



23º Congresso Nacional de Iniciação Científica

TÍTULO: UMA PERSPECTIVA DOS INSTITUTOS FEDERAIS EM MINAS GERAIS SOB A ANÁLISE HISTÓRICA DE BOLSAS OFERTADAS PELO CNPQ PARA A ÁREA DA COMPUTAÇÃO ENTRE 2015 A 2022

CATEGORIA: CONCLUÍDO

ÁREA: CIÊNCIAS EXATAS E DA TERRA

SUBÁREA: Computação e Informática

INSTITUIÇÃO: INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE MINAS GERAIS - IFMG

AUTOR(ES): LUCAS LAET LIRA

ORIENTADOR(ES): RENATO MIRANDA FILHO, CARLOS ALEXANDRE SILVA

COLABORADOR(ES): PEDROGABRIELCRUZ00@GMAIL.COM, GIOVANNYMATHEUS28@GMAIL.COM

1. RESUMO

O Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) é a principal agência de fomento à pesquisa do país e as áreas de Engenharia e Computação compõem os principais eixos formadores e geradores de produtos no campo de tecnologia. Tendo em vista a importância do fomento desta agência, sobretudo, na área de Computação, este artigo busca identificar a relevância dos Institutos Federais de Minas Gerais no recebimento de recursos no âmbito científico nacional. Além disso, foi realizada uma comparação com universidades federais mineiras, por meio de uma análise de dados referente à captação de bolsas de pesquisas em um cenário regido por diferentes governanças políticas entre o período de 2015 a 2022. Nota-se uma discrepância entre os fomentos por gestões governamentais e uma mudança na disponibilidade de bolsas em suas categorias, enquanto no âmbito específico das bolsas destinadas aos Institutos Federais, é verificada uma presença relevante na captação, onde estes tiveram destaque nesse aspecto, em comparação a universidades tradicionais do estado de Minas Gerais.

2. INTRODUÇÃO

A Educação Profissional no Brasil tem origem no século XIX, com a criação do Colégio de Fábricas por D. João VI e ao longo dos anos vem passando por constante transformação econômica, política e social (Macedo, 2017). Em continuidade ao processo de expansão da rede federal de educação profissional, os Institutos Federais de Educação Profissional e Tecnológica (IFs) são criados a partir da lei nº 11.892, de 29 de dezembro de 2008¹, onde suas unidades vêm ganhando relevância e participação a nível nacional com relação à pesquisa e abrangência de alunos².

No estado de Minas Gerais existem cinco Institutos Federais, que juntos, representam uma estrutura de 56 campi, suas reitorias e polos de inovação. Tendo a história de alguns de seus atuais campi iniciada nos primórdios da instauração de escolas técnicas, como é o caso do campus Inconfidentes, fundado em 1918 como uma das Escolas Agrícolas e, que hoje faz parte do IF do Sul de Minas.

Os IFs tem sido destaque ao longo dos anos, onde mesmo se tratando de instituições recentes, é possível citar alguns exemplos de destaques como a classificação do Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) no ranking do Centro de Classificações Universitárias Mundiais CWUR² de 2022/2023, cuja metodologia considera qualidade da educação, emprego de ex-alunos, qualidade do corpo docente e desempenho da pesquisa. Em 2019, em relação ao ensino médio, os Institutos Federais ficaram à frente de escolas públicas

1. LEI Nº 11.892, DE 29 DE DEZEMBRO DE 2008. https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/11892.htm

2. Center for World University Rankings. <https://cwur.org/2022-23.php>

e particulares no Programa Internacional de Avaliação de Estudantes (Pisa 2018³), realizado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), onde é aplicada uma avaliação a jovens estudantes na faixa etária de 15 a 16 anos, que contempla as áreas de matemática, ciências e leitura, com os IFs superando as escolas concorrentes nas 3 áreas do conhecimento avaliadas.

A boa qualidade das instituições federais, sobretudo em relação à pesquisa, também deriva dos subsídios oriundos de agências de fomento como o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq⁴). O CNPq, fundado em 1951, tem como uma das principais atribuições fomentar a pesquisa científica, tecnológica e de inovação do país. É primordial a relação do CNPq com a pesquisa nacional, sendo esta agência detentora do maior banco de currículos acadêmicos da América Latina, e que ao longo de sua existência ofereceu milhares de bolsas de diversas modalidades (Domingos, M., 2011).

