Investigando a Relação entre a Nota no Vestibular com o Desempenho em Introdução à Programação

Leonardo Silva
Instituto Federal de Pernambuco
Garanhuns, Pernambuco
leonardo.silva@garanhuns.ifpe.edu.br

Joanne Gabriela dos Santos Silva
Instituto Federal de Pernambuco
Garanhuns, Pernambuco
mukelectric@gmail.com

Milena Siqueira Santos
Instituto Federal de Pernambuco
Garanhuns, Pernambuco
milenaifpe@gmail.com

RESUMO

Há uma demanda em identificar estudantes que irão necessitar de suporte pedagógico em introdução à programação. Com isto, é possível definir ações pedagógicas que visam combater a evasão e retenção, reconhecidamente altas neste componente curricular. Uma das linhas de investigação nesta temática analisa como diferentes dados dos estudantes podem ajudar a explicar o seu desempenho em programação. Dentre elas, destaca-se o desempenho no vestibular, um dado de fácil acesso à instituição de ensino. Existem análises em nível internacional sobre o impacto dessa variável, no entanto, observa-se uma lacuna de dados do contexto educacional brasileiro, que o presente estudo visa mitigar. É de interesse ao escopo deste estudo investigar a relação entre a nota do vestibular para o ingresso em um curso técnico de informática integrado ao ensino médio com o desempenho em programação. Os dados acadêmicos de 292 estudantes matriculados em um curso técnico de computação foram analisados, e suas pontuações no vestibular foram correlacionadas com o desempenho em introdução à programação. Identificou-se uma correlação positiva moderada, o que sugere um potencial uso dessa variável em modelos preditivos. Um procedimento de meta-análise também foi realizado para sumarizar os resultados de outros estudos sobre essa temática, contribuindo assim para construção do corpo de conhecimento sobre o tema.

CCS CONCEPTS

• **Social and professional topics** → **Computing education**; **CS1**.

PALAVRAS-CHAVE

Aprendizado de programação, Predição de desempenho, Nota no vestibular.

1 INTRODUÇÃO

O aprendizado de programação é um processo complexo que em muitos casos resulta em elevados índices de reprovação e evasão [18]; [23]. Dentre as múltiplas estratégias para mitigar esse problema, uma delas busca identificar estudantes com maior probabilidade de necessitar de apoio pedagógico, antes que evadam ou reprovem. Essa linha de investigação visa a construção de modelos de predição, o que poderia subsidiar ações que visem suporte pedagógico preventivo para contribuir com o êxito educacional destes alunos, como atividades de reforço e monitoria, logo no início do ano letivo.

Múltiplas variáveis podem ser incluídas nos modelos preditivos, como gênero, faixa etária, conhecimento prévio em programação, entre outros [8]. A nota do vestibular é uma outra opção, que vem sendo estudada pela facilidade no acesso à esse tipo de informação [22]. Diversas pesquisas foram direcionadas para investigar como a nota do vestibular pode explicar problemas futuros de aprendizagem de alguns estudantes [3, 12, 19]. No entanto, observa-se a necessidade de explorar essa temática no contexto educacional brasileiro. Sabe-se que a educação é impactada por fatores regionais, sociais e culturais, que variam entre regiões [15], portanto, se mostra necessário ter dados representativos sobre estudantes de computação brasileiros.

Além o problema destacado anteriormente, não foi identificada uma sumarização do conhecimento já publicado na literatura sobre a temática investigada. Isto é necessário, pois ao agregar os resultados de múltiplos estudos acadêmicos é possível obter uma visão holística que fornece melhores indicativos para a tomada de decisão [5]. No contexto de investigação, ao sumarizar os dados de múltiplos estudos pode-se melhor avaliar se a nota do vestibular é de fato uma variável associada ao desempenho em introdução à programação.

Este estudo tem por objetivo preencher as lacunas mencionadas acima e prover evidências estatísticas sobre a relação entre o desempenho acadêmico em introdução à programação com a nota de ingresso na instituição de ensino. Assim, a investigação foi orientada por duas questões de pesquisa: *i) a nota em introdução à programação está correlacionada com a nota de ingresso do estudante no vestibular?;* e *ii) quais as evidências estatísticas sobre a influência da nota de vestibular no desempenho em introdução à programação?*.

