunto de dados é crítica porque ajuda os cientistas de dados a iluminarem padrões previamente desconhecidos, relacionamentos ou outros achados. E, pode ser utilizada como uma ferramenta pedagógica para facilitar o aprendizado. Ao invés de começar com conceitos abstratos e fórmulas sem contexto, os docentes podem iniciar pelo universo dos dados por meio de técnicas exploratórias.

## **2.1 FORMAÇÃO DOCENTE NA EDUCAÇÃO ESTATÍSTICA E O USO DE TECNOLOGIAS NO ENSINO DA ESTATÍSTICA**

A abordagem tradicional da formação matemática, que se baseava principalmente em fórmulas, algoritmos e exercícios repetitivos para fixação de conteúdos, já não atende mais às necessidades educacionais atuais. Espera-se que a educação vá além da memorização e da aplicação mecânica de regras. É necessário desenvolver o pensamento crítico e a aplicação prática dos conceitos matemáticos em situações reais. De acordo com Lopes (2013, p. 905) “para ensinar estatística, não é suficiente entender a teoria matemática e os procedimentos estatísticos; é preciso fornecer ilustrações reais aos estudantes e saber como usá-las para envolver os alunos[...]”.

É essencial que o processo formativo capacite os educadores a criar atividades didáticas que promovam uma aprendizagem estatística significativa e contextualizada. Os futuros professores precisam estar preparados para desenvolver práticas pedagógicas que permitam aos estudantes compreenderem a Estatística como uma ferramenta poderosa

para a análise de situações do mundo real, e não apenas como um conjunto de procedimentos técnicos. Isso implica que os professores devem ser capazes de elaborar atividades que estimulem o pensamento crítico dos alunos, incentivando-os a questionar, analisar e interpretar dados em diferentes contextos (BARBOSA, *et al.*, 2020).

Nessa jornada de desenvolvimento é fundamental reconhecer o professor como um agente ativo em seu processo de aprendizagem. A formação docente deve estar estruturada de maneira a motivar o professor a investir continuamente em sua carreira. Isso envolve proporcionar oportunidades de desenvolvimento profissional que sejam relevantes e que inspirem o docente a buscar o aprimoramento constante de suas habilidades e conhecimentos promovendo uma reflexão crítica sobre a prática pedagógica e encorajar o professor a se envolver ativamente em seu crescimento profissional levando o docente a investir em sua carreira (LOPES, 2008).

Por meio de experiências, os professores terão a oportunidade de se tornarem protagonistas em seu próprio processo de aprendizagem, desenvolvem habilidades essenciais. Ao refletirem sobre essas práticas em um contexto formativo, os docentes podem avaliar a eficácia dessas abordagens pedagógicas e decidir, de forma fundamentada, como e quando incluir tais procedimentos em seu repertório profissional. Essa reflexão crítica sobre a prática pedagógica permite que os docentes adaptem e personalizem suas estratégias para atender melhor às necessidades e contextos específicos de suas turmas (BARBOSA, *et al.*, 2020).

As diretrizes GAISE (2016) foram desenvolvidas pela *American Statistical Association (ASA)* a fim de fornecer uma ferramenta de ensino e avaliação da estatística que enfatiza a compreensão conceitual em vez da mera memorização de procedimentos que vem influenciando o ensino da Estatística no Brasil e são baseadas em algumas recomendações como: importância dos alunos entenderem os conceitos fundamentais da estatística, ao invés de apenas memorizarem procedimentos; usar dados reais em situações reais promovendo o uso de dados reais em situações de aprendizagem, para que os alunos veem relevância da estatística em suas vidas diárias e em contextos diversos, como economia, saúde e ciência; enfatizar a análise de dados e a interpretação de resultados; desenvolver habilidades de comunicação estatística utilizando gráficos e

tabelas para apoiar suas conclusões; promover o pensamento crítico e a tomada de decisões informadas: preparar os estudantes para serem cidadãos bem informados e capazes de analisar informações estatísticas de maneira crítica e utilizar tecnologia para facilitar o aprendizado, como o incentivo ao uso de ferramentas tecnológicas e softwares estatísticos que facilitem a visualização e análise de dados.