Este trabalho contribui com a linha de estudo sobre o fomento da pesquisa científica através do CNPq, analisando historicamente o contexto de MG e os IFs, de modo a agregar na discussão assim como levantado por Neiva et al., 2022 – onde apesar da especificidade de modalidade das bolsas, o tema do investimento na produção científica é o cerne das análises. Ademais, no âmbito dos Institutos Federais, o contexto de Mesquita et al., 2022, agregou à visão do potencial dessas instituições, onde o cenário de Roraima instiga outras linhas de pesquisa no Brasil, não só ligadas a iniciação científica, mas também outras modalidades relacionadas a teses, entre outros.

A organização do conteúdo se dá pela seguinte estrutura: Na seção **3** são descritas as análises aplicadas e propósitos do artigo. Na seção **4** é dissertado sobre os tratamentos dos dados e preparo destes para cada finalidade estudada. A seção **5** apresenta o desenvolvimento com aplicação dos critérios de pesquisa definidos e visualização dos resultados. Na seção **6** ocorrem as constatações das descobertas da partir do Desenvolvimento. E por fim, a seção **7** apresenta as conclusões, interpretações e visão de trabalhos futuros para o tema.

3. OBJETIVOS

Este trabalho tem como objetivo geral promover uma análise do conjunto de dados sobre bolsas contempladas pelo CNPq, para os cursos que podem ser relacionados ao eixo de tecnologia da informação, no estado de Minas Gerais e a respeito dos seus Institutos Federais. Fora analisado o período entre os anos de 2015 a 2022.

Como objetivos específicos foram analisadas as relações entre o governo vigente no ano de referência com o seu investimento em bolsas através do CNPq em MG; entre a distribuição de bolsas nos IFs e demais instituições contempladas em Minas Gerais, bem como comparada a captação de bolsas pelos IFs com Universidades

3. PISA 2018. <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/pisa/resultados>

4. <https://www.gov.br/CNPq/pt-br>

Federais de MG averiguando a correlação entre número de bolsas e a quantidade de alunos das instituições.

Em dessarte, este trabalho também tem como finalidade avaliar as gestões governamentais no intervalo de tempo analisado, no que tange ao fomento em pesquisas. Durante o período analisado, o Brasil passou por três diferentes gestões governamentais: o governo Dilma, que decorreu de 2015 até o primeiro semestre de 2016, o governo Temer, decorrido entre o segundo semestre de 2016 até o final de 2018, e por último, o governo Bolsonaro, ocorrendo de 2019 a 2022.

4. METODOLOGIA

Neste trabalho utilizou-se o conjunto de dados histórico, disponibilizado pelo CNPq⁵, no período compreendido entre 2015 e 2022. Assim, formando uma base única, com 35 atributos e aproximadamente 1 milhão e 400 mil registros. Esses registros descrevem todas as bolsas distribuídas a nível nacional, em todas as áreas de conhecimento no intervalo dos 8 anos selecionados. Após isso, realizou-se um pré-processamento do conjunto de dados, em que foram removidos registros com dados faltantes e/ou cujo escopo extrapolava o objeto de estudo deste trabalho.

Visando identificar grupos de dados com características similares na base utilizada, para a análise específica do fomento de bolsas em MG, foi selecionado o algoritmo *K-means* (Hartigan, 1975). O mesmo aplica um método comumente utilizado nas áreas de mineração de dados e aprendizado de máquina, onde seu conceito é dividir os dados em K (regulável) grupos que tenham distâncias próximas entre seus valores, através de uma medida de similaridade entre amostras, que é obtida por meio de uma função de distância euclidiana – ou seja, a base de dados é segregada em K partições, onde estas serão definidas por conjuntos de dados parecidos.

