

Impacto da COVID-19 na Formação de Estudantes de Pós-Graduação no Brasil: Uma Análise a Partir de Dados da Plataforma Lattes do CNPq

Pedro L. C. Mucci¹, Alberto H. F. Laender¹

¹Departamento de Ciência da Computação
Universidade Federal de Minas Gerais
Belo Horizonte – MG – Brasil

pedromucci@dcc.ufmg.br, laender@dcc.ufmg.br

Abstract. *The COVID-19 pandemic had a global social and economic impact, including in the development of scientific research, due to social distancing measures and severe financial resource reduction. Using data from the Lattes Platform, this work compares the academic production of Brazilian postgraduates before, during and after the pandemic crises, analyzing the expressive decrease occurred in the completion of master's and doctoral degrees in Brazil. The results highlight the extent of the impact on academic trajectories, highlighting the differential impact across specific knowledge areas.*

Resumo. *A pandemia da COVID-19 teve um impacto social e econômico global, inclusive no desenvolvimento de pesquisas científicas, devido às medidas de distanciamento social e à redução de recursos financeiros. Usando dados da Plataforma Lattes, este trabalho compara a produção acadêmica de pós-graduandos brasileiros antes, durante e após a pandemia, analisando a diminuição expressiva nas conclusões de mestrados e doutorados. Os resultados ressaltam a extensão do impacto nas trajetórias acadêmicas, inclusive indicando o diferencial de impacto em áreas específicas do conhecimento.*

1. Introdução

A pandemia da COVID-19, causada pelo vírus SARS-CoV-2, emergiu como um evento de significância global que marcou a primeira metade do século XXI. Iniciada em Wuhan, China, no final de 2019 [Sohrabi et al. 2020], a disseminação rápida e generalizada do vírus resultou em uma crise de saúde pública sem precedentes. Além das graves consequências para a saúde humana, a pandemia desencadeou uma série de impactos sociais, econômicos e políticos.

A ciência foi uma das atividades que mais sofreram com as consequências da pandemia, principalmente no que diz respeito à interrupção das pesquisas em andamento, especialmente aquelas referentes a campos não diretamente relacionados à COVID-19. Muitos laboratórios tiveram que redirecionar recursos e pessoal para o estudo do vírus e seu combate [Emanuel et al. 2020], o que levou a atrasos e, em alguns casos, à interrupção de projetos de pesquisa cruciais em outras áreas médicas.

Além disso, as restrições de viagens e o fechamento temporário de instituições acadêmicas e laboratórios dificultaram a colaboração entre cientistas e a coleta de dados

em estudos de campo, afetando a continuidade de pesquisas em várias disciplinas. O financiamento para pesquisa também enfrentou desafios, já que os governos passaram a direcionar recursos substanciais para combate à pandemia, diminuindo o repasse de recursos para outras áreas. A revisão e publicação de artigos científicos também foram afetadas, devido à necessidade de acelerar a disseminação de informações relativas à COVID-19, muitas vezes levando a processos de revisão menos rigorosos e à disseminação de resultados ainda preliminares ou inadequadamente avaliados. Isso gerou preocupações sobre a qualidade e a confiabilidade das informações científicas em circulação.

No Brasil, para enfrentar os efeitos negativos da pandemia sobre a produção científica, foram implementadas diversas iniciativas. O governo federal, agências de financiamento e instituições de pesquisa mobilizaram esforços para apoiar projetos de pesquisa relacionados à COVID-19, provendo financiamento adicional e infraestrutura específica para a realização de estudos relacionados ao vírus e à sua disseminação. Além disso, foram adotadas medidas para flexibilizar prazos de relatórios e prestações de contas de projetos de pesquisa, a fim de acomodar as interrupções causadas pela pandemia. Universidades e instituições de pesquisa promoveram a adaptação ao trabalho remoto e investiram em recursos tecnológicos para viabilizar a continuidade das atividades acadêmicas e de pesquisa, mas os impactos na produtividade mantiveram-se.

Diante desse cenário, o presente artigo visa analisar o impacto da pandemia da COVID-19 na pesquisa científica brasileira em diferentes áreas do conhecimento, tendo como indicadores dados sobre a conclusão de cursos de pós-graduação *stricto sensu* (mestrado e doutorado). A partir dos dados analisados, é possível compreender como as diferentes áreas do conhecimento foram afetadas.

O restante deste artigo está estruturado da seguinte forma. A Seção 2 aborda trabalhos relacionados. A seguir, a Seção 3 descreve a metodologia adotada, enquanto a Seção 4 apresenta os resultados obtidos. Finalmente, a Seção 5 apresenta as conclusões finais e provê uma breve análise das perspectivas futuras deste trabalho.