Os recursos tecnológicos devem ser empregados para ajudar os estudantes a visualizar conceitos e desenvolver a compreensão de ideias abstratas por simulações. Sobre essa temática Lopes (2013, p. 908) comenta que “Independentemente das ferramentas utilizadas, um *software* é importante para visualizar o uso da tecnologia não apenas como uma maneira para calcular números, mas como uma forma de explorar ideias conceituais e melhorar a aprendizagem dos alunos”.

No ensino da Estatística, por exemplo, a programação permite que os alunos manipulem grandes conjuntos de dados, visualizem padrões e realizem análises complexas, facilitando a compreensão de conceitos abstratos. E, ainda oferece ferramentas para explorar conceitos em diferentes áreas, como matemática, ciências e até humanidades. Ao programar, os estudantes aprendem a criar algoritmos, analisar dados e automatizar tarefas, habilidades que são cada vez mais valorizadas no mundo moderno.

É necessário que os professores estejam preparados para assumir essa responsabilidade. A formação é essencial para acompanhar tamanha onda de mudanças e desenvolvimento. “Na educação o foco, além de ensinar, é ajudar a integrar ensino e vida, reflexão e ação, a ter uma visão de totalidade. Educar é ajudar a integrar todas as dimensões da vida, a encontrar nosso caminho intelectual” (MORAN, MASETTO e BEHRENS, 2000, p. 12).

### 3 METODOLOGIA

Este estudo se configurou como uma pesquisa exploratória com dados quantitativos e qualitativos visando explorar um tópico para entender suas

complexidades. E, como o seu planejamento é bastante flexível possibilita ao pesquisador utilizar uma variedade de técnicas, como revisão de literatura, entrevistas e observações.

A coleta de dados foi aplicado nos meses maio e junho de 2024 por meio de questionários constituídos de diversas perguntas ordenadas e respondidas pelos participantes sem a intervenção da pesquisadora. Houve perguntas abertas (chamadas livres ou não limitadas), fechadas e de múltiplas escolhas (perguntas fechadas que apresentam uma série de possíveis respostas). Nos questionários, também, usamos a escala de *Likert* no qual os respondentes escolheram somente um dos pontos fixos estipulados em uma linha num sistema com um ponto neutro no meio da escala havendo ancoras nas extremidades. Esses questionários foram preenchidos pelos participantes e, também pela pesquisadora como Diário.

Os participantes do estudo consistiram em 19 alunos da Pós-graduação (Mestrado e Doutorado) em Ensino de Ciências e Matemática da Universidade Cruzeiro do Sul, situados na cidade de São Paulo. Por se tratar de uma pesquisa envolvendo seres humanos o projeto foi aprovado no Comitê de Ética e Pesquisa conforme CAAE 75690823.0.0000.8084, número do parecer 6.525.189

No decorrer do curso de mestrado/doutorado, os participantes realizaram uma disciplina eletiva estruturada em 7 encontros/aulas presenciais semanais com 4 horas/aula dia que ocorreram nos meses de maio e junho de 2024.

O objetivo do 1º encontro foi fornecer uma compreensão abrangente da Ciência de Dados, abordando os tipos de análise de dados, a conceituação e sua transformação, a Análise Exploratória de Dados (AED) e as etapas envolvidas. E, ainda, o que são variáveis, incluindo seus tipos.

No 2º Encontro o objetivo foi preparar os alunos a utilizar a linguagem de programação *Python* e a ferramenta Colab para a análise manipulando os dados utilizando a biblioteca *Pandas*, criar *dataframes*, agrupar dados com a função *groupby* e calcular distribuições de frequência. Criar tabelas de frequência utilizando *Python*, utilizando conceitos estatísticos como frequência absoluta, frequência acumulada e frequência relativa.



Em seguida, no 3º encontro o objetivo foi entender e aplicar conceitos de estatística descritiva utilizando a biblioteca Pandas em Python. Calcular e interpretar medidas de posição (média aritmética, moda, mediana, quartis) e medidas de dispersão (variância, desvio padrão). Carregar dados de arquivos Excel, manipular dataframes, selecionar colunas específicas, e identificar valores extremos dentro dos dados.