Para responder as questões de pesquisa acima foram utilizados os dados acadêmicos de 292 estudantes matriculados no componente curricular de introdução à programação de um curso de computação de nível técnico integrado ao ensino médio. O ingresso à instituição é feito por meio de um processo seletivo (vestibular) e os dados, que são públicos, foram analisados em comparação ao desempenho dos alunos após um ano letivo.

Este trabalho está organizado no seguinte formato: na Seção 2 são apresentados estudos sobre o tema realizados em outros contextos educacionais. Na Seção 3 detalha-se o procedimento metodológico, seguido dos resultados na Seção 4, e da análise sobre os impactos no processo de ensino e aprendizagem, apresentados na Seção 5. Por fim, as limitações do estudo são apresentadas na Seção 6.

Fica permitido ao(s) autor(es) ou a terceiros a reprodução ou distribuição, em parte ou no todo, do material extraído dessa obra, de forma verbatim, adaptada ou remixada, bem como a criação ou produção a partir do conteúdo dessa obra, para fins não comerciais, desde que sejam atribuídos os devidos créditos à criação original, sob os termos da licença CC BY-NC 4.0.

EduComp'21, Abril 27–30, 2021, Jataí, Goiás, Brasil

© 2021 Copyright mantido pelo(s) autor(es). Direitos de publicação licenciados à Sociedade Brasileira de Computação (SBC).

2 REVISÃO DA LITERATURA

Pesquisadores e professores argumentam que muitos estudantes ingressam no curso de computação com limitações em conhecimentos correlatos (e.g., matemática, interpretação textual, entre outros) que podem impactar o seu desempenho acadêmico [16]; [10]. Avalia-se que a nota do vestibular pode refletir tais dificuldades e ser utilizada como preditor para o desempenho dos estudantes no componente curricular de programação.

Esta é uma temática com décadas de interesse, motivada por encontrar formas de prever o desempenho do estudante previamente a situações que possam levá-lo à evasão ou retenção [8]. Um dos primeiros estudos destinados à essa investigação foi realizado por Bauer et al. [1]. Os autores correlacionaram as notas de ingresso na Universidade de Michigan com o desempenho em atividades de programação e identificaram a existência de uma correlação moderada entre essas variáveis (0,49).

Katz et al. também analisaram como as notas dos estudantes em um exame padronizado pelo governo americano e utilizado para o ingresso em algumas universidades, denominado *Scholastic Assessment Test* (SAT), correlacionou-se com o desempenho em introdução à programação [11]. O resultado sugere uma fraca correlação de 0,25.

Mais recentemente, Lamber realizou uma estratégia similar e também correlacionou as notas no SAT com o desempenho dos estudantes em programação [13]. A correlação identificada (0,16) foi considerada muito fraca, corroborando os resultados encontrados em [11].

Gomes e Mendes analisaram o desempenho de estudantes portugueses que utilizaram diferentes linguagens de programação: C e Python [10]. Diferentes níveis de correlação foram identificados entre entre a nota de ingresso na instituição de ensino com o desempenho em programação, que variou em razão da linguagem de programação utilizada. Estudantes que utilizaram a linguagem Python apresentaram uma correlação de 0,46, enquanto que o grupo que utilizou a linguagem C apresentou correlação de 0,28.

Silva et al. utilizaram as notas em um exame admissional para uma universidade do México, e correlacionaram com o desempenho em introdução à programação, resultando em uma correlação moderada de 0,41 [24].

Stanko et al. utilizaram notas de um exame nacional de ingresso em universidades da Rússia e correlacionaram com as notas em diferentes disciplinas, dentre elas em introdução à programação [25]. A partir disto construíram um modelo para predizer o desempenho do estudante a partir desta nota de ingresso.

Ramos identificou uma correlação moderada de 0,45 entre As notas admissionais de um exame nacional das Filipinas com o desempenho em introdução à programação [19].

Pereira et al. avaliaram o impacto de diferente variáveis sócioeconômicas, culturais e demográficas no aprendizado de programação [17]. O objetivo também é auxiliar professores com dados que possam identificar previamente estudantes em necessidade de assistência. Verificou-se que ter filhos e o gênero foram as únicas variáveis correlacionadas ao desempenho. Os autores analisaram a origem da vaga de ingresso no vestibular (cotista ou não cotista), mas não utilizaram a nota em seus cálculos.