5. DESENVOLVIMENTO

Com o intuito de trazer a perspectiva de dentro de Minas Gerais e no escopo das bolsas na área da computação, foram então selecionados os registros que possuem “Sigla UF Origem” igual a “MG” e “Área” igual a “Ciência da Computação”, “Tecnologias Digitais” ou “Tecnologia da Informação de Comunicação”, que englobam os cursos e disciplinas de computação identificados na coluna “Subárea”. Após essa redução, a base atualizada passou a contar com 4.626 bolsas, onde as colunas (também identificadas pelo nome de atributo) preservadas seguiram os critérios de seleção relacionados a:

- **Período:** utilizado para identificar os anos que foram concedidas as bolsas, a fim de analisar a evolução do investimento científico. Atributo: Ano referência.
- **Destino das bolsas:** utilizado para auxiliar no agrupamento de dados identificativos. Atributos: Instituição Origem e Sigla Instituição Destino.

5. Dataset CNPq. https://basedosdados.org/dataset/bolsas-e-auxilios-pagos?external_link=Baixar+ano+a+ano.

- **Áreas do conhecimento:** permite entender para quais áreas do conhecimento as bolsas estão sendo concedidas. Atributos: Linha de Fomento, Modalidade, Categoria/Nível, Programa CNPq, Área e Subárea.
- **Fomento:** permite analisar qualitativamente os valores pagos por bolsa. Atributo: Valor Pago.

5.1 Aplicação do algoritmo *K-means* para análise do fomento científico de bolsas pelo CNPq em Minas Gerais

Com o escopo da base definido, fez-se uma nova filtragem referente às modalidades entre todas as bolsas de pesquisa contempladas em MG, especificamente em Iniciação Científica (*IC*), Iniciação Científica Júnior (*ICJ*), Mestrado (*GM*) e Doutorado (*GD*). A motivação para o uso dessas modalidades se deu por serem as mais frequentes na base de dados.

Para submeter os dados ao *K-means*, foi necessário realizar a conversão dos dados categóricos, ou seja, os textos como “ICJ - Iniciação Científica Júnior”, “GM – Mestrado”, entre outros presentes na coluna das modalidades das bolsas, em dados numéricos - pois é dessa forma que o modelo consegue interpretá-los. Para isso, foi utilizada a técnica de *One-Hot Encoding*, que consiste em criar uma tabela onde as colunas representam cada valor único submetido (cada modalidade de bolsa) e a opção por esta técnica de pré-processamento de dados deu-se por ela não atribuir pesos diferentes, categorizando os mesmos de forma binária – preenchendo 1 na coluna compatível com a informação da modalidade de determinada bolsa e 0 nas demais para todas as linhas/registros - conforme exemplo abaixo na tabela 1.

IC	ICJ	GM	GD
0	0	0	1
1	0	0	0
0	0	0	1
...			

Tabela 1. Exemplo de tabela criada pelo algoritmo One-hot encoding.

Utilizou-se o método de *Elbow* (Humaira and Rasyidah, 2020) para verificar o número ideal de *clusters* (partições/agrupamentos de dados) *K* a ser utilizado. O valor ideal de *K* para o algoritmo é identificado onde a distorção começa a diminuir mais rapidamente (Fig. “*Distorção do algoritmo elbow X número de clusters*”). Como resultado, foram obtidos dados agrupados através de intervalos do atributo de Valor Pago, ou seja, faixas próximas de valores das bolsas, conforme a Figura 1 – segunda imagem.

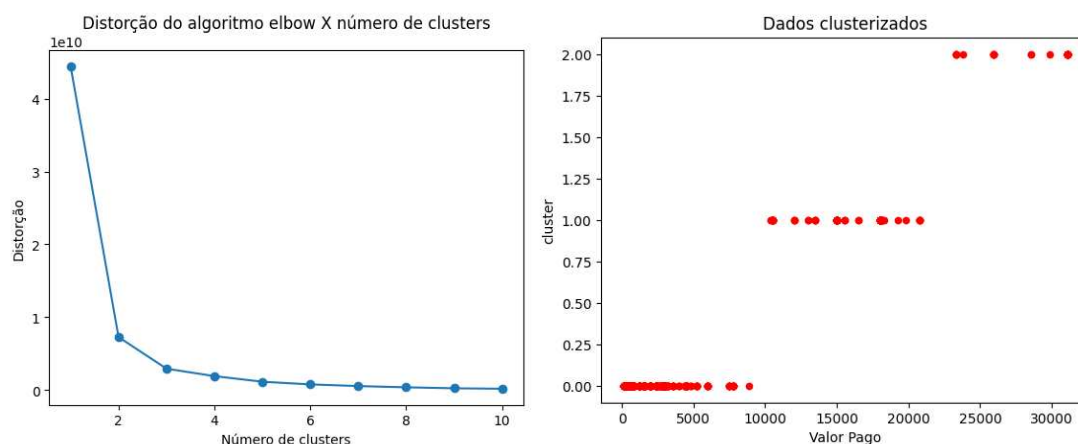


Figura 1. *K-means* aplicado aos dados.