Logo após, no 4º encontro o objetivo foi utilizar as bibliotecas *Seaborn* e *Matplotlib* para criar e analisar diversos tipos de gráficos, aplicando conceitos estatísticos. Criar gráficos de barra vertical e horizontal, gráficos de setor (pizza), histogramas, *box plots* e diagramas de dispersão. Compreender e aplicar conceitos estatísticos como distribuição de dados, medidas de tendência central e dispersão.

No 5º encontro o objetivo foi realizar a limpeza de dados. Identificar e tratar valores ausentes, observar e alterar tipos de dados, remover espaços em branco e substituir valores negativos.

E, no penúltimo encontro o objetivo foi analisar e interpretar relações entre variáveis qualitativas e quantitativas utilizando medidas de estatística descritiva e gráficos apropriados.

No 7º encontro os alunos em grupos apresentaram as conclusões de uma análise exploratória de dados referente a uma base de dados real e governamental que poderia ser sobre cultura, educação, saúde e entre outros. E, ao final, os alunos foram convidados a responder o Questionário Final.

No primeiro encontro os estudantes que aceitaram participar da pesquisa assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) em que foram informados a respeito da proposta de trabalho colaborativo e dos instrumentos utilizados para a coleta de dados. Responderam o Primeiro questionário denominado Perfil. Esse questionário constou com várias questões sobre tempo de atuação, formação acadêmica, tecnologias utilizadas, dificuldades tecnológicas e conhecimento de linguagens de programação.

No sétimo encontro os alunos foram convidados a responderem o questionário final que englobou diversos aspectos relacionados à percepção dos alunos sobre as aulas.

Após a coleta de dados obtidos os resultados o próximo passo foi a análise e interpretação e de acordo com Marconi e Lakatos (2017, p. 183) “na interpretação dos dados de pesquisa, é importante que eles sejam colocados de forma sintética e de maneira clara e acessível”. Utilizamos a Análise de Conteúdo que envolve análise de texto e pode ser usado em conjunto com dados quantitativos para contextualizar as descobertas.

## **4 RESULTADOS E DISCUSSÃO**

Os formulários foram respondidos de forma anônima. Cada questionário foi feito em formulários distintos e, sem qualquer identificação do respondente.

O questionário denominado Perfil incluiu diversas questões sobre tempo de atuação, formação acadêmica, tecnologias utilizadas, dificuldades tecnológicas e conhecimento de linguagens de programação.

Referente à atuação e aos níveis de ensino a maioria dos alunos tem uma experiência significativa em docência, com sete respondentes possuindo mais de 20 anos de experiência e ainda podemos afirmar que aproximadamente 78% dos professores têm mais de 10 anos de experiência.

Muitos atuam no Ensino Médio (14 docentes), seguido pelo Ensino Superior e Educação Fundamental - Últimos Anos, ambos com nove respondentes. Essa variedade foi benéfica para a pesquisa, pois permitiu que a coleta de dados tivesse diferentes perspectivas educacionais, enriquecendo a análise e as conclusões.

No que se refere a formação acadêmica destacou-se com 40% a licenciatura em Matemática. E, Pedagogia é a segunda formação mais comum, em torno de 25%. As demais foram Licenciatura em Física, em Ciências Biológicas, Química, Administração e Sistema de Informação, com um respondente cada.