2. Trabalhos Relacionados

Os efeitos da pandemia da COVID-19 na produção acadêmica e científica têm sido amplamente estudados por diferentes autores ao longo dos últimos anos, com base em diferentes recortes de análise. Nesta seção apresentamos uma breve visão dessa temática, de modo a caracterizar a sua relevância.

Um dos fatores analisados na literatura, por exemplo, é o impacto diferencial a partir do gênero de pesquisadores. Os dados sugerem que a pandemia afetou a produtividade científica das mulheres, especialmente aquelas com crianças, devido ao fechamento de escolas e creches. Isso resultou em menos submissões de artigos científicos por mulheres e uma lacuna de gênero significativa [Rangel et al. 2021] nas Plataformas *preprint*.

A partir de uma outra análise, um trabalho mais recente [Pimentel 2022] apresenta um estudo bibliométrico em que são observados indicadores de quantidade e qualidade de produção científica no período pré-pandêmico (2018-2019), durante a pandemia (2020-2021) e no período posterior ao estado de emergência determinado pelo governo brasileiro (2022). Os dados dessa pesquisa apontaram para um declínio no número de publicações, mas sem afetar a sua qualidade. No entanto, o autor optou por restringir o estudo à

produção científica de docentes orientadores de um único programa de pós-graduação, de modo que os resultados podem não representar a totalidade do cenário brasileiro.

No cenário internacional, alguns estudos também apontam o impacto da COVID-19. Um trabalho, que explora esse impacto nas Ciências da Vida a partir de dados da plataforma PubMed, aponta um aumento de 6,5 vezes nas publicações sobre a COVID-19 e uma redução de 10 a 12% nas publicações em áreas não relacionadas, aumentando para 16 a 19% quando ponderadas pelo fator de impacto [Sohrabi et al. 2021].

Um outro estudo [Riccaboni and Verginer 2022] examina os variados efeitos da COVID-19 na pesquisa básica e em laboratórios com animais, bem como na pesquisa clínica, nos próprios pesquisadores, na captação de recursos, no *networking*, na trajetória futura dos pesquisadores e nas publicações acadêmicas. Além disso, o estudo destaca como a crise promoveu altos níveis de colaboração internacional e a adoção de plataformas de telecomunicação remota, ressaltando como tais adaptações influenciaram os padrões de trabalho e ensino atuais.

Finalmente, em um estudo sobre a produção científica em periódicos brasileiros de acesso aberto, foi realizada uma análise temporal de dados extraídos da Plataforma Lattes do CNPq, confirmando a sua viabilidade como fonte de dados [Dias et al. 2020]. Ademais, o pré-processamento dos dados como abordado no presente trabalho foi descrito em outros estudos anteriores [Cota et al. 2021, Dores 2017].

3. Metodologia

De acordo com a metodologia adotada, o foco da etapa inicial envolveu a obtenção dos dados provenientes da Plataforma Lattes, relativos aos currículos dos mestres e doutores ali registrados. Esses currículos, referentes a um total de 1.041.197 pesquisadores, foram extraídos em formato XML, compreendendo um volume de dados equivalente a 21,9 gigabytes (GB). Os dados foram coletados no início de 2024, o que pode limitar as informações disponíveis, principalmente sobre os pesquisadores que concluíram suas pesquisas no final de 2023 e não atualizaram seus dados na Plataforma.

O processamento desses dados foi conduzido mediante a implementação de um *parser* desenvolvido em Python, utilizando a biblioteca *xml.etree.ElementTree*. Para otimizar a manipulação eficaz dos arquivos compactados disponibilizados pela Plataforma, foi empregada a biblioteca *ZipFile*. Após o processamento pelo *parser*, os dados relevantes foram compiladas em um arquivo CSV, proporcionando uma estrutura mais acessível e condensada para análises subsequentes. As análises propriamente ditas foram conduzidas utilizando a linguagem de programação Python, destacando-se o emprego das bibliotecas *Pandas* e *Matplotlib* como principais ferramentas. Essa abordagem permitiu a geração de gráficos e a extração das informações desejadas de maneira eficiente e sistemática.

Os dados coletados para a análise foram aqueles referentes a pesquisadores com título de mestre e doutor obtidos entre os anos 2013 e 2023. Esse período foi escolhido a fim de observar o crescimento contínuo e não impactado por grandes fatores, como a pandemia, e que não fosse um período de tempo grande que dificultasse a análise. Além disso, as conclusões que não tinham uma grande área do conhecimento associada, ou que tinham grande área de conhecimento definida como “Outros”, não foram utilizadas para a análise, de modo que as grandes áreas consideradas foram as mesmas adotadas pelo

CNPq: Ciências Agrárias, Ciências Biológicas, Ciências da Saúde, Ciências Exatas e da Terra, Ciências Humanas, Ciências Sociais Aplicadas, Engenharias e Linguística, Letras e Artes.