5.1.1 Distribuição das bolsas de MG entre os governos federais

A partir do resultado obtido pelo método elbow, visualizado no gráfico de distorção da figura 1, o conjunto de dados foi dividido em duas partes: dados entre os intervalos de ano 2015 a 2018, considerando o governo de Dilma/Temer (**Governo 1**), e 2019 a 2022, considerando o governo de Bolsonaro (**Governo 2**). A Tabela 2 apresenta a relação dos governos 1 e 2, seus *clusters* que se dividem entre as faixas de valores de bolsas contempladas e a distribuição percentual da quantidade dessas bolsas separada por modalidades em cada *cluster*.

Governo	Cluster	Faixa	IC	ICJ	GM	GD
1	0	R\$ 100,00 a R\$ 8.845,00	69,9%	22,3%	5,7%	2,1%
	1	R\$ 10.376,00 a R\$ 20.752,00	0,0%	0,0%	80,0%	20,0%
	2	R\$ 23.346,00 a R\$ 31.128,00	0,0%	0,0%	2,8%	97,2%
2	0	R\$ 0,00 a R\$ 7.782,00	63,5%	30,2%	3,6%	2,7%
	1	R\$ 9.000,00 a R\$ 20.752,00	0,0%	0,0%	76,1%	23,9%
	2	R\$ 22.000,00 a R\$ 36.316,00	0,0%	0,0%	0,0%	100,0%

Tabela 2. Distribuição das bolsas por governos com *clusters* agrupados por faixas de valores de bolsas - Algoritmo *K-means*.

Observando as modalidades de bolsas em cada cluster entre os dois governos, constata-se que:

- As bolsas de Iniciação Científica Júnior aumentaram em quantidade no Governo 2.
- As bolsas pagas de valor intermediário (cluster 1) para a modalidade de Doutorado tiveram um aumento de 3,9% no Governo 2 em relação ao mesmo cluster do Governo 1.
- Não houve pagamentos de bolsas de valor alto (cluster 2) para Mestrado no Governo 2, assim como a distribuição de bolsas nessa modalidade diminuiu de forma geral (cerca de 8,8%).

5.2 A Relação de distribuição das bolsas em Minas Gerais e nos Institutos Federais

A partir das 4.626 bolsas obtidas no levantamento, a primeira averiguação dar-se-á pelo contraste entre a distribuição das bolsas da área de computação para os IFs como um conjunto único e as bolsas da área de computação para as demais instituições contempladas em MG. Uma série histórica apresentando a quantidade das unidades (de bolsas) e os anos decorrentes, permite visualizar a informação de tal contraste, conforme apresentado na Figura 2 de acordo com os dados quantitativos da Tabela 3.

