

**FUNDAÇÃO GETULIO VARGAS
ESCOLA DE POLÍTICAS PÚBLICAS E GOVERNO**

PATRÍCIA DE CASTRO NEVES LINHARES

**MONITORAMENTO E AVALIAÇÃO DE POLÍTICAS PÚBLICAS: ANÁLISE
COMPARADA DE BOLSAS DE DOUTORADO PLENO NO PAÍS E NO
EXTERIOR CONCEDIDAS PELA CAPES**

BRASÍLIA

2021

PATRÍCIA DE CASTRO NEVES LINHARES

**MONITORAMENTO E AVALIAÇÃO DE POLÍTICAS PÚBLICAS: ANÁLISE
COMPARADA DE BOLSAS DE DOUTORADO PLENO NO PAÍS E NO
EXTERIOR CONCEDIDAS PELA CAPES**

Dissertação apresentada à Escola de Políticas Públicas e Governo da Fundação Getúlio Vargas como requisito para obtenção do título de mestre em Políticas Públicas e Governo.

Área de concentração: Políticas Públicas e Governo

Orientador: Prof. Dr. Ricardo Corrêa Gomes

BRASÍLIA

2021

Linhares, Patrícia de Castro Neves.

Monitoramento e avaliação de políticas públicas: análise comparada de bolsas de doutorado pleno no país e no exterior concedidas pela CAPES / Patrícia de Castro Neves Linhares. – 2021
106 f.

Orientador: Ricardo Corrêa Gomes.

Dissertação (mestrado MPPG) - Fundação Getulio Vargas, Escola de Políticas Públicas e Governo.

1. Educação e Estado. 2. Políticas públicas - Brasil. 3. Bolsas de estudo. 4. Brasil. Coordenação do Aperfeiçoamento de Pessoal de Nível Superior. I. Gomes, Ricardo Corrêa. II. Dissertação (mestrado MPPG) – Escola de Políticas Públicas e Governo. III. Fundação Getulio Vargas. IV. Título.

CDU: 378.3(81)

PATRÍCIA DE CASTRO NEVES LINHARES

**MONITORAMENTO E AVALIAÇÃO DE POLÍTICAS PÚBLICAS:
ANÁLISE COMPARADA DE BOLSAS DE DOUTORADO PLENO
NO PAÍS E NO EXTERIOR CONCEDIDAS PELA CAPES**

Dissertação apresentada à Escola de Políticas Públicas e Governo da Fundação Getúlio Vargas como requisito para obtenção do título de mestre em Políticas Públicas e Governo.

Área de concentração: Políticas Públicas e Governo

Data de aprovação: 15/06/2021.

Prof. Dr. Ricardo Corrêa Gomes
Orientador
FGV - EAESP

Prof. Dr. Fernando Luiz Abrucio
FGV - EAESP

**Prof. Dr. Antônio Sérgio Araújo
Fernandes**
UFBA-NPGA

Com gratidão, dedico este trabalho a Deus, Autor da vida, Criador de todas as coisas visíveis e invisíveis, a quem devo tudo o que sou e tenho. E ao meu esposo, companheiro amável, que me incentivou e muito me compreendeu e ajudou.

AGRADECIMENTOS

Ao Soberano Deus Pai Altíssimo a quem sirvo, por seu amor, sua fidelidade e sua misericórdia infinita. A Deus Filho, Jesus Cristo, meu Senhor e Salvador, o caminho, a verdade e a vida. A Deus Espírito Santo, o Paráclito, melhor consultor que inspira a sabedoria, a ciência, a fortaleza, o entendimento. À Maria Santíssima, Mãe de Jesus e nossa, que é bendita entre todas as mulheres, pela intercessão e por passar à frente dos meus passos, das minhas decisões.

Aos meus pais, Davi e Luzia, que me sustentaram, criaram e educaram. A minha mãe amiga, que me conduziu pelos caminhos retos e até hoje é meu exemplo de simplicidade, humildade, caridade, amor incondicional e doação.

Ao meu esposo querido, Marcelo, e as minhas amadas filhas, Helena e Luíza, família linda que Deus me concedeu e que tanto me alegra, que sempre estiveram ao meu lado, na alegria e na tristeza, na saúde e na doença.

As minhas amigas e irmãs de coração, Patrícia Santiago, Marybel Camargo, Ilka Mendonça, e a mais nova, Luciana Barbosa, pelas palavras de carinho e incentivo, pelo apoio emocional e espiritual.

Ao meu orientador, Prof. Ricardo Corrêa Gomes, primeiro professor no curso, que desde o início me apoiou e aceitou com generosidade me orientar, pela compreensão, atenção e competência.

Aos colegas parceiros da Turma 1 do MPPG/FGV, pelas partilhas, ora alegres, ora de dificuldades, nos quais sempre encontrei ajuda e amizade, bem como aos colegas da CAPES, especialmente meu chefe imediato, Rubens Sousa, pela compreensão, pelo total apoio e incentivo.

Ao Programa de Pós-Graduação do MPPG/FGV-Brasília, na pessoa do Prof. Benjamin Miranda Tabak, Coordenador do curso, que sempre esteve próximo aos discentes, disposto para ajudar em tudo. Aos demais professores e monitores, pelos conhecimentos compartilhados e pelas aulas preciosas, e à equipe de funcionários, pela atenção dispensada nas demandas dos alunos.

*"Comece fazendo o que é necessário, depois o que é possível, e de repente você
estará fazendo o impossível."
(São Francisco de Assis)*

RESUMO

A CAPES dispõe de várias ações por meio das quais executa as políticas públicas educacionais direcionadas à pós-graduação, sendo as ações de investimento na formação de recursos humanos de alto nível no país e no exterior – concessão de bolsas de pós-graduação no Brasil e no exterior – as que possuem maiores dotações orçamentárias anualmente, de acordo com as especificações contidas em seus orçamentos anuais. No debate acerca das políticas públicas educacionais implementadas pela CAPES, alguns assuntos relacionados aos programas de concessão de bolsas têm sido discutidos na comunidade acadêmica, a saber: acompanhamento e monitoramento dos programas; aprimoramento de metas e de indicadores de desempenho dos programas; entre outros. Sob a perspectiva das três dimensões do desempenho – eficácia, eficiência e efetividade, este estudo pretende contribuir com discussões sobre o aperfeiçoamento das práticas de monitoramento e avaliação dos programas de concessão de bolsas no âmbito da CAPES, utilizando-se de um estudo comparado entre as bolsas de doutorado pleno no país e no exterior. Os dados para a pesquisa foram extraídos de sistema mantido pela CAPES que armazena informações do Sistema Nacional de Pós-Graduação – SNPG, e do currículo *lattes* dos bolsistas, usando ferramentas de TI próprias para esse tipo de extração. A análise dos dados foi feita utilizando a ferramenta de BI *QlikSense* e o software estatístico R. Os resultados obtidos indicam que: o investimento em bolsas no país é maior, mas o valor individual percebido pelos bolsistas é maior para as bolsas de estudo no exterior; a produção científica dos bolsistas no país é significativamente maior que a dos bolsistas no exterior, em relação ao valor individual da bolsa e a concessão de valor maior de bolsa não resulta necessariamente em mais artigos publicados.

Palavras-chave: Políticas públicas. Monitoramento. Avaliação. Bolsa de estudo. Doutorado pleno.

ABSTRACT

CAPES Foundation has several actions through which it executes educational public policies aimed at graduate education, with investment actions in the training of high-level human resources in the country and abroad - granting of graduate scholarships in Brazil and abroad - those with the highest annual budget allocations, according to the specifications contained in their annual budgets. In the debate about educational public policies implemented by CAPES, some issues related to scholarship award programs have been discussed in the academic community, such as follow-up and monitoring of programs; improvement of program performance targets and indicators; among others. From the perspective of the three dimension of performance – effectiveness, efficiency and effectiveness, this research aims to contribute to discussions on improving the practices for monitoring and evaluating scholarship award programs under CAPES, using a comparative study among the scholarships in Brazil and abroad. The data for the research were extract from a CAPES' information System and from the lattes curriculum of the scholarship holders, using Information Technology tools for this extraction. The data analysis was performed using the BI tool Qlik Sense and the statistical software R. The results of the analysis indicate that: investment in scholarships in Brazil is greater, but the individual value perceived by scholarship holders is greater for those students who study abroad; the scientific production of scholarship holders in Brazil is significantly higher than that of the students who study abroad, in relation to the individual scholarship amount; and a higher scholarship amount does not necessarily results in higher number of published scientific articles.

Keywords: Public policies. Evaluation. Monitoring. Scholarships. PhD program.

LISTA DE ILUSTRAÇÕES

Figura 1 – Legenda dos ícones utilizados nas ilustrações	39
Figura 2 – Desenho da primeira etapa para a obtenção dos dados	40
Figura 3 – Desenho da segunda etapa para a obtenção dos dados	40
Figura 4 – Desenho da terceira etapa para a obtenção dos dados	41
Figura 5 – Desenho da quarta e quinta etapas para a obtenção dos dados . . .	42
Figura 6 – Modelo do Banco de Dados	43
Figura 7 – Desenho da sexta etapa para a obtenção dos dados	44
Figura 8 – Desenho da oitava e nona etapas para a obtenção dos dados	46
Figura 9 – Distribuição de bolsas no país por programa CAPES - DPB	48
Figura 10 – Distribuição de bolsas no exterior por programa CAPES - DRI	49
Figura 11 – <i>Status</i> do doutorado dos bolsistas	50
Figura 12 – <i>Status</i> do doutorado dos bolsistas no exterior	50
Figura 13 – Distribuição dos bolsistas segundo o endereço profissional	51
Figura 14 – Quantidade de bolsas nas grandes áreas do conhecimento	53
Figura 15 – Quantidade de bolsas nas grandes áreas do conhecimento – DRI .	54
Figura 16 – Quantidade de bolsas nas grandes áreas do conhecimento – DPB .	54
Figura 17 – Distribuição de artigos publicados	55
Figura 18 – Distribuição de artigos publicados segundo tipo de bolsa	56
Figura 19 – Artigos publicados no período de 2010 a 2020	57
Figura 20 – Artigos publicados no período de 2010 a 2020 segundo tipo de bolsa	57
Figura 21 – Artigos publicados segundo grande área do conhecimento	58
Figura 22 – Artigos publicados segundo grande área do conhecimento e tipo de bolsa	58
Figura 23 – Artigos publicados segundo grande área do conhecimento e tipo de bolsa	59
Figura 24 – Artigos publicados segundo grande área do conhecimento e tipo de bolsa - DPB	60
Figura 25 – Artigos publicados segundo grande área do conhecimento e tipo de bolsa - DRI	60
Figura 26 – Distribuição do valor da bolsa (em reais)	61
Figura 27 – Distribuição do valor da bolsa (em reais) segundo tipo de bolsa . . .	62
Figura 28 – Distribuição do valor da bolsa (em reais) segundo grande área do conhecimento	62
Figura 29 – Distribuição do valor da bolsa (em reais) segundo grande área do conhecimento - DPB	63

Figura 30 – Distribuição do valor da bolsa (em reais) segundo grande área do conhecimento - DRI	63
Figura 31 – Distribuição do valor da bolsa DPB (em reais) por região	64
Figura 32 – Distribuição do valor da bolsa DPB (em reais) por região e grande área do conhecimento	65
Figura 33 – Quantidade de artigos publicados e valor da bolsa (em reais)	66
Figura 34 – Quantidade de artigos publicados e valor da bolsa (em reais) segundo tipo de bolsa	66
Figura 35 – Quantidade de artigos publicados e valor da bolsa DPB (em reais) segundo região	67
Figura A.1–Capítulos de livros publicados por ano	83
Figura A.2–Capítulos de livros por ano e tipo de bolsa	83
Figura A.3–Capítulos de livros por grande área de conhecimento	84
Figura A.4–Distribuição de capítulos de livros e valor da bolsa (em reais) por tipo de bolsa	84
Figura A.5–Livros publicados por ano	85
Figura A.6–Livros por ano e tipo de bolsa	85
Figura A.7–Livros por grande área de conhecimento	86
Figura A.8–Distribuição de livros publicados e valor da bolsa (em reais) segundo tipo de bolsa	86
Figura A.9– <i>Softwares</i> desenvolvidos por ano	87
Figura A.10– <i>Softwares</i> desenvolvidos por ano e tipo de bolsa	87
Figura A.11– <i>Softwares</i> desenvolvidos por grande área de conhecimento	88
Figura A.12– Distribuição de <i>softwares</i> desenvolvidos e valor da bolsa (em reais) segundo tipo de bolsa	88
Figura A.13– Patentes registradas por ano	89
Figura A.14– Patentes registradas por ano e tipo de bolsa	89
Figura A.15– Patentes registradas por grande área de conhecimento	90
Figura A.16– Distribuição de patentes registradas e valor da bolsa (em reais) segundo tipo de bolsa	90
Figura A.17– Produtos tecnológicos desenvolvidos por ano	91
Figura A.18– Produtos tecnológicos desenvolvidos por ano e tipo de bolsa . . .	91
Figura A.19– Produtos tecnológicos desenvolvidos por grande área de conhecimento	92
Figura A.20– Distribuição de produtos tecnológicos desenvolvidos e valor da bolsa (em reais) segundo tipo de bolsa	92

LISTA DE TABELAS

Tabela 1 – Cotações moedas estrangeiras	48
Tabela 2 – Países de vínculo profissional dos bolsistas	51
Tabela 3 – Execução da Ação 0487 no período de 2010 a 2014	69
Tabela 4 – Execução da Ação 09GK no período de 2010 a 2014	70
Tabela 5 – Execução da Ação 0935 no período de 2010 a 2014	70
Tabela 6 – Meta financeira da Ação 0487 no período de 2010 a 2014	71
Tabela 7 – Meta financeira da Ação 09GK no período de 2010 a 2014	71
Tabela 8 – Meta financeira da Ação 0935 no período de 2010 a 2014	71
Tabela 9 – Estatísticas do valor da bolsa segundo o tipo de bolsa	72
Tabela A.1–Estatísticas da quantidade de artigos segundo o tipo de bolsa	81
Tabela A.2–Estatísticas da quantidade de artigos segundo a área de conhecimento	81
Tabela A.3–Estatísticas da quantidade de artigos segundo o tipo de bolsa e a área de conhecimento	81
Tabela A.4–Estatísticas do valor da bolsa segundo a grande área	82
Tabela A.5–Estatísticas do valor da bolsa segundo tipo de bolsa e a área de conhecimento	82
Tabela A.6–Estatísticas do valor da bolsa no país segundo a região	82

LISTA DE QUADROS

Quadro 1 – Definição de Eficácia, Eficiência e Efetividade	28
Quadro 2 – Programa de Governo 1375 sob responsabilidade da CAPES . . .	68
Quadro 3 – Programa de Governo 2032 sob responsabilidade da CAPES . . .	68

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Problematização	19
1.2	Hipótese	20
1.3	Objetivos	20
1.3.1	Objetivo Geral	20
1.3.2	Objetivos Específicos	20
1.4	Justificativa do estudo	21
2	REFERENCIAL TEÓRICO	23
2.1	Políticas Públicas – conceito, funções e ciclo	23
2.2	Monitoramento e Avaliação de Políticas Públicas	25
2.3	Indicadores sociais para as atividades de monitoramento e avaliação	29
3	METODOLOGIA	33
3.1	Extração de dados do sistema da CAPES	34
3.2	Extração e normalização de dados do Currículo <i>Lattes</i>	36
4	ANÁLISE E RESULTADOS	47
4.1	Programas de concessão de bolsa (DPB e DRI)	47
4.2	Titulação	49
4.3	Endereço profissional	50
4.4	Grande Área do conhecimento	53
4.5	Artigos	54
4.6	Investimento em bolsa	60
4.7	Investimento em bolsa no país por região	64
4.8	Relação entre quantidade de artigos e investimento em bolsa	65
4.9	Desempenho dos Programas de bolsa da CAPES	67
5	CONCLUSÕES	73
	REFERÊNCIAS	76
	APÊNDICE A – FIGURAS E TABELAS	81
	APÊNDICE B – <i>SCRIPTS</i> UTILIZADOS NO PROCESSO ETL	93

B.1	<i>Script em Bash - Captura de nome e ID das telas paginadas da plataforma Lattes</i>	93
B.2	<i>Script em Python - Relacionar Nome e ID 1</i>	93
B.3	<i>Script em Python - Relacionar Nome e ID 2</i>	96
B.4	<i>Script em Python – Download do XML 1</i>	98
B.5	<i>Script em Python – Download do XML 2</i>	98
B.6	<i>Script em Python – Descompactar XML</i>	100
B.7	<i>Script em Python – Carga no banco MySQL</i>	102
B.8	<i>Script em Python – Gerar CSV</i>	105
B.9	<i>Script em Python – Verificação do arquivo zip</i>	107

LISTA DE ABREVIATURAS E SIGLAS

BEX	Doutorado Pleno CSF
BI	<i>Business Intelligence</i>
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CF/88	Constituição Federal de 1988
CGU	Controladoria-Geral da União
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
CsF	Programa Ciência sem Fronteiras
DINTER	Programa de Doutorado Interinstitucional
DOC-PLENO	Programa de Doutorado Pleno do Exterior
DPB	Diretoria de Programas e Bolsas no País
DPE	Programa de Doutorado Pleno no Exterior
DRI	Diretoria de Relações Internacionais
DS	Programa de Demanda Social
ETL	Extract, Transform and Load
FAP	Fundação de Apoio à Pesquisa
MEC	Ministério da Educação
MINDIN	Programa MINTER/DINTER CAPES/SETEC
MRE	Ministério das Relações Exteriores
PBE-DPM	Bolsa Especial para Doutorado em Pesquisa Médica
PDRAE	Plano Diretor da Reforma do Aparelho do Estado
PPA	Plano Plurianual
PROEX	Programa de Excelência Acadêmica

PROEX-TAXAS	Programa de Excelência Acadêmica – Taxas
PPG	Programa de Pós-Graduação
SCPB	Sistema de Conciliação e Pagamento de Bolsas
SIMEC	Sistema Integrado de Monitoramento, Execução e Controle do Ministério da Educação
SNPG	Sistema Nacional de Pós-Graduação
TCU	Tribunal de Contas da União
TI	Tecnologia da Informação
TIC	Tecnologias de Informação e Comunicação
Art.	Artigo

1 INTRODUÇÃO

As políticas públicas são formuladas para atender a uma necessidade ou demanda da sociedade (BRAGA; GOMES, 2016) ou para solucionar um problema. Easton (1965) define políticas públicas como um conjunto de elementos interligados, sendo estes elementos a formulação da política, seus resultados e o ambiente em que ela foi implementada. Desenvolvidas por meio de ações, decisões, atividades, projetos e programas, as políticas públicas devem ser monitoradas e avaliadas, a fim de subsidiar os gestores públicos com informações acerca de sua operacionalização, seu funcionamento, seus resultados e efeitos.

Segundo Abrucio (2010), a Constituição Federal de 1988 (CF/88), no campo das políticas públicas, traz cinco grandes pilares, quais sejam: o primeiro é a busca pela universalização das políticas, a fim de promover a plena garantia dos direitos sociais; o segundo refere-se à democratização da ação do Estado, sob os aspectos da participação no plano deliberativo e do controle do poder público; o terceiro trata-se da profissionalização da burocracia, viabilizado pelos concursos e pelas carreiras públicas, sendo esta uma condição primordial para garantir a qualidade das ações governamentais; e os dois últimos pilares estão concatenados, participando da dinâmica federativa.

“Trata-se da descentralização, preferencialmente em prol da municipalização das políticas, e a preocupação com a interdependência federativa, na forma de medidas de combate à desigualdade, de preocupações em torno da cooperação intergovernamental e da definição de um raio importante de ações federais como agente nacional.” (ABRUCIO, 2010, p.45).

A crescente demanda por políticas públicas e a escassez de recursos públicos aumenta a necessidade de eficiência na gestão pública, de forma a potencializar os efeitos dos investimentos nas ações governamentais e a minimizar seus custos operacionais. Para Jannuzzi (2016), são crescentes o empenho e a preocupação referentes à produção de informação e conhecimento para conduzir a gestão de políticas e programas sociais, mediante às atividades de monitoramento e avaliação da ação pública em várias esferas. Nesse contexto, faz-se necessário o desenvolvimento de sistemas de monitoramento e avaliação que assegurem melhores resultados na gestão dos recursos públicos.

A nova gestão pública dentro da reforma do aparelho do Estado teve como instrumento fundamental o Plano Diretor da Reforma do Aparelho do Estado - PDRAE (1995), que enfatizava a eficiência e a busca por resultados e buscou fortalecer a

capacidade de governo do Estado, substituindo a administração pública burocrática – rígida e ineficiente, focada em si, pela administração pública gerencial – flexível e eficiente, direcionada para o cidadão (COSTA, 2008). Posteriormente, inovações foram trazidas pelo Plano Plurianual (PPA) 2000-2003, tais como o foco na obtenção de resultados do programa e a sua avaliação, definição clara dos objetivos, público-alvo, valores, metas físicas, e o uso de indicadores para a mensuração dos resultados (COUTO; CARDOSO JÚNIOR, 2020).

Segundo Mazza (2020), o modelo de administração pública gerencial com foco no controle de resultados foi implementado pelo princípio da eficiência, de maneira que os atos praticados pela administração pública devem ser pautados com a maior qualidade, competência e eficácia possíveis no interesse da sociedade. O princípio da eficiência é o mais recente princípio constitucional que rege a administração pública, tendo sido legitimado com a promulgação da Emenda Constitucional nº 19, de 04 de junho de 1998, a qual modificou o regime e dispôs sobre princípios e normas da Administração Pública dentre outras providências. Este princípio imputa ao gestor o dever de gerir a coisa pública com legalidade, moralidade, economicidade, efetividade e transparência, com vistas ao cumprimento das metas pré-estabelecidas.

No universo das políticas públicas estão inseridas as políticas públicas educacionais, que fazem parte das políticas sociais. Para Azevedo (2004), a educação é uma política social, e sob esta concepção, afirma que é preciso interpretá-la no espaço reservado às políticas públicas, a fim de perceber as intervenções provenientes do contexto social em que elas são elaboradas. De acordo com Castro et al. (1988) não há consenso sobre a definição do termo política social, apesar de a bibliografia sobre o tema ser ampla e variada. Fazem parte da política social no Brasil, por exemplo, segundo a ideia de “área de atuação”, utilizada por Castro et al. (1988): educação, emprego, habitação, saneamento, saúde, seguridade social, urbanismo.

No cenário brasileiro, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, doravante CAPES, se destaca na formulação e execução de políticas públicas educacionais, sobretudo na implementação de ações de fomento à pós-graduação *stricto sensu* no Brasil e no exterior.

A CAPES, fundação pública vinculada ao Ministério da Educação (MEC), possui a missão primordial de subsidiar este Ministério no tocante à formulação de políticas públicas direcionadas à formação de pessoal qualificado para o atendimento das demandas dos setores público e privado. Dessa forma, desempenha função fundamental e expressiva na expansão e consolidação da pós-graduação *stricto sensu* (mestrado e doutorado) em todo o Brasil, e, desde 2007, somou-se à sua missão a atuação na formação inicial e continuada de professores para a educação básica.

Essa fundação tem o papel de expandir e consolidar a pós-graduação brasileira,

promover a cooperação acadêmica internacional e a formação de profissionais da educação básica. Para atingir seus objetivos, a instituição promove parcerias com universidades, centros universitários, institutos de pesquisa, docentes, discentes e pesquisadores, nas esferas nacional e internacional.

A CAPES tem contribuído, desde a sua criação em 1951, com a promoção de quadros de alto nível técnico científico para o país, sendo essa uma condição imprescindível para a manutenção do desenvolvimento sustentável e de longo prazo da nação brasileira. Com base nos dados apresentados no Relatório de Gestão do exercício de 2018 desta fundação pública, é possível constatar a amplitude do investimento realizado, a partir do qual foram concedidas, por exemplo, 101.228 (cento e uma mil, duzentas e vinte e oito) bolsas de estudo no país em 2018, sendo 47.506 (quarenta e sete mil, quinhentas e seis) de mestrado, 44.530 (quarenta e quatro mil, quinhentas e trinta) de doutorado, 7.075 (sete mil e setenta e cinco) de pós-doutorado, e 2.117 (duas mil, cento e dezessete) de outras modalidades, tais como iniciação científica, pesquisador visitante, Idiomas sem Fronteiras e outras, perfazendo um investimento de mais de R\$ 2,1 bilhões, beneficiando discentes, docentes e pesquisadores de mais de 3,6 mil Programas de Pós-Graduação (PPGs) *stricto sensu* recomendados pela CAPES (BRASIL, 2019b).

No que se refere às bolsas no exterior, também no ano de 2018, foram beneficiados 8.179 (oito mil cento e setenta e nove) estudantes, pesquisadores e professores, com destino a diversos países, sendo mais de quatro mil bolsas da modalidade de doutorado-sanduíche, quase duas mil bolsas de graduação-sanduíche e de doutorado pleno. Ademais, 684 (seiscentos e oitenta e quatro) cidadãos estrangeiros, entre discentes e docentes, que vieram ao Brasil desenvolver seus estudos e suas pesquisas. Cabe mencionar que no ano de 2018 a implementação do Programa CAPES-Print teve início, com a seleção de projetos institucionais de internacionalização da pós-graduação de 36 (trinta e seis) Instituições de Ensino Superior no Brasil. A estimativa do programa é conceder mais de 5.500 (cinco mil e quinhentas) bolsas de estudo no exterior por ano, no período de 2018 a 2022.

1.1 Problemática

É possível pensar que os discentes de instituições de ensino no exterior produzam mais trabalhos científicos, devido aos indicadores e índices educacionais dessas instituições serem, em sua maioria, melhores que os das instituições de ensino no Brasil, sobretudo nos países desenvolvidos.

Entende-se que o alto investimento em bolsas no exterior é justificado por fatores como a internacionalização dos estudantes, a promoção da formação internacional

de redes de pesquisa, e a melhoria da qualidade da produção acadêmica atrelada à pós-graduação brasileira, entre outros elementos.

Neste contexto, tem-se como pergunta de pesquisa: Qual o impacto da concessão de bolsa de doutorado pleno no exterior em relação à concessão da mesma modalidade de bolsa no país para a pós-graduação no Brasil?

O problema principal a ser analisado neste trabalho trata da investigação sobre os montantes de recursos financeiros investidos em bolsas de doutorado pleno pela CAPES, no Brasil e no exterior, e na produção acadêmica e científica (artigos, capítulos de livros, livros e outros) realizada pelos beneficiários dessas bolsas.

1.2 Hipótese

A bolsa de doutorado pleno no exterior é mais onerosa para os cofres públicos e não implica mais publicação acadêmica e científica em relação à bolsa de doutorado pleno no país.

1.3 Objetivos

1.3.1 Objetivo Geral

O objetivo geral da pesquisa é investigar qual o impacto para o país da concessão de bolsas de estudo da modalidade de doutorado pleno no exterior em relação à concessão de bolsas da mesma modalidade no Brasil, utilizando banco de dados da CAPES bem como de extração de informações constantes dos currículos *lattes* dos bolsistas, perpassando pela análise dos processos de monitoramento e avaliação desses programas de concessão de bolsa de pós-graduação *stricto sensu*, no país e no exterior, implementados pela CAPES.

1.3.2 Objetivos Específicos

- I. Conhecer os processos de monitoramento e avaliação inseridos no ciclo de políticas públicas;
- II. Fazer um levantamento dos indicadores utilizados na avaliação dos programas de concessão de bolsa de doutorado pleno e na aferição de resultados;
- III. Comparar custos e benefícios da concessão de bolsas de doutorado pleno no exterior e no país; e

- IV. Propor melhorias para a entidade no que tange à avaliação dos programas de concessão de bolsas.

1.4 Justificativa do estudo

A CAPES dispõe de várias ações por meio das quais executa as políticas públicas educacionais direcionadas à pós-graduação, sendo as ações de investimento na formação de recursos humanos de alto nível no país e no exterior – concessão de bolsas de pós-graduação no Brasil e no exterior – as que possuem maiores dotações orçamentárias anualmente, de acordo com as especificações contidas em seus orçamentos anuais.

Os programas de concessão de bolsas da CAPES foram definidos no PPA 2012-2015, com as seguintes classificações:

"Programa 2032 - Educação Superior: Graduação, Pós-Graduação, Ensino, Pesquisa e Extensão.

Objetivo 0803 - Apoiar a formação de pessoal qualificado em nível superior para fortalecer o sistema nacional de educação, contribuindo para a melhoria da educação básica e para o fortalecimento e o crescimento da ciência, da tecnologia e da inovação, visando ao desenvolvimento sustentável do Brasil.

Metas 2012-2015: elevar o número de mestres e doutores titulados por ano, de forma a alcançar a meta do PNE 2011-2020; elevar o percentual de mestres e doutores no corpo docente em efetivo exercício nas instituições de educação superior, de forma a alcançar a meta do PNE 2011-2020.

Iniciativa 0390 - Concessão de bolsas, auxílios e outros mecanismos, no país e no exterior, para a formação, valorização e capacitação de recursos humanos e para promover cooperação internacional no Sistema Nacional de Pós-Graduação, em áreas de interesse nacional e regional, inclusive formação de professores, garantida equidade étnico-racial e de gênero". (BRASIL, 2012a)

No debate acerca das políticas públicas educacionais implementadas pela CAPES, alguns assuntos relacionados aos programas de concessão de bolsas têm sido discutidos na comunidade acadêmica, a saber: acompanhamento e monitoramento dos programas; possibilidades legais e vedações para o acúmulo de bolsas e para o acúmulo com vínculo empregatício; aprimoramento de metas e de indicadores de desempenho dos programas; custo-benefício da concessão de bolsa no exterior em relação à concessão de bolsa de mesma modalidade no país; entre outros tópicos avaliados pelos órgãos de controle.

O presente estudo pretende contribuir com análises sobre o custo-benefício da concessão de bolsas de doutorado pleno no exterior em relação à concessão de bolsas de doutorado pleno no país, e assim, com discussões sobre o aperfeiçoamento

das práticas de monitoramento e avaliação dos programas de concessão de bolsas no âmbito da CAPES.

Além disso, essa pesquisa poderá trazer benefícios a outras entidades e outros órgãos públicos, federais e estaduais, que também fomentam a pós-graduação brasileira por meio da concessão de bolsas de estudo, como o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), as Fundações de Apoio à Pesquisa (FAPs) e o Ministério das Relações Exteriores (MRE), na medida em que o objeto desse estudo poderá ser aproveitado para a revisão de seus próprios controles internos.