Após observar os dados processados, o próximo passo foi a escolha da métrica usada para detalhar as conclusões tiradas dos gráficos. A métrica escolhida foi a taxa de crescimento ano a ano do número de concluintes de mestrado e doutorado, definida por: $\text{Crescimento (em \%)} = \frac{\text{Quantidade de conclusões no ano N} - \text{Quantidade de conclusões no ano N-1}}{\text{Quantidade de conclusões no ano N-1}} \times 100$.

4. Resultados

A seguir, são apresentados os quatro gráficos gerados a partir dos dados obtidos, sendo que os dois primeiros possibilitam observar os números brutos de conclusões de mestrado e doutorado, e os dois seguintes apresentam as respectivas taxas de crescimento. Os gráficos estão divididos em duas imagens com duas sub imagens cada que destacam características específicas. Por fim, é apresentada uma tabela com os resultados de uma análise mais profunda dos dados.

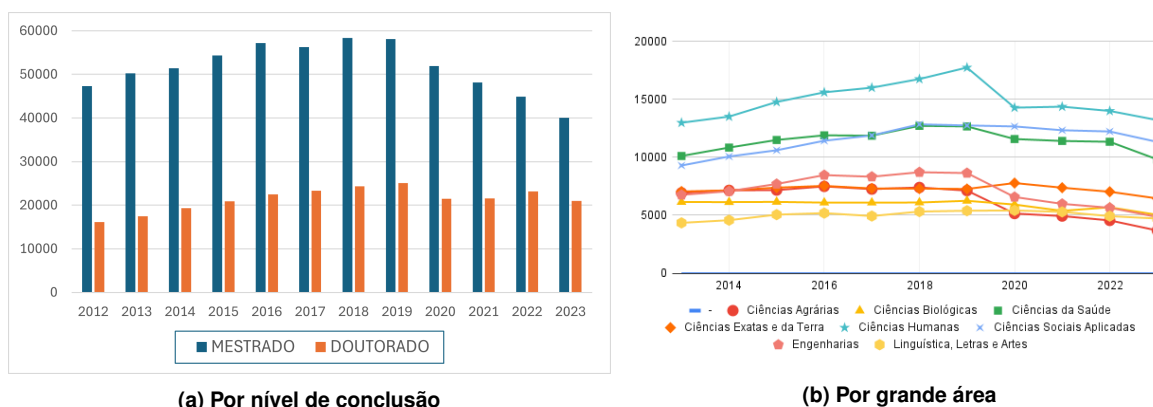


Figura 1. Conclusões de pós-graduação por ano (2013-2023)

Na Figura 1a, observa-se que a curva de crescimento do número de conclusões segue o comportamento esperado. O gráfico tem 2012 como ano inicial, a partir do qual ocorre um crescimento constante até chegar ao ápice de conclusões em 2019. A partir do ano pandêmico (2020), observa-se que o número de conclusões diminui consideravelmente, sendo essa queda constante até o ano de 2023. Ainda na Figura 1a, podem ser observadas as proporções de cada um dos dois tipos de formação ao longo dos anos: mestrado (azul) e doutorado (vermelho). Nesse gráfico, nota-se que entre os anos de 2013 e 2023 as conclusões de mestrado representam, em média, 70,71% das conclusões totais e os doutorados, 29,29%.

Já na Figura 1b, os mesmos dados são apresentados sob a perspectiva das grandes áreas do conhecimento. Nota-se, assim, as áreas que mais formaram pós-graduandos (Ciências Humanas, Ciências Sociais Aplicadas e Ciências da Saúde) e as que tiveram o menor número de conclusões (Linguística, Letras e Artes e Ciências Biológicas no período 2013-2018, e Ciências Agrárias e Linguística, Letras e Artes no período 2019-2023).