Observa-se que diversos estudos investigaram a correlação entre a nota de ingresso nas instituições de ensino com o desempenho em programação. Diferentes graus de correlação foram encontrados, o que pode ser explicado por múltiplos fatores, como sociais e culturais, ou mesmo associados à linguagem de programação utilizada. No entanto, observa-se uma carência de estudos com dados no contexto educacional brasileiro, objetivo este que o presente trabalho busca mitigar.

3 METODOLOGIA

A metodologia utilizada para conduzir este estudo consistiu das seguintes etapas: i) definição das questões de pesquisa; ii) coleta de dados; iii) cálculos estatísticos e v) análise dos resultados.

3.1 Definição das Questões de pesquisa

O presente estudo tem por objetivo central verificar a relação entre o desempenho em programação com a nota de ingresso na instituição de ensino. Portanto, as seguintes questões de pesquisa foram elaboradas: *i) a nota em introdução à programação está correlacionada com a nota de ingresso do estudante no vestibular?*; e *ii) quais as evidências estatísticas sobre a influência da nota de vestibular no desempenho em introdução à programação?*.

Assim, observa-se a formação de dois objetivos, no qual o primeiro é prover dados sobre estudantes brasileiros de computação, e o segundo visa sumarizar o corpo de conhecimento sobre a temática agregando os diferentes resultados por meio de um processo de meta-análise.

3.2 Descrição da Amostra e Contexto Pedagógico

A amostra deste estudo é composta por 292 estudantes matriculados em um curso técnico de computação integrado ao ensino médio do Instituto Federal de Pernambuco - Campus Garanhuns. Neste curso é ofertada qualificação profissional para que os egressos atuem no segmento de desenvolvimento de software. A formação ocorre concomitantemente ao período em que cursam o ensino médio, de forma integrada ao seu currículo. Esta modalidade de ensino profissionalizante possui duração de 4 anos, na qual os estudantes possuem componentes curriculares tradicionais do ensino médio, como matemática, português, física, dentre outros, mas também componentes de computação, como introdução à programação, análise orientada a objetos, programação para a Internet, testes de software, entre outros.

O componente curricular de introdução à programação está presente no primeiro ano de curso e dispõe de uma carga horária semanal de 2 horas, totalizando 80 horas ao término do ano letivo. O conteúdo curricular contempla: pensamento algorítmico, variáveis, tipos de dados, estruturas de controle de fluxo e repetição, vetores e funções, ministrados com a linguagem de programação Java. As aulas ocorreram em um laboratório de informática, no qual em alguns casos havia a disponibilidade de um computador por estudante e em outros era necessário que dois estudantes o compartilhassem. O professor deste componente curricular, em geral, utilizou uma metodologia de ensino expositiva, intercalada com atividades práticas de laboratório. Quatro avaliações foram realizadas ao longo do ano, sendo uma por bimestre e com direito a

recuperação. As atividades avaliativas foram realizadas com papel e caneta e os alunos deveriam construir algoritmos para problemas que eram passados. Nos anos analisados, os índices de reprovação no componente curricular foi de aproximadamente 50%.

Os dados coletados representam um período de análise de cinco anos de estudantes matriculados em turmas diferentes. Os estudantes apresentavam faixa etária entre 13 a 16 anos, e há uma distribuição equivalente entre os gêneros. Aproximadamente 60% dos estudantes recebem bolsas de estudo em razão de sua vulnerabilidade social e parte deles não possui computador em casa, sendo o campus a única oportunidade para prática da programação. Em geral, a grande maioria destes alunos (aproximadamente 90%) não tiveram contato com programação previamente ao curso.

3.3 Coleta de dados

O desempenho em programação foi mensurado por meio da média final dos estudantes no componente curricular de introdução à programação, compartilhados pelo setor de registro acadêmico do campus Garanhuns do Instituto Federal de Pernambuco, que autorizou seu uso para fins científicos. As notas dos estudantes no vestibular da instituição foram coletadas no site oficial da instituição e são abertas de domínio público. Essa nota é resultante de uma avaliação composta por 50 questões das áreas de matemática, língua portuguesa e conhecimentos gerais, além de uma redação.