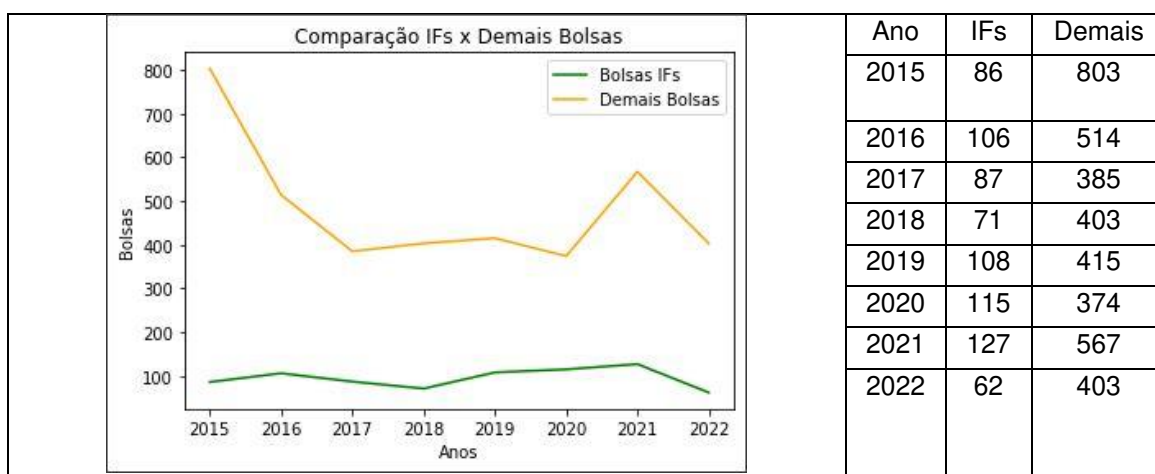


Figura 2 & Tabela 3. Comparativo entre a distribuição de bolsas de pesquisa do CNPq na área da computação para os IFs de MG e para as demais instituições de MG & relação anos x quantidade de bolsas captadas pelos IFs e demais instituições.

5.3 A representatividade dos IFs em relação a outras Universidades em Minas Gerais

Os Institutos Federais em Minas Gerais possuem um grande histórico de transformação e crescimento de suas unidades, totalizando um número de 52.236 alunos com matrículas em curso, segundo dados de 2021⁶. A Tabela 4 descreve os Institutos Federais do estado de Minas Gerais, o quantitativo de campi/polos de inovação e de cursos relacionados com Tecnologia da Informação (TI)/Computação, bem como a quantidade de cursos por modalidade, distribuída entre Médio-Técnico (T), Superior (S), Especialização (E) e Formação Inicial e Continuada (F) para cada instituto.

6. <https://www.gov.br/mec/pt-br/pnp>

Instituto	Campi/Polos	Cursos TI	T	S	E	F
IF do Triângulo Mineiro (IFTM)	14	51	20	26	1	4
IF do Sul de Minas Gerais (IF Sul de Minas)	8	45	28	15	2	0
IF do Sudeste de Minas Gerais (IF Sudeste MG)	10	38	17	8	4	9
IF do Norte de Minas Gerais (IFNMG)	11	27	17	10	0	0
IF de Minas Gerais (IFMG)	19	12	7	5	0	0

Tabela 4. IFs de MG - Campi/Polos - Cursos de TI - Modalidades de Cursos (T-S-E-F).

Em destarte a relação de alunos e proporção dos IFs, também foram selecionadas algumas das instituições mais tradicionais do estado, sendo elas o Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), Universidade Federal de Viçosa (UFV), Universidade Federal de Uberlândia (UFU) e a Universidade Federal de Ouro Preto (UFOP), escolhidas por sua presença significativa na captação de bolsas e produção científica no período (2015 - 2022). O objetivo foi avaliar a correlação entre as quantidades de bolsas e alunos ativos, possibilitando uma melhor interpretação a respeito do crescimento e a produção científica fomentada. Portanto, construiu-se uma base agregando a quantidade de bolsas em Tecnologia da Informação (TI)/Computação e a quantidade de alunos, conforme a distribuição apresentada na Figura 3 e Tabela 5.