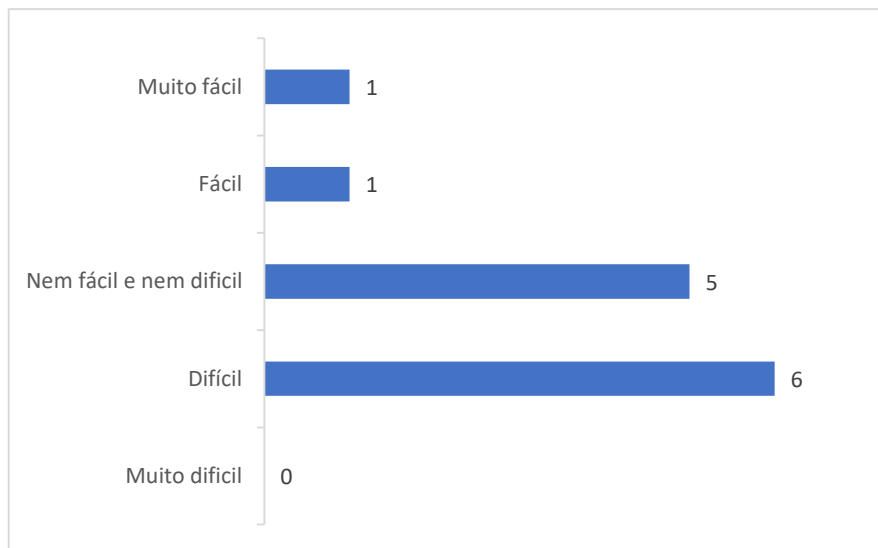
No quesito tecnologia ao serem questionados: “Quais ferramentas ou recursos tecnológicos você utiliza com mais frequência em suas aulas?” A predominância foi o uso de computadores indicando que é a ferramenta tecnológica mais comum nas salas de aula, possivelmente devido à sua versatilidade e ampla gama de aplicações educacionais. Softwares educacionais também foram bem votados o que sugere uma integração significativa de ferramentas digitais específicas para o ensino e aprendizagem. No item “Outros” o aluno indicou como “Atividades com Sistemas embarcados (Arduino)”.

Em relação quais *softwares*/aplicativos já usou em suas aulas foram citadas uma variedade de ferramentas desde a programação (Python, Scratch) até a colaboração e comunicação (Teams, Google Forms). E, houve a presença de plataformas de aprendizado interativo e personalizável (Seneca, Mesalva) indicando um movimento em direção ao ensino adaptativo.

Diante do pergunta: “Quais desafios ou obstáculos você enfrenta ao usar a tecnologia na sala de aula?” A maioria apontou a infraestrutura como a principal dificuldade para a implementação de tecnologia em sala de aula. Problemas como acesso à internet, falta de equipamentos modernos e velocidade da rede indicando uma necessidade crítica de melhorias na infraestrutura tecnológica. Ainda emergiu a discussão sobre a dificuldades dos alunos em relação a certas tecnologias indicando que, embora menos frequentes, existem barreiras no uso da tecnologia pelos alunos, como falta de acesso a dispositivos em casa ou falta de habilidades digitais. A dificuldade dos docentes, também foi citada, demonstrando a necessidade de formação contínua para facilitar a integração eficaz da tecnologia no ensino.

Referente às percepções dos professores em relação as dificuldade sobre a linguagem de programação Python a Figura 1 - Nível de dificuldade linguagem Python apresenta as respostas dos alunos à pergunta sobre como eles avaliam o seu nível de dificuldade na utilização da linguagem Python durante os encontros.

Figura 1 - Nível de dificuldade linguagem Python



Fonte: Autores (2024)

Seis participantes classificaram a utilização da linguagem como "Difícil". Esta avaliação sugere que muitos participantes encontraram desafios consideráveis ao trabalhar com *Python* durante os encontros. Ainda, cinco pessoas a consideram "Nem fácil nem difícil", sugerindo que a linguagem não foi percebida como extremamente simples, mas também não apresentou grandes desafios. E, duas pessoas consideraram Fácil ou Muito fácil. A ausência de respostas para "Muito difícil" reforça a reputação como uma linguagem acessível para a maioria das pessoas.

Na categoria *Feedback* a pergunta "Houve algum assunto tratado nos encontros que teve mais dificuldade? Se sim, qual?" Cinco alunos indicaram que não encontraram dificuldades significativas com os assuntos tratados nos encontros, sendo um ponto positivo. Três alunos apontaram a programação como a área de maior dificuldade evidenciando que, embora o interesse e a participação foram bons, a programação representou um desafio para uma fração dos estudantes, possivelmente devido à complexidade inerente ao aprendizado de novas linguagens de codificação. Dois alunos identificaram a estatística como um dos temas mais desafiadores. Este dado sugere que, apesar dos esforços para reforçar os conceitos estatísticos, ainda existiram lacunas de compreensão.