2 REFERENCIAL TEÓRICO

2.1 Políticas Públicas – conceito, funções e ciclo

Para uma melhor compreensão sobre os conceitos, as definições, e tudo o que se refere ao monitoramento e à avaliação de políticas públicas, faz-se necessário, primeiramente, entender o que são políticas públicas, como e para que são elaboradas e as etapas que compõem seu ciclo de vida, sendo a avaliação um dos estágios deste ciclo.

As políticas públicas são objetos de estudo de várias áreas do conhecimento, como Ciência Política, Administração, Sociologia, Economia, Direito, Educação e outras. Não há definição única sobre o termo política pública (ALMEIDA; GOMES, 2018; SOUZA, 2006), são diversas as definições que perpassam pelo campo da ciência política, com as definições mais clássicas, até os conceitos mais atuais.

Lynn e Gould (1980) define políticas públicas como um conjunto específico de ações do governo que produzirão efeitos específicos. Na concepção de Silva (2014), política pública é uma forma de regulação ou intervenção na sociedade, articulando diferentes sujeitos que apresentam interesses e expectativas diversas. A política pública “constitui um conjunto de ações e omissões do Estado decorrente de decisões e não decisão, constituída por jogo de interesses, tendo como limites e condicionamentos os processos econômicos, políticos e sociais.” (SILVA, 2014, p. 37).

Depreende-se do conceito apresentado por Silva (2014) que as políticas públicas são uma forma de interferência ou não interferência, ou uma regulação, elaboradas com a função de atender a uma demanda ou necessidade da sociedade, para solucionar um problema, ou corrigir assimetrias, e que envolve diversos sujeitos, tais como governantes, gestores, agentes públicos, comunidade, tendo como princípio norteador a busca pela equidade social. São as políticas públicas que dão visibilidade e materialidade ao Estado e, por isso, são definidas como “o Estado em ação” (AZEVEDO, 2004, p. 60).

Segundo Frey (2000), as políticas públicas podem ser classificadas em quatro tipos: distributiva, redistributiva, regulatória e constitutiva. As políticas públicas distributivas têm a função de distribuir bens, serviços ou recursos a um grupo específico da população, por exemplo, a política de cotas na educação. As políticas públicas redistributivas se diferenciam das distributivas por retirarem recursos financeiros do orçamento de todos para redistribuir bens, serviços ou quantias a um grupo específico da população, como ocorre no sistema previdenciário.

Já as políticas públicas regulatórias são aquelas que estipulam regras para

padrões de comportamento, e apresentam-se na forma de normativos, ou seja, as portarias, os decretos, as leis, como o Código de Trânsito Brasileiro, a título de exemplo. Por fim, as políticas públicas constitutivas ocupam-se dos procedimentos e estabelecem as regras gerais para as negociações das políticas públicas distributiva, redistributiva e regulatória, a exemplo da distribuição de competências aos diversos entes da federação constante da Constituição Federal de 1988.

Importante ressaltar que nem sempre as chamadas “políticas governamentais” são públicas. É preciso considerar a quem se destinam os resultados, os benefícios da política, bem como se o seu processo de formulação foi submetido ao debate público, para ser considerada política pública.

As políticas públicas são executadas por meio de ações e programas que precisam ser estruturados em uma sequência de diferentes fases que se correlacionam. Segundo Rua (2009, p. 37), o ciclo político pode ser entendido como uma “abordagem para o estudo das políticas públicas que identifica fases sequenciais e interativas-iterativas no processo de produção de uma política”. Conforme descreve a autora, são estágios deste ciclo:

- I. **Formação da agenda**, momento em que são estudadas as questões públicas advindas das demandas, necessidades e dos problemas sociais, incluindo as situações emergenciais, e definidas as prioridades em função dos recursos disponíveis, da urgência e relevância do problema, e de outras situações circunstanciais;
- II. **Formulação da política**: fase em que são apresentadas as possíveis alternativas, soluções, propostas para sanar as questões já estabelecidas na agenda. Nesta etapa acontece o desenho de programas, que não são determinados apenas pelos conhecimentos técnicos de especialistas, mas também pelo poder político dos setores envolvidos, o que pode expressar interesses diversos que precisam coadunar para o amparo do maior número de beneficiários possível;
- III. **Processo de tomada de decisão**, o qual se refere à decisão pela alternativa mais adequada para a resolução da questão pública. É nesta fase que são definidos o prazo, os custos e o público-alvo das ações políticas, e que as decisões são formalizadas transformando-se em leis e programas. Para Peters (1995), a legitimidade pode talvez ser conquistada, em parte, por meio da eficácia, e se processos amplamente burocráticos de tomada de decisão forem capazes de produzir resultados valorizados pela população, então as instituições provavelmente serão aceitas como tomadores de decisão apropriados;
- IV. **Implementação da política**, que ocorre quando a política pública começa a ser executada, de maneira que as decisões deixam de ser meramente formais

e se materializam em ações, intervindo na realidade social. Para Rua (2009), a implementação vem acompanhada do monitoramento, que pode ser entendido como um conjunto de atividades sistematizadas de acompanhamento dos processos adotados, dos resultados preliminares e intermediários obtidos e do comportamento do ambiente da política, visando facilitar a conquista dos objetivos pretendidos com a política em implementação;

- V. **Avaliação**: compreende a definição de critérios, indicadores e padrões para medir o desempenho da política pública, a fim de subsidiar a análise de resultados e impactos gerados, e assim, a decisão dos gestores sobre a continuidade ou descontinuidade da política, ou ainda a sua reformulação ou seu redirecionamento. Na visão de Frey (2000, p. 29), a etapa da avaliação é “imprescindível para o desenvolvimento e a adaptação contínua das formas e instrumentos de ação pública”.

2.2 Monitoramento e Avaliação de Políticas Públicas

Segundo o Guia prático de análise *ex post* – Avaliação de Políticas Públicas (2018b), da Presidência da República, “a verificação da efetividade de uma política passa, necessariamente, por seu monitoramento e sua avaliação”. Para Rua (2009), a avaliação assumiu o status de instrumento estratégico em todo o ciclo da política pública a partir de meados da década de 1980, no cenário da reforma da administração pública e do processo de modificações das relações entre o Estado e a sociedade, em que o foco nos processos foi substituído pelo foco nos resultados. A reforma do Estado do final dos anos 70 e início dos anos 80 ocorreu em um panorama de crise, segundo Abrucio (1998, p. 178), apresentando características de “escassez de recursos públicos, enfraquecimento do poder estatal e de avanço de uma ideologia privatizante”.

Na visão de Sano (2003, p. 40), a mudança de foco nos processos para o resultado proporciona uma maior flexibilidade na gestão das ações, na medida em que o foco nos resultados “introduz a questão das metas a serem atingidas por cada provedor de bem público, passando este fator a ser o balizador da avaliação”. Dessa forma, é preciso aferir também, de acordo com Guerra et al. (2012), em que proporção os objetivos elencados em documentos do modelo orçamentário brasileiro, tais como o Plano Plurianual, a Lei de Diretrizes Orçamentárias e os Planos Diretores, estão sendo alcançados, a fim de possibilitar a mensuração do nível de eficiência de um gestor público. Também para as autoras Grau e Bozzi (2008), as funções de monitoramento e de avaliação ganharam relevância nessas mesmas circunstâncias, do surgimento de uma gestão pública orientada para resultados, com destaque para uma atuação mais efetiva do Estado.

Soares e Cunha (2015, p. 17) indicam quatro motivos para monitorar e avaliar políticas públicas:

"(1) Permitir à sociedade conhecer e participar do andamento de programas e projetos, assim como verificar os resultados que eles têm alcançado, o que possibilita maior controle social; (2) Prover informações estratégicas para os gestores governamentais fazerem correções e aperfeiçoarem as políticas; (3) Possibilitar aos órgãos de controle, como tribunais de contas, controladorias, auditorias e outros mais, analisar o emprego adequado dos recursos públicos; (4) Atingir resultados mais positivos na elaboração e execução das políticas públicas, com melhoria nas condições de vida dos cidadãos."

Os processos de acompanhamento e monitoramento de programas e ações têm por objetivo a harmonização entre aquilo que foi planejado e o que está sendo executado. Trata-se de uma espécie de avaliação de eficiência, sendo possível interferir na ação enquanto esta se encontra em andamento, buscando corrigir rumos ao passo em que os desvios são identificados, e que ocorre durante a fase de implementação da política pública. De acordo com Faria (1998, p. 45), corrigir rumos, neste contexto, "significa (ou pode significar) modificar cenários jurídico-administrativo, financeiro, organizacional, bem como requalificar recursos humanos a cargo dos trabalhos."

Para Garcia (2000), o monitoramento é um processo ordenado e contínuo, capaz de prover informações sintéticas e tempestivas, que viabiliza uma célere avaliação situacional e permite a adoção de medidas confirmativas ou corretivas das ações monitoradas. Segundo Rua (2009, p. 38), o monitoramento "é um instrumento de gestão das políticas públicas e o seu objetivo é facilitar a consecução dos objetivos pretendidos com a política".

Assim, o monitoramento deve ocorrer concomitante e continuamente com a execução das ações, com o objetivo maior de garantir que os objetivos traçados no planejamento sejam atingidos, mas também com a intenção de monitorar a evolução da execução, os insumos e bens envolvidos, o tempo gasto, enfim, toda a atividade que está em andamento, podendo, inclusive, alterar o planejamento inicial.

Já o acompanhamento, na visão de Rua (2009), é caracterizado pelo desenvolvimento de atividades de supervisão e fiscalização, sem natureza interativa, em que os dados obtidos são rasos, limitando-se a serem registrados e informados, e seus resultados dificilmente são compartilhados e escassamente são utilizados como ferramenta de gestão.

Grau e Bozzi (2008) salientam que o monitoramento e a avaliação devem ser tratados como instrumentos diferentes, apesar de serem muito próximos. Enquanto o monitoramento utiliza indicadores de desempenho pré-estabelecidos, que podem ser quantitativos ou qualitativos, a avaliação é específica, e visa medir o grau de confiabilidade do processo, o resultado e os efeitos causados pela implementação da

política pública. Cohen e Franco (2004) destacam que o monitoramento é uma atividade de gerência que ocorre durante uma fase da execução, e a avaliação pode ser realizada a qualquer tempo, isto é, antes, durante ou algum tempo depois da implementação de um projeto.

Para Rua (2009), o monitoramento e a avaliação têm por objetivo auxiliar os gestores da política pública na tomada de decisões, fornecendo informações referentes aos arranjos necessários para que a política alcance os resultados esperados. Existe um vínculo entre o monitoramento e a avaliação segundo Jannuzzi (2005), que se manifesta no decorrer do tempo, visando fornecer informações relativas à forma de implementação dos programas (tarefa do monitoramento) e dos resultados e desejos esperados (tarefa da avaliação).

Segundo Ala-Harja e Helgason (2000) a avaliação pode ser de diversos tipos, tais como avaliação de programas, avaliação de desempenho organizacional, avaliações gerenciais e auditorias de desempenho. Weiss (1998) define a avaliação como a atividade que busca medir as operações e os resultados de uma política ou de um programa, fazendo uma comparação entre critérios pré-estabelecidos, sejam eles explícitos ou implícitos, a fim de contribuir para o aperfeiçoamento da política ou do programa avaliado.

Dessa forma, pode-se entender que a avaliação é uma ferramenta utilizada para melhorar a ação do governo, na medida em que serve como base para a tomada de decisões, referente à continuidade ou ao cancelamento de determinada ação ou programa de governo, bem como para avaliar o alcance dos objetivos, e qual o real impacto na sociedade, se a ação contribuiu, de fato, para solucionar o problema e atender a demanda da sociedade.

A avaliação de políticas públicas consiste em uma avaliação formal, a qual "permite julgar processos, produtos, efeitos ou impactos de políticas, programas ou projetos públicos." (RUA, 2009, p. 110). Esta autora discorre que, a avaliação perpassa pelos seguintes estágios: o primeiro, com foco na mensuração dos fatos analisados, o próximo estágio, referente às formas de alcançar os resultados, progredindo para um julgamento das interferências quanto à sua eficácia, eficiência, efetividade, sustentabilidade e outros pontos, como por exemplo, a equidade.

Segundo a literatura, as três dimensões do desempenho (3E's) são:

- I. Eficácia: quando uma atividade atinge o objetivo ou a meta. Para Sander (1982, p.12), "etimologicamente, eficácia significa capacidade ou potencialidade para alcançar resultados almejados";
- II. Eficiência: fazer as coisas da maneira certa, mais adequada, buscando a otimiz-

ção de resultados. De acordo com Sander (1982, p. 11), a eficiência “é o critério administrativo que revela a capacidade real de produzir o máximo com o mínimo de recursos, energia e tempo”; e

- III. Efetividade: refere-se aos benefícios advindos à organização das ações certas feitas da maneira certa. É o impacto causado em relação ao objetivo. Segundo Sander (1982, p. 13), “efetividade é o critério de desempenho que mede a capacidade de produzir a solução ou resposta desejada pelos participantes da comunidade”.

Rua (2009) define as três dimensões do desempenho conforme descrito no Quadro 1 abaixo:

Quadro 1 – Definição de Eficácia, Eficiência e Efetividade

Dimensões de desempenho	Definição
Eficácia	Relaciona atividades com seus produtos iniciais, intermediários e finais (metas e objetivos).
Eficiência	Relaciona produtos com seus custos (financeiros, humanos, tempo).
Efetividade	Relaciona produtos com seus efeitos na realidade que se quer transformar, ou seja, consequências.

Fonte: Adaptado de Rua (2009)

Nesse contexto, a avaliação não é uma ação isolada, mas uma das fases do processo de planejamento das políticas - materializadas em programas e projetos. Assim, a avaliação também contribui para o aperfeiçoamento das ações na esfera operacional, podendo indicar recomendações para redução de custos, melhor aplicação dos insumos e de pessoal, melhoramento dos processos envolvidos, providências relativas ao controle das ações, a fim de evitar fraudes e desperdícios.

A avaliação pode ainda auxiliar os cidadãos a conhecer o que está sendo feito com os recursos financeiros públicos e a considerar alternativas às políticas implementadas (ALA-HARJA; HELGASON, 2000). Para os autores, questões relativas à legitimidade do setor público e à necessidade de fortalecimento da responsabilidade do governo pedem por mecanismos de controle novos e mais sofisticados, como por exemplo a avaliação. Assim, o problema do desempenho e da *accountability* podem ser consideradas motivações para se realizarem estudos de avaliação.

As democracias têm lidado com o aumento da busca por *accountability*, na visão de Filgueiras (2018). Em sociedades em que existam altos níveis de confiança e satisfação, provavelmente ocorre uma passividade popular que gere baixas demandas de *accountability*. Por sua vez, a desconfiança da cidadania e a insatisfação da população com os governos estimulam processos de mudança e criam demandas

por *accountability*, sobretudo em contextos de alto nível de midiaticização da política. Segundo o autor, o Brasil é o país menos satisfeito com o regime democrático, com relação à satisfação com a democracia. Essa insatisfação dos brasileiros decorre de problemas estruturais relacionados à baixa qualidade dos serviços públicos, a uma forte percepção de corrupção, de não equidade no tratamento e de existência de privilégios, bem como a uma série de particularismos que ainda estruturam a relação da sociedade brasileira com o Estado. Nesse cenário, as demandas por *accountability* crescem.

As rotinas de *accountability* podem ser instrumento político poderoso, declara Simon (1957). O controle da administração pública brasileira é feito no contexto de procedimentos administrativos e criminais, que visam à correção da aplicação de recursos públicos, e, também, à correção quanto aos princípios da administração pública inscritos no art. 37 da CF/88. Para Balbe (2013), a institucionalização do controle interno permitiu maior capacidade de fiscalização, controle e monitoramento das políticas públicas, de forma a fazer avançar o processo de *accountability*.

2.3 Indicadores sociais para as atividades de monitoramento e avaliação

Para Jannuzzi (2005), o interesse por assuntos relacionados aos indicadores sociais e sua aplicação nas operações relativas ao planejamento governamental e ao ciclo de formulação e avaliação das políticas públicas têm aumentado no Brasil, em todas as esferas de governo e nos fóruns de discussão do tema. São várias as circunstâncias que justificam esse crescente interesse no uso de indicadores na administração pública, tais como: mudanças institucionais vivenciadas pela administração pública em todo o país, com ênfase na fixação da sistemática do planejamento plurianual, no aprimoramento dos controles internos, com a alteração do foco das auditorias de avaliação de conformidade legal para a avaliação de desempenho dos programas executadas pelos Tribunais de Contas, na reforma gerencial da gestão pública em torno dos anos 1990 (GARCIA, 2000; COSTA; CASTANHAR, 2003).

Jannuzzi menciona, ainda, o aperfeiçoamento do controle social do Estado brasileiro nos últimos 20 anos. Os gastos públicos passaram a ser mais fiscalizados pela sociedade, mídia e sindicatos, os quais começaram a exigir o uso mais eficiente, eficaz e efetivo dos recursos públicos, o que ocasionou a reorganização das atividades relativas ao planejamento em bases mais técnicas. Para Peters (1995), não se pode negar que a gestão no setor público é uma tarefa difícil e que pode ocupar o tempo e a energia dos servidores públicos mais experientes. Ademais, as demandas por maior economia e eficiência no governo colocaram ainda mais pressão sobre os funcionários públicos para administrar suas organizações da forma mais eficaz possível.

Outro fator que tem contribuído para a divulgação do uso de indicadores é o

acesso mais fácil às informações mais elaboradas, sejam elas de natureza administrativa e estatística, oportunizado pelas novas tecnologias de informação e comunicação. Nesta ótica, Jannuzzi (2001, p. 35) ressalta que “os indicadores sociais são instrumentos para o efetivo empoderamento da sociedade civil, de controle e direcionamento das atividades do poder público”. Instrumentos importantes para o monitoramento e a avaliação, Pfeiffer (2000, p. 157) frisa que os indicadores têm a missão de “caracterizar mais detalhadamente os objetivos e resultados, estabelecendo o que e quanto se pretende alcançar, e fornecem uma base para o acompanhamento e a avaliação do planejado”.

A definição do termo indicador pode ser baseada sob diferentes pontos de vistas, segundo Minayo (2009, p. 84). De forma geral, os pesquisadores, por exemplo, definem indicadores como:

"parâmetros quantificados ou qualitativos que servem para detalhar se os objetivos de uma proposta estão sendo bem conduzidos (avaliação de processo) ou foram alcançados (avaliação de resultados). Como uma espécie de sinalizadores da realidade, a maioria dos indicadores dá ênfase ao sentido de medida e balizamento de processos de construção da realidade ou de elaboração de investigações avaliativas".

A autora acrescenta que as instituições públicas utilizam indicadores principalmente sob o sentido de especificações quantitativas e qualitativas para aferir o alcance de objetivos, metas e resultados pré-estabelecidos. No aspecto de sua utilidade, além de ser um instrumento para medir, para estabelecer parâmetros e para avaliar, os indicadores são importantes ferramentas de gestão, uma vez que permitem ao gestor atuar sobre dimensões-chave de sistemas e de processos, monitorando acontecimentos que devem ser alterados, incentivados ou potencializados, a contar do início de uma intervenção até a conquista do resultado planejado.

No âmbito das políticas públicas, Jannuzzi (2005) considera que os indicadores sociais são medidas que permitem operacionalizar um conceito abstrato ou uma demanda de interesse programático. De acordo com o autor, “Os indicadores apontam, indicam, aproximam, traduzem em termos operacionais as dimensões sociais de interesse definidas a partir de escolhas teóricas ou políticas realizadas anteriormente.” (JANNUZZI, 2005, p. 138). Eles servem para auxiliar nas atividades de planejamento e de formulação de políticas sociais.

Segundo o Guia prático de análise *ex ante* – Avaliação de Políticas Públicas (2018a), da Presidência da República, os indicadores são informações que viabilizam desde o acompanhamento do andamento até a medição do cumprimento dos objetivos de uma política. Ou seja, eles são imprescindíveis em todas as fases do processo de formulação e implementação das políticas públicas.

"Eles se referem à quantidade, à qualidade, ao cronograma e aos custos

observados comparativamente. Uma vez fixados durante a elaboração ex ante, por meio de um modelo lógico, os indicadores definem quais dados devem ser coletados para que se possa realizar o acompanhamento da implementação e a avaliação sobre a obtenção ou não dos resultados esperados" (BRASIL, 2018a, p. 101).

Para Kusek e Rist (2004), o gestor pode descobrir, por meio dos indicadores, se os projetos, programas ou políticas estão na direção correta ou se é necessário promover reformulações nas metas planejadas. Na visão dos autores, porém, os indicadores são relevantes somente quando são medidos em relação a um objetivo estabelecido, isto é, os objetivos e as metas a serem atingidos devem ser muito bem definidos no procedimento de definição de indicadores.

Neste cenário, Cohen e Franco (2004) propõem diferenciação entre os termos objetivo, meta e indicador. Consoante os autores, o objetivo é a condição que se pretende alcançar. Já a meta é a quantificação do objetivo, isto é, a fase desejável a ser alcançada em um período estabelecido. Por fim, um indicador é a medida que propicia a verificação do nível de alcance do objetivo.

Jannuzzi (2005) sugere doze propriedades desejáveis na construção de indicadores, quais sejam: relevância para a agenda política, validade de representação do conceito, confiabilidade da medida, cobertura populacional, sensibilidade às ações previstas, especificidade ao programa, transparência metodológica na sua construção, comunicabilidade ao público, factibilidade operacional para sua obtenção, periodicidade na sua atualização, desagregabilidade populacional e territorial, comparabilidade da série histórica.

Considerando os estudos de Soares e Cunha (2015), os indicadores podem ser classificados conforme descrito no seguinte rol exemplificativo:

- I. **Tangíveis** (de fácil observação e aferição) ou **Intangíveis** (necessitam de outros indicadores para captarem questões complexas), segundo a capacidade de perceber questões objetivas ou subjetivas da realidade;
- II. **Quantitativos** (retratados de forma numérica) ou **Qualitativos** (em que há a atribuição de valor não numérico, tais como “sim”, “não”, “bom”, “ruim”, “atende”, “não atende”, etc.);
- III. Quanto à perspectiva da gestão de determinado programa, podem ser: **Indicadores de Estrutura**, que se ocupam em apurar se as estruturas física, de pessoal, de equipamentos, etc. são adequadas aos objetivos do programa; **Indicadores de Processos**, os quais verificam se o andamento do programa social está se desenvolvendo de acordo com o que foi planejado; **Indicadores de Resultado**, que visam medir efeitos, impactos e produtos das ações desenvolvidas, possuindo

relação direta com objetivos e diretrizes do programa social, e **Indicadores de Cobertura**, que estão atrelados aos beneficiários, e possuem o objetivo de verificar se o programa está realmente atendendo quem deveria ser o seu destinatário.

Batista e Popinigis (1999) por sua vez caracteriza os indicadores da seguinte forma:

- I. **Indicadores Estratégicos**: permitem acompanhar o desempenho da organização, mensurando o grau de aproximação do que está sendo executado com a consecução de sua Visão;
- II. **Indicadores de Produtividade (Eficiência)**: utilizados para medir a relação entre os recursos despendidos e os produtos/ serviços gerados;
- III. **Indicadores de Qualidade (Eficácia)**: concentram-se na medida de satisfação dos usuários e nas características dos produtos/serviços;
- IV. **Indicadores de Efetividade (Impacto)**: têm como foco os efeitos dos produtos/-serviços; e
- V. **Indicadores de Capacidade**: aferem a capacidade de resposta de um processo por meio da relação entre os produtos/serviços gerados por unidade de tempo.

Cada fase do ciclo de formulação e avaliação de políticas públicas ou de programas sociais abrange o uso de um conjunto de indicadores de diversas naturezas e propriedades, em conformidade com as necessidades peculiares das atividades envolvidas em cada fase. Por exemplo, no estágio de implementação e execução dos programas, "são necessários indicadores de monitoramento, que devem primar pela sensibilidade, especificidade e, sobretudo, pela periodicidade com que estão disponíveis" (JANNUZZI, 2005, p. 152).

Para Cohen e Franco (2004), monitoramento e avaliação de programas são expressões utilizadas para denominar procedimentos de acompanhamento de programas, com foco na análise da sua eficiência, eficácia e efetividade. Monitoramento e avaliação são processos minuciosos, sistematizados, contínuos, com vistas a auxiliar o gestor público fornecendo informações sobre o ritmo e a forma de implementação dos programas (indicadores de monitoramento) e dos resultados e efeitos desejados (indicadores de avaliação).

3 METODOLOGIA

A pesquisa é do tipo exploratória com abordagem quantitativa. No que se refere aos procedimentos técnicos, foram utilizadas pesquisa bibliográfica em literatura nacional e internacional sobre políticas públicas, monitoramento e avaliação no setor público e pesquisa documental, em que foi feita a análise de regulamentos, manuais, relatórios de gestão, portarias, e outros documentos da CAPES, dos dados divulgados na página eletrônica desta fundação pública e na Plataforma Sucupira (ferramenta que coleta dados relativos ao Sistema Nacional de Pós-Graduação - SNPG) e no portal operacional e de gestão do MEC, qual seja o Sistema Integrado de Monitoramento, Execução e Controle do Ministério da Educação (SIMEC), bem como de relatórios de auditoria de contas e outros pareceres exarados pela CGU e pelo TCU (Tribunal de Contas da União).

Os dados sobre as bolsas de doutorado pleno no país e no exterior concedidas pela CAPES foram coletados na plataforma Sucupira, que é uma ferramenta utilizada para coletar informações, realizar análises e avaliações, disponibilizando em tempo real e com mais transparência as informações, os processos e procedimentos que a CAPES realiza no Sistema Nacional de Pós-Graduação para toda a comunidade acadêmica, em comunicação com o Sistema de Conciliação e Pagamento de Bolsas (SCPB), sendo ambos os sistemas mantidos pela CAPES.

Para efetuar as análises estatísticas da proposição do trabalho, fez-se necessária a análise objetiva dos currículos dos bolsistas. Por isso, foram coletadas também informações do currículo *lattes* dos bolsistas, com a elaboração de *script Bash* para baixar todos os nomes e números de identificação (IDs) dos currículos na plataforma do CNPq, e com a posterior associação do ID *lattes* para cada nome dos bolsistas listados na extração da CAPES. Após o *download* do XML dos currículos utilizando rotina *Python*, foram extraídas as informações dos currículos consideradas relevantes para o estudo, por exemplo, o número de artigos publicados, que foi considerado como a variável dependente nesta pesquisa, na intenção de verificar se o valor da bolsa concedido é capaz de explicar a quantidade de artigos publicados.

Foi feita uma análise estatística descritiva dos dados, mostrada em quadros e gráficos, e uma investigação de comparação entre os dois grupos de bolsistas, quais sejam no país e no exterior, relacionando variáveis, a fim de averiguar a relação de custo-benefício das bolsas de doutorado pleno no exterior comparada às bolsas de doutorado pleno no país, utilizando medidas de posição e dispersão de acordo com Morettin e Bussab (2017):

I. Média:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (3.1)$$

II. Variância:

$$S^2 = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n - 1} \quad (3.2)$$

III. Desvio-padrão:

$$S = \sqrt{\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n - 1}} \quad (3.3)$$

IV. Mediana: é calculada selecionando a posição ordenada do conjunto de dados. Essa posição é dada por:

$$Med = \frac{n + 1}{2} \quad (3.4)$$

em que n é o tamanho da amostra. Caso o resultado seja um valor com decimal, deve ser feita a média dos valores entre as duas posições em que o resultado se encontra. Por exemplo, se a mediana resultar na posição 3,5 deve-se calcular a média das observações nas posições 3 e 4.

Além dessas medidas, também foi utilizado o coeficiente de variação (CV), dado por:

$$CV = \frac{S}{\bar{X}} * 100 \quad (3.5)$$

Essa métrica fornece uma visão sobre a distribuição dos dados. Quanto maior o CV, mais dispersos os dados são – ou seja, mais distantes da média eles estão. Um CV de até 10% é considerado homogêneo, com dados que são próximos da média. Um CV entre 10% e 20% é considerado heterogêneo, mas razoável. Valores de CV acima de 20% é bastante heterogêneo, com dados muito dispersos em relação à média.

3.1 Extração de dados do sistema da CAPES

A plataforma Sucupira é uma ferramenta utilizada para coletar informações, realizar análises e avaliações, disponibilizando em tempo real e com mais transparência as informações, os processos e procedimentos que a CAPES realiza no Sistema Nacional de Pós-Graduação para toda a comunidade acadêmica. Propicia, ainda, a parte gerencial-operacional de todos os processos e permite maior participação das pró-reitorias e dos coordenadores de Programas de Pós-Graduação, proporcionando, de forma geral, maior transparência e confiabilidade dos dados fornecidos pelos Programas.

Com acesso restrito aos Pró-Reitores, Coordenadores de PPGs, Coordenadores de área e gestores da Diretoria de Avaliação da CAPES, esse sistema fornece informações diversas sobre a pós-graduação brasileira, tais como o processo de avaliação dos cursos de mestrado e doutorado, dos cursos avaliados e reconhecidos pela CAPES, bem como de suas notas, dos projetos de cooperação entre instituições, o cadastro de discentes e docentes das Instituições de Ensino Superior, a classificação Qualis de um periódico, o cadastro de discentes com bolsa da CAPES, do CNPq ou de Fundações de Apoio à Pesquisa, a avaliação e as notas (Conceito CAPES) dos cursos de mestrado e doutorado, o catálogo de trabalhos de conclusão de curso, entre outras.

A alimentação da plataforma Sucupira é feita, em grande parte manualmente, por cada Programa de Pós-Graduação, responsável pelas informações internas do PPG, o que exige bastante atenção daqueles que fazem o *input* dos dados, a fim de evitar o preenchimento incompleto, equivocado, repetido ou a falta do preenchimento de campos dos formulários. Porém há a integração com outros sistemas, como a plataforma *Lattes* e o Sistema de Controle de Bolsas e Auxílios (SCBA) da CAPES para parte das informações requeridas pela ferramenta.

A extração dos dados abrangeu as bolsas com início da vigência nos anos de 2010, 2011, 2012, 2013 e 2014. Considerando que o prazo regular médio de duração do curso de doutorado é de 48 meses, que nem todos os bolsistas receberam financiamento para todo o período do curso de doutorado e que é possível ao bolsista, de forma excepcional, solicitar e receber extensão da concessão da bolsa por tempo superior aos 48 meses, o término da vigência das bolsas variou entre os anos de 2011 a 2019.

A escolha do início da vigência das bolsas nos anos de 2010 a 2014 foi devida ao prazo existente entre o ano de titulação até o ano de 2020 para a produção de trabalhos acadêmicos resultantes do doutorado. Acredita-se que o prazo médio de 4 anos desde o término do doutorado é um período razoável para a publicação de artigos científicos.