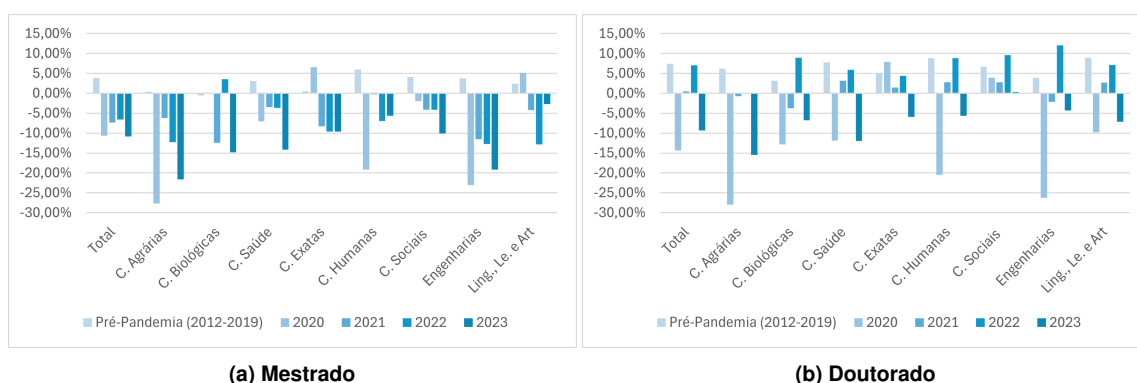


Figura 2. Taxa de crescimento das conclusões por grande área (2013-2023)

Ao se comparar os gráficos das Figuras 2a e 2b, é possível observar que em 2020 o doutorado foi mais impactado pela pandemia (-14,35%) do que o mestrado (-10,35%), mas em 2021 e 2022 voltou a se reerguer (0,52% e 7,09%, respectivamente), enquanto o mestrado sofreu novas quedas (-7,32% e -6,62%). Nesse contexto, as Ciências Agrárias foi a grande área mais afetada, com uma redução de 27,68% nas conclusões de mestrado e de 27,92% no doutorado. Notavelmente, essa grande área apresentou a menor diferença entre os impactos no mestrado e no doutorado, apenas 0,24%. Em contraste, nas Ciências Biológicas e nas Letras observa-se maior diferença de impacto entre os dois níveis (13,02% e 14,94%, respectivamente), sendo os doutorados mais afetados do que os mestrados.

Em segundo lugar, as Engenharias sofreram uma redução de 23,91% nas conclusões. As Ciências Exatas e da Terra tiveram um desempenho positivo em 2020, com crescimento de 7,02%, mas não sustentaram esse crescimento, apresentando uma queda de 5,16% em 2021. As Ciências Humanas experimentaram uma queda brusca de 19,49% em 2020, mas essa tendência de declínio não se manteve nos anos seguintes, com um ligeiro crescimento de 0,59% em 2021. A Linguística permaneceu relativamente estável, com um pequeno crescimento de 0,5% em 2020 e uma queda de 2,33% em 2021. Em 2020, as Ciências Biológicas destacaram-se com um comportamento atípico: enquanto as conclusões de doutorado diminuiram 12,86%, as de mestrado mantiveram-se estáveis, com leve crescimento de 0,16%. No entanto, em 2021, a tendência se inverteu, com queda de 12,46% no mestrado e de 3,71% no doutorado.

A Tabela 1 apresenta uma visão geral dos dados a partir da média da progressão da taxa de crescimento ao longo dos anos pré-pandêmicos. A média de crescimento das conclusões de mestrado naquele período foi de 3,82%, enquanto que no caso dos doutorados essa média foi de 7,43%. A média geral de crescimento das conclusões (mestrado e doutorado) foi de 4,52%. No período pandêmico, o percentual de crescimento não apenas deixou de seguir a tendência positiva observada anteriormente, como teve uma queda brusca próxima de 12,00%. Em 2021, as conclusões sofreram mais uma queda substancial, dessa vez de 5,03% em relação ao ano anterior. Finalmente, os anos de 2022 e 2023 apresentam quedas de 2,38% e 10,31%, respectivamente. No entanto, para uma análise mais precisa dessa redução, são necessários dados mais atualizados provenientes da Plataforma Lattes.

Tabela 1. Conclusões de pós-graduação por ano - Taxa de Crescimento

Período	Mestrado (%)	Doutorado (%)	Total (%)
2013-2019	3,82	7,43	4,52
2020	-10,6	-14,35	-11,73
2021	-7,32	0,52	-5,03
2022	-6,62	7,09	-2,38
2023	-10,8	-9,35	-10,31

Por fim, uma página em React ([link](#)) foi gerada para otimizar a visualização dos dados, proporcionando maior qualidade estética e interatividade. Na página gerada, é possível selecionar um período entre 1946 e 2023 para observar as conclusões de mestrado e/ou doutorado, filtrar por grande área, e escolher o tipo de gráfico de saída (barras ou barras empilhadas).