Diante da questão: “O que mais gostou?” a maioria dos alunos (8) expressou um forte entusiasmo pelo aprendizado da linguagem Python refletindo uma certa satisfação com a introdução a essa ferramenta e sua aplicabilidade. A abordagem didática (4), também, foi valorizada sugerindo um ambiente de aprendizado positivo e de apoio. A praticidade e a aplicação dos conceitos ensinados foram mencionadas indicando que os alunos não só aprenderam conceitos teóricos, mas também viram como aplicá-los em situações reais. O desenvolvimento de uma linguagem computacional em colaboração com os colegas foi um aspecto apreciado indicando que a colaboração e a interação com os pares foram partes importantes da experiência de aprendizado. A integração de diferentes áreas do conhecimento, como estatística e programação, foi reconhecida e valorizada destacando a importância e o valor percebido dessa abordagem multidisciplinar. Enfim, a análise das respostas revela que as aulas foram bem recebidas.

Sobre o que “menos gostaram” durante as aulas várias respostas indicaram que a curta duração do curso foi um desafio significativo. Alguns alunos mencionaram dificuldades específicas relacionadas à execução de comandos e ao entendimento dos conceitos de programação. Houve também respostas que expressaram uma satisfação geral com o curso, indicando que, para alguns participantes, a experiência foi completamente positiva.

Ao serem questionados “O que você achou da experiência na utilização do Python nas aulas de estatística?” Muitas respostas refletiram um alto grau de apreciação e uma percepção muito positiva sobre o uso da linguagem nas aulas. Diversos alunos destacaram o valor e a novidade da experiência. A utilidade prática do Python, também, foi destacada por vários participantes e refletem a percepção de uma ferramenta aplicável em contextos reais, especialmente para a análise de grandes volumes de dados. Embora, a maioria das respostas fosse positiva, alguns alunos relataram dificuldades significativas encontraram barreiras no aprendizado e na aplicação da linguagem.

A proposta do Trabalho Final era realizar uma AED de uma base de dados disponível em algum site governamental, ou seja, os dados deveriam ser extraídos de uma base real e pública, como por exemplo <http://dados.prefeitura.sp.gov.br/>. Assim utilizaram os conceitos utilizados em aula para responder o problema ou objetivo

determinado pelo grupo. No último dia de aula, ocorreram as apresentações dos grupos para mostrar as análises efetuadas e as principais conclusões.

Foram 6 grupos e conforme combinado cada grupo realizou a Análise Exploratória utilizando a linguagem Python de uma base de dados escolhida pelo grupo. A fim de preservar a identidade dos alunos não citaremos os seus nomes. Somente os citaremos por grupos como G1, G2, G3, G4, G5 e G6.

O G1 comparou os resultados de uma unidade escolar em relação às regiões pertencentes à mesma Diretoria Regional de Educação (DRE) referente às notas das disciplinas de Matemática e Língua Portuguesa. Realizaram a busca de dados abertos da Prefeitura de São Paulo do ano de 2022.

A análise do G2 foi norteada pelas seguintes questões: Qual o valor total que cada Diretoria Regional de Educação (DRE) recebeu de verba? Qual a proporção do valor total recebido? Quantas unidades educacionais existem por DRE? Qual a média dos valores recebidos por DRE? E, para isso a Base de Dados analisada foi Programa de Transferência de Recursos Financeiros (PTRF) do ano de 2023, assim durante a apresentação os alunos destacaram as análises e conclusões sobre a distribuição e uso dos recursos financeiros transferidos às unidades educacionais, considerando as especificidades de cada Diretoria Regional de Educação.

O G3 analisou a quantidade de alunos por turma no Ensino Fundamental e Ensino Médio no Brasil. Utilizaram os dados do INEP sobre a quantidade de alunos por turma no ano de 2023. A questão de pesquisa que norteou foi: Como se distribui a quantidade de alunos por turma no Ensino Fundamental e no Ensino Médio nas diferentes regiões do Brasil? Durante a análise foram verificados: Número máximo e mínimo de alunos por turma; Estados com maior e menor número de alunos por nível de ensino e o número de alunos por sala por nível de ensino.

O G4 abordou a demanda e matrículas na Rede Municipal de Ensino da Cidade de São Paulo, proporcionando informações detalhadas sobre o acesso e permanência de crianças, jovens e adultos nas instituições educacionais. Foram utilizadas as informações da Secretaria Municipal de Educação disponíveis no Portal de Dados Abertos.