Foram extraídas as seguintes informações dos bolsistas: nome; número do Cadastro de Pessoa Física (CPF); e-mail; endereço do currículo *lattes*; nome do programa de bolsa CAPES; nível da bolsa; diretoria responsável pela bolsa; sistema de origem da bolsa; instituição de destino; país de destino; UF de destino; grande área; código da grande área; área do conhecimento; código da área do conhecimento; área de avaliação; código da área de avaliação; instituição de origem; UF de origem; código do programa SNPG; curso do programa; código do curso do programa; código do processo de bolsa; ano e mês do início da bolsa; ano e mês do término da bolsa; informação sobre titulação ou outra situação SNPG (titulado, desligado, abandono, em andamento, etc.); ano de titulação SNPG; valor financeiro da rubricas da bolsa (bolsa, mensalidade,

taxa escolar, auxílio deslocamento, auxílio seguro saúde, adicional localidade, auxílio instalação, mensalidade de licença maternidade, reembolsos, outros créditos, outros débitos, capital, adicional dependente, auxílio seguro saúde do dependente).

Vale destacar que um mesmo estudante pode ter recebido mais de uma bolsa de programas de concessão diferentes, em períodos diferentes, o que gerou dados para a ocorrência de até quatro bolsas CAPES para um mesmo bolsista.

Ressalta-se que, referente às bolsas no exterior, somente foram consideradas para o estudo aquelas que são financiadas com recursos apenas da CAPES, ou seja, foram excluídas as bolsas que são custeadas em parceria e/ou cooperação com entidades estrangeiras, devido à dificuldade de mensurar a contrapartida oferecida por essas instituições, e para que se compare apenas o montante de recursos públicos federais gasto na concessão das bolsas no país e no exterior.

3.2 Extração e normalização de dados do Currículo *Lattes*

Criada e mantida pelo CNPq, a plataforma *Lattes* é um repositório dinâmico de currículos em um único sistema de informações, que permite aos cidadãos incluir, atualizar, consultar, gerir de forma remota os dados de seu currículo. Foi idealizada para contribuir com as ações de planejamento, gestão e operacionalização de agências de fomento à pesquisa, federais e estaduais, de instituições de ensino e de institutos de pesquisa. A plataforma passou a ser fonte estratégica para as atividades de planejamento e gestão bem como para a formulação de políticas públicas do Ministério da Ciência, Tecnologia, Inovação e Comunicações e, também, para outros órgãos do governo dessas áreas (CNPQ, 2020).

O currículo *lattes* é atualmente o modelo nacional de registro de informações acadêmicas de estudantes e pesquisadores, e abrange a possibilidade de registro de dados profissionais e atividades complementares. Tornou-se componente indispensável para a análise de mérito de pedidos de financiamento nas áreas de ciência, tecnologia e outras.

Atualmente a plataforma *Lattes* retém o cadastro dos currículos dos membros alvo do estudo, com a ocorrência de cerca de 7 milhões de currículos cadastrados em seu banco de dados no mês de referência de dezembro de 2020. Desses, cerca de 54 mil é o alvo da pesquisa proposta.

O primeiro desafio encontrado para a realização desta pesquisa foi a obtenção dos currículos dos bolsistas em formato válido e estruturado, de maneira que fosse possível a montagem de um banco de dados pesquisável e indexado. Inicialmente foram consideradas as seguintes fontes de dados para a obtenção das informações

desejadas: a Capes e a plataforma pública *Lattes*/CNPq.

A Capes, enquanto entidade que concede e implementa as bolsas, e efetiva o pagamento dos valores diretamente aos beneficiários, possui o controle das bolsas concedidas, bem como dos dados pessoais dos bolsistas beneficiados, contudo, não detém informações consolidadas acerca da produção científica dos discentes, do endereço profissional, e mesmo da titulação, já que diversos dados devem ser informados pelos Programas de Pós-Graduação no sistema de coleta da CAPES.

O CNPq, como entidade pública responsável pela manutenção da plataforma *Lattes*, pode disponibilizar dados da base de currículos às instituições, desde que atendidos os critérios estabelecidos. Para as pessoas físicas, o acesso poderá ser concedido sob a responsabilidade da instituição à qual o cidadão esteja vinculado, após análise e aprovação do Formulário de Habilitação, Termo de Compromisso, e outros documentos pelo CNPq. Diante das dificuldades encontradas para solicitar o acesso diretamente ao CNPq, essa possibilidade de obtenção dos dados foi afastada.

Uma vez definida a fonte de dados dos currículos dos bolsistas, a coleta de dados deparou-se com um enorme obstáculo, qual seja a não possibilidade de acesso rápido ou de forma contínua aos cerca de 54 mil currículos objeto da pesquisa, tendo em vista que a plataforma não permite um *download* geral em massa dos dados. O procedimento necessário para a obtenção de um único currículo é o que segue:

- I. Acessar a plataforma no seguinte endereço: <<http://buscatextual.cnpq.br/buscatextual/busca.do?metodo=apresentar>>;
- II. Digitar o nome do bolsista no campo específico;
- III. Buscar o currículo;
- IV. Vencer o desafio de captcha implementado na plataforma, o qual valida o acesso humano;
- V. Abrir o currículo; e
- VI. Fazer o *download* do currículo em formato XML, pois foi observado que o XML fornecido pela plataforma contém os dados de forma estruturada, facilitando o trabalho necessário para normalizar os dados, caso eles fossem coletados diretamente de uma página HTML, por exemplo.

Ademais, para cada pesquisa feita, podem aparecer homônimos ou resultados duplicados, sendo necessário o discernimento para eliminar os resultados não desejados.

A execução da tarefa descrita acima pode demorar até um minuto, e dependendo do resultado, podem ser necessários vários minutos para um único bolsista. Fazendo-se um cálculo aproximado com 54.200 elementos e considerando o gasto de um minuto para cada busca, isso daria um esforço aproximado de 903 horas ou 37 dias, o que é humanamente inviável.

Desta forma, os esforços foram direcionados para a obtenção dos dados de forma automatizada, usando ferramentas de TI próprias para esse tipo de extração de informações.

A automatização fomentou o levantamento das seguintes necessidades macro para que o objetivo fosse atendido: acessar a plataforma *Lattes*; elencar os currículos dos bolsistas CAPES alvo da pesquisa; fazer o *download* dos currículos; extrair as informações de cada currículo; normalizar as informações obtidas de forma que pudessem ser aplicadas fórmulas estatísticas que apresentem a realidade do objetivo da pesquisa.

Para o atendimento das necessidades elencadas, foram construídas rotinas automatizadas que podem ser demonstradas, de forma simplificada, nas etapas ilustradas abaixo. Cabe ressaltar que essas codificações não são o objetivo do trabalho, portanto elas não foram executadas exatamente como descritas abaixo. Todo o processo, por exigir decisões subjetivas e apresentar erros aleatórios nas extrações, foi executado de forma assíncrona e os erros foram sendo tratados caso a caso, de modo a atingir o objetivo final que é a obtenção dos currículos *lattes* normalizados e indexados.

O processo de ETL (sigla em inglês de *Extract, Transform and Load* – Extração, Transformação e Carga) até a apresentação dos dados contou com as seguintes ferramentas de TI:

- I. Linguagem de programação *Bash*;
- II. Linguagem de programação *Python*;
- III. Base de dados MySQL;
- IV. Ferramenta *Adminer* para administrar a base de dados em MySQL;
- V. Ferramenta *Docker*, para executar o banco de dados localmente, bem como a ferramenta *Adminer*;
- VI. Excel e Arquivos CSV;
- VII. Ferramenta de BI *QlikSense*;
- VIII. *Software* estatístico e linguagem de programação R.

Para a melhor compreensão das ilustrações elaboradas, utiliza-se a seguinte legenda:

Figura 1 – Legenda dos ícones utilizados nas ilustrações



Fonte: Elaboração própria

O processo foi dividido nas seguintes etapas:

- I. Primeira etapa: foi elaborado um *script Bash* (Anexo B) para fazer o *download* de todos os nomes e ID (Número Identificador) dos currículos *lattes*. Utilizando a página de pesquisa da plataforma, a busca foi feita com o termo * (asterisco) nas bases de “Doutores” e “Demais pesquisadores”, a fim de que o sistema retornasse todas as ocorrências possíveis. O retorno dessa pesquisa foi uma tela paginada. A rotina leu cada uma das paginações disponibilizadas, extraiu e salvou o nome e o ID disponíveis no *link* de cada currículo em um arquivo texto, o que gerou uma saída em torno de 7 milhões de registros. O ID disponibilizado na página de pesquisa da plataforma constitui-se de um código alfanumérico iniciado pela letra K. Esse código é de extrema importância, pois ele é necessário para fazer o *download* do XML dos currículos, como descrito nas próximas etapas.

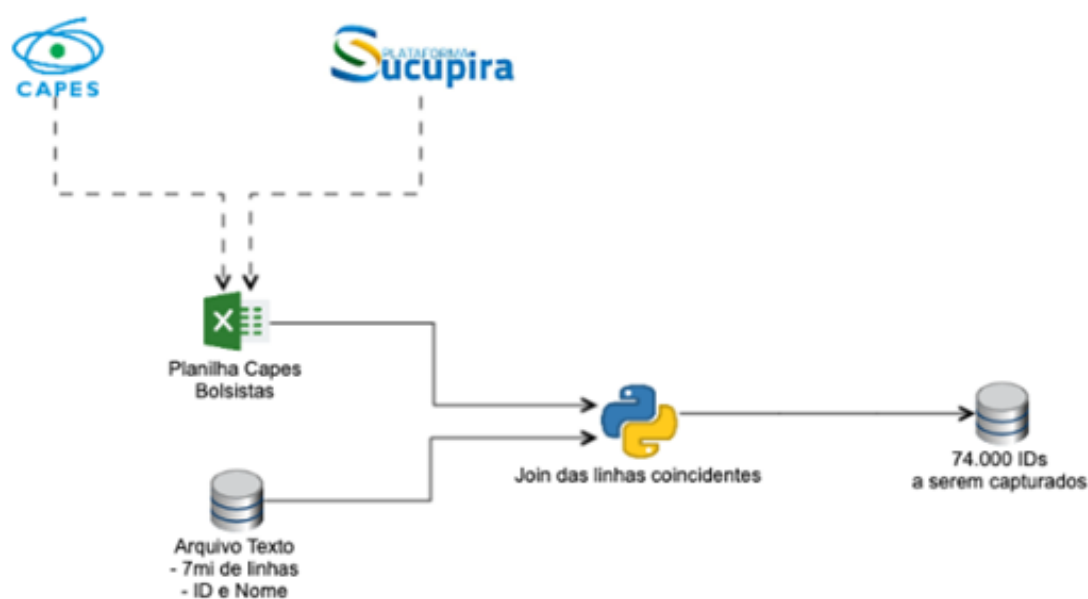
Figura 2 – Desenho da primeira etapa para a obtenção dos dados



Fonte: Elaboração própria

II. Segunda etapa: elaboração de rotina *Python* para fazer o cruzamento desses nomes com os cerca de 54 mil nomes alvos do estudo, relacionando para cada nome o ID correspondente. Para cada um dos 7 milhões de registros, foi verificado se há um nome correspondente entre os 54 mil bolsistas, o que gerou uma saída relacionando os nomes e IDs dos currículos para *download*. Essa lista conteve cerca de 74 mil nomes. O número diverge do domínio da pesquisa, pois nesta etapa foram levantados todos os nomes de possíveis bolsistas candidatos. Porém nesse universo coletado, existem milhares de homônimos que deverão ser devidamente excluídos do estudo nas etapas posteriores.

Figura 3 – Desenho da segunda etapa para a obtenção dos dados

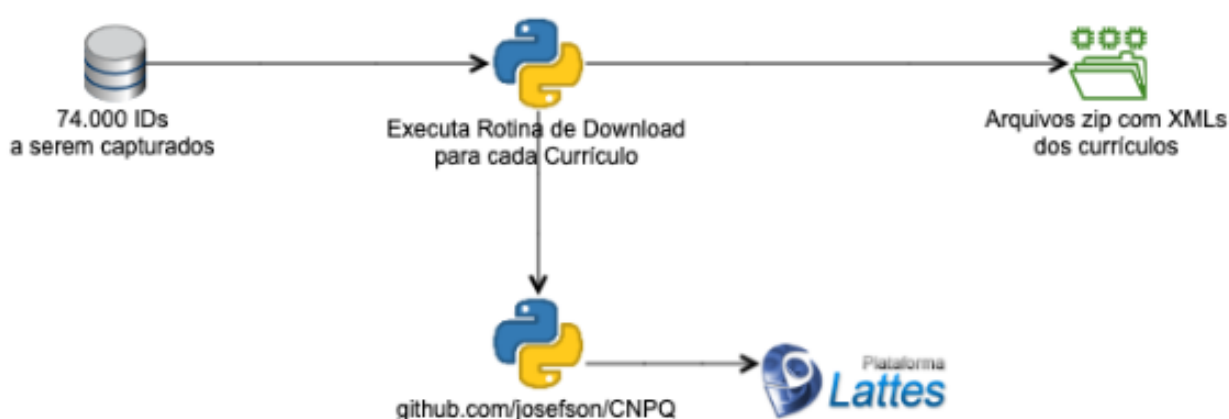


Fonte: Elaboração própria

III. Terceira etapa: elaboração de nova rotina *Python* que possibilita o *download*

do XML do currículo. Essa rotina utilizou uma biblioteca escrita também em linguagem *Python*, disponibilizada de forma pública (SOUZA, 2018), que recebe como parâmetro o ID do currículo extraído na etapa anterior e salva o resultado compactado na forma de um arquivo zip. Esse arquivo zip contém o XML estruturado do currículo do bolsista. A biblioteca destacada acima executa, de forma automatizada, o passo a passo humano necessário para baixar o currículo de um bolsista no formato XML. Assim, o produto desta etapa resultou em cerca de 74.000 XMLs salvos.

Figura 4 – Desenho da terceira etapa para a obtenção dos dados



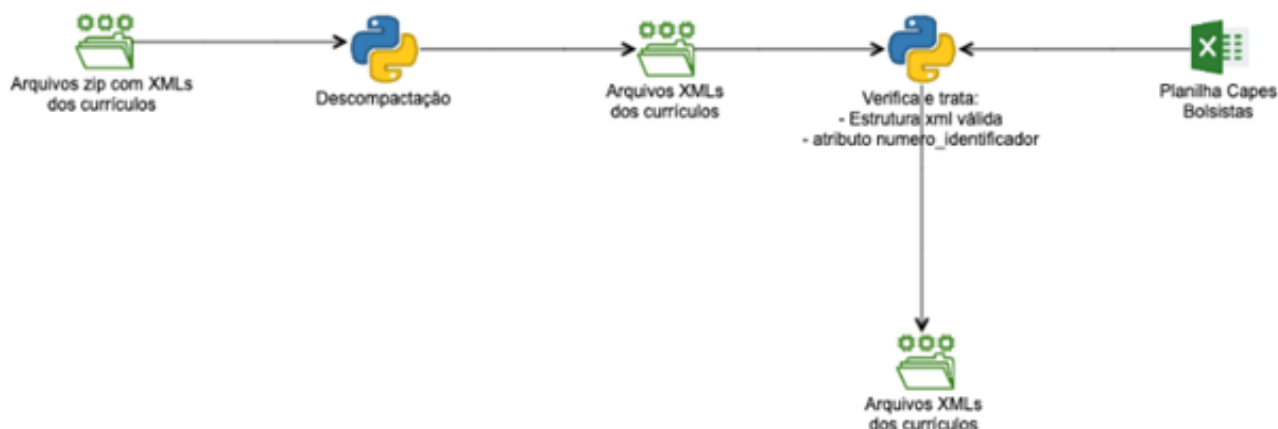
Fonte: Elaboração própria

- IV. Quarta etapa: foi feita a descompactação dos arquivos zip utilizando rotina *Python*, gerando um arquivo XML descompactado para cada bolsista. Caso o arquivo não aceitasse ser descompactado, então ele era dado como corrompido e feita nova captura.
- V. Quinta etapa: ainda utilizando rotina em *Python*, foi feita a validação e, também, a sanitização de cada XML como descrito abaixo:

a) Verificação da existência do atributo “nome do bolsista” no XML. Caso o atributo não tivesse sido localizado, o XML era dado como inválido e capturado novamente; e

b) Verificação da existência do atributo “número identificador”, que passa a ser o atributo chave de toda a base de dados dessa pesquisa. Caso o mesmo não exista, então o nome do arquivo zip (sem a extensão) é salvo nesse campo dentro do XML. Foi identificado que todos os *downloads* trazem o “número identificador” no nome do arquivo, mas nem sempre esse número identificador está informado no campo estruturado do XML, embora o XML esteja íntegro e com as informações corretas do currículo.

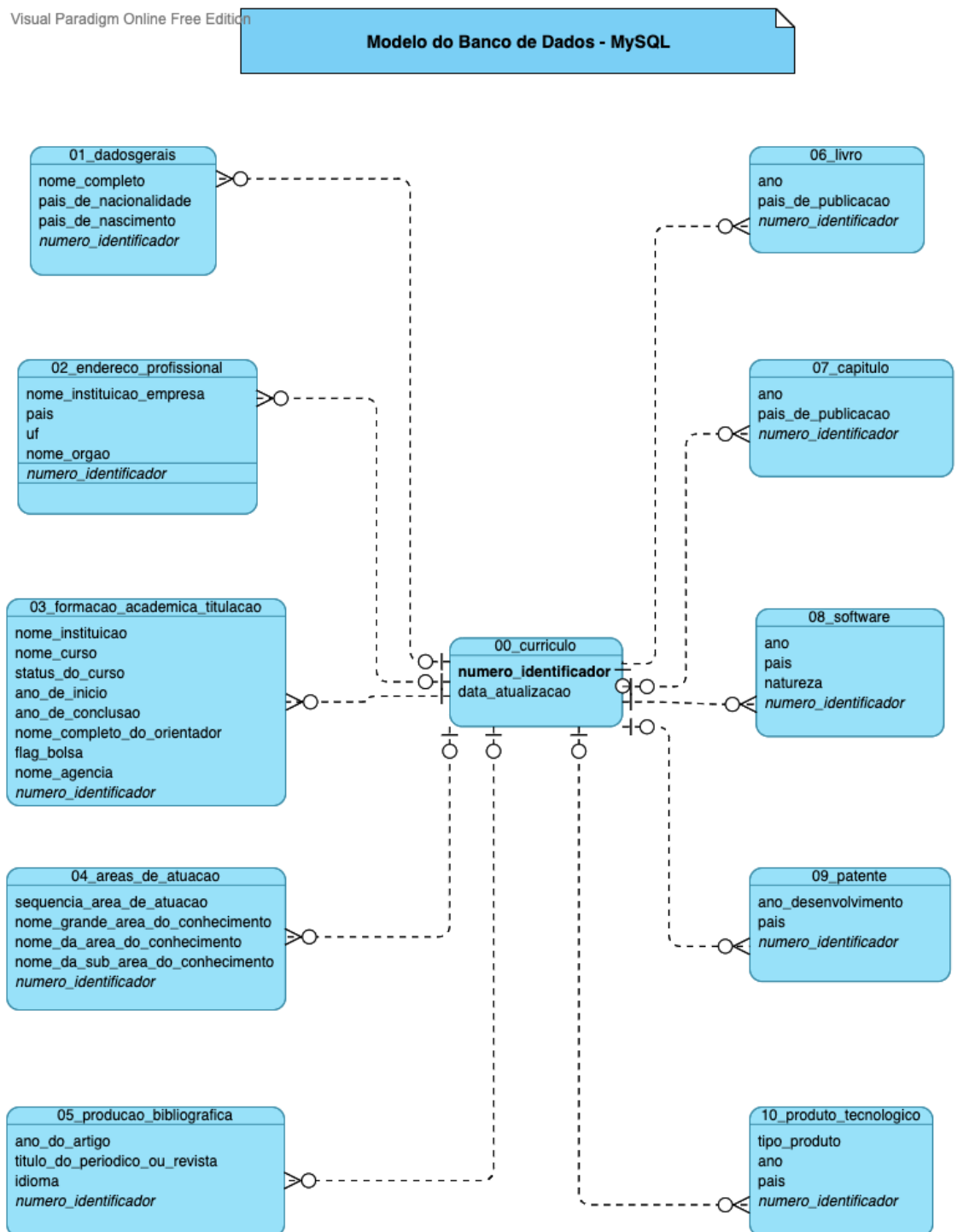
Figura 5 – Desenho da quarta e quinta etapas para a obtenção dos dados



Fonte: Elaboração própria

VI. Sexta etapa: estruturação do banco de dados. Usando o produto da análise anterior, modelou-se 11 tabelas com os seus atributos. Todas as tabelas dependentes do mesmo atributo chave, qual seja o “número identificador”. Todos os atributos foram definidos na forma de texto. Usando a ferramenta *Adminer*, foram criadas as tabelas e os campos correspondentes no banco de dados MySQL. A Figura 6 apresenta o modelo estruturado para receber os dados dos XMLs. Por fim, foi elaborada rotina em *Python* para ler cada arquivo XML e registrar as informações no banco de dados MySQL (para cada um dos currículos armazenados, a rotina carrega o XML do currículo, lista os atributos desejados, salva os atributos no banco de dados). Após a carga ter sido efetuada no banco de dados, foi feita, através da ferramenta *Adminer*, a exportação de cada tabela para um arquivo CSV.

Figura 6 – Modelo do Banco de Dados



Fonte: Elaboração própria

Figura 7 – Desenho da sexta etapa para a obtenção dos dados



Fonte: Elaboração própria

VII. Sétima etapa: após a análise do XML de alguns currículos aleatórios, foram extraídas as seguintes informações dos componentes listados no XML, utilizando rotina *Python*:

1. Currículo: número de identificação do currículo *lattes*; última data de atualização do currículo;
2. Dados Gerais: nome completo; país de nacionalidade; país de nascimento;
3. Endereço Profissional: nome da instituição, nome do órgão, país e UF;
4. Formação Acadêmica (apenas doutorado): nome da instituição de ensino; nome do curso; status do curso; ano de início; ano de conclusão; nome do orientador; se recebeu bolsa de estudos; nome da agência de fomento da bolsa;
5. Áreas de Atuação: sequência da área de atuação; grande área do conhecimento; nome da área do conhecimento; nome da subárea do conhecimento;
6. Produção Bibliográfica: ano de publicação; nome da revista ou periódico; idioma;
7. Livros Publicados: ano de publicação; país de publicação;
8. Capítulos de Livros Publicados: ano de publicação; país de publicação;
9. *Softwares* Desenvolvidos: natureza; ano de desenvolvimento; país de desenvolvimento;
10. Patentes Registradas: ano de desenvolvimento; país de desenvolvimento;
11. Produtos Tecnológicos Desenvolvidos: tipo do produto; ano de desenvolvimento; país de desenvolvimento.

VIII. Oitava etapa: usando a ferramenta de BI *QlikSense*, elaborou-se uma rotina para ler todos os CSVs gerados na etapa anterior, bem como a leitura da planilha em

excel fornecida pela Capes. Essa rotina além de carregar os dados, trata-os da seguinte forma:

- a) Filtra a planilha da Capes para considerar apenas linhas com a grande área informada;
- b) Une a planilha Capes e os arquivos CSVs com as informações dos currículos, Vinculando-os com o “número identificador”;
- c) Calcula o valor total dos recursos financeiros recebidos por bolsista;
- d) Sumariza em um único campo todas as rubricas de valor das bolsas;
- e) Define valor moeda estrangeira para cada tipo de moeda paga pelas bolsas, permitindo, assim, um recálculo para a moeda Real (R\$) sempre que necessário;
- f) Delimita o intervalo de 2010 a 2021, para as seguintes tabelas: artigo científico, livro, capítulos de livros, *software*, patente e produto tecnológico.

A ferramenta *QlikSense* foi utilizada para construir possíveis visões com dimensões e medidas gerais das informações coletadas, bem como para gerar uma planilha unificada com todos os dados normalizados dos currículos e da planilha fornecida pela Capes com as informações das bolsas.

IX. Nona etapa: Por fim, foi utilizado o *software* estatístico R para a construção dos gráficos e tabelas apresentados neste trabalho.

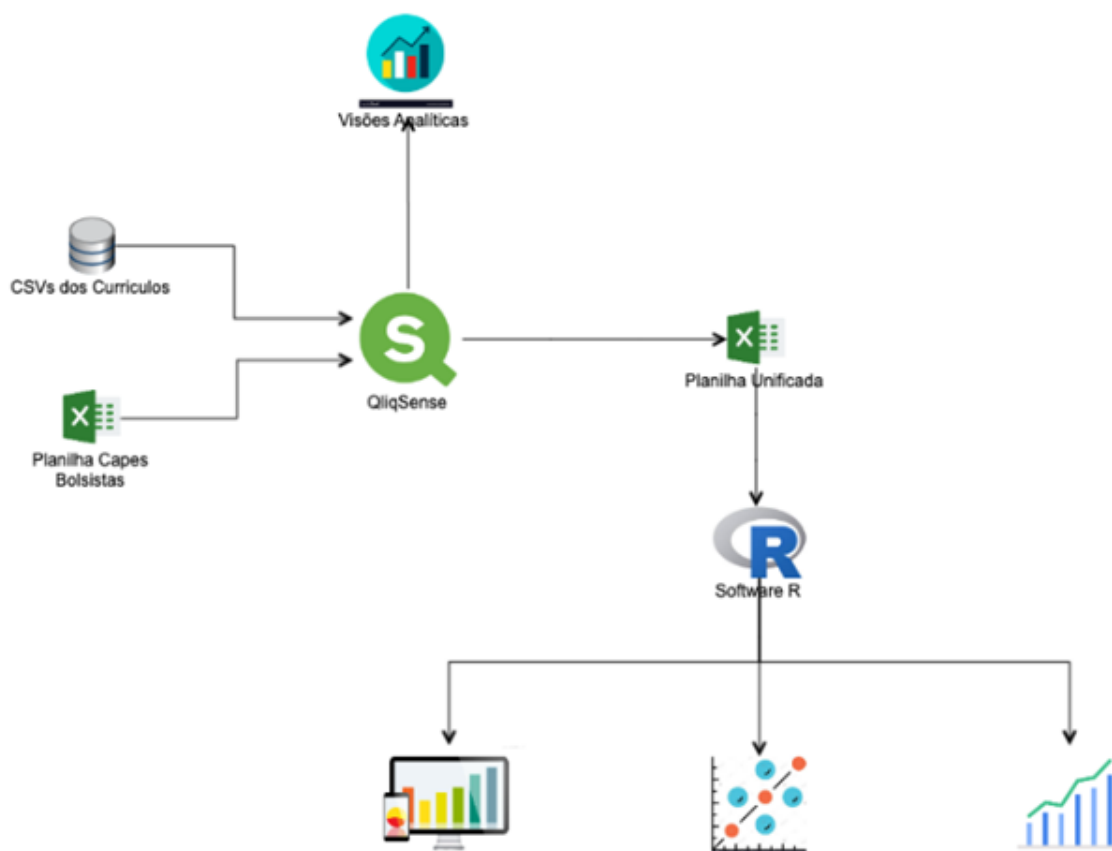
Vale mencionar que alguns dados foram verificados de forma manual diretamente na plataforma *Lattes*, por exemplo, bolsistas com nomes diferentes, mas com o mesmo número identificador do currículo. Foi preciso confirmar a quem o número pertencia e buscar o número daquele que estava errado. Outra situação foi a das pessoas homônimas, em que em alguns casos, foi possível descobrir quem era o bolsista listado na planilha, considerando informações tais como grande área, área do conhecimento, instituição de origem e outras.

Foi necessário excluir da pesquisa alguns registros das bolsas fornecidos pela CAPES, a saber:

- I. Daqueles bolsistas em que não foi possível armazenar o XML dos currículos, mesmo após novas tentativas de *download*. O arquivo ou estava corrompido ou em branco, sem dados, nitidamente com algum erro em sua estrutura.
- II. Daqueles bolsistas que não possuíam o endereço do currículo *lattes* na planilha.
- III. Dos homônimos que não possuíam endereço do currículo *lattes* na planilha, e que não foi possível identificar o nome correto.
- IV. Daqueles bolsistas sem a informação da grande área do conhecimento.

- V. Dos bolsistas com valor recebido de bolsa igual a zero, provavelmente por terem sido casos de estorno ou devolução.
- VI. Das produções acadêmico-científicas (artigos, capítulos de livros, livros, *software*, patentes, produtos tecnológicos) anteriores a 2010 e posteriores a 2021.

Figura 8 – Desenho da oitava e nona etapas para a obtenção dos dados



Fonte: Elaboração própria

4 ANÁLISE E RESULTADOS

Das 54.209 (cinquenta e quatro mil duzentos e nove) bolsas listadas na extração da CAPES, 791 (setecentos e noventa e um) não possuíam o endereço do currículo *lattes*. Destes, foi possível resgatar os números de identificação de 404 currículos utilizando os nomes completos, sendo que o restante foi excluído da pesquisa.

Foram identificados 33 (trinta e três) nomes duplicados, com número identificador do currículo diferente. Destes, 29 (vinte e nove) foram identificados corretamente de forma manual por meio da análise do currículo e dos dados extraídos da CAPES, e os outros também foram excluídos da pesquisa.

Não foi possível fazer o *download* do XML ou abrir o arquivo salvo de 112 (cento e doze) currículos, provavelmente devido a algum erro no XML ou na plataforma *Lattes*. Estes também foram excluídos do estudo. Por fim, alguns bolsistas estavam com a grande área do conhecimento, a área do conhecimento e a área de avaliação em branco na extração da CAPES, mas apenas 2 (dois) foram excluídos da pesquisa por não ter sido possível resgatar as informações dos currículos.

Do total de bolsas aptas para a pesquisa - 53.693 (cinquenta e três mil seiscentos e noventa e três), 1.428 (mil quatrocentos e vinte e oito) são bolsas no exterior, de programas de concessão de bolsas da Diretoria de Relações Internacionais (DRI), quais sejam Programa de Doutorado Pleno do Exterior (DOC-PLENO), Doutorado Pleno CSF (BEX) e DPE-Programa de Doutorado Pleno no Exterior. O restante das bolsas foi concedido pela Diretoria de Programas e Bolsas no País (DPB), por meio dos seguintes programas: Bolsa Especial para Doutorado em Pesquisa Médica (PBE-DPM), Programa de Demanda Social (DS), Programa de Excelência Acadêmica (PROEX), Programa de Excelência Acadêmica – Taxas (PROEX/TAXAS), Programa MINTER/DINTER CAPES/SETEC (MINDIN) e Programa de Doutorado Interinstitucional (DINTER).

4.1 Programas de concessão de bolsa (DPB e DRI)

Desde o ano de 2013, a bolsa de doutorado pleno no Brasil possui valor fixo de R\$2.200,00 por mês, conforme Portaria CAPES nº 1 de 3 de janeiro de 2020. No período de 2010 a 2012, o valor da bolsa variou de R\$1.800,00 a R\$2.000,00. A bolsa de doutorado pleno no exterior também possuiu valor fixo, contudo, em moeda estrangeira, a depender a instituição de destino do bolsista.