5. Conclusões

Diante do exposto, torna-se evidente que a pandemia da COVID-19 teve impactos significativos na pesquisa científica brasileira, como refletido nos dados analisados sobre a conclusão de cursos de pós-graduação. A interrupção de projetos, o redirecionamento de recursos, as restrições de viagens e as mudanças nos processos de publicação e revisão científica contribuíram para uma dinâmica desafiadora no cenário acadêmico.

A análise temporal revelou uma queda expressiva no número de conclusões de mestrado e doutorado durante o ápice da pandemia. Em 2020, os efeitos negativos foram mais acentuados no doutorado, enquanto em 2021 foram mais expressivos no mestrado. Hipotetiza-se que o impacto nas conclusões de mestrado, em 2021, reflete principalmente a paralisação de 2020, o que impediu o início de novos trabalhos que se encerrassem no ano seguinte. Além disso, outra hipótese é que os trabalhos de mestrado por serem mais curtos que os de doutorado, tiveram maior possibilidade de aceleração em seu desenvolvimento para que fossem encerrados ainda em 2020, enquanto os doutorandos tiveram maior flexibilização de prazos.

As análises também mostraram uma adaptação desigual das diferentes áreas do conhecimento às adversidades impostas. Não coincidentemente, identificou-se maior impacto em grandes áreas do conhecimento que dependem fortemente de trabalhos de campo e/ou atividades presenciais, como Ciências Agrárias, Engenharias e Ciências Humanas. Paralelamente, houve um menor impacto em áreas que lidam bem com o trabalho a distância, como Ciências Exatas e da Terra e Linguística. A grande área de Ciências Biológicas apresentou comportamentos únicos, potencialmente por seu campo de trabalho relacionado ao vírus da COVID-19 e a necessidade de estudos sobre o tema. Os dados sugerem, ainda, que as medidas implementadas para diminuir esses impactos, como flexibilização de prazos e apoio financeiro, embora importantes, não foram suficientes para evitar a diminuição na produção científica.

Em relação às perspectivas futuras, uma das limitações deste trabalho é depender da atualização manual por parte dos pós-graduandos de seus perfis na Plataforma Lattes, uma vez que quanto mais frequentes forem as atualizações mais fidedignas serão as análises. Além disso, a análise por grande área pode ser superficial, uma vez que existem impactos diferentes dentro de cada uma delas.

Agradecimentos

Este trabalho foi parcialmente financiado pelo CNPq (Processo 308528/2019-0) e pela FAPEMIG (Processo APQ-02302-17).

Referências

- Cota, J. M. M. C., Laender, A. H. F., and Prates, R. O. (2021). Science tree: a platform for exploring the Brazilian academic genealogy. *Journal of the Brazilian Computer Society*, 27(1).
- Dias, P. M., Dias, T. M. R., and Moita, G. F. (2020). Evolução da Produção Científica em Periódicos de Acesso Aberto no Brasil: Uma Análise Temporal. *Ciência da Informação*, 48(3).
- Dores, W. J. (2017). Um Estudo Sobre a Genealogia Acadêmica Brasileira. Master's thesis, Universidade Federal de Minas Gerais, Belo Horizonte, MG.
- Emanuel, E. J., Persad, G., Upshur, R., Thome, B., Parker, M., Glickman, A., Zhang, C., Boyle, C., Smith, M., and Phillips, J. P. (2020). Fair allocation of scarce medical resources in the time of covid-19. *New England Journal of Medicine*, 382(21):2049–2055.
- Pimentel, D. E. (2022). Impacto da Pandemia da COVID-19 na Produção Científica de Pesquisadores de um Programa de Pós-graduação. Trabalho de Conclusão de Curso, Universidade Federal de Mato Grosso do Sul, Campo Grande, MS.
- Rangel, É. B., Roza, B. d. A., Schirmer, J., Castro, M. d., and Boin, I. (2021). Impacto da COVID-19 na produção científica: um alerta para a disparidade de gêneros. *Brazilian Journal of Transplantation*, 24(3):59–61.
- Riccaboni, M. and Verginer, L. (2022). The impact of the COVID-19 pandemic on scientific research in the life sciences. *PloS one*, 17(2):e0263001.
- Sohrabi, C., Alsafi, Z., O'Neill, N., Khan, M., Kerwan, A., Al-Jabir, A., Iosifidis, C., and Agha, R. (2020). World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International Journal of Surgery*, 76:71–76.
- Sohrabi, C., Mathew, G., Franchi, T., Kerwan, A., Griffin, M., Del Mundo, J. S. C., Ali, S. A., Agha, M., and Agha, R. (2021). Impact of the coronavirus (COVID-19) pandemic on scientific research and implications for clinical academic training – a review. *International Journal of Surgery*, 86:57–63.