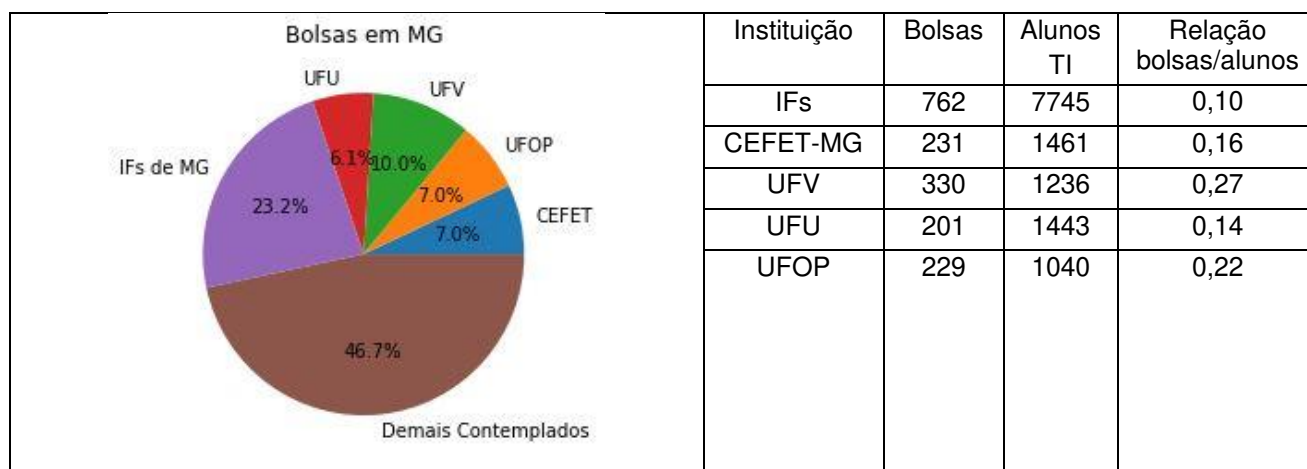


Figura 3 & Tabela 5. Comparação entre IFs e UFs e relação bolsas CNPq captadas x alunos dos cursos TI das instituições.

6. RESULTADOS

Ao longo do período analisado, houve uma queda drástica no valor distribuído em bolsas de Minas Gerais, começando a partir de 2016 e se mantendo em queda até 2022. O investimento em bolsas em MG em 2021 é o mais baixo no período analisado, sendo quatro vezes menor que o valor de 2015. No período do governo Jair Bolsonaro o investimento em bolsas foi de R\$ 10.674.994,88 milhões, cerca de R\$ 1.703.781,29 milhões abaixo de um ano de investimento no governo Dilma, que fora cerca de R\$ 12.378.776,17 milhões em 2015. A queda abrupta no período de 2016, pode ser atribuída ao governo transitório de Michel Temer, que assumiu a

presidência após o *impeachment* de Dilma Rousseff em agosto do mesmo ano. Onde fora encaminhada a proposta de Emenda à Constituição (PEC) 241/2016 *pec-241*, que congelou os investimentos também em educação, proposta a qual teve continuidade no governo posterior.

Percebe-se um contraste com relação a evolução do volume de bolsas durante o período analisado, entre as bolsas contempladas para as demais instituições e as bolsas contempladas para os IFs mineiros, onde, ocorre um decaimento de 52% entre 2015 a 2017 sob as demais bolsas. Após este efeito, esse grupo manteve-se com um volume consideravelmente inferior ano a ano, configurando uma queda de 49,81% da quantidade de bolsas distribuídas pelo CNPq, quando é feita a análise do período inicial de 2015 e do período final de 2022.

Não obstante, os IFs de MG tiveram uma média de crescimento de 12,29% na quantidade de bolsas recebidas pelo CNPq durante o intervalo de 7 anos. Esse comportamento dos IFs, que aparenta ser divergente do processo e contexto de redução dos investimentos em educação, dada a PEC-241, a pandemia do covid-19 e o cancelamento de editais de bolsas, pode ser compreendido pelo crescimento orgânico das novas instituições, devido a ofertas de novos cursos e/ou faculdades nas áreas, mesmo considerando que alguns IFs já existiam como escolas técnicas previamente. Outro fator influenciador é a interiorização da educação, onde muitos campi dos Institutos, impactam regiões onde não ocorria educação pública superior ou técnica (Rosinke et al, 2020).

Percebe-se que os Institutos compõem uma parcela majoritária das bolsas distribuídas em comparação as universidades selecionadas em MG. A quantidade de alunos de TI dos IFs é até sete vezes maior do que algumas destas universidades. Tal fato, pode ser relacionado com a oferta do ensino técnico nos IFs. Além disso, fazendo uso da correlação de Pearson (Figueiredo Filho and Silva Júnior, 2009), verificou-se que a quantidade de alunos está fortemente correlacionada com a quantidade de bolsas ($p = 0,97$). A correlação estatística verifica a relação entre duas variáveis, onde próxima ou igual a 1 significa alta, próxima ou igual a 0 baixa e próxima ou igual a -1 inversamente correlacionadas.

7. CONSIDERAÇÕES FINAIS

Embora as bolsas de estudo tenham um papel fundamental na promoção de oportunidades acadêmicas e no desenvolvimento científico, o investimento governamental tem diminuído nos últimos 8 anos, sobretudo após a gestão governamental que se encerra em 2016 e o surgimento de diretrizes como a PEC-241. Os cortes em investimentos na educação pública foram alvos de críticas sob governos posteriores, e o decaimento do fomento a bolsas reflete parte desse cenário de depressão no incentivo à produção científica relacionado ao CNPq.