A fim de que os valores das bolsas no país e no exterior pudessem ser comparados, foi necessário fazer a transformação de todos os valores em moedas estrangeiras

para a moeda Real (R\$), utilizando a cotação das moedas disponível no site do Banco Central do Brasil - Bacen, com a mesma data de referência para todas as moedas, qual seja a data de 31/3/2021.

Ressalta-se que foram consideradas duas casas decimais para a conversão dos valores, de acordo com a Tabela 1.

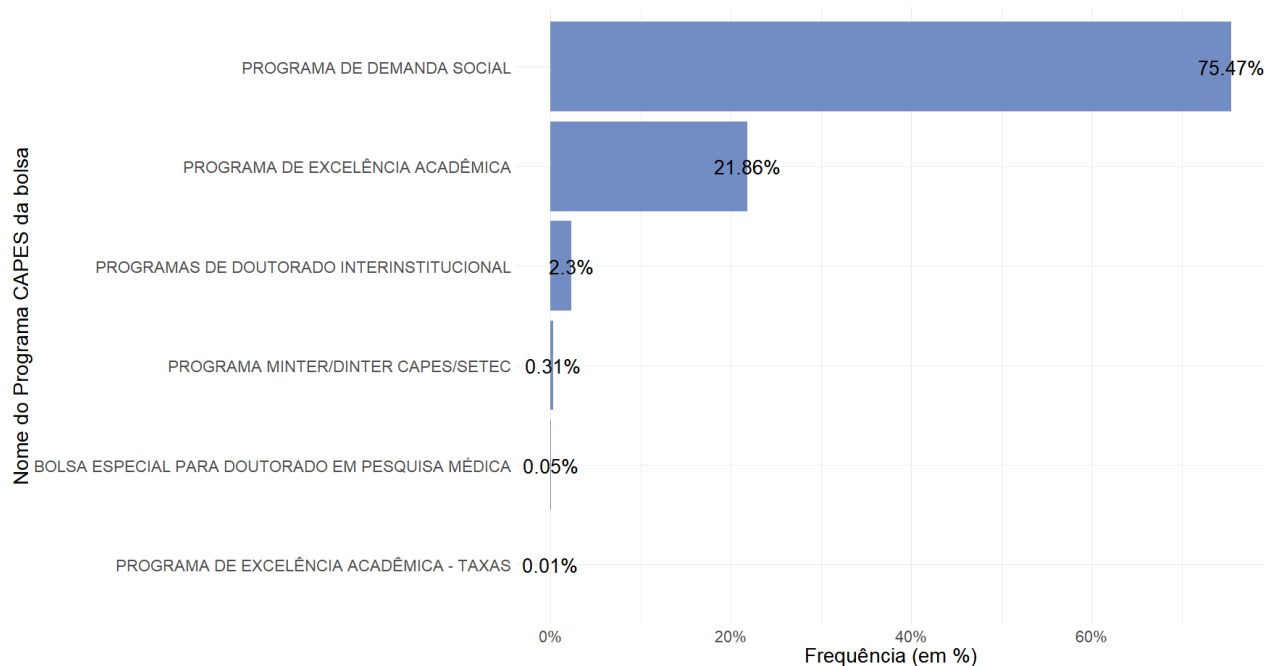
Tabela 1 – Cotações moedas estrangeiras

Moeda	Cotação em R\$
AS (dólar australiano)	4,33
EUR (euro)	6,68
USD (dólar americano)	5,69
JPY (iene japonês)	0,05
CAN (dólar canadense)	4,53
CHF (franco suíço)	6,04
DKK (coroa dinamarquesa)	0,89
SEK (coroa sueca real)	0,65
LS (libra esterlina)	7,85

Fonte: Elaboração própria

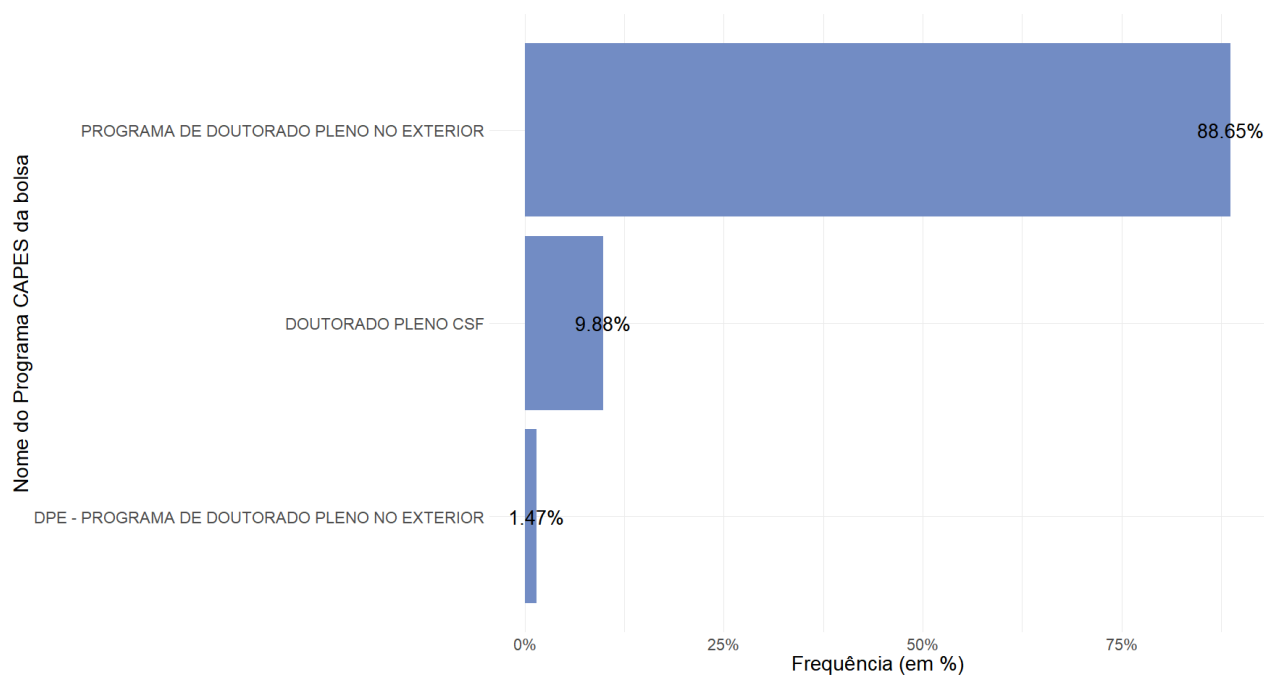
Das bolsas implementadas pela DPB para a realização do doutorado no Brasil, o programa de Demanda Social segue com o maior número de concessões, enquanto o Programa de Doutorado Pleno no Exterior se destaca com o maior número de concessões para a realização do doutorado no exterior, conforme Figuras 9 e 10.

Figura 9 – Distribuição de bolsas no país por programa CAPES - DPB



Fonte: Elaboração própria

Figura 10 – Distribuição de bolsas no exterior por programa CAPES - DRI

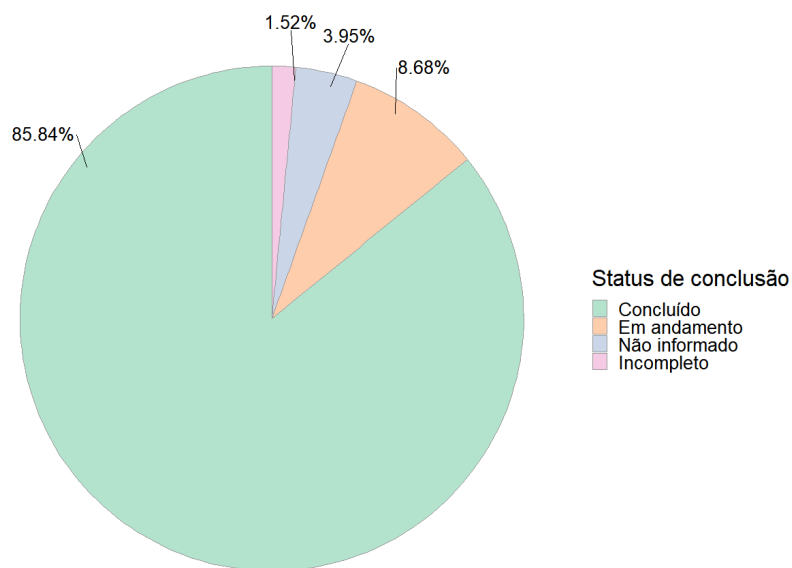


Fonte: Elaboração própria

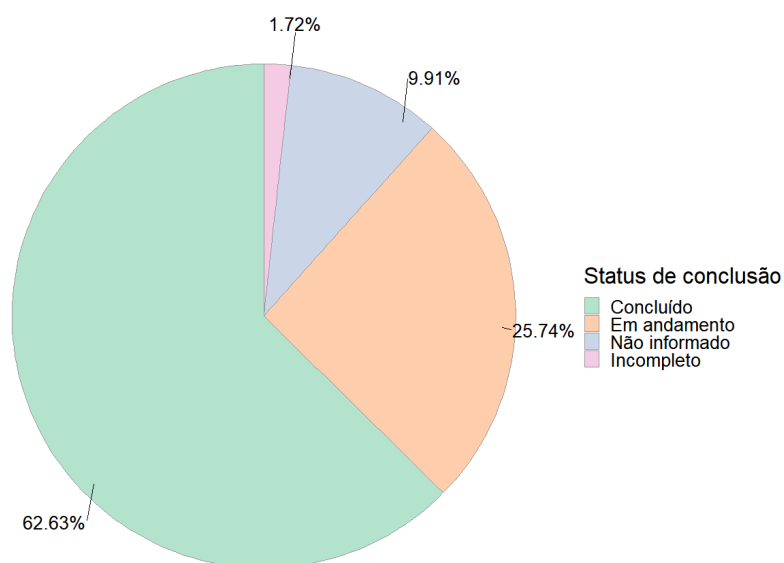
4.2 Titulação

De acordo com as informações coletadas dos currículos *lattes* dos bolsistas, 85,8% dos estudantes concluíram o curso de doutorado, 8,6% cadastraram o doutorado como curso em andamento, 3,9% não informaram e apenas 1,5% não concluíram o curso.

Dos estudantes que receberam bolsa no exterior, o percentual do número de doutorado em andamento aumenta para 25,7%, e o percentual dos que não informaram aumenta para 9,9%, conforme Figura 12.

Figura 11 – *Status* do doutorado dos bolsistas

Fonte: Elaboração própria

Figura 12 – *Status* do doutorado dos bolsistas no exterior

Fonte: Elaboração própria

4.3 Endereço profissional

O endereço profissional cadastrado no currículo dos bolsistas também foi extraído. A Figura 13 mostra que os bolsistas estão vinculados profissionalmente a diversas instituições pelo mundo.

Figura 13 – Distribuição dos bolsistas segundo o endereço profissional



Fonte: Elaboração própria

Dos 53.693 currículos, um total de 32.977 (61,42%) cadastrou endereço profissional no Brasil, 19.661 (36,62%) não preencheu o campo de endereço profissional, e o restante (1,96%) cadastrou endereço no exterior. Na Tabela 2 é possível verificar a quantidade de bolsistas e os países em que consta o vínculo profissional em ordem crescente da maior quantidade de bolsistas das duas diretorias.

Tabela 2 – Países de vínculo profissional dos bolsistas

País (endereço profissional)	Bolsa no país (DPB)	Bolsa no exterior (DRI)	Total geral
Brasil	32.518	459	32.977
Não informado	19.020	641	19.661
Estados Unidos	211	93	304
Portugal	39	42	81
França	50	20	70
Alemanha	47	22	69
Canadá	44	23	67
Colômbia	43	-	43
Grã-Bretanha	18	24	42
Espanha	21	19	40
Inglaterra	14	15	29
Holanda	11	16	27
Austrália	16	7	23
Chile	19	2	21
Suécia	16	4	20
Argentina	16	1	17
Dinamarca	9	8	17
Itália	10	7	17

Tabela 2 - continuação da página anterior

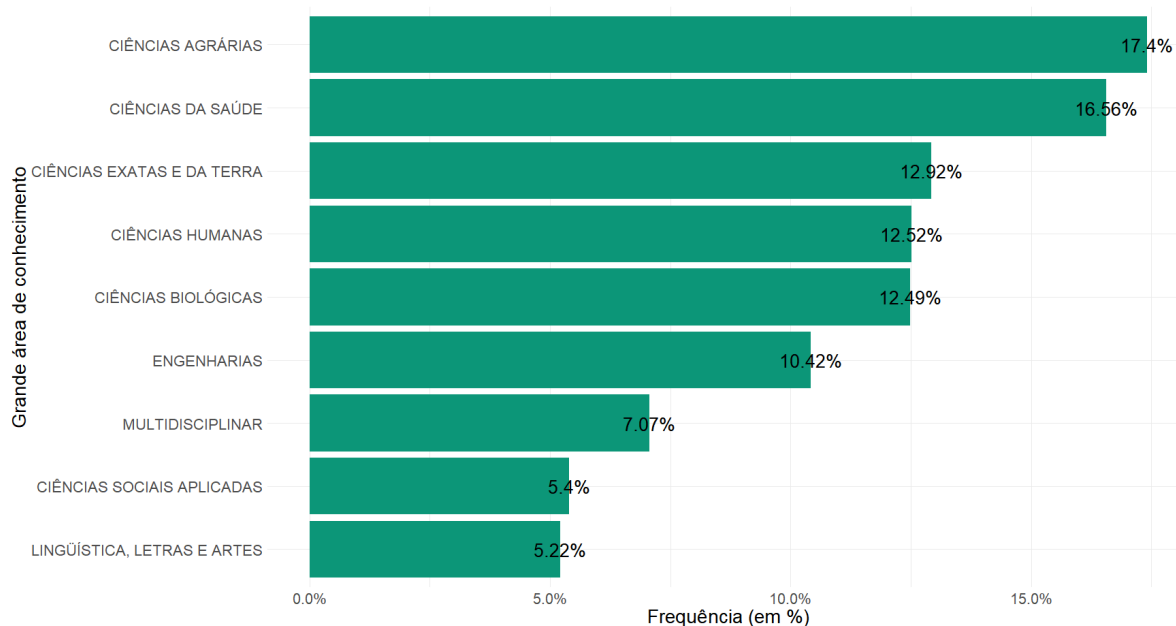
Irlanda	7	8	15
Peru	15	-	15
Bélgica	11	3	14
Uruguai	12	-	12
Áustria	6	3	9
Escócia	5	4	9
México	8	-	8
China	6	-	6
Cuba	6	-	6
Equador	6	-	6
Moçambique	6	-	6
República Tcheca	6	-	6
Polônia	5	-	5
África do Sul	4	-	4
Finlândia	3	1	4
Nova Zelândia	3	1	4
Suíça	4	-	4
Angola	3	-	3
Irlanda do Norte	3	-	3
Israel	2	1	3
Costa Rica	2	-	2
Estônia	2	-	2
Japão	1	1	2
Nigéria	2	-	2
Porto Rico	1	1	2
Tailândia	2	-	2
Afeganistão	1	-	1
Arábia Saudita	-	1	1
Brunei	1	-	1
Cabo Verde	1	-	1
Egito	1	-	1
Emirados Árabes	1	-	1
Etiópia	1	-	1
Hungria	1	-	1
Índia	1	-	1
Marrocos	1	-	1
Mônaco	1	-	1
Noruega	-	1	1
Paquistão	1	-	1
Ucrânia	1	-	1
Total geral	52.265	1.428	53.693

Fonte: Elaboração própria

4.4 Grande Área do conhecimento

A grande área do conhecimento de Ciências Agrárias foi a área com o maior número de bolsistas, com 17,4% das bolsas concedidas no período. A grande área de Linguística, Letras e Artes, com 5,2% das bolsas concedidas, foi a área com o menor número de bolsas no mesmo período.

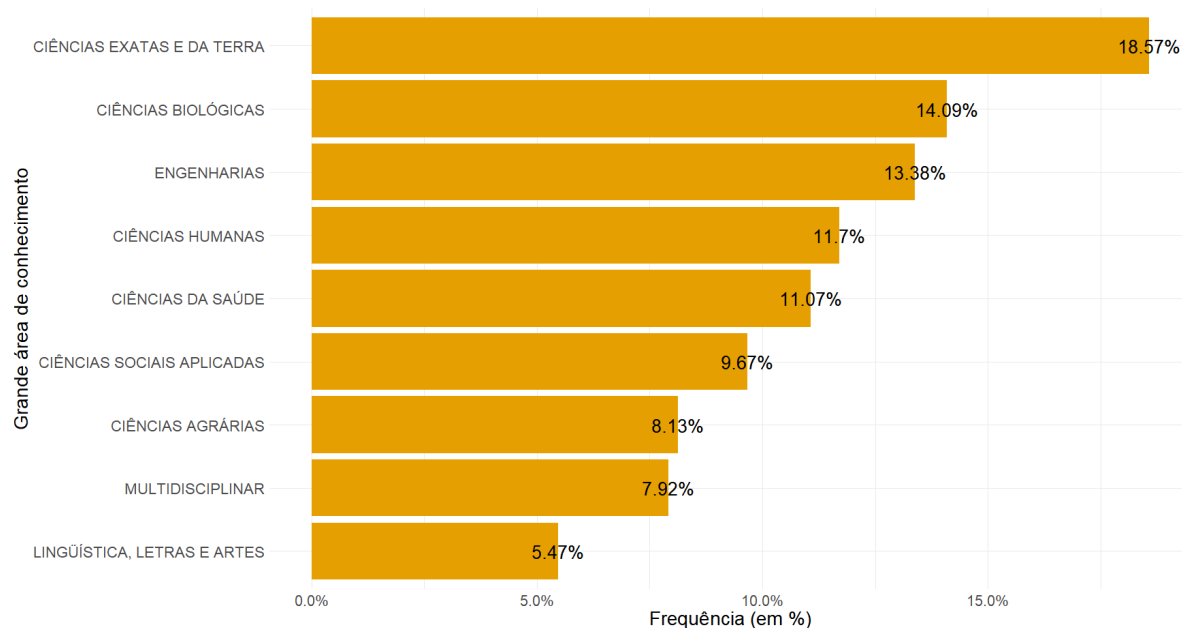
Figura 14 – Quantidade de bolsas nas grandes áreas do conhecimento



Fonte: Elaboração própria

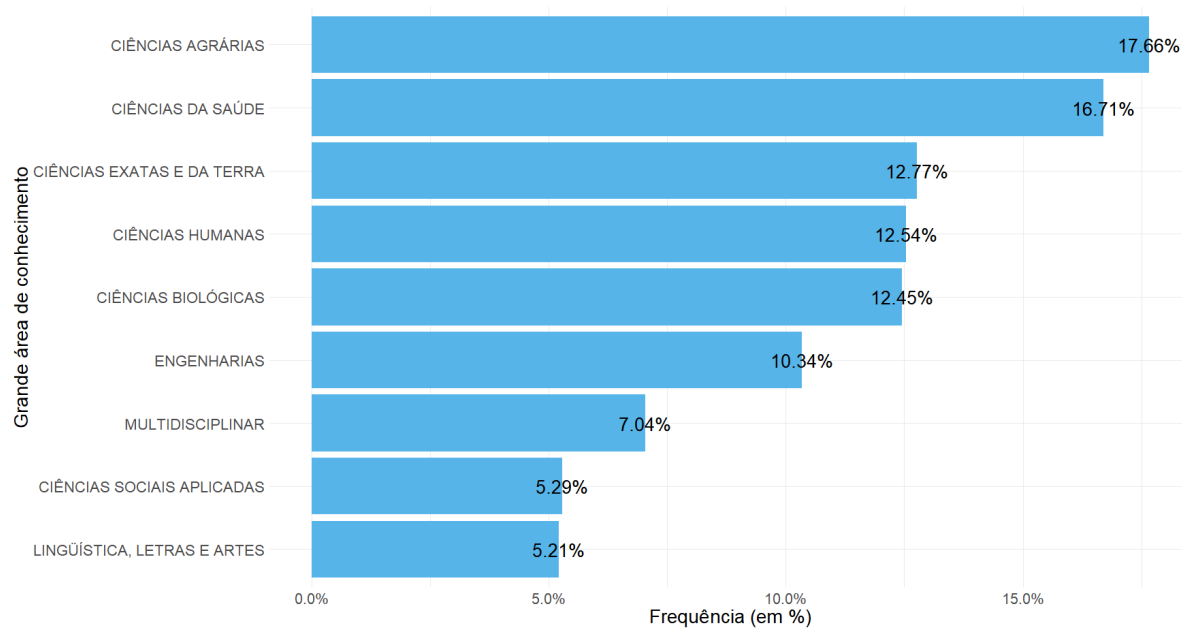
Há variação das áreas com maior número de bolsas concedidas quando são analisadas as bolsas no exterior apenas. Das bolsas concedidas pela DRI, a grande área do conhecimento com um maior número de bolsistas é a de Ciências Exatas e da Terra, com 18,5% das bolsas concedidas, conforme Figura 15. No caso das bolsas concedidas pela DBP, não há variação das áreas com maior e menor concessão, de acordo com a Figura 16.

Figura 15 – Quantidade de bolsas nas grandes áreas do conhecimento – DRI



Fonte: Elaboração própria

Figura 16 – Quantidade de bolsas nas grandes áreas do conhecimento – DPB



Fonte: Elaboração própria

4.5 Artigos

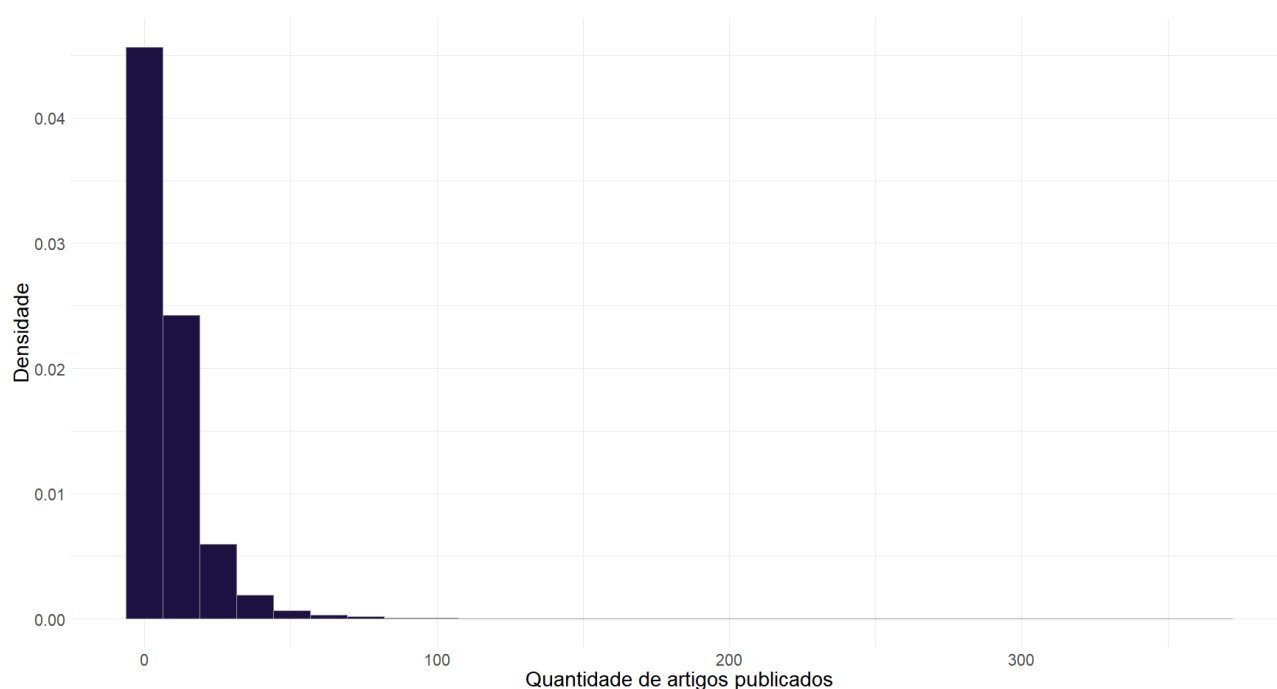
A distribuição de artigos publicados é bem assimétrica, ou seja, não segue o comportamento de uma distribuição normal, que tem formato de sino com o pico igual à média e à mediana. Na Figura 17 é possível verificar que há uma concentração grande de pequenos valores, mas com a presença de valores extremos (*outliers*) que

influenciam na distribuição. Isso se reflete na diferença entre a média de publicações, de 8.69, e a mediana, de 5.

Isto quer dizer que 50% das bolsas produziram até cinco artigos ao longo do período analisado, mas existem outros bolsistas que produziram um volume grande de artigos. Essa diferença significativa entre média e mediana também é um indício dessa influência, pois em uma distribuição normal há simetria e essas métricas são iguais. O coeficiente de variação, calculado pela razão entre o desvio-padrão e a média, foi de 137,7%, o que confirma a altíssima variabilidade dos dados.

Por fim, a produção varia de 0 a 366 artigos publicados. Do total de bolsas, 6.399 (11,4%) não publicaram nenhum artigo durante o período analisado, o que representa a ausência de produção científica de nosso interesse durante o período do curso de doutorado.

Figura 17 – Distribuição de artigos publicados

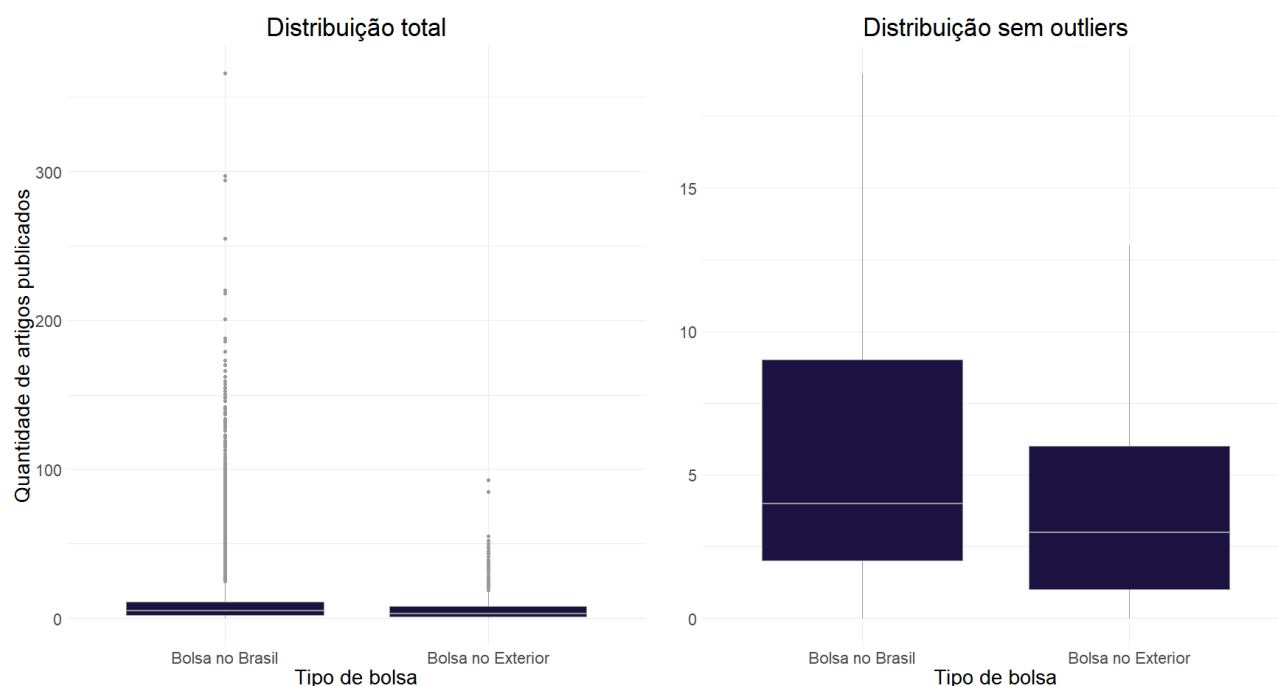


Fonte: Elaboração própria

Esse comportamento dos *outliers* fica ainda mais evidente quando a quantidade de artigos publicados é observada segundo o tipo de bolsa recebida pelo estudante. É possível notar que boa parte das observações é considerada um valor extremo para a distribuição. Quando desconsiderados, é possível ver com mais clareza a diferença entre os tipos de bolsa, situação em que as bolsas no país têm resultados superiores aos das bolsas no exterior.

Contabilizando 52.170 bolsas de origem da DPB e 1.427 de origem da DRI, a média de artigos publicados pelos bolsistas no país é de 8,78 e a mediana é de 5

Figura 18 – Distribuição de artigos publicados segundo tipo de bolsa



Fonte: Elaboração própria

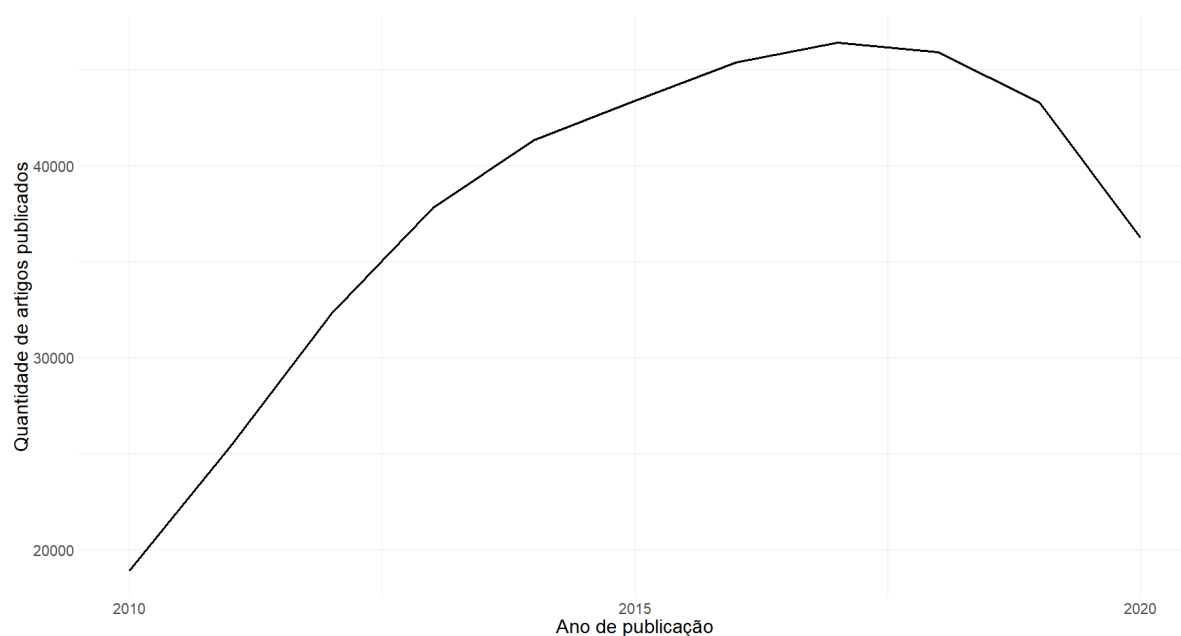
artigos, enquanto para os bolsistas no exterior a média é 5,66 e a mediana é 3, uma clara diferença entre a quantidade produzida pelos bolsistas. Em ambos os casos, o coeficiente de variação está em torno de 138%, confirmando a grande dispersão dos dados verificada no gráfico de distribuição total.