3.4 Cálculos estatísticos

Para responder a questão de pesquisa i) realizou-se um cálculo de correlação estatística entre as médias finais no componente curricular de programação com as pontuações alcançadas no vestibular. Para sumarizar o corpo de evidências apontado na questão de pesquisa ii), realizou-se um procedimento estatístico de meta-análise que consiste na sumarização dos resultados de múltiplos estudos [4]. Os benefícios são diversos: a) aumenta-se o tamanho da amostra e consequentemente a representação mais adequada de uma determinada população e b) possibilita-se muitas vezes calcular um tamanho de efeito com maior poder estatístico [2].

4 RESULTADOS

As notas dos estudantes na disciplina de introdução à programação apresentaram média de 4,3 (de um total de 10 pontos), mediana de 4,6 e desvio padrão de 2,99. As pontuações no vestibular apresentaram média de 38,77 (de um total de 100 pontos), mediana de 38,97 e desvio padrão de 13,28. Um gráfico boxplot é apresentado na Figura 1.

4.1 Cálculo de correlação

Pela Figura 2, observa-se a existência de *outliers* e de uma relação não linear entre as variáveis. Portanto, a correlação de Pearson foi descartada [6] e optou-se pelo uso do cálculo não paramétrico de Spearman. O resultado indicou haver uma correlação positiva de 0,36 com um p-value <0,01. Considera-se que para um nível de significância de 5% o resultado é estatisticamente significante. Para um intervalo de confiança de 95% foi identificado um intervalo de confiança de 0,25 a 0,45.

Tabela 1: Dados utilizados no cálculo da meta-análise.

Trabalho	Tamanho da amostra	Correlação
Este estudo	292	0,36
[10]	138	0,28
[13]	341	0,16
[11]	65	0,25
[1]	68	0,49
[19]	69	0,45
[24]	1168	0,41

4.2 Cálculo de meta-análise

O procedimento de meta-análise adotado sumarizou os resultados dos estudos descritos na Tabela 1. Para realizar o cálculo utilizou-se o método de Hunter–Schmidt, pois ele se mostra adequado em razão do uso de dados correlacionais [9]. Os cálculos foram realizados com a ferramenta estatística R, utilizando o pacote *metaphor* e a função *rma*.

A estimativa de efeito sumarizada foi de **r = 0,3650** (p-value <0,001), com intervalo de confiança de [0,2623 a 0,4677].

5 ANÁLISE DOS RESULTADOS

5.1 a) O desempenho em introdução à programação está correlacionado com a nota de ingresso do estudante no vestibular?

Observa-se que os estudantes apresentaram um desempenho heterogêneo no componente curricular de introdução à programação (desvio padrão de 2,99). Observa-se na Figura 2 que uma quantidade significativa de estudantes obteve desempenho satisfatório em programação apresentando baixo resultado no vestibular e vice-versa, demonstrando que a relação entre variáveis não é totalmente linear.

A correlação de 0,36 sugere que a associação entre as variáveis analisadas é moderada [21]. O resultado encontrado corrobora outros achados na literatura acadêmica e se aproxima de outros estudos, como [1]; [10]; [24] e [19]. O que sugere que a relação entre a nota no vestibular com o desempenho varia entre fraca a moderada, a discussão acerca disto será aprofundada na Seção 5.2.

Em razão da variabilidade dos resultados entre estudos não é possível concluir se a nota do vestibular será uma variável preditora adequada em modelos de regressão. É possível que seja, mas em determinados contextos educacionais que precisam ser melhor investigados para identificá-los. Portanto, recomenda-se que investigações futuras explorem essa correlação para ampliar a amostragem, em especial com dados do contexto educacional brasileiro.

Ainda sim, ao comparar a correlação com outras variáveis que também já foram associadas com o desempenho em programação observa-se que o desempenho no vestibular é uma das mais altas. Por exemplo, gênero (0,07 [17] e 0,08 [28]), experiência prévia em programação (-0,2 [27] e 0,02 [17]), estratégias de aprendizagem