Alguns dos impactos podem ser identificados através da falta de mão de obra qualificada - onde o Brasil, segundo a pesquisa "Talent Shortage" de 2023,

conduzida pela empresa de consultoria *manPowergroup*, possui uma escassez de 80% e *TI e Análise de dados* são as habilidades mais buscadas, representando cerca de 27% da demanda. Quanto a outro fator denominado "fuga de cérebros", para outros países – se trata de profissionais que migram para outros países por razões diversas, onde Silva. 2023, em “*fuga de cérebros: êxodo de capital intelectual especializado no brasil na última década e os impactos para o desenvolvimento nacional*”, traz que as 3 maiores motivações são baixos salários, falta de oportunidades e de reconhecimento.

A partir do contexto obtido e analisado a respeito da participação das instituições em Minas Gerais, foi possível identificar que os IFs possuem uma presença significativa na captação de investimentos a pesquisa, apesar da relação bolsas por alunos ser a menor das comparadas. Há crescimento desde 2015, que apesar de não consistente percentualmente a cada ano, possui plena relação com o potencial de alavancar o fomento em Minas Gerais. O menor investimento em bolsas do CNPq durante o governo de 2019-2022 retrata um desafio na projeção desse crescimento.

Para trabalhos futuros, pretende-se estender a análise sobre a área da computação em âmbito nacional, bem como realizar o estudo da relação de cursos ofertados na modalidade “Educação a Distância” (EAD) pelos IFs com o crescimento do fomento à pesquisa científica em suas instituições.

8. FONTES CONSULTADAS

- Macedo, P. C. S. (2017). Educação profissional e desenvolvimento territorial: a expansão dos institutos federais de educação, ciência e tecnologia. *Revista Brasileira da Educação Profissional e Tecnológica*, 2(13):94–106.
- PISA 2018. <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/pisa/resultados>.
- Domingos, M. (2011). A trajetória do CNPq. *Acervo*, 17(2):19–40.
- NEIVA, Silmara Cássia Pereira Couto et al. Perspectivas da ciência brasileira: um estudo sobre a distribuição de bolsas de pesquisa em produtividade do CNPq ao longo do ano de 2019. *Revista Interdisciplinar Científica Aplicada*, v. 16, n. 3, p. 51-71, 2022.
- DE MESQUITA, Patrícia Paiva; SBARAINI, Fabiana Leticia; FERREIRA, Daniele Sayuri Fujita. Iniciação científica e tecnológica no Instituto Federal de Educação, ciência e tecnologia de Roraima: um panorama entre 2015 e 2019: Scientific and technological initiation in the Federal Institute of Education, science and technology of Roraima: an overview between 2015 and 2019. *Brazilian Journal of Development*, v. 8, n. 7, p. 54151-54165, 2022.
- Times Higher Education World University Rankings. <https://www.topuniversities.com/university-rankings/world-university-rankings>.
- Center for World University Rankings. <https://cwur.org/2022-23.php>.
- CNPq. <https://www.gov.br/CNPq/pt-br>.
- de Mesquita, P. P., Sbaraini, F. L., and Ferreira, D. S. F. (2022). Iniciação científica e tecnológica no instituto federal de educação, ciência e tecnologia de roraima: um panorama entre 2015 e 2019.
- Hartigan, J. A. (1975). *Clustering algorithms*. John Wiley & Sons, Inc.
- Figueiredo Filho, D. B. and Silva Júnior, J. A. (2009). Desvendando os mistérios do coeficiente de correlação de pearson (r). *Revista Política Hoje*, 18(1):115–146.
- Silva, W. P. B. (2023). Fuga de cérebros: êxodo de capital intelectual especializado no brasil na última década e os impactos para o desenvolvimento nacional.
- Dataset CNPq. https://basedosdados.org/dataset/bolsas-e-auxilios-pagos?external_link=Baixar+ano+a+ano.
- Rosinke, J. G., de Carvalho, E. T., Rosinke, G. C. L., and da Silva, G. J. S. (2020). A participação dos institutos federais na interiorização da educação superior presencial no brasil. *Research, Society and Development*, 9(2):24.