No caso dos bolsistas no país, o valor máximo de artigos publicados por bolsa foi de 366, e 6.091 bolsas não produziram artigos publicados, o que equivale a 11,7% das bolsas de origem da DPB. Já para os bolsistas no exterior, o volume máximo foi de 93 e 308 não publicaram artigos (21,6% das bolsas de origem da DRI).

A Figura 19 mostra o quantitativo de artigos publicados por todos os bolsistas no período de 2010 a 2020, e a Figura 20 demonstra a publicação ao longo do mesmo período por tipo de bolsa – país e exterior.

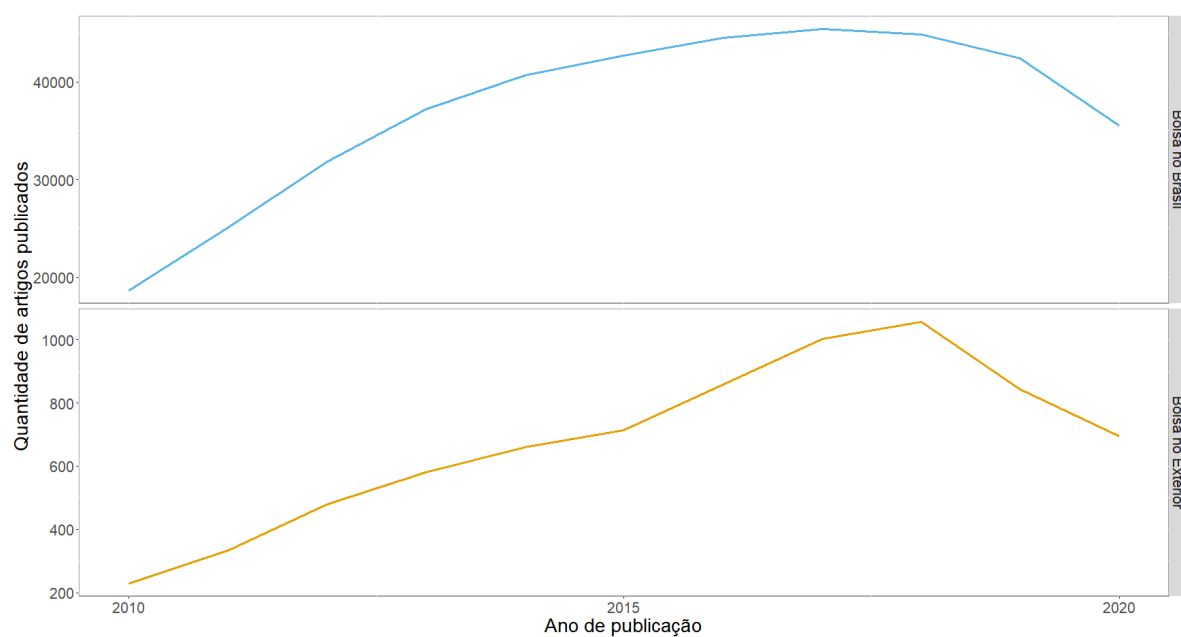
Do total de artigos publicados pelos bolsistas, a maior concentração de publicação está na área de Ciências da Saúde, com 25,6% da produção total, seguida da área de Ciências Agrárias, com 22,9% dos artigos publicados. A área de Letras, Linguística e Artes é a área que menos produziu (Figura 21).

Figura 19 – Artigos publicados no período de 2010 a 2020



Fonte: Elaboração própria

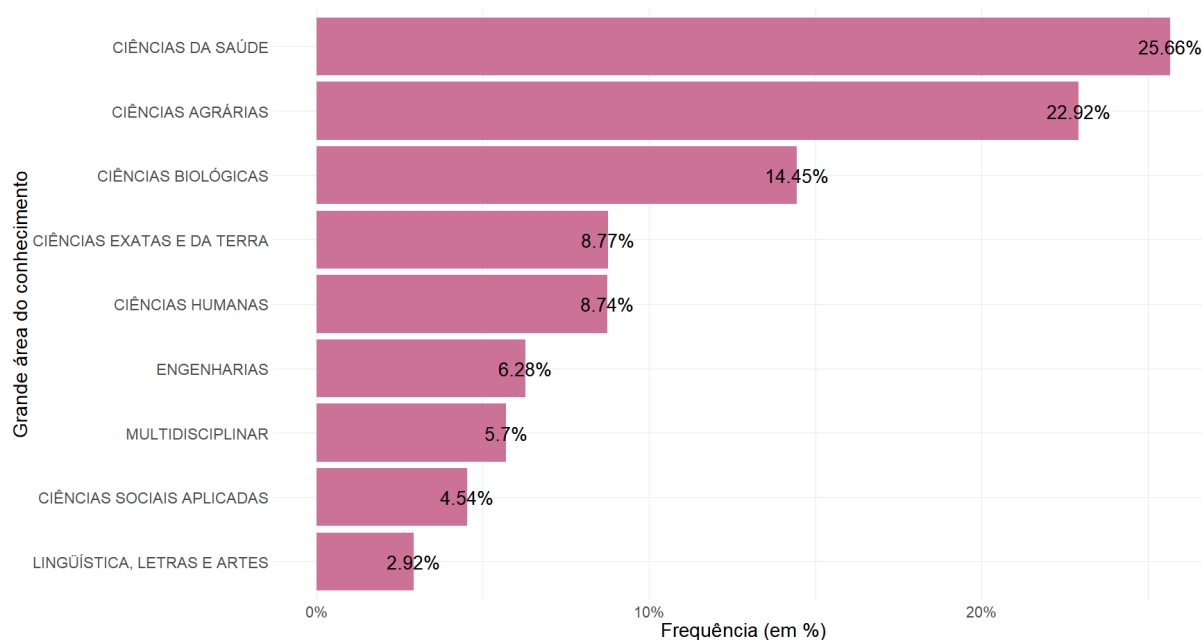
Figura 20 – Artigos publicados no período de 2010 a 2020 segundo tipo de bolsa



Fonte: Elaboração própria

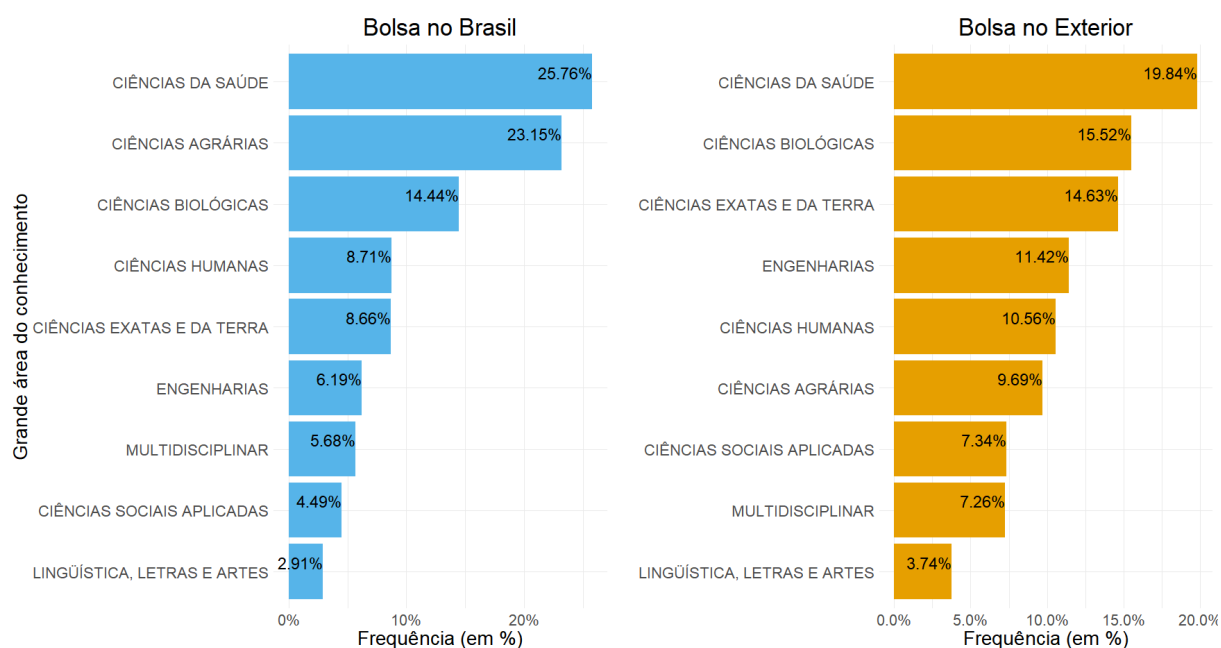
Separando o tipo de bolsa, na Figura 22, nota-se que a área de Ciências da Saúde permanece com o maior percentual de publicação para os dois tipos – país e exterior, mas no caso das bolsas no exterior, a área de Ciências Biológicas aparece como a segunda área com maior percentual de produção, seguida da área de Ciências Exatas e da Terra, diferindo da classificação das áreas com maior percentual de publicação oriundas das bolsas no país.

Figura 21 – Artigos publicados segundo grande área do conhecimento



Fonte: Elaboração própria

Figura 22 – Artigos publicados segundo grande área do conhecimento e tipo de bolsa

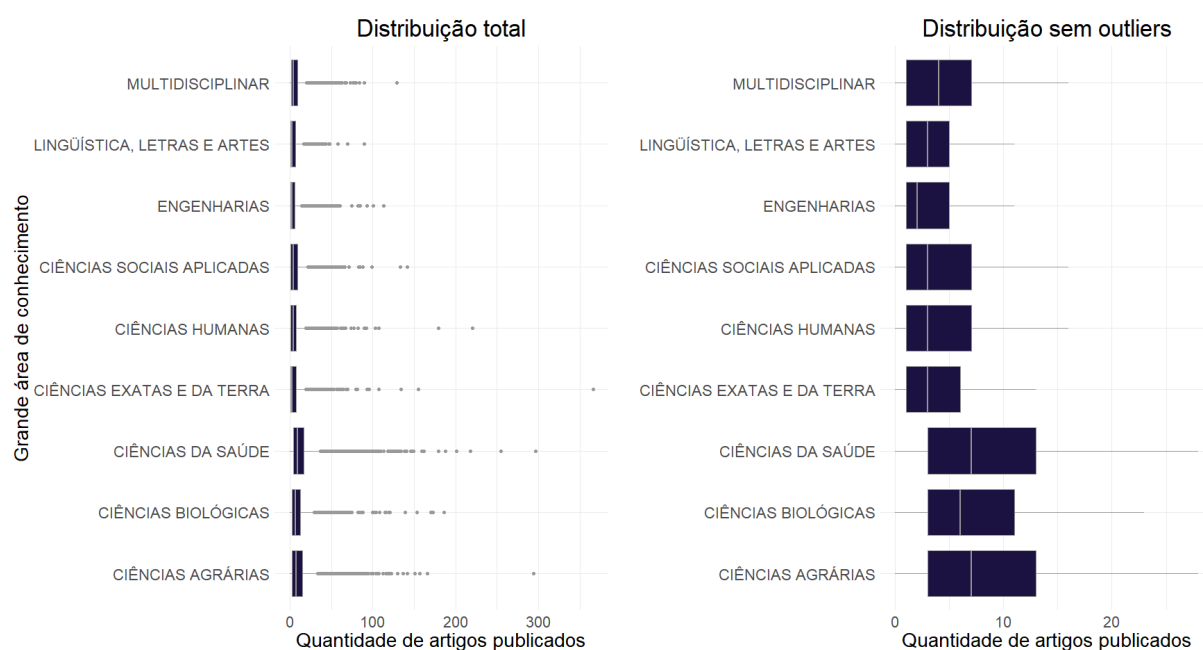


Fonte: Elaboração própria

Na separação das grandes áreas do conhecimento, a distribuição de artigos produzidos também apresenta uma grande quantidade de *outliers*. Ao retirá-los para observar melhor o comportamento da distribuição, é possível ver que há um grande destaque das áreas de Ciências da Saúde, Ciências Biológicas e Ciências Agrárias (Figura 23).

A média dessas áreas é de 13,5, 10,1 e 11,5, respectivamente, enquanto as outras áreas têm média inferior a 10. A mediana também é superior, resultando em 9, 7 e 8, enquanto o restante teve mediana entre 3 e 4. Com isso, também se pode observar a grande influência dos valores extremos na distribuição.

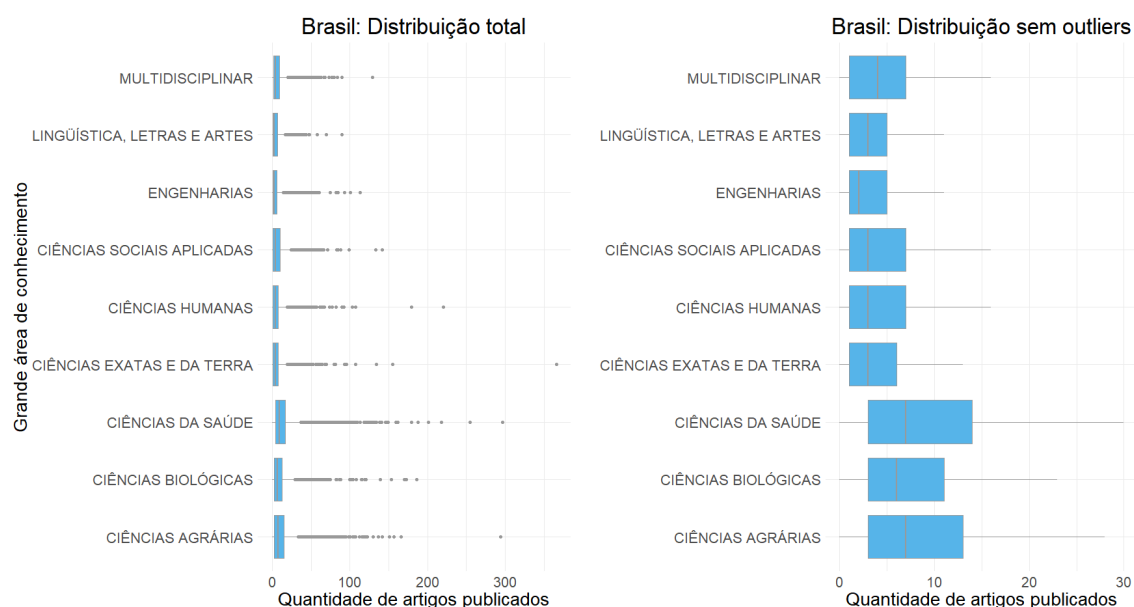
Figura 23 – Artigos publicados segundo grande área do conhecimento e tipo de bolsa



Fonte: Elaboração própria

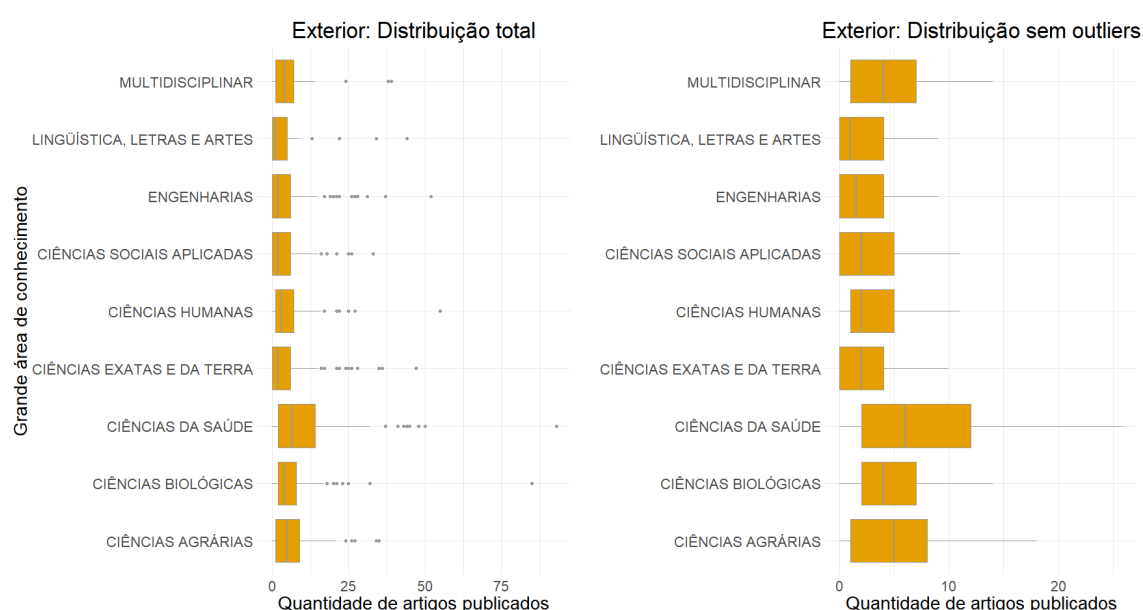
O comportamento comparativo das áreas entre bolsistas com bolsa no Brasil (Figura 24) e com bolsa no exterior (Figura 25) é bastante semelhante ao comportamento geral dos dados. No exterior, há um destaque maior para as Ciências da Saúde.

Figura 24 – Artigos publicados segundo grande área do conhecimento e tipo de bolsa - DPB



Fonte: Elaboração própria

Figura 25 – Artigos publicados segundo grande área do conhecimento e tipo de bolsa - DRI

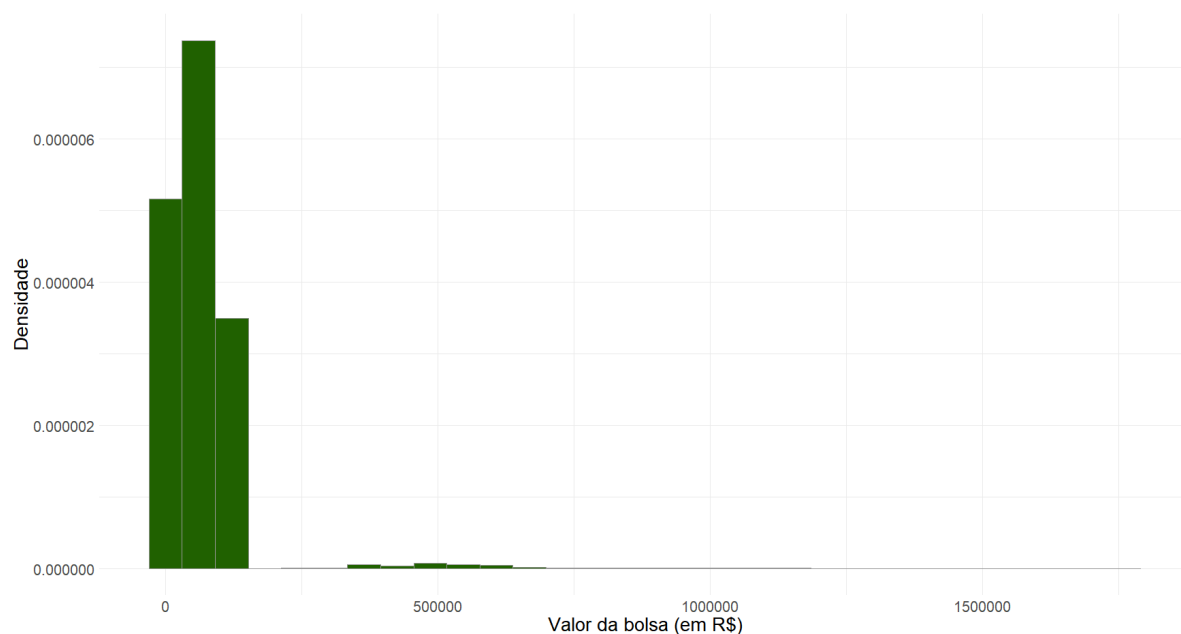


Fonte: Elaboração própria

4.6 Investimento em bolsa

A distribuição do valor da bolsa tem um comportamento bastante assimétrico. No total, existem bolsas que variam de R\$200,00 até R\$1.760.881,00, com média de R\$69.384,71 e mediana de R\$57.200,00. Há uma variabilidade grande dos dados, observada na Figura 26 e pelo coeficiente de variação, calculado em 147,0%.

Figura 26 – Distribuição do valor da bolsa (em reais)

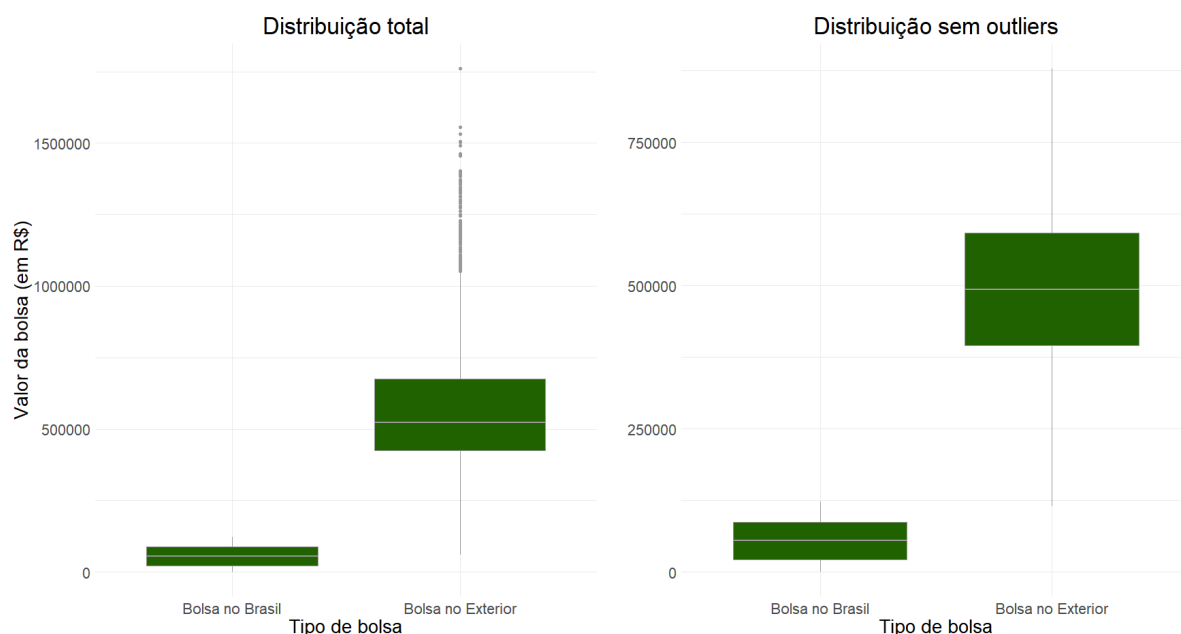


Fonte: Elaboração própria

Quando essa variável é observada segundo o tipo de bolsa (Figura 27), vê-se uma diferença grande na distribuição: enquanto as bolsas no Brasil ficam entre o mínimo de R\$200,00 e o máximo de R\$123.200,00, com média de R\$55.091,00 e mediana de R\$55.000,00, as bolsas no exterior vão de R\$60.906,00 a R\$1.760.881,00, com média de R\$591.964,00 e mediana de R\$522.583,00. Isso evidencia a discrepância entre os tipos de bolsa e a influência dos gastos com as bolsas no exterior no total do gasto com bolsas individuais. A disparidade entre as bolsas no Brasil, no entanto, é maior do que das bolsas no exterior, com coeficiente de variação de 62,0% para bolsas no país contra 44,0% das bolsas no exterior.

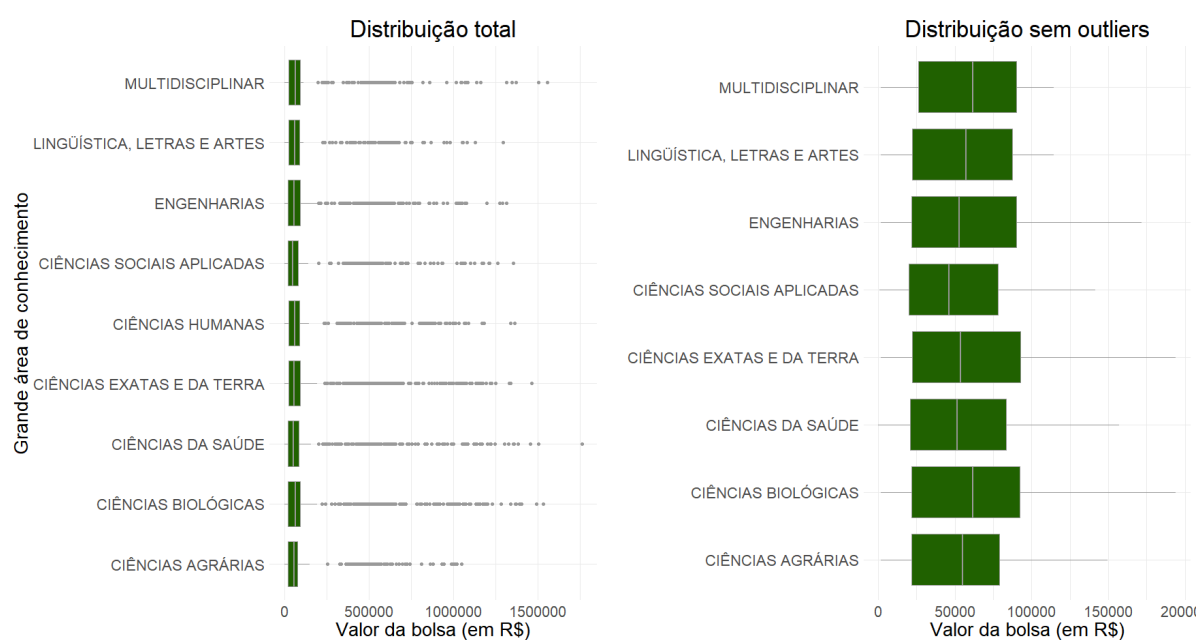
Com relação às áreas de conhecimento (Figura 28), vê-se que há uma presença significativa dos *outliers* em todas, mas não há muita divergência entre as áreas. Ciências Sociais Aplicadas e Ciências Exatas e da Terra são as mais afetadas por valores extremos, o que se reflete na distribuição como um todo. Essas áreas têm uma média de R\$75.057,00 e R\$76.868,00 respectivamente, mas as medianas são de R\$49.400,00 e R\$57.200,00. O coeficiente de variação dessas áreas também está entre os maiores – 170,0% para Ciências Sociais Aplicadas e 157,0% para Ciências Exatas e da Terra.

Figura 27 – Distribuição do valor da bolsa (em reais) segundo tipo de bolsa



Fonte: Elaboração própria

Figura 28 – Distribuição do valor da bolsa (em reais) segundo grande área do conhecimento



Fonte: Elaboração própria

Observando a distribuição das bolsas, no país e no exterior, por grande área (Figuras 29 e 30), fica clara a disparidade em relação aos valores concedidos para cada tipo de bolsa. A mediana das bolsas no país está entre R\$46.200,00 (em Ciências Sociais Aplicadas) e R\$61.600,00 (em Ciências Biológicas e Multidisciplinar), enquanto as bolsas no exterior têm mediana entre R\$476.501,00 (Ciências Agrárias) e

R\$570.309,00 (Linguística, Letras e Artes).