Apresentaram uma abordagem estruturada sobre a demanda registrada e as matrículas na Educação Infantil, Fundamental e EJA em São Paulo.

A análise do G5 foi norteadada pela questão: Como a rede UniCEU (Universidade nos Centros Educacionais Unificados) contribui para a formação inicial e continuada dos estudantes, promovendo a expansão do acesso ao ensino superior na cidade de São Paulo? Os dados utilizados na análise referem-se aos polos UniCEU, cursos e vagas oferecidas. Foi apresentado o Total de vagas oferecidas por Instituições de Ensino Superior (IES) parceiras, como UNIFESP (Universidade Federal de São Paulo), UFABC (Universidade Federal do ABC) e FIOCRUZ (Fundação Oswaldo Cruz), a distribuição de vagas por polo UniCEU, a quantidade de vagas por modalidade de curso (Graduação, Tecnológico, Especializações, etc.) e a identificação dos cursos com menor e o maior número de vagas.

O G6 analisou o perfil de repasses do PDDE (Programa Dinheiro Direto na Escola) em 2023 para as escolas públicas do Brasil. O banco de dados escolhido refere-se aos repasses do PDDE para as escolas públicas do Brasil. Durante a análise o grupo identificou o perfil de repasses do PDDE por localização (urbana e rural) e por esfera administrativa.

Todos os grupos utilizaram diferentes bases de dados educacionais para identificar padrões e fornecer *insights* sobre diferentes aspectos do sistema educacional, como matrículas, repasses financeiros, desempenho escolar e distribuição de alunos por turma. E, apesar da disponibilidade de bases de dados sobre outros assuntos nos *sites*, os grupos escolheram voluntariamente a temática e ainda assim cada grupo se focou em áreas específicas refletindo a diversidade de temas que podem ser abordados através da análise exploratória de dados em educação. Todos importaram e analisaram os dados utilizando o ambiente do Google Colab e foram empregadas bibliotecas *Python*, como *Pandas*, *Seaborn* e *Matplotlib*.

## 5 CONCLUSÃO

Embora tenha havido momentos em que foi necessário revisar conceitos estatísticos e enfrentar dificuldades com a programação, ao final na apresentação dos Trabalhos os grupos demonstraram um uso consistente de ferramentas de análise de dados e visualização. Explorando diversos aspectos do sistema educacional de forma ativa e engajada.

À medida que foram sendo trabalhados os conceitos e as ferramentas, o desejo de aprofundar o conhecimento se tornou evidente, com muitos docentes mostrando iniciativa em praticar e experimentar além do que era proposto. Como resultado, o ambiente de aprendizado se tornou colaborativo e interativo.

Essa nova oportunidade permitiu não apenas que aprendessem ou relembassem os conceitos tradicionais da Estatística Descritiva, mas também aplicassem esse conhecimento utilizando uma tecnologia moderna que ainda facilitou a construção de gráficos e a interpretação de dados de maneira mais intuitiva e visual. Além disso, essa iniciativa promoveu a inovação no ensino da estatística, tornando-o mais acessível e relevante.

Ao introduzir *Python*, uma das linguagens de programação mais utilizadas e poderosas para análise de dados, os alunos puderam explorar de forma prática como manipular, transformar e visualizar grandes volumes de dados, ou seja, a familiaridade com Python, também, lhes proporcionou uma atualização importante, dado que esta linguagem é amplamente adotada em diversas indústrias para análise de dados, inteligência artificial e machine learning.

Para os docentes que nunca tiveram contato com a programação, participar de encontros focados em bibliotecas como Pandas e Seaborn na Análise Exploratória de Dados pode ser uma experiência transformadora. O uso dessas bibliotecas desmistifica a programação, mostrando que é possível dominar habilidades que antes poderiam parecer distantes ou até mesmo complexas.



Esse tipo de formação não só proporciona uma compreensão técnica das ferramentas, mas também desperta a curiosidade e o interesse dos docentes por uma nova área. Muitos docentes, ao verem as possibilidades que a programação oferece, começam a enxergar o aprendizado de código como algo dentro de seu alcance, o que pode abrir portas para novas oportunidades pedagógicas e profissionais.