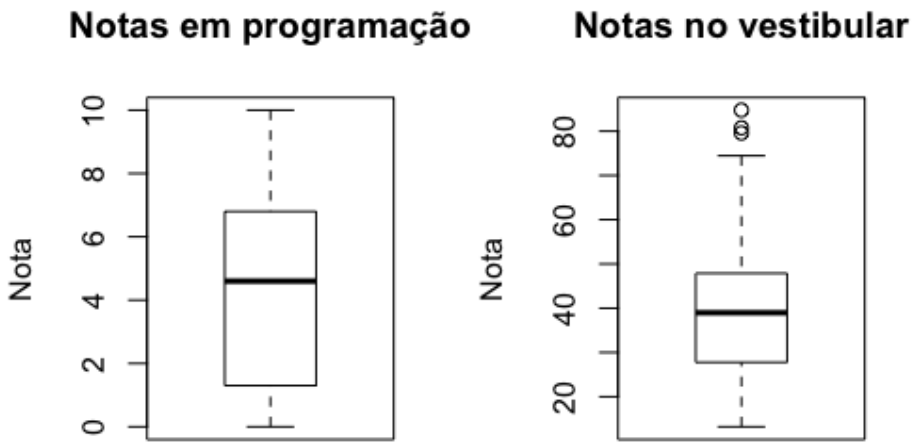

Figura 1: Gráfico boxplot com as notas em introdução à programação e no vestibular.

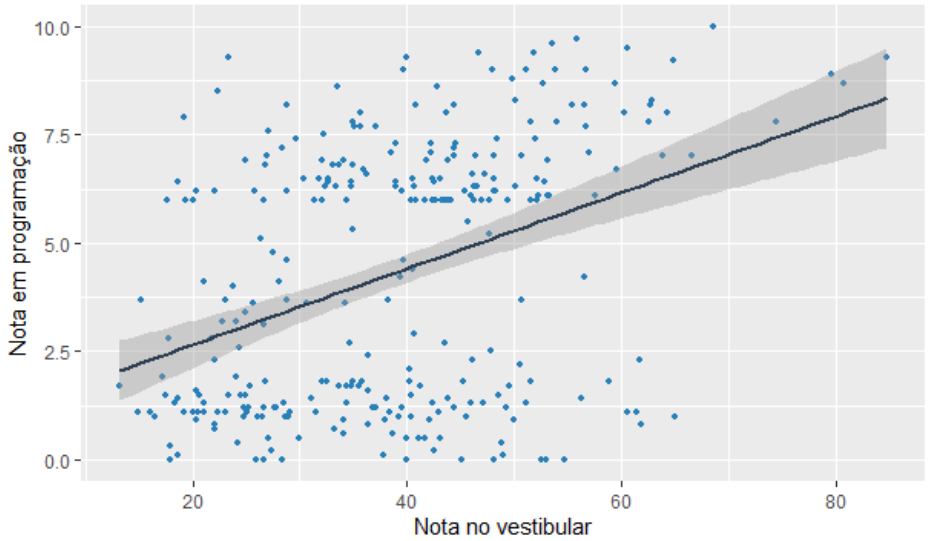

Figura 2: Correlação entre a nota do vestibular e o desempenho em programação.

(0,22 [27]) e matemática (0,2 [27], 0,37 [26] e 0,49 [10]). Esse resultado corrobora a necessidade de melhor explorar o uso da variável nota do vestibular em modelos preditivos.

Pondera-se que o processo de ensino e aprendizagem de programação é complexo e influenciado por múltiplas variáveis. Por exemplo, sabe-se que o aprendizado dos estudantes é impactado pelo seu conhecimento em componentes correlatos (e.g., matemática, resolução de problemas), suas habilidades cognitivas e metacognitivas, pelas estratégias pedagógicas dos professores, dentre outras [20]. Os dados apresentados devem ser interpretados como possíveis indicadores e não afirmações que implicam em uma relação causa e efeito. Como descrito acima, o maior uso do corpo de conhecimento produzido neste estudo é a indicação de que a nota do vestibular é uma variável a ser considerada em modelos de predição multivariados.

A considerar o exposto nos parágrafos anteriores, um olhar mais amplo sobre a temática de predição dos alunos com maior probabilidade de necessitar de suporte pedagógico, aponta para a necessidade de construir modelos de regressão que considerem múltiplas variáveis, buscando assim maximizar seu poder preditivo. Professores e instituições de ensino podem fazer uso de modelos como um suporte adicional à sua prática docente. A partir disso, podem-se desenvolver estratégias pedagógicas que aumentem a possibilidade de êxito dos estudantes.

Por fim, a análise deste resultado deve considerar o contexto deste estudo que envolve estudantes ingressos em um curso técnico na modalidade integrada ao ensino médio. Estudos futuros podem ampliar de forma sistemática a busca na literatura por outros estudos, como também coletar dados do ensino superior brasileiro.