Figura 29 – Distribuição do valor da bolsa (em reais) segundo grande área do conhecimento - DPB

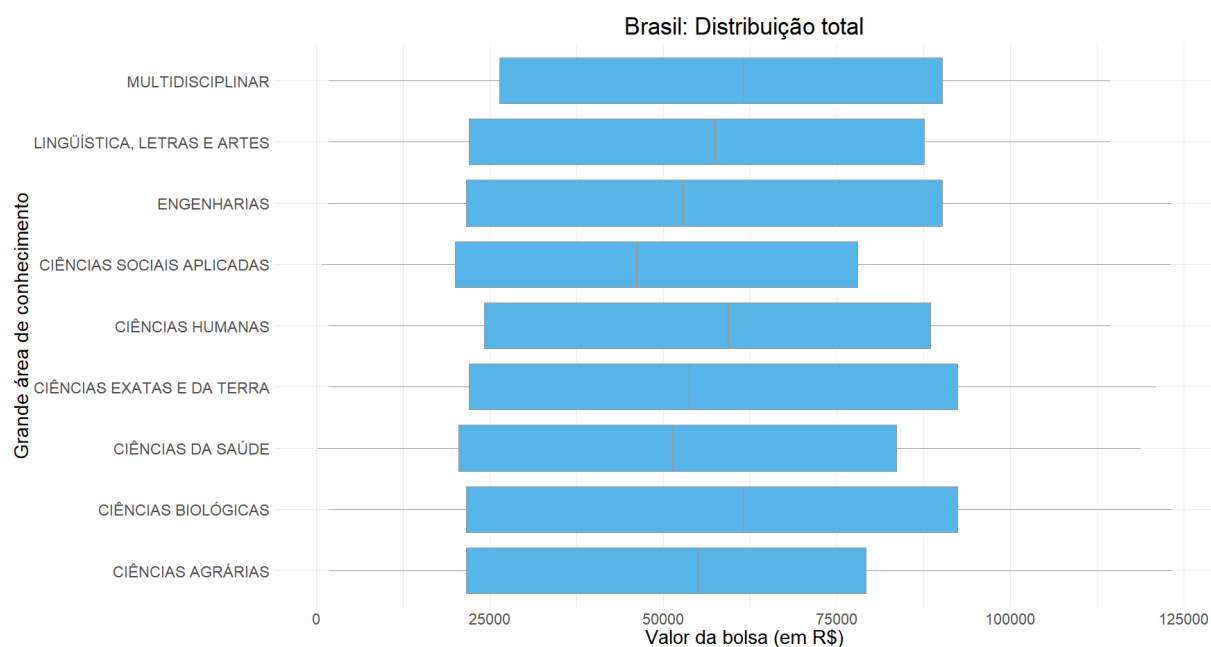
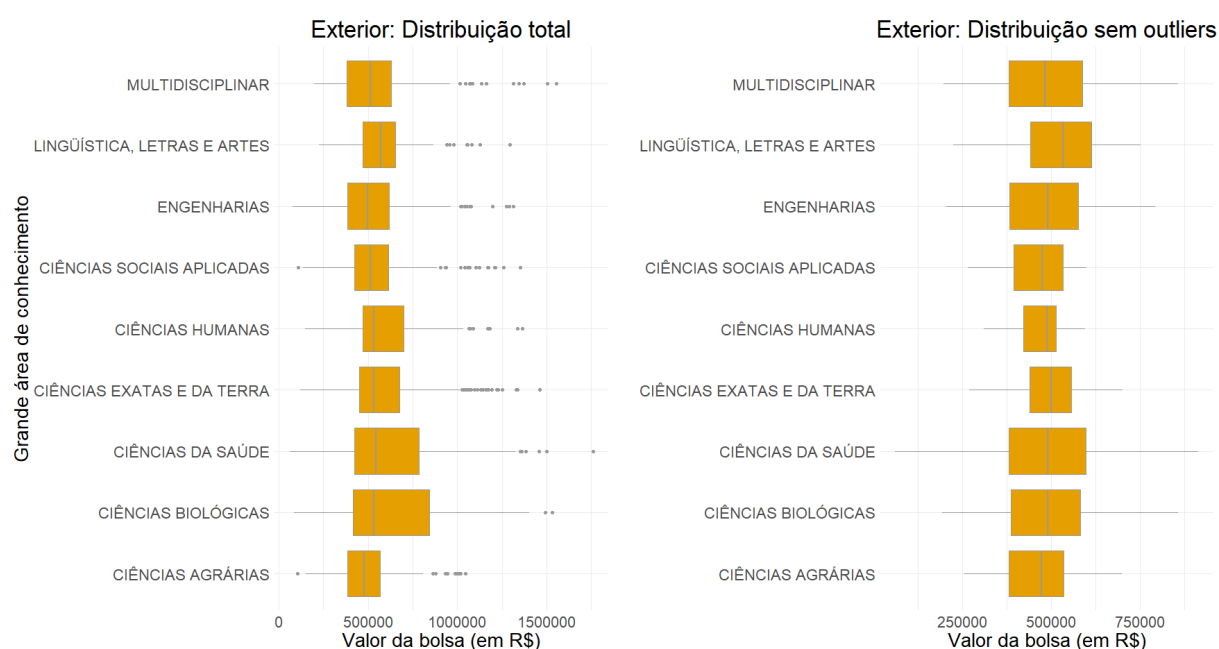


Figura 30 – Distribuição do valor da bolsa (em reais) segundo grande área do conhecimento - DRI



As bolsas no exterior, apesar de terem coeficiente de variação menor, apresentam *outliers*, enquanto as bolsas no Brasil têm uma distribuição sem valores extremos. Pela primeira vez, apareceram tanto *outliers* superiores quanto inferiores. Os superiores

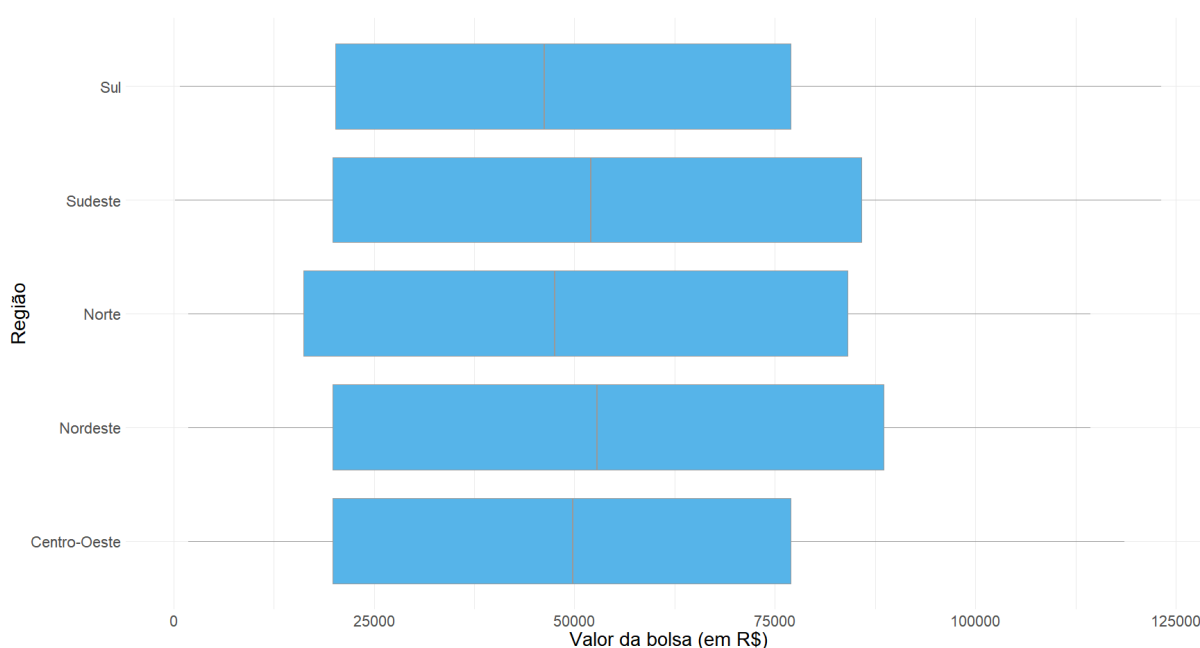
são valores maiores do que os verificados na distribuição, enquanto os inferiores são menores. Assim, existem bolsas que tiveram investimento muito superior a outras, mas também bolsas com valor inferior ao observado na distribuição, indicando a disparidade no investimento de bolsas no exterior.

4.7 Investimento em bolsa no país por região

Considerando ainda os recursos financeiros investidos nas bolsas no país, nota-se que essas bolsas têm uma distribuição bastante semelhante quando observadas as regiões do Brasil, segundo a instituição onde o bolsista realizou o doutorado.

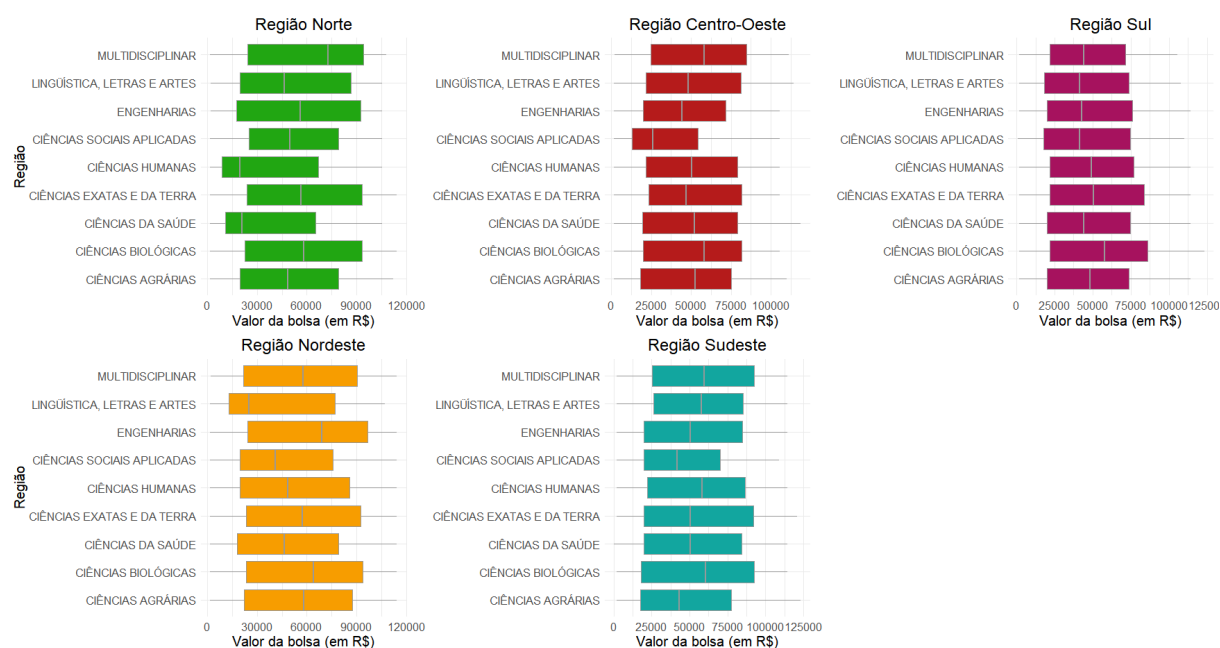
É possível observar que as regiões Nordeste e Sudeste têm um comportamento muito próximo, com média de R\$61.091,00 e mediana de R\$53.800,00 para o Nordeste, enquanto o Sudeste tem média de R\$60.235,00 e mediana de R\$52.800,00. Em relação às áreas de conhecimento, percebe-se que o Sul é a região com menos diferenças na distribuição dos valores investidos nas bolsas entre as áreas.

Figura 31 – Distribuição do valor da bolsa DPB (em reais) por região



Fonte: Elaboração própria

Figura 32 – Distribuição do valor da bolsa DPB (em reais) por região e grande área do conhecimento



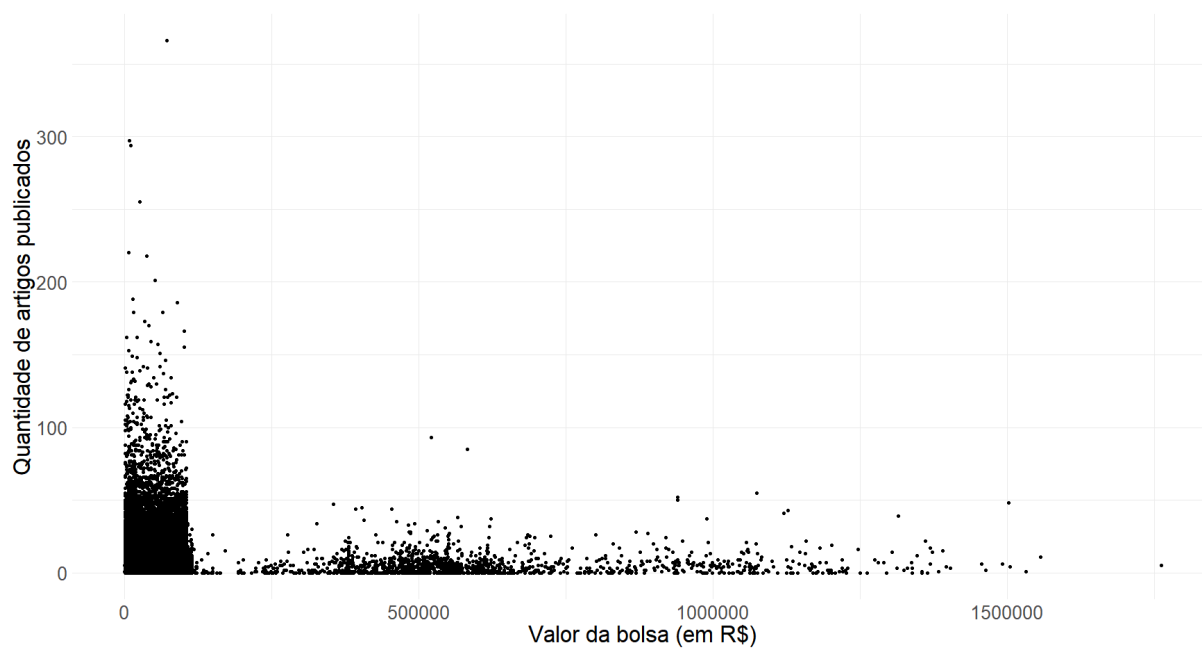
Fonte: Elaboração própria

4.8 Relação entre quantidade de artigos e investimento em bolsa

Comparando o valor da bolsa e a quantidade de artigos produzidos, observa-se que a dispersão dos dados parece estar concentrada no eixo inferior do gráfico, embora haja uma maior distribuição em relação ao eixo Y quanto menor for o valor no eixo X. Também não há um comportamento linear dos dados, com a nuvem de dados apresentando um comportamento de crescimento ou decréscimo entre os eixos. A presença de um comportamento linear determina que há relação entre as variáveis – ou seja o crescimento da variável no eixo X influencia no crescimento/decréscimo linear da variável no eixo Y. Assim, pode-se entender que não há relação linear entre o valor da bolsa (eixo X) e a quantidade de artigos publicados (eixo Y) pelas bolsas analisadas (Figura 33).

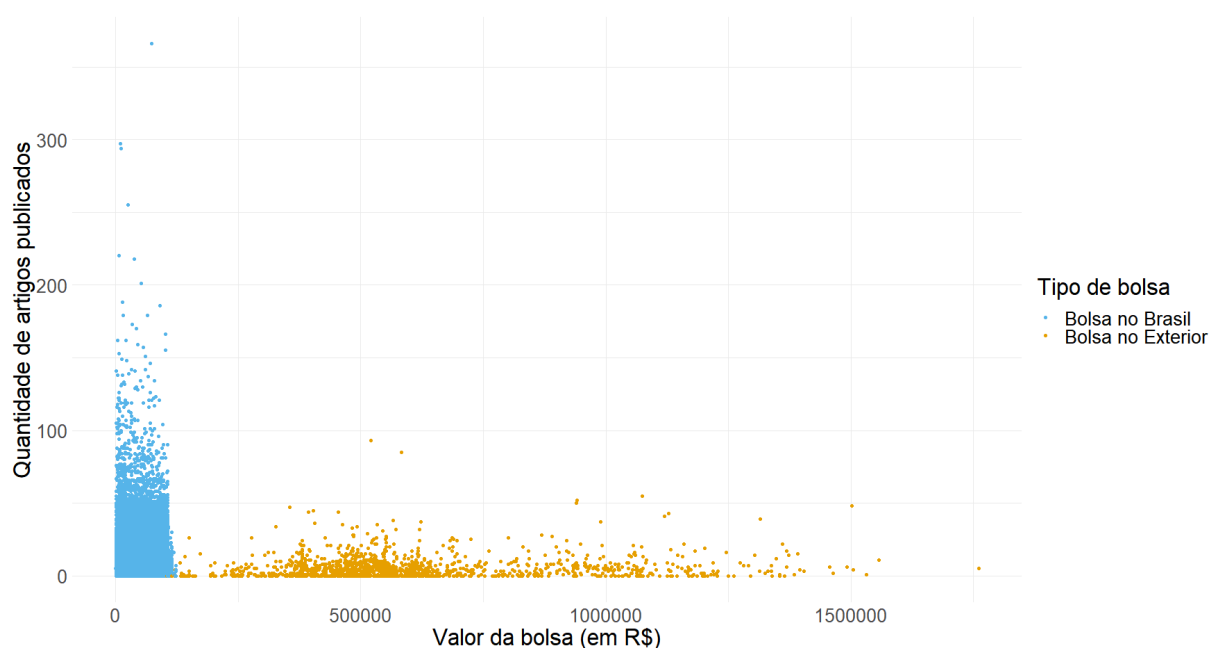
Separando os dados segundo o tipo de bolsa (Figura 34), é possível visualizar que, ainda que não exista relação linear entre as variáveis, bolsas no Brasil resultam em mais trabalhos publicados e uma menor aplicação de recursos, enquanto as bolsas no exterior custam financeiramente mais sem atingir o mesmo patamar de produção que as bolsas no país alcançam. Essa diferença é ainda mais visível calculando o custo de produção pela razão entre o valor da bolsa e a quantidade de artigos publicados. Em média, a produção de um artigo no Brasil custa uma bolsa total de R\$15.585,00, enquanto uma publicação de um bolsista no exterior custa uma bolsa completa de R\$195.832,00.

Figura 33 – Quantidade de artigos publicados e valor da bolsa (em reais)



Fonte: Elaboração própria

Figura 34 – Quantidade de artigos publicados e valor da bolsa (em reais) segundo tipo de bolsa

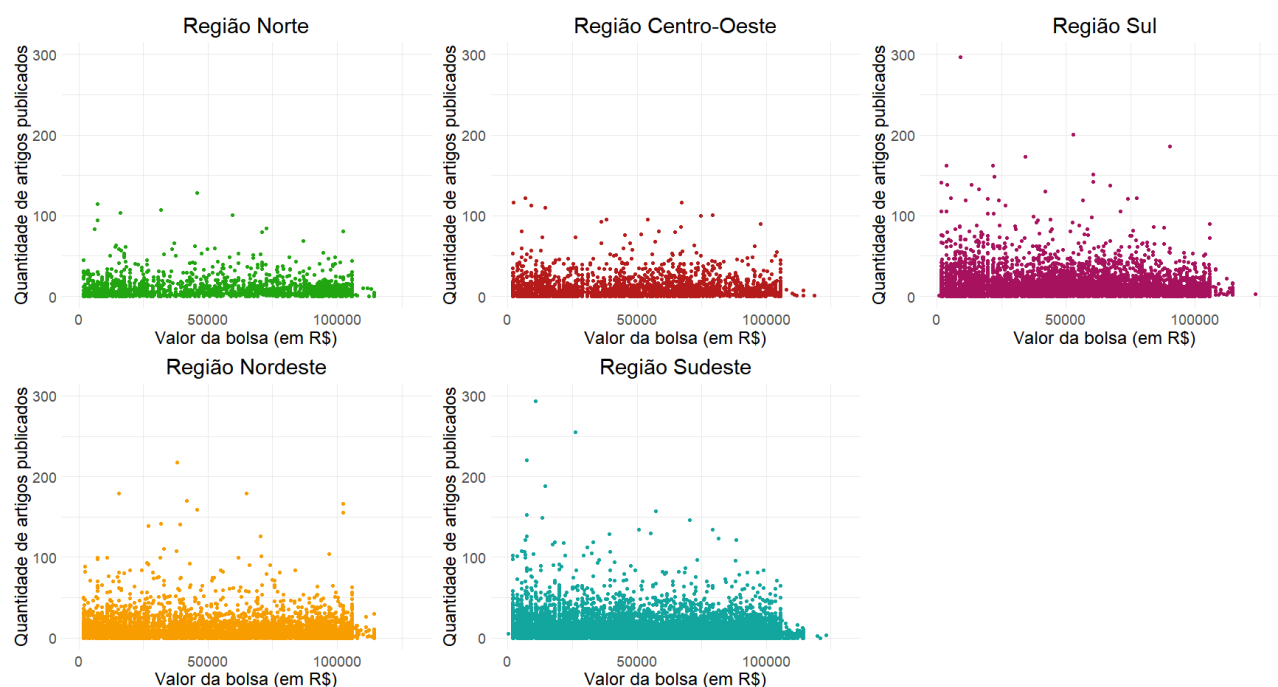


Fonte: Elaboração própria

O mesmo comportamento pode ser observado entre os outros tipos de produções (capítulos de livros, livros, patentes, softwares e produtos tecnológicos – ver Anexo A).

Para as bolsas no país, é possível analisar a relação de publicação de artigos pelo valor da bolsa, por região, e constatar que não há uma diferença muito grande quanto ao custo da bolsa, mas as regiões Nordeste, Sudeste e Sul conseguem alcançar um número maior de publicações, e concentram uma quantidade maior de bolsas – 6.214, 14.991 e 6.862, respectivamente, contra 1.792 bolsas no Norte e 2.583 no Centro-Oeste. No entanto, Sul e Centro-Oeste são as regiões que mais produzem artigos por bolsa – 12,60 e 10,40, respectivamente, contra 9,69 no Nordeste, 9,86 no Norte e 9,05 no Sudeste. Um total de 19.728 bolsas no país não teve sua UF informada na planilha e, por isso, não foram consideradas nesta análise.

Figura 35 – Quantidade de artigos publicados e valor da bolsa DPB (em reais) segundo região



Fonte: Elaboração própria

4.9 Desempenho dos Programas de bolsa da CAPES

O PPA, previsto no art. 165 da CF/88, é o instrumento de planejamento governamental de médio prazo que define as políticas públicas do governo para o período de quatro anos, sendo fundamental para organizar a ação pública, com a intenção de cumprir os fundamentos e os objetivos da República. O PPA define as prioridades do governo e designa, de maneira regionalizada, as diretrizes, os objetivos e as metas da Administração Pública, e é estruturado em programas e ações que devem resultar em serviços e bens para a sociedade em diversas áreas.

De acordo com os relatórios de gestão dos exercícios de 2010 e 2011 da CAPES, o Programa de Governo 1375 estava sob a responsabilidade da CAPES, e traduzia a missão e a visão desta entidade pública. Em linhas gerais, tem-se que a missão da CAPES é subsidiar o Ministério da Educação na formulação de políticas públicas educacionais direcionadas à pós-graduação, por meio de ações que estimulem a formação de recursos humanos altamente qualificados para a docência, a pesquisa e o atendimento das demandas por pessoal qualificado dos setores público e privado, bem como pela avaliação dos cursos de pós-graduação *stricto sensu* oferecidos no Brasil, a fim de promover o crescimento e o aperfeiçoamento do SNPG como um todo. Algumas especificações do Programa 1375 estão descritas no Quadro 2.

Quadro 2 – Programa de Governo 1375 sob responsabilidade da CAPES

Identificação do Programa de Governo	
Código no PPA: 1375	Desenvolvimento do Ensino da Pós-Graduação e da Pesquisa Científica
Tipo de Programa	Finalístico
Objetivo Geral	Formar pessoal de alto nível no país e no exterior, com vistas à produção do conhecimento científico, para a solução dos grandes desafios educacionais, econômicos e sociais do Brasil.
Objetivos Específicos	Acompanhamento e a avaliação dos Programas de Pós-Graduação realizada pela CAPES, cujo resultado traduz-se em diagnósticos que subsidiam esta Fundação na formulação de políticas de pós-graduação, fomentando o ensino de pós-graduação nos seguintes aspectos: na concessão de bolsas de estudo no país e no exterior; a manutenção dos programas de pós-graduação; o acesso ao acervo bibliográfico nacional e internacional; e, no cumprimento de acordos internacionais.

Fonte: Adaptado do Relatório de Gestão CAPES 2011 (BRASIL, 2012b)

Importante mencionar que o Relatório de Gestão do exercício de 2012 da CAPES não faz menção ao Programa 1375, mas foi possível identificar que as bolsas contempladas neste estudo estavam inseridas no Programa 2032, do tipo Temático. No Quadro 3 tem-se as especificações deste Programa.

Quadro 3 – Programa de Governo 2032 sob responsabilidade da CAPES

Identificação do Programa de Governo	
Código no PPA: 2032	Educação Superior - Graduação, Pós-Graduação, Ensino, Pesquisa e Extensão
Tipo de Programa	Temático
Iniciativa	Concessão de bolsas, auxílios e outros mecanismos, no país e no exterior, para a formação, valorização e capacitação de recursos humanos e para promover cooperação internacional no sistema nacional de pós-graduação, em áreas de interesse nacional e regional, inclusive formação de professores, garantida equidade étnico-racial e de gênero
Objetivo	Apoiar a formação de pessoal qualificado em nível superior para fortalecer o sistema nacional de educação, contribuindo para a melhoria da educação básica e para o fortalecimento e o crescimento da ciência, da tecnologia e da inovação, visando ao desenvolvimento sustentável do Brasil

Fonte: Adaptado do Relatório de Gestão CAPES 2012 (BRASIL, 2013)

Um programa de governo é composto por ações, que são definidas como

“operação da qual resultam produtos (bens ou serviços) que contribuem para atender ao objetivo de um programa”, segundo o Manual Técnico do Orçamento 2019 (BRASIL, 2019a). Assim, os Programas contemplam diversas ações na CAPES, das quais se destacam neste estudo as Ações sob os códigos 0487, 09GK e 0935, pois são nessas ações em que estão inseridas as bolsas analisadas por esta pesquisa. Nas Tabelas 3, 4 e 5 é possível verificar as execuções física e orçamentária de cada uma dessas ações nos anos de 2010 a 2014.

Tabela 3 – Execução da Ação 0487 no período de 2010 a 2014

Ação 0487	2010	2011	2012	2013*	2014
Bolsas concedidas (doutorado pleno)	19.940	22.327	27.589	32.157	39.954
Execução orçamentária	R\$ 841.851.352,23	R\$ 1.154.261.451,91	R\$ 1.273.683.316,05	R\$ 1.953.283.265,94	R\$ 2.289.780.730,38

* Não incluídas as 910 bolsas de doutorado pleno concedidas pelo Programa CsF e o respectivo orçamento executado para essas bolsas.

** O valor da execução orçamentária é referente a toda a Ação.

Fonte: Adaptado dos Relatórios de Gestão CAPES (2011, 2012b, 2013, 2014 e 2015)

Na Ação 0487 estão inseridas as bolsas no país contempladas neste estudo oriundas dos seguintes programas de concessão: DS, MINTER/DINTER CAPES/SE-TEC (MINDIN), DINTER, Programa Especial para Doutorado em Pesquisa Médica, Doutorado Pleno CSF (BEX) e parte das bolsas concedidas pelo Programa PROEX. Importante ressaltar que no ano de 2013 teve início a concessão de bolsas pelo programa Ciências sem Fronteiras (CsF), o que contribuiu para incrementar o número de bolsas concedidas bem como o montante investido. Não foram incluídas no cômputo das bolsas concedidas um total de 910 (novecentas e dez) bolsas de doutorado pleno implementadas pelo CsF no ano de 2013. O valor da execução orçamentária de todo o programa CsF declarado no Relatório de Gestão da CAPES de 2013 foi de R\$ 1.349.154.222,00.

Os valores da execução financeira apenas das bolsas de doutorado pleno não estão expressos nos relatórios de gestão da CAPES, não sendo possível verificar o valor exato dos recursos financeiros despendidos com essa modalidade de bolsa. Contudo, é possível verificar que há um crescimento da concessão no período estudado, da mesma forma, há incremento dos valores executados com a Ação como um todo.

Como pode ser visto na Tabela 4, na Ação 09GK há um aumento no número de bolsas de doutorado pleno concedidas e no orçamento executado entre os anos de 2010 e 2011, porém a análise fica prejudicada devido à falta de informações da Ação nos anos de 2012 a 2014 nos relatórios de gestão da CAPES. O relatório de gestão de 2012 não menciona a Ação 09GK, e os relatórios dos anos de 2013 e 2014 trazem apenas informações relativas aos Restos a Pagar. É possível que essa Ação tenha sido descontinuada, e outra Ação tenha incorporado os programas de concessão de bolsas da Ação 09GK. As bolsas no país dos programas PROEX e PROEX/TAXAS

estavam abarcadas por esta Ação e, de acordo com os relatórios de gestão, as bolsas desse programa foram incorporadas pela Ação 0487. Assim como descrito no Quadro da Ação 0487, também o valor da execução orçamentária da Ação 09GK refere-se ao gasto com toda a Ação, não apenas com as bolsas de doutorado pleno.

Tabela 4 – Execução da Ação 09GK no período de 2010 a 2014

Ação 09GK	2010	2011	2012	2013*	2014*
Bolsas concedidas (doutorado pleno)	2.002	2.243	-	-	-
Execução orçamentária	R\$ 86.797.659,00	R\$ 117.142.618,64	-	-	-

* A Ação 09GK foi citada apenas no tópico de Restos a Pagar nos anos de 2013 e 2014.

** O valor da execução orçamentária é referente a toda a Ação.

Fonte: Adaptado dos Relatórios de Gestão CAPES (2011, 2012b, 2013, 2014 e 2015)

No âmbito da Ação 0935, o relatório de gestão de 2013 não expressa o valor total da execução orçamentária, mas foi possível resgatar a quantidade de bolsas no exterior do programa de doutorado pleno, como pode ser visto na Tabela 5. No ano de 2014, é possível que a Ação 0935 tenha sido descontinuada e que outras Ações tenham assumido as bolsas inseridas nesta Ação, em que estão inseridos os seguintes programas do estudo: DOC-PLENO e DPE - Programa de Doutorado Pleno no Exterior.

Tabela 5 – Execução da Ação 0935 no período de 2010 a 2014

Ação 0935	2010	2011	2012	2013	2014*
Bolsas concedidas (doutorado pleno)	481	425	507	449	-
Execução orçamentária	R\$ 80.044.230,47	R\$ 141.385.842,12	R\$ 549.213.253,36	-	-

* A Ação 0935 foi citada apenas no tópico de Restos a Pagar no ano de 2014.

** O valor da execução orçamentária é referente a toda a Ação.

Fonte: Adaptado dos Relatórios de Gestão CAPES (2011, 2012b, 2013, 2014 e 2015)

Também na Ação 0935 o entendimento e a análise dos dados restam prejudicados tendo em vista a falta de informações claras e expressas nos relatórios de gestão da CAPES. Ainda assim, é possível confirmar que a concessão de bolsa no país é significativamente maior que a concessão no exterior e sempre permaneceu crescente no período de 2010 a 2014. Com as bolsas no exterior esse crescimento constante não é verificado, pois houve redução da concessão em 2011, em relação ao ano anterior, e também em 2013, quando comparado com a quantidade de bolsas concedidas em 2012.

Nas Tabelas 6, 7 e 8 podem ser verificados os valores financeiros previsto e executado para a Ação 0487 nos anos de 2010 a 2014, para a Ação 09GK nos anos de 2010 e 2011 e para a Ação 0935 nos anos de 2010 a 2012, conforme dados disponíveis nos relatórios de gestão da CAPES.

Tabela 6 – Meta financeira da Ação 0487 no período de 2010 a 2014

Ação 0487	2010	2011	2012	2013*	2014
Previsão	R\$ 868.801.241,00	R\$ 1.158.452.243,00	R\$ 1.655.517.621,00	R\$ 3.024.856.065,51	R\$ 4.003.861.071,00
Execução	R\$ 841.851.352,23	R\$ 1.154.261.451,91	R\$ 1.273.683.316,05	R\$ 3.018.364.931,58	R\$ 4.000.334.238,75
% Executado	96,9	99,6	76,9	99,8	99,9

* Foram incluídas as bolsas concedidas pelo Programa CsF

Fonte: Adaptado dos Relatórios de Gestão CAPES (2011, 2012b, 2013, 2014 e 2015)

Tabela 7 – Meta financeira da Ação 09GK no período de 2010 a 2014

Ação 09GK	2010	2011	2012	2013	2014
Previsão	R\$ 86.797.659,00	R\$ 117.176.840,00	-	-	-
Execução	R\$ 86.797.659,00	R\$ 117.142.618,64	-	-	-
% Executado	100	99,97	-	-	-

Fonte: Adaptado dos Relatórios de Gestão CAPES (2011, 2012b, 2013, 2014 e 2015)

Tabela 8 – Meta financeira da Ação 0935 no período de 2010 a 2014

Ação 0935	2010	2011	2012	2013	2014
Previsão	R\$ 118.822.823,00	R\$ 126.894.236,00	R\$ 639.819.591,00	-	-
Execução	R\$ 80.044.230,47	R\$ 141.385.842,12	R\$ 549.213.253,36	-	-
% Executado	67,4	111,4	85,8	-	-

Fonte: Adaptado dos Relatórios de Gestão CAPES (2011, 2012b, 2013, 2014 e 2015)

Nota-se que o percentual executado das Ações 0487 (bolsas no país e parte das bolsas do Programa CsF) e 09GK (bolsas no país) é muito próximo do total, com exceção do ano de 2012, em que foram executados 76,9% do orçamento na Ação 0487. Já na Ação 0935 (bolsas no exterior), verifica-se a execução mais distante do total nos anos de 2010 e 2012, contudo a execução da meta foi superada em 11,4% em 2011.