É importante estimular essa curiosidade nos docentes, mostrando que a programação não é exclusiva de áreas técnicas, mas uma competência com potencial para enriquecer o ensino em diversas disciplinas.

Apesar do tempo limitado, como alguns professores mencionaram, foi possível “abrir uma pequena fresta numa janela” e mostrar que a linguagem de programação pode, de fato, fazer parte da formação de professores. Mesmo com os desafios e a necessidade de ajustes, a experiência proporcionou uma visão clara do potencial da programação na formação.

Enfim, a possibilidade de integrar essas habilidades na formação docente mostrou-se viável e relevante, destacando a importância da inovação tecnológica. Embora, ainda haja muito a ser feito, essa introdução à programação na educação estatística plantou as sementes para uma futura integração mais robusta e eficaz dessas ferramentas no contexto educacional, fortalecendo o ensino e a aprendizagem de maneira significativa.

As pesquisas futuras sobre tecnologias na formação de professores representam uma área promissora para o desenvolvimento de práticas educacionais que estejam alinhadas com as demandas do século XXI. À medida em que o mundo digital continua a evoluir rapidamente, é crucial que os professores estejam preparados não apenas para utilizar essas tecnologias, mas também para integrá-las de forma eficaz em suas práticas pedagógicas.

## REFERÊNCIAS

AMARAL, Fernando. **Introdução a ciência de dados: mineração de dados e Big Data**. Rio de Janeiro: Alta Books, 2016.

BARBOSA, Geovane C. *et al.* Análise de Trajetórias de Professores que Ensinam Probabilidade e Estatística com Auxílio do Software Iramuteq. **JIEEM**, 13, 2020. 420-428.

BOSSI, Katia; SCHIMIGUEL, Juliano. Utilização das tecnologias digitais no ensino da estatística: uma visão baseada em análise de dissertações e teses. **REVISTA SOCIEDADE CIENTÍFICA**, 25 dezembro 2023.

BOSSI, Katia M. L. **Educação Estatística mediado por linguagem computacional na formação de professores**. Tese (Doutorado em Ensino de Ciências e Matemática) – Universidade Cruzeiro do Sul. São Paulo. 2024.

CAMPOS, Celso R. **A educação estatística: uma investigação acerca dos aspectos relevantes à didática da estatística em cursos de graduação**. Tese(doutorado) Universidade Estadual Paulista, Instituto de Geociências e Ciências Exatas. Rio Claro. 2007.

CAZORLA, Irene M.; KATAOKA, Veronica Y.; SILVA, Claudia B. D. Trajetoria e perspectiva da educação estatística no Brasil: um olhar a partir do GT12. In: LOPES, Celi E.; COUTINHO, Cileda D. Q. E. S.; ALMOULOU, Saddo A. **Estudos e reflexões em educação estatística**. Campinas: Mercado de Letras, 2010. p. 19 - 44.

GAISE. Guidelines for Assessment and Instruction in Statistics Education College Report. **GAISE College Report ASA Revision Committee**, 2016.

LIMA, Maria I. F. S. **Formação Continuada do Professor e Uso das Tecnologias Digitais**. Anais Educon 2020. São Cristóvão/SE: [s.n.]. 2020. p. 3-15.

LOPES, Celi. Educação Estatística no Curso de Licenciatura em Matemática. **Bolema**, Rio Claro, dezembro 2013. 901-915.

LOPES, Celi E. O ensino da estatística e da probabilidade na educação básica e a formação dos professores. **Cadernos Cedes**, FapUNIFESP (SciELO), 28, Abril 2008. 57-73.

MARCONI, Marina D. A.; LAKATOS, Eva M. **Fundamentos de Metodologia Científica**. 8. ed. São Paulo: Atlas, 2017.

MORAN, José M.; MASETTO, Marcos T.; BEHRENS, Marilda A. **Novas tecnologias e mediação pedagógica**. Campinas: Coleção Papirus Educação, 2000.

TOLEDO, Sezilia E. R. G. O. D.; LOPES, Celi. Violência em Números. **Revista da Sociedade Brasileira de Educação Matemática**, São Paulo, 2021. 1-19.