Sugere-se que estudos futuros calculem a correlação por área da avaliação do vestibular, por exemplo, o desempenho em matemática, português, entre outros, diferente deste estudo que utilizou a nota geral. Como reportado por Leeper e Silver, a amostra analisada por eles obteve correlação estatisticamente significante apenas com os resultados na área de línguas [14].

5.2 b) Quais as evidências estatísticas sobre a influência da nota de vestibular no desempenho em introdução à programação?

O cálculo da meta-análise considerou todas as correlações analisadas e resultou em uma estimativa de efeito de 0,3650. Considera-se que em um contexto educacional, como este estudo, o tamanho de efeito encontrado é significativo [7]. Esse conhecimento pode ser utilizado por pesquisadores, professores e gestores educacionais, como um suporte para identificar estudantes que podem apresentar situação de vulnerabilidade acadêmica em introdução à programação.

Como descrito anteriormente, a variável investigada neste estudo apresenta potencial para ser incluída em modelos preditivos, como os de regressão. O uso desses modelos é um auxiliar para a instituição de ensino identificar os estudantes com maior probabilidade de necessitarem apoio pedagógico, antes que enfrentem maiores dificuldades durante o aprendizado de programação.

Pondera-se que os dados correlacionais não implicam em causalidade e um melhor entendimento sobre a educação em programação deve considerar múltiplas fontes de evidências, como os dados apresentados neste estudo, somados aos resultados de experimentos e estudos qualitativos.

6 LIMITAÇÕES DO ESTUDO

Uma limitação de estudos que utilizam dados correlacionais, como este, é que uma correlação não implica em causalidade. Assim, limita-se a capacidade de afirmação que a nota no vestibular apresenta um impacto no desempenho em programação. Ainda sim, os possuem a sua validade quanto ao objeto de análise. Sugere-se que o estudo em questão seja replicado por outras instituições de ensino para construir um corpo de conhecimento mais sólido acerca do impacto desta variável em estudantes de computação brasileiros.

A amostra utilizada também é limitada quanto à generalização dos resultados, tendo em vista o contexto educacional que envolveu estudantes do ensino médio técnico. Sugere-se que estudos futuros também avaliem os dados de estudantes de nível superior.

7 CONCLUSÃO

Este estudo analisou a relação entre as pontuações que estudantes obtiveram no vestibular de ingresso na instituição de ensino com o seu posterior desempenho em introdução à programação. Para alcançar esse objetivo foram realizados cálculos estatísticos de correlação e meta-análise, comparando os resultados à outros publicados na literatura.

Observou-se uma correlação positiva moderada, que aproxima-se dos resultados encontrados por outros estudos. A meta-análise realizada apresenta a sumarização de um conjunto de evidências e fortalece o argumento que a nota do vestibular pode ser um dos preditores para o desempenho em programação. Ressalvas são feitas em relação a outros fatores que também impactam o aprendizado.

Propõe-se em trabalhos futuros que construam modelos de regressão multivariados, incluindo a nota no vestibular como uma das suas variáveis. Também sugere-se que outros contextos educacionais, como dados do ensino superior brasileiro sejam utilizados.

REFERÊNCIAS

[1] Roger Bauer, William A Mehrens, and John F Vinsonhaler. 1968. Predicting performance in a computer programming course. *Educational and Psychological Measurement* 28, 4 (1968), 1159–1164.

[2] Lawrence D Cohn and Betsy J Becker. 2003. How meta-analysis increases statistical power. *Psychological methods* 8, 3 (2003), 243.

[3] James S Cole and Robert M Gonyea. 2010. Accuracy of self-reported SAT and ACT test scores: Implications for research. *Research in Higher Education* 51, 4 (2010), 305–319.

[4] Thomas D Cook, Harris Cooper, David S Cordray, Heidi Hartmann, Larry V Hedges, and Richard J Light. 1992. *Meta-analysis for explanation: A casebook*. Russell Sage Foundation.

[5] Philip Davies. 1999. What is evidence-based education? *British journal of educational studies* 47, 2 (1999), 108–121.

[6] Joost CF de Winter, Samuel D Gosling, and Jeff Potter. 2016. Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological methods* 21, 3 (2016), 273.

[7] Joseph A Durlak. 2009. How to select, calculate, and interpret effect sizes. *Journal of pediatric psychology* 34, 9 (2009), 917–928.