Os dados apresentados nos relatórios de gestão da CAPES revelam a atuação desta entidade pública no fortalecimento do SNPG por meio da concessão de bolsas de estudo e de recursos de custeio e capital aos programas de pós-graduação no Brasil, favorecendo a formação de pessoal qualificado para atender as demandas da pesquisa, da ciência e da educação brasileiras, contribuindo também com a produção acadêmico-científica.

Também é possível confirmar o fomento à pós-graduação e a efetividade das ações da CAPES neste estudo, que mostra que os bolsistas têm produção acadêmica significativa e relevante para o país, além de corroborar com os dados dos relatórios de gestão no que tange à aplicação de recursos com bolsas no país quando comparado com o investimento em bolsas no exterior. O montante destinado à concessão de bolsas no Brasil é maior, porém o valor médio individual é maior para as bolsas de estudo no exterior, como consta na Tabela 9.

Tabela 9 – Estatísticas do valor da bolsa segundo o tipo de bolsa

Tipo de bolsa	Mínimo	Média	Mediana	Desvio-padrão	Máximo	Coef. de Variação
Bolsa no Brasil	R\$200,00	R\$55.090,64	R\$55.000,00	R\$34.161,34	R\$123.200,00	62,01%
Bolsa no Exterior	R\$60.905,76	R\$591.964,42	R\$522.583,08	R\$260.380,60	R\$1.760.881,20	43,99%

Fonte: Elaboração própria

5 CONCLUSÕES

A etapa de avaliação das políticas públicas é imprescindível para verificar se essas estão atendendo às necessidades da população, ou solucionando os problemas existentes, uma vez que são esses seus objetivos. Para tratar do tema políticas públicas é necessário englobar também a avaliação, o desempenho, os resultados alcançados, as melhorias dos procedimentos de acompanhamento e monitoramento dessas políticas.

Com efeito, observa-se um grande volume de discussão acerca da importância da etapa de avaliação de programas, projetos e ações, com vistas ao acompanhamento contínuo e sistemático do desempenho das políticas públicas e com o objetivo de racionalizar o uso de recursos públicos.

Com as transformações experimentadas pela sociedade, e considerando o advento de novos meios de Tecnologia de Informação e Comunicação (TIC), também a Administração Pública deve se atualizar, utilizando-se dessas novas tecnologias para manter o atendimento aos anseios da sociedade de maneira perene e satisfatória (BRAGA; GOMES, 2016). Para os autores, nesse cenário de constantes transformações, a chamada “Governança Responsiva”, também denominada *“New Public Governance”* (OSBORNE, 2006), é proposta como um novo modelo de Administração Pública.

Por “Governança Responsiva” entende-se como sendo um “modelo orientado por princípios de responsabilização, transparência e participação, no qual a efetividade se vincula a seu atributo-chave: a responsividade” (BRAGA; GOMES, 2016, p. 487), cujo conceito está ligado à efetividade governamental. A busca pela eficiência, eficácia e efetividade na utilização dos recursos financeiros, humanos, materiais, naturais se faz cada vez mais indispensável no contexto de insuficiência de recursos.

Sob a ótica da sociedade, o bom uso dos meios de TIC contribui para uma maior participação dos cidadãos no ciclo de políticas públicas, na medida em que elas facilitam o acesso a informações, de forma a conhecer em que e como os recursos públicos estão sendo utilizados, bem como qual o retorno desse investimento para a sociedade, além de outros estágios das políticas. Sob o ponto de vista da Administração Pública, os novos modelos eletrônicos contribuem para direcionar os rumos das políticas públicas, dar transparência das ações executadas à sociedade, responsabilizar os agentes públicos, aperfeiçoar o processo de tomada de decisão, entre outros.

Convém salientar que a CAPES, por meio da plataforma Sucupira, e o CNPq, por meio da plataforma *Lattes*, utilizam dos meios de TIC para coletar e manter dados de relevância para o SNPG brasileiro, sem os quais este estudo não teria sido pos-

sível. Contudo, o acesso às informações pode ser ainda mais ampliado e facilitado à sociedade, bem como é possível aperfeiçoar os procedimentos de coleta de dados e a integração entre os bancos das diversas entidades do governo.

Outra dificuldade encontrada neste estudo foi a apresentação de informações não tão claras e a ausência de dados mais detalhados nos relatórios de gestão da CAPES, o que prejudicou a análise da relação entre a quantidade de bolsa concedida e os valores executados com as bolsas de doutorado pleno apenas.

A CAPES tem contribuído de forma efetiva para o crescimento do SNPG no Brasil com a execução das metas previstas nos PPAs ao longo dos anos, promovendo o aumento da qualidade e do desenvolvimento da pós-graduação, expandindo o fomento aos cursos de maneira crescente no período de 2010 a 2014, como evidencia os dados dos relatórios de gestão e os dados desta pesquisa.

De acordo com os dados dos relatórios de gestão, tem-se que o volume de recursos financeiros investidos pelas ações de fomento da CAPES é maior com os programas de concessão no Brasil, mas quando é feito um cálculo aproximado do valor proporcional por bolsa, verifica-se que o gasto individual com a bolsa para estudo no exterior é maior que o gasto por bolsista no país. Isso também é atestado com as análises feitas por este estudo.

Os resultados apresentados nesta pesquisa podem contribuir com a política de concessão de bolsa da CAPES, no sentido de aperfeiçoar os critérios para a alocação e distribuição dos recursos financeiros para o fomento das bolsas, bem como no processo de avaliação e dos indicadores de desempenho das políticas públicas educacionais de fomento à pós-graduação.

A CAPES utiliza diversos indicadores de desempenho para avaliar as políticas públicas por ela implementadas, tais como o número de alunos matriculados nos cursos de mestrado e doutorado, o índice de mestre e doutores titulados, o número de bolsistas no exterior, o número de artigos indexados, entre outros. No indicador de produção científica indexada, conta-se o quantitativo geral da produção intelectual dos programas de pós-graduação, não havendo separação das publicações feitas por beneficiários de bolsa de estudo. Este estudo pode contribuir para o aperfeiçoamento do indicador de desempenho de produção científica, por meio de comparação da produção dos dois grupos – bolsistas e não bolsistas.

As análises dos dados apontam que a hipótese deste estudo foi atendida: a concessão de bolsa de maior valor não implica mais publicações. Ou seja, não necessariamente quanto maior o valor recebido maior será o número de artigos publicados. Os bolsistas que mais publicaram no período estudado não receberam bolsas tão altas, porém tiveram publicação significativa, o que é atestado pelas análises feitas utilizando

a quantidade de artigos produzidos e o valor da bolsa, sendo essa constatação reforçada pelas análises feitas com as demais produções acadêmicas (livro, capítulos de livros, *software*, patente e produto tecnológico), mas que podem ser objeto de futuras pesquisas.

Importante destacar que esta pesquisa não contemplou a qualidade e o impacto das publicações, mas tão somente a quantidade. Futuros estudos poderão avaliar a qualidade das publicações, de acordo com o fator de impacto das revistas e dos periódicos nos quais os artigos foram publicados, por exemplo.

Diante dos resultados apontados, é natural refletir sobre o destino dos recursos públicos e a contrapartida para a sociedade. Convém ponderar a relação de custo-benefício das bolsas no país e no exterior, a fim de aperfeiçoar o SNPG brasileiro com a melhor utilização possível de recursos financeiros dos cofres públicos.

É nítido que a produção acadêmico-científica produzida pelos bolsistas não é o único elemento de contrapartida para o país, havendo outros aspectos a serem levados em consideração como retorno positivo para o Brasil. Contudo, se o único indicador para a concessão de bolsas fosse a quantidade de artigos publicados, com o mesmo montante investido em bolsa no exterior seria possível incrementar o valor da bolsa no país ou conceder mais bolsas no país, a fim de atingir um número maior de estudantes.

Pensar em “Governança Responsiva” como forma de alcançar mais efetividade com os programas da CAPES de concessão de bolsa pode ser uma forma de garantir uma utilização mais efetiva dos recursos públicos.

REFERÊNCIAS

- ABRUCIO, F. L. Os avanços e os dilemas do modelo pós-burocrático: a reforma da administração pública à luz da experiência internacional recente. *Reforma do Estado e administração pública gerencial*, FGV Rio de Janeiro, v. 7, p. 173–200, 1998. Citado na página 25.
- ABRUCIO, F. L. A dinâmica federativa da educação brasileira: diagnóstico e propostas de aperfeiçoamento. *Educação e federalismo no Brasil: combater as desigualdades, garantir a diversidade*. Brasília: UNESCO, v. 20102010, p. 01–29, 2010. Citado na página 17.
- ALA-HARJA, M.; HELGASON, S. Em direção às melhores práticas de avaliação. *Revista do Serviço Público*, v. 51, n. 4, p. 5–60, 2000. Citado 2 vezes nas páginas 27 e 28.
- ALMEIDA, L. d. A.; GOMES, R. C. Processo das políticas públicas: revisão de literatura, reflexões teóricas e apontamentos para futuras pesquisas. *Cadernos EBAPE. BR*, SciELO Brasil, v. 16, n. 3, p. 444–455, 2018. Citado na página 23.
- AZEVEDO, J. M. L. d. *A educação como política pública*. 3 ed.. ed. [S.l.]: Autores Associados, 2004. Citado 2 vezes nas páginas 18 e 23.
- BALBE, R. D. S. Controle interno e o foco nos resultados. In: BELO HORIZONTE: FÓRUM. [S.l.], 2013. Citado na página 29.
- BATISTA, F. F.; POPINIGIS, F. Elaboração de indicadores de desempenho institucional. *São Paulo: Instituto Serzedelo Correa*, 1999. Citado na página 32.
- BRAGA, L. V.; GOMES, R. C. Participação eletrônica e suas relações com governo eletrônico, efetividade governamental e accountability. *Organizações & Sociedade*, SciELO Brasil, v. 23, n. 78, p. 487–506, 2016. Citado 2 vezes nas páginas 17 e 73.
- BRASIL. Plano diretor da reforma do aparelho do estado. *Brasília: Câmara da Reforma do Estado*, 1995. Disponível em <<http://www.biblioteca.presidencia.gov.br/publicacoes-oficiais/catalogo/fhc/plano-diretor-da-reforma-do-aparelho-do-estado-1995.pdf>>. Acesso em: 22 jul. 2019. Citado na página 17.
- BRASIL. Relatório de gestão do exercício de 2010. *Brasília: Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES*, 2011. Disponível em: <<https://cutt.ly/9nk7OHI>>. Acesso em: 10 abr. 2021. Citado 3 vezes nas páginas 69, 70 e 71.
- BRASIL. Lei nº 12.593, de 18 de janeiro de 2012. 2012a. Disponível em: <<https://cutt.ly/NngwXf2>>. Acesso em: 30 out. 2020. Citado na página 21.
- BRASIL. Relatório de gestão do exercício de 2011. *Brasília: Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES*, 2012b. Disponível em: <<https://cutt.ly/Vnk7iGa>>. Acesso em: 10 abr. 2021. Citado 4 vezes nas páginas 68, 69, 70 e 71.

BRASIL. Relatório de gestão do exercício de 2012. *Brasília: Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES*, 2013. Disponível em: <<https://cutt.ly/nnk7LrM>>. Acesso em: 10 abr. 2021. Citado 4 vezes nas páginas 68, 69, 70 e 71.

BRASIL. Relatório de gestão do exercício de 2013. *Brasília: Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES*, 2014. Disponível em: <<https://cutt.ly/Nnk7EZH>>. Acesso em: 10 abr. 2021. Citado 3 vezes nas páginas 69, 70 e 71.

BRASIL. Relatório de gestão do exercício de 2014. *Brasília: Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES*, 2015. Disponível em: <<https://cutt.ly/Onk708Y>>. Acesso em: 10 abr. 2021. Citado 3 vezes nas páginas 69, 70 e 71.

BRASIL. Avaliação de políticas públicas - guia prático de análise ex ante. *Brasília: Casa Civil da Presidência da República. Instituto de Pesquisa Econômica Aplicada – IPEA*, Vol 1, p. 192 p., 2018a. Disponível em: <<https://cutt.ly/Onh5hym>>. Acesso em: 05 ago. 2019. Citado 2 vezes nas páginas 30 e 31.

BRASIL. Avaliação de políticas públicas - guia prático de análise ex post. *Brasília: Casa Civil da Presidência da República. Instituto de Pesquisa Econômica Aplicada – IPEA*, Vol 2, p. 301 p., 2018b. Disponível em <<https://cutt.ly/qnlxZs9>>. Acesso em: 05 ago. 2019. Citado na página 25.

BRASIL. Manual técnico de orçamento – mto 2019. *Brasília: Ministério da Economia*, 2019a. Disponível em: <<https://www1.siop.planejamento.gov.br/mto/lib/exe/fetch.php/mto2019:mto2019-versao9.pdf>>. Acesso em: 15 mai. 2021. Citado na página 69.

BRASIL. Relatório de gestão do exercício de 2018. *Brasília: Ministério da Educação. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES*, 2019b. Disponível em: <<https://cutt.ly/DnlWKvu>>. Acesso em: 02 out. 2020. Citado na página 19.

CASTRO, J. A. de et al. A cf/88 e as políticas sociais brasileiras. In: *A Constituição Brasileira de 1988 Revisitada: Recuperação histórica e desafios atuais das políticas públicas nas áreas econômica e social*. [S.l.]: IPEA, 1988. Citado na página 18.

CNPQ. *Sobre a plataforma Lattes*. 2020. Disponível em: <http://lattes.cnpq.br/>. Acesso em: 30 mar. 2021. Citado na página 36.

COHEN, E.; FRANCO, R. *Avaliação de projetos sociais*. 6 ed.. ed. [S.l.]: Petrópolis, RJ: Vozes, 2004. Citado 3 vezes nas páginas 27, 31 e 32.

COSTA, F. L. d. Brasil: 200 anos de estado; 200 anos de administração pública; 200 anos de reformas. *Revista de Administração Pública*, SciELO Brasil, v. 42, n. 5, 2008. Disponível em <<http://www.scielo.br/pdf/rap/v42n5/a03v42n5.pdf>>. Acesso em 22 jul. 2019. Citado na página 18.

COSTA, F. L. da; CASTANHAR, J. C. Avaliação de programas públicos: desafios conceituais e metodológicos. *Brazilian Journal of Public Administration*, v. 37, n. 5, p. 969–a, 2003. Disponível em

<<http://bibliotecadigital.fgv.br/ojs/index.php/rap/article/view/6509/5093>>. Acesso em: 10 ago. 2019. Citado na página 29.

COUTO, L. F.; CARDOSO JÚNIOR, J. C. P. A função dos planos plurianuais no direcionamento dos orçamentos anuais: avaliação da trajetória dos ppas no cumprimento da sua missão constitucional e o lugar do ppa 2020-2023. Instituto de Pesquisa Econômica Aplicada (Ipea), 2020. Citado na página 18.

EASTON, D. A framework for political analysis. *Englewood Cliffs, HJ: Prentice-Hall*, 1965. Citado na página 17.

FARIA, R. M. Avaliação de programas sociais: evoluções e tendências. In: *Avaliação em Políticas Públicas Sociais: uma questão em debate*. [S.l.]: São Paulo: Editora Cortez, 1998. Citado na página 26.

FILGUEIRAS, F. d. B. Burocracias do controle, controle da burocracia e accountability no brasil. In: P., G. L. e. V. O. R. (Ed.). *Burocracia e políticas públicas no Brasil: intersecções analíticas*. [S.l.]: Instituto de Pesquisa Econômica Aplicada (Ipea), 2018. Disponível em <<http://repositorio.ipea.gov.br/handle/11058/8617>>. Acesso em: 25 set. 2020. Citado na página 28.

FREY, K. Políticas públicas: um debate conceitual e reflexões referentes à prática da análise de políticas públicas no brasil. *Planejamento e políticas públicas*, n. 21, 2000. Disponível em <<http://www.ipea.gov.br/ppp/index.php/PPP/article/view/89/158>>. Acesso em: 13 ago. 2019. Citado 2 vezes nas páginas 23 e 25.

GARCIA, R. C. Reorganização do processo de planejamento do governo federal: o ppa 2000-2003. *Texto para Discussão nº 726*, Instituto de Pesquisa Econômica Aplicada (Ipea), 2000. Disponível em <<https://cutt.ly/Onglq4p>>. Acesso em: 12 ago. 2019. Citado 2 vezes nas páginas 26 e 29.

GRAU, N. C.; BOZZI, S. O. Fortalecimento de los sistemas de monitoreo y evaluación (m&e) em américa latina: Informe comparativo de 12 países. *Natal, RNA: SEARH*, 2008. Tradução de Ernesto Montes-Bradley y Estayes. Citado 2 vezes nas páginas 25 e 26.

GUERRA, L. C. B. et al. Análise epistemológica da nova administração pública à luz de kuhn e popper. *Revista Eletrônica do Mestrado Profissional em Administração da Universidade Potiguar (RAUnP)*, v. 4, n. 1, p. 43–53, 2012. Citado na página 25.

JANNUZZI, P. d. M. *Indicadores sociais no Brasil: conceitos, fontes de dados e aplicações*. [S.l.]: Campinas: Alínea, 2001. Citado na página 30.

JANNUZZI, P. d. M. Indicadores para diagnóstico, monitoramento e avaliação de programas sociais no brasil. Escola Nacional de Administração Pública (ENAP), 2005. Disponível em <<https://cutt.ly/OngltKG>>. Acesso em: 07 ago. 2019. Citado 5 vezes nas páginas 27, 29, 30, 31 e 32.

JANNUZZI, P. d. M. *Monitoramento e avaliação de programas sociais: uma introdução aos conceitos e técnicas*. [S.l.]: Campinas: Alínea, 2016. Citado na página 17.

KUSEK, J. Z.; RIST, R. C. *Ten steps to a results-based monitoring and evaluation system: a handbook for development practitioners*. [S.l.]: World Bank Publications, 2004. Citado na página 31.

LYNN, L. E.; GOULD, S. G. *Designing public policy: a casebook on the role of policy analysis*. [S.l.]: Goodyear Publishing Company, 1980. Citado na página 23.

MAZZA, A. *Manual de direito administrativo*. [S.l.]: Saraiva Educação SA, 2020. Citado na página 18.

MINAYO, M. C. d. S. Construção de indicadores qualitativos para avaliação de mudanças. *Revista Brasileira de Educação Médica*, v. 33, p. 83–91, 2009. Disponível em <<https://cutt.ly/4nh23Cn>>. Acesso em: 06 ago. 2019. Citado na página 30.

MORETTIN, P. A.; BUSSAB, W. O. *Estatística básica*. [S.l.]: Saraiva Educação SA, 2017. Citado na página 33.

OSBORNE, S. P. The new public governance? *Public Management Review*, 2006. Citado na página 73.

PETERS, B. G. *The politics of bureaucracy*. [S.l.]: Routledge, 1995. Citado 2 vezes nas páginas 24 e 29.

PFEIFFER, P. O quadro lógico: um método para planejar e gerenciar mudanças. *Revista do Serviço Público*, v. 51, n. 1, 2000. Disponível em <<https://revista.enap.gov.br/index.php/RSP/article/view/320/326>>. Acesso em: 18 mai. 2019. Citado na página 30.

RUA, M. d. G. *Políticas Públicas*. [S.l.], 2009. Disponível em:<<https://cutt.ly/hnlWuMK>> Acesso em: 02 ago. 2019. Citado 5 vezes nas páginas 24, 25, 26, 27 e 28.

SANDER, B. Administração da educação no brasil: é hora da relevância. *Educação Brasileira*, v. 4, n. 9, 1982. Citado 2 vezes nas páginas 27 e 28.

SANO, H. *Nova Gestão Pública e accountability: o caso das organizações sociais paulistas*. Tese (Doutorado) — Fundação Getúlio Vargas, 2003. Disponível em: Acesso em: <<http://www.bresserpereira.org.br/documents/mare/os/03.10.15-DISSERTACAO-HIRO.pdf>>. Acesso em: 18 mar. 2021. Citado na página 25.

SILVA, M. O. d. Silva e. Avaliação de políticas públicas e programas sociais: teoria e prática. In: _____. [S.l.]: Veras Editora, 2014. cap. Avaliação de políticas e programas sociais: aspectos conceituais e metodológicos. Citado na página 23.

SIMON, H. A. *Administrative behavior*. [S.l.]: Simon and Schuster, 1957. Citado na página 29.

SOARES, M.; CUNHA, E. Programa esporte e lazer da cidade (pelc) - ensino a distância (ead). *Curso de Formação em Políticas Públicas de Esporte e Lazer: Módulo Avaliação*, 2015. Citado 2 vezes nas páginas 26 e 31.

SOUZA, C. Políticas públicas: uma revisão da literatura. *Sociologias*, SciELO Brasil, n. 16, p. 20–45, 2006. Citado na página 23.

SOUZA, J. F. *CNPQ*. [S.l.]: GitHub, 2018. Disponível em: <https://github.com/josefson/CNPQ>. Acesso em: 15 out. 2020. Citado na página 41.

WEISS, C. H. *Evaluation*. [S.l.]: Upper Saddle River: Prentice-Hall, 1998. Citado na página 27.

APÊNDICE A – FIGURAS E TABELAS

Tabela A.1 – Estatísticas da quantidade de artigos segundo o tipo de bolsa

Tipo de bolsa	Mínimo	Média	Mediana	Desvio-padrão	Máximo	Coef. de Variação
Bolsa no Brasil	0	8,78	5	12,06	366	137,36%
Bolsa no Exterior	0	5,66	3	7,83	93	138,33%

Fonte: Elaboração própria

Tabela A.2 – Estatísticas da quantidade de artigos segundo a área de conhecimento

Tipo de bolsa	Mínimo	Média	Mediana	Desvio-padrão	Máximo	Coef. de Variação
CIÊNCIAS AGRÁRIAS	0	11,46	8	13,06	294	114,03%
CIÊNCIAS BIOLÓGICAS	0	10,07	7	11,94	186	118,59%
CIÊNCIAS DA SAÚDE	0	13,48	9	16,76	297	124,33%
CIÊNCIAS EXATAS E DA TERRA	0	5,90	3	9,18	366	155,63%
CIÊNCIAS HUMANAS	0	6,07	4	8,46	220	139,39%
CIÊNCIAS SOCIAIS APLICADAS	0	7,31	4	10,50	142	143,55%
ENGENHARIAS	0	5,24	3	7,67	113	146,27%
LINGÜÍSTICA, LETRAS E ARTES	0	4,87	3	6,22	90	127,54%
MULTIDISCIPLINAR	0	7,02	4	9,21	129	131,19%

Fonte: Elaboração própria

Tabela A.3 – Estatísticas da quantidade de artigos segundo o tipo de bolsa e a área de conhecimento

Tipo de bolsa	Grande área	Mínimo	Média	Mediana	Desvio-padrão	Máximo	Coef. de Variação
Bolsa no Brasil	CIÊNCIAS AGRÁRIAS	0	11,51	8	13,11	294	113,85%
	CIÊNCIAS BIOLÓGICAS	0	10,18	7	12,03	186	118,07%
	CIÊNCIAS DA SAÚDE	0	13,54	9	16,83	297	124,25%
	CIÊNCIAS EXATAS E DA TERRA	0	5,96	4	9,27	366	155,56%
	CIÊNCIAS HUMANAS	0	6,10	4	8,51	220	139,51%
	CIÊNCIAS SOCIAIS APLICADAS	0	7,47	4	10,66	142	142,83%
	ENGENHARIAS	0	5,26	3	7,68	113	146,05%
	LINGÜÍSTICA, LETRAS E ARTES	0	4,90	3	6,19	90	126,36%
Bolsa no Exterior	MULTIDISCIPLINAR	0	7,08	4	9,29	129	131,17%
	CIÊNCIAS AGRÁRIAS	0	6,74	5	7,15	35	106,00%
	CIÊNCIAS BIOLÓGICAS	0	6,23	4	7,72	85	123,90%
	CIÊNCIAS DA SAÚDE	0	10,14	7	12,33	93	121,56%
	CIÊNCIAS EXATAS E DA TERRA	0	4,46	2	6,55	47	146,93%
	CIÊNCIAS HUMANAS	0	5,11	3	6,59	55	128,94%
	CIÊNCIAS SOCIAIS APLICADAS	0	4,30	2	5,67	33	131,96%
	ENGENHARIAS	0	4,83	2	7,38	52	152,96%
	LINGÜÍSTICA, LETRAS E ARTES	0	3,87	1	6,89	44	178,04%
	MULTIDISCIPLINAR	0	5,19	4	6,18	39	119,25%

Fonte: Elaboração própria

Tabela A.4 – Estatísticas do valor da bolsa segundo a grande área

Grande área	Mínimo	Média	Mediana	Desvio-padrão	Máximo	Coef. de Variação
CIÊNCIAS AGRÁRIAS	R\$ 1.800,00	R\$ 59.150,51	R\$ 55.600,00	R\$ 63.604,07	R\$ 1.046.393,84	107,53%
CIÊNCIAS BIOLÓGICAS	R\$ 1.800,00	R\$ 75.122,93	R\$ 63.800,00	R\$ 117.970,94	R\$ 1.531.837,84	157,04%
CIÊNCIAS DA SAÚDE	R\$ 200,00	R\$ 63.080,79	R\$ 52.800,00	R\$ 94.246,17	R\$ 1.760.881,20	149,41%
CIÊNCIAS EXATAS E DA TERRA	R\$ 1.800,00	R\$ 76.868,03	R\$ 57.200,00	R\$ 120.738,29	R\$ 1.463.280,48	157,07%
CIÊNCIAS HUMANAS	R\$ 1.800,00	R\$ 70.676,86	R\$ 61.400,00	R\$ 99.746,34	R\$ 1.363.706,54	141,13%
CIÊNCIAS SOCIAIS APLICADAS	R\$ 800,00	R\$ 75.057,41	R\$ 49.400,00	R\$ 128.130,46	R\$ 1.354.259,68	170,71%
ENGENHARIAS	R\$ 1.800,00	R\$ 71.859,86	R\$ 56.000,00	R\$ 103.579,40	R\$ 1.312.811,68	144,14%
LINGÜÍSTICA, LETRAS E ARTES	R\$ 1.800,00	R\$ 71.031,38	R\$ 59.800,00	R\$ 100.337,72	R\$ 1.294.929,38	141,26%
MULTIDISCIPLINAR	R\$ 1.800,00	R\$ 74.039,46	R\$ 63.800,00	R\$ 106.288,30	R\$ 1.556.243,49	143,56%

Fonte: Elaboração própria

Tabela A.5 – Estatísticas do valor da bolsa segundo tipo de bolsa e a área de conhecimento

Tipo de bolsa	Grande área	Mínimo	Média	Mediana	Desvio-padrão	Máximo	Coef. de Variação
Bolsa no Brasil	CIÊNCIAS AGRÁRIAS	R\$ 1.800,00	R\$ 53.398,89	R\$ 55.000,00	R\$ 32.629,58	R\$ 123.200,00	61,11%
	CIÊNCIAS BIOLÓGICAS	R\$ 1.800,00	R\$ 57.585,54	R\$ 61.600,00	R\$ 35.867,32	R\$ 123.200,00	62,29%
	CIÊNCIAS DA SAÚDE	R\$ 200,00	R\$ 52.753,03	R\$ 51.400,00	R\$ 33.610,89	R\$ 118.800,00	63,71%
	CIÊNCIAS EXATAS E DA TERRA	R\$ 1.800,00	R\$ 56.116,39	R\$ 53.800,00	R\$ 35.501,15	R\$ 121.000,00	63,26%
	CIÊNCIAS HUMANAS	R\$ 1.800,00	R\$ 56.854,62	R\$ 59.400,00	R\$ 33.821,84	R\$ 114.400,00	59,49%
	CIÊNCIAS SOCIAIS APLICADAS	R\$ 800,00	R\$ 49.985,35	R\$ 46.200,00	R\$ 32.300,55	R\$ 123.200,00	64,62%
	ENGENHARIAS	R\$ 1.800,00	R\$ 55.172,93	R\$ 52.800,00	R\$ 35.236,88	R\$ 123.200,00	63,87%
	LINGÜÍSTICA, LETRAS E ARTES	R\$ 1.800,00	R\$ 56.129,61	R\$ 57.400,00	R\$ 33.678,20	R\$ 114.400,00	60,00%
Bolsa no Exterior	MULTIDISCIPLINAR	R\$ 1.800,00	R\$ 58.409,81	R\$ 61.600,00	R\$ 33.358,58	R\$ 114.400,00	57,11%
	CIÊNCIAS AGRÁRIAS	R\$ 102.559,14	R\$ 515.908,61	R\$ 476.500,62	R\$ 172.443,43	R\$ 1.046.393,84	33,43%
	CIÊNCIAS BIOLÓGICAS	R\$ 83.096,76	R\$ 641.641,49	R\$ 533.352,34	R\$ 302.399,76	R\$ 1.531.837,84	47,13%
	CIÊNCIAS DA SAÚDE	R\$ 60.905,76	R\$ 632.741,67	R\$ 545.126,63	R\$ 326.702,15	R\$ 1.760.881,20	51,63%
	CIÊNCIAS EXATAS E DA TERRA	R\$ 120.262,24	R\$ 598.478,06	R\$ 531.601,08	R\$ 258.112,08	R\$ 1.463.280,48	43,13%
	CIÊNCIAS HUMANAS	R\$ 147.661,40	R\$ 612.227,07	R\$ 531.674,28	R\$ 233.353,08	R\$ 1.363.706,54	38,12%
	CIÊNCIAS SOCIAIS APLICADAS	R\$ 109.839,76	R\$ 576.135,23	R\$ 513.855,79	R\$ 245.683,40	R\$ 1.354.259,68	42,64%
	ENGENHARIAS	R\$ 76.766,56	R\$ 543.200,26	R\$ 498.329,10	R\$ 221.088,13	R\$ 1.312.811,68	40,70%
	LINGÜÍSTICA, LETRAS E ARTES	R\$ 224.514,80	R\$ 590.492,09	R\$ 570.308,69	R\$ 210.598,93	R\$ 1.294.929,38	35,66%
	MULTIDISCIPLINAR	R\$ 197.360,60	R\$ 582.210,62	R\$ 514.168,48	R\$ 277.191,27	R\$ 1.556.243,49	47,61%

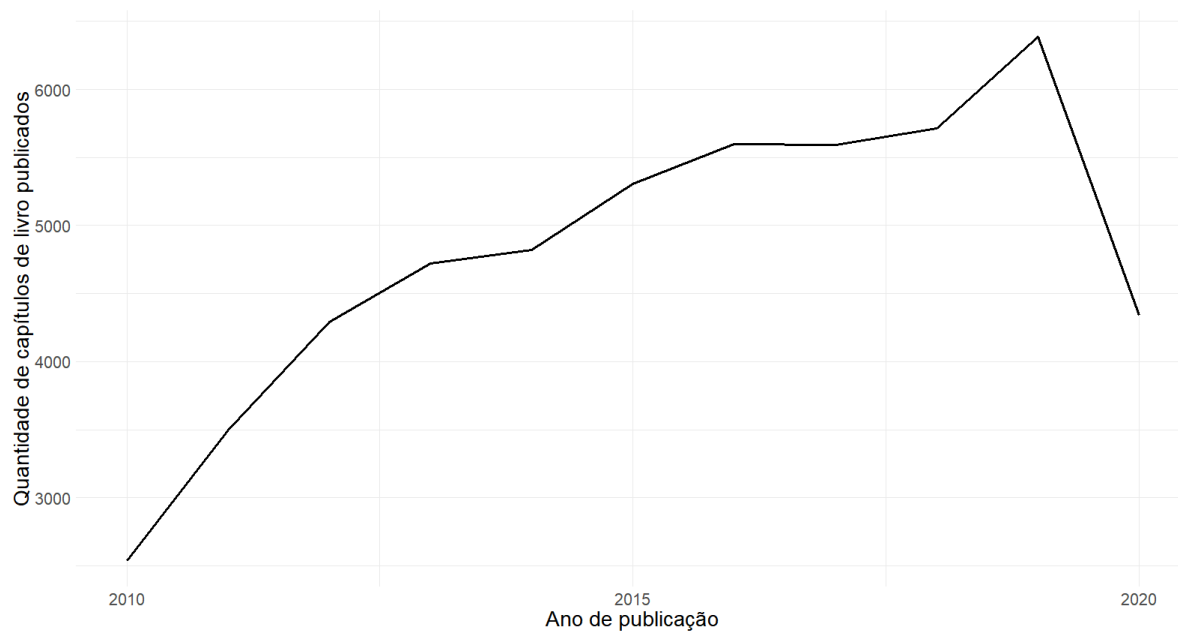
Fonte: Elaboração própria

Tabela A.6 – Estatísticas do valor da bolsa no país segundo a região

Região	Mínimo	Média	Mediana	Desvio-padrão	Máximo	Coeficiente de Variação
Centro-Oeste	R\$ 1.800,00	R\$ 57.408,59	R\$ 50.600,00	R\$ 71.620,46	R\$ 1.056.041,49	124,76%
Nordeste	R\$ 1.800,00	R\$ 61.090,69	R\$ 53.800,00	R\$ 70.261,97	R\$ 1.219.172,34	115,01%
Norte	R\$ 1.800,00	R\$ 59.256,99	R\$ 48.400,00	R\$ 78.446,42	R\$ 1.502.164,84	132,38%
Sudeste	R\$ 200,00	R\$ 60.234,90	R\$ 52.800,00	R\$ 77.395,73	R\$ 1.504.755,34	128,49%
Sul	R\$ 800,00	R\$ 58.612,89	R\$ 47.600,00	R\$ 85.968,55	R\$ 1.531.837,84	146,67%

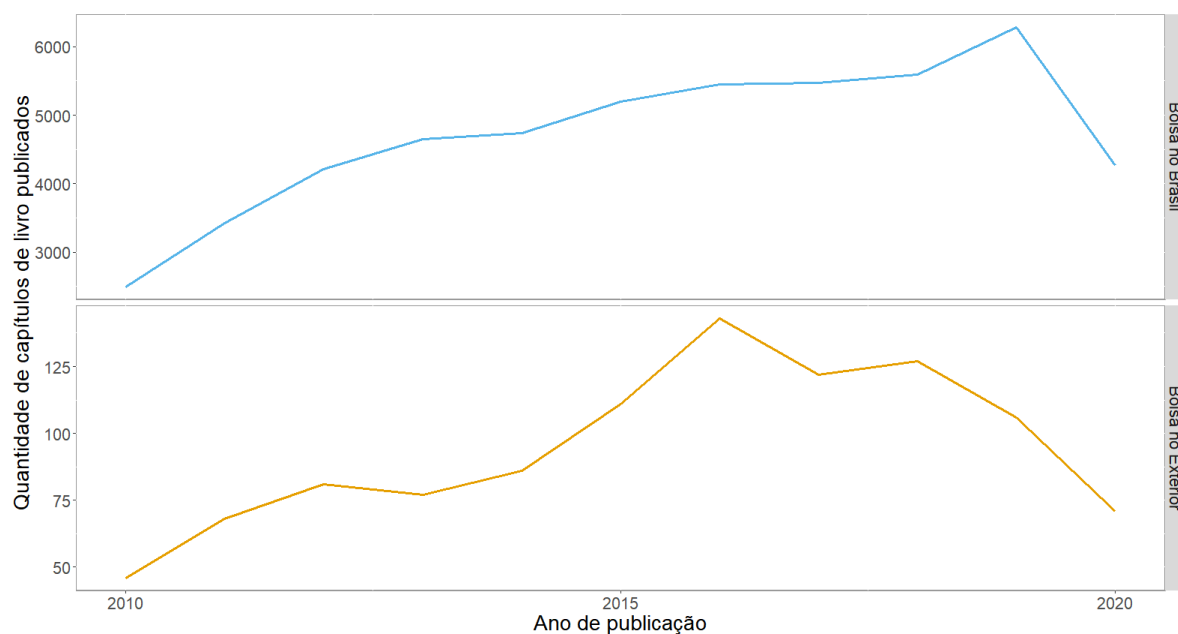
Fonte: Elaboração própria

Figura A.1 – Capítulos de livros publicados por ano



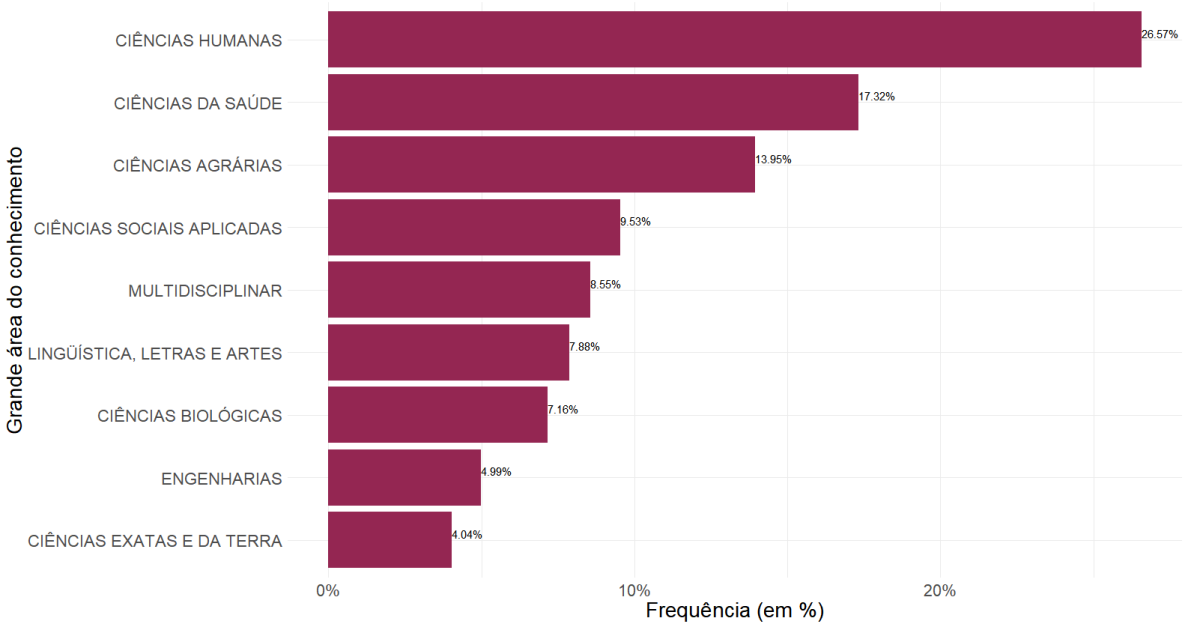
Fonte: Elaboração própria

Figura A.2 – Capítulos de livros por ano e tipo de bolsa



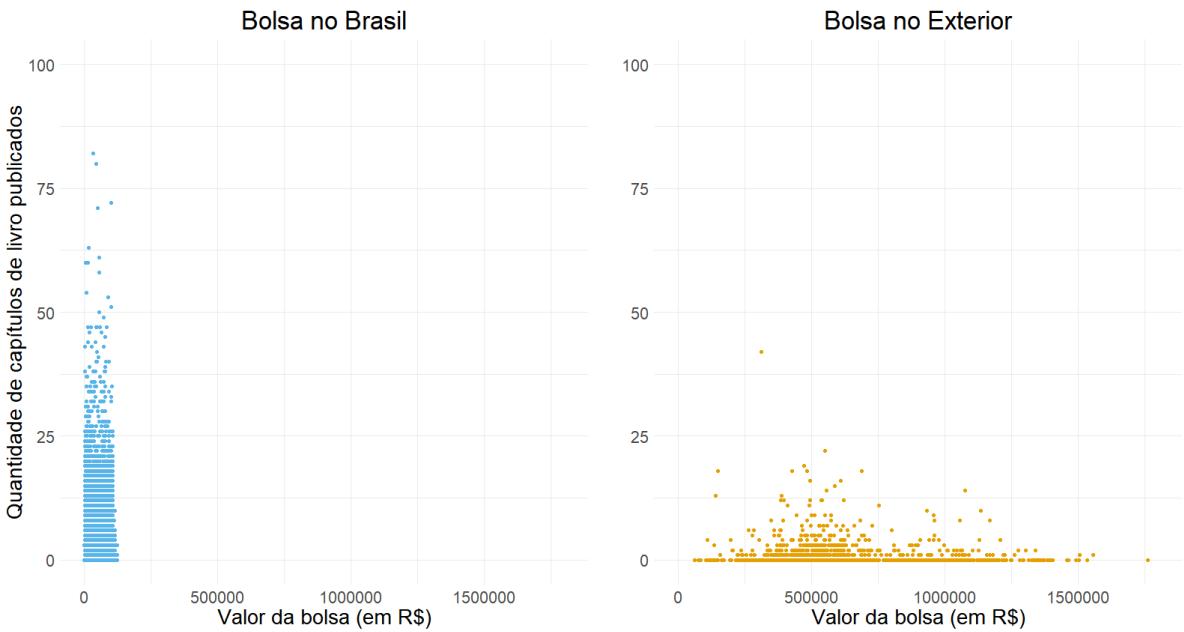
Fonte: Elaboração própria

Figura A.3 – Capítulos de livros por grande área de conhecimento



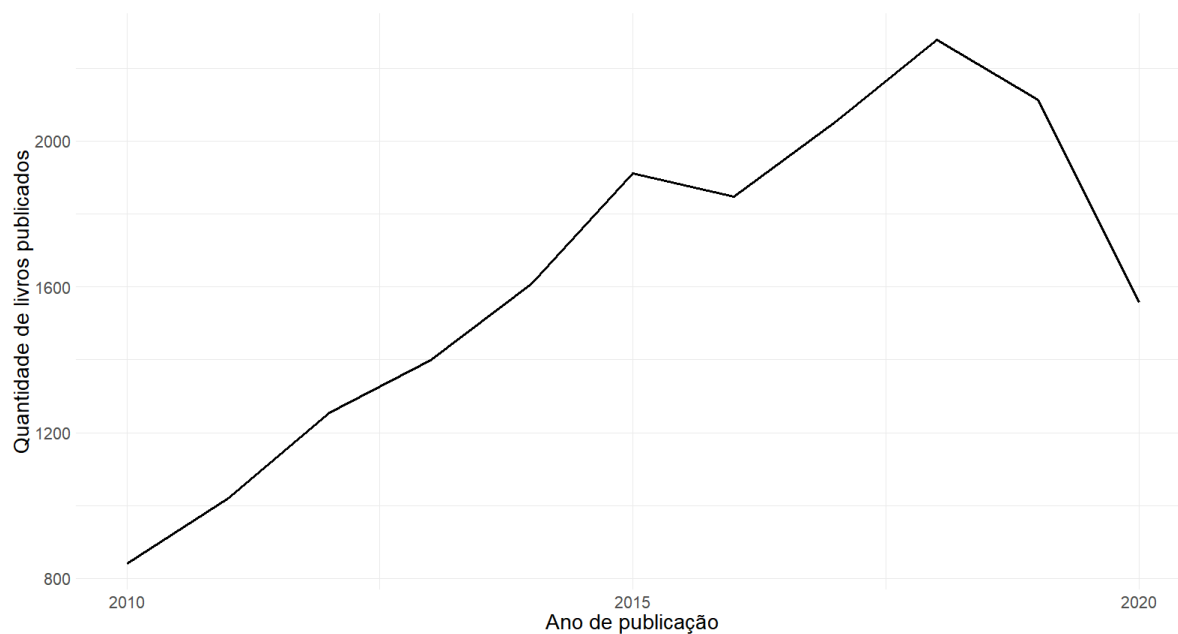
Fonte: Elaboração própria

Figura A.4 – Distribuição de capítulos de livros e valor da bolsa (em reais) por tipo de bolsa



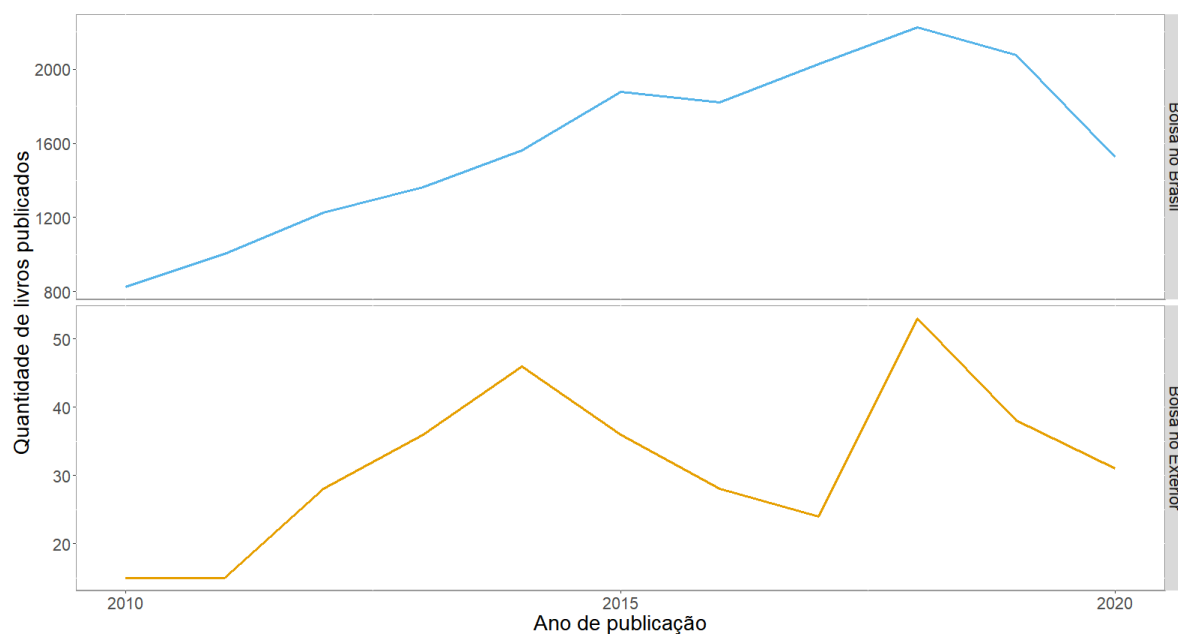
Fonte: Elaboração própria

Figura A.5 – Livros publicados por ano



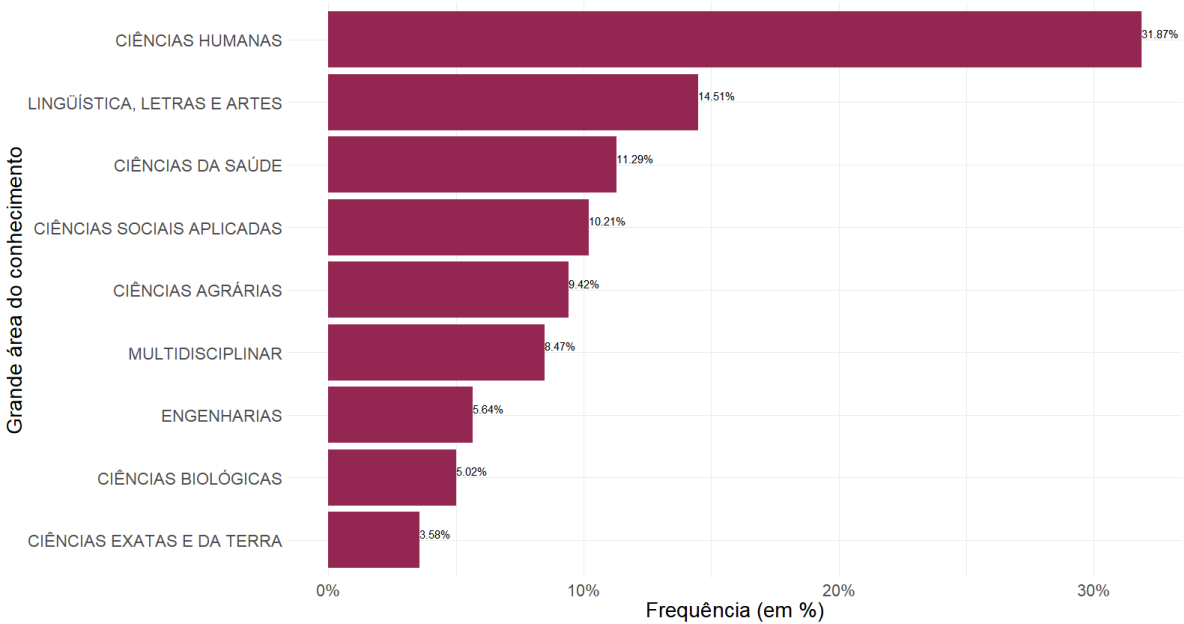
Fonte: Elaboração própria

Figura A.6 – Livros por ano e tipo de bolsa



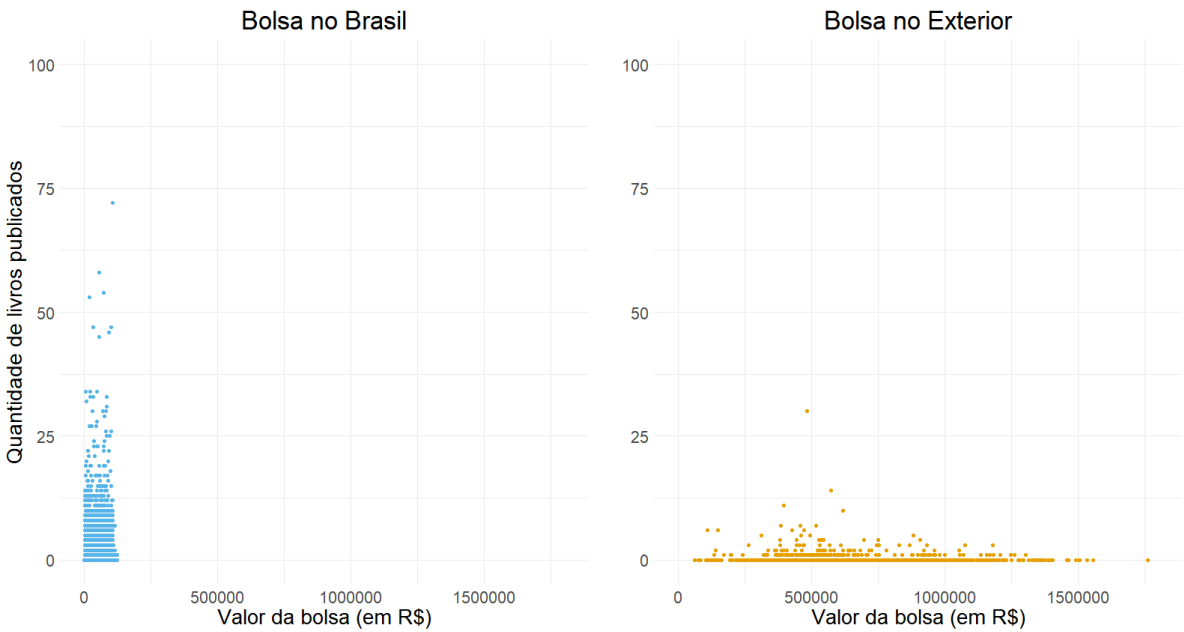
Fonte: Elaboração própria

Figura A.7 – Livros por grande área de conhecimento



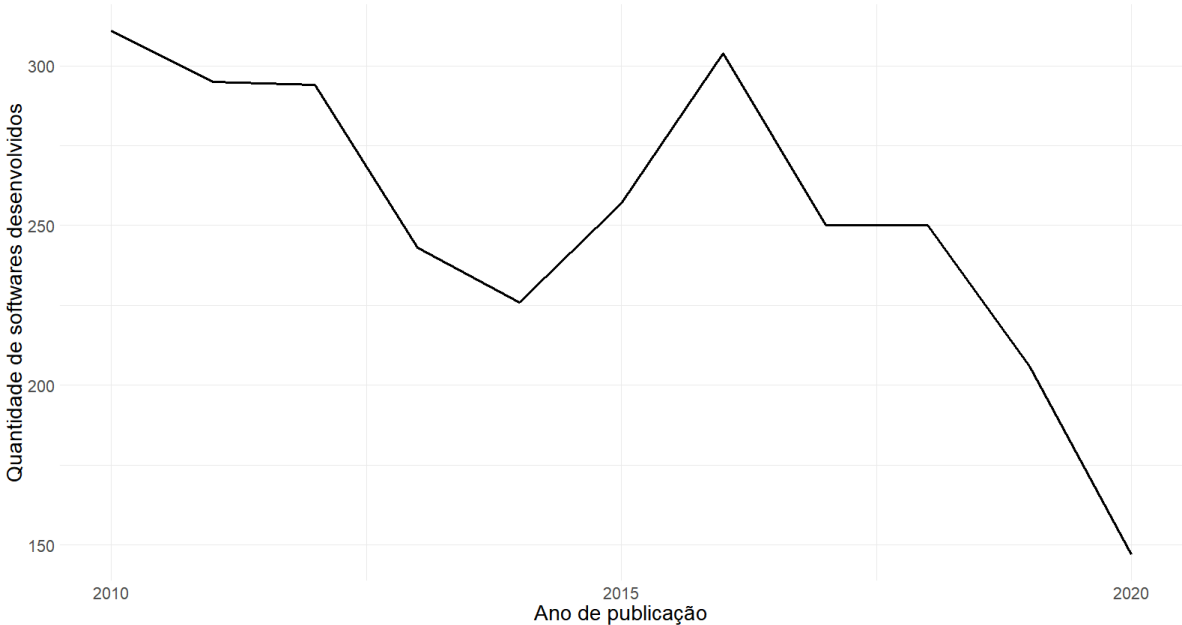
Fonte: Elaboração própria

Figura A.8 – Distribuição de livros publicados e valor da bolsa (em reais) segundo tipo de bolsa



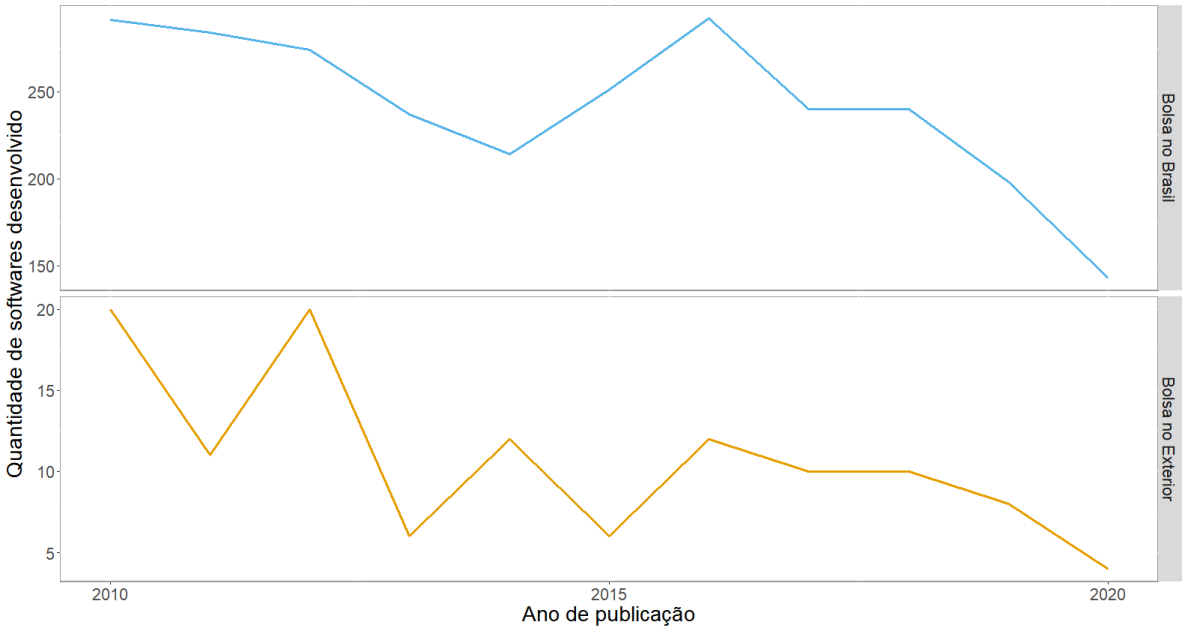
Fonte: Elaboração própria

Figura A.9 – Softwares desenvolvidos por ano



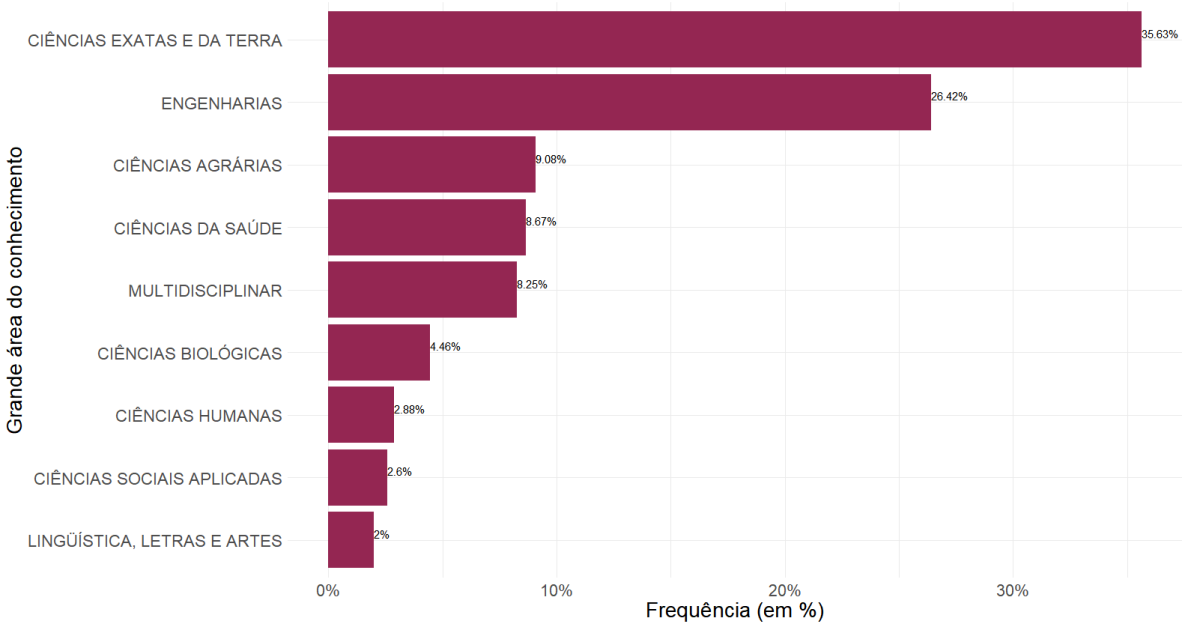
Fonte: Elaboração própria

Figura A.10 – Softwares desenvolvidos por ano e tipo de bolsa



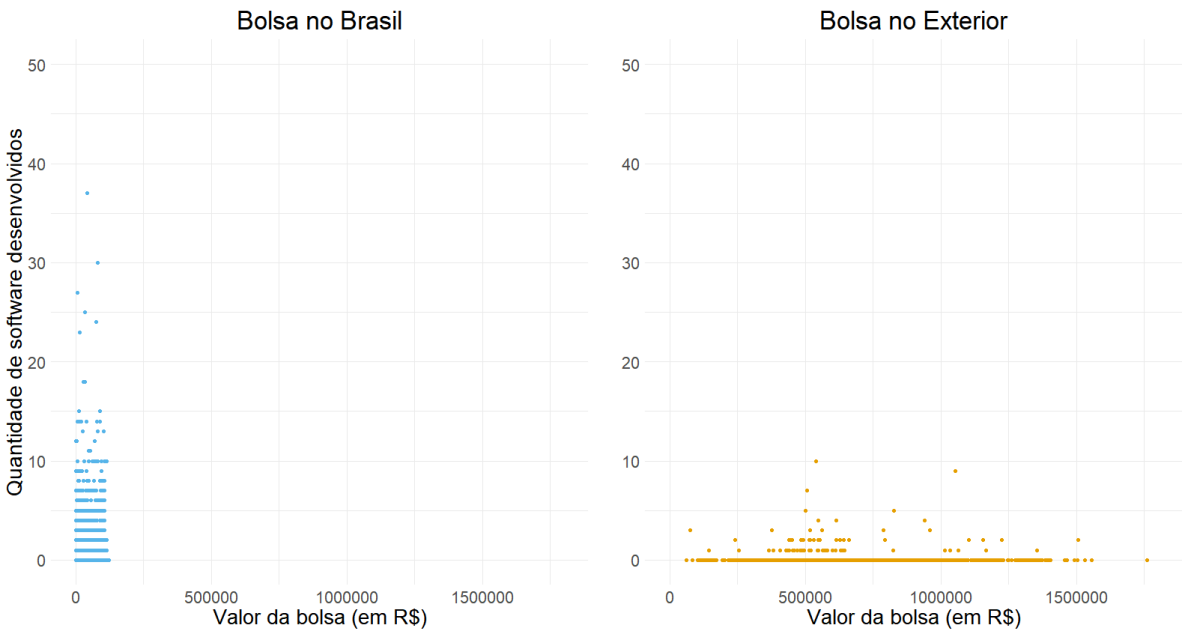
Fonte: Elaboração própria

Figura A.11 – *Softwares* desenvolvidos por grande área de conhecimento