[8] AF ElGamal. 2013. An educational data mining model for predicting student performance in programming course. *International Journal of Computer Applications* 70, 17 (2013), 22–28.

[9] Andy P Field. 2005. Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychological methods* 10, 4 (2005), 444.

[10] Anabela Gomes and António Mendes. 2008. A study on student's characteristics and programming learning. In *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 2895–2904.

[11] Sandra Katz, John Aronis, David Allbritton, Christine Wilson, and Mary Lou Soffa. 2003. A study to identify predictors of achievement in an introductory computer science course. In *Proceedings of the 2003 SIGMIS conference on Computer personnel research: Freedom in Philadelphia–leveraging differences and diversity in the IT workforce*. ACM, 157–161.

[12] Jennifer L Kobrin, Brian F Patterson, Emily J Shaw, Krista D Mattern, and Sandra M Barbuti. 2008. Validity of the SAT® for Predicting First-Year College Grade Point Average. Research Report No. 2008-5. *College Board* (2008).

[13] Lynn Lambert. 2015. Factors that predict success in CS1. *Journal of Computing Sciences in Colleges* 31, 2 (2015), 165–171.

[14] RR Leeper and JL Silver. 1982. Predicting success in a first programming course. *ACM SIGCSE Bulletin* 14, 1 (1982), 147–150.

[15] Simon Marginson and Thi Kim Anh Dang. 2017. Vygotsky's sociocultural theory in the context of globalization. *Asia Pacific Journal of Education* 37, 1 (2017), 116–129.

[16] Rodrigo Pessoa Medeiros, Geber Lisboa Ramalho, and Taciana Pontual Falcão. 2018. A systematic literature review on teaching and learning introductory programming in higher education. *IEEE Transactions on Education* 99 (2018), 1–14.

[17] André Pereira, Leandro Carvalho, and Eduardo Souto. 2019. Analisando a influência de atributos demográficos no desempenho de estudantes em uma disciplina de introdução à programação. In *Anais do XXVII Workshop sobre Educação em Computação*. SBC, 360–369.

[18] Yizhou Qian and James Lehman. 2017. Students' misconceptions and other difficulties in introductory programming: A literature review. *ACM Transactions on Computing Education (TOCE)* 18, 1 (2017), 1–24.

[19] Maria Cristina M Ramos. 2018. Correlation between Entrance Exam Scores (Stanine) and Academic Performance. In *Proceedings of the 2018 2nd International Conference on Algorithms, Computing and Systems*. ACM, 110–114.

[20] Anthony V Robins. 2019. Novice programmers and introductory programming. *The Cambridge Handbook of Computing Education Research, Cambridge Handbooks*

in Psychology (2019), 327–376.

[21] Patrick Schober, Christa Boer, and Lothar A Schwarte. 2018. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia* 126, 5 (2018), 1763–1768.

[22] William E Sedlacek and Javaune Adams-Gaston. 1992. Predicting the academic success of student-athletes using SAT and noncognitive variables. *Journal of Counseling & Development* (1992).

[23] Leonardo S Silva. 2018. Análise do aprendizado em programação de estudantes do ensino técnico integrado do Instituto Federal de Pernambuco. In *Anais do V Encontro Nacional de Computação dos Institutos Federais*. SBC.

[24] Gabriela Silva-Maceda, P David Arjona-Villicaña, and F Edgar Castillo-Barrera. 2016. More time or better tools? A large-scale retrospective comparison of pedagogical approaches to teach programming. *IEEE Transactions on Education* 59, 4 (2016), 274–281.

[25] T Stanko, O Zhirosh, D Johnston, and S Gartsev. 2017. On possibility of prediction of academic performance and potential improvements of admission campaign at IT university. In *2017 IEEE Global Engineering Education Conference (EDUCON)*. IEEE, 862–866.

[26] Michael V Stein. 2002. Mathematical preparation as a basis for success in CS-II. *Journal of Computing Sciences in Colleges* 17, 4 (2002), 28–38.

[27] Christopher Watson, Frederick WB Li, and Jamie L Godwin. 2014. No tests required: comparing traditional and dynamic predictors of programming success. In *Proceedings of the 45th ACM technical symposium on Computer science education*. 469–474.

[28] Laurie Honour Werth. 1986. Predicting student performance in a beginning computer science class. *ACM SIGCSE Bulletin* 18, 1 (1986), 138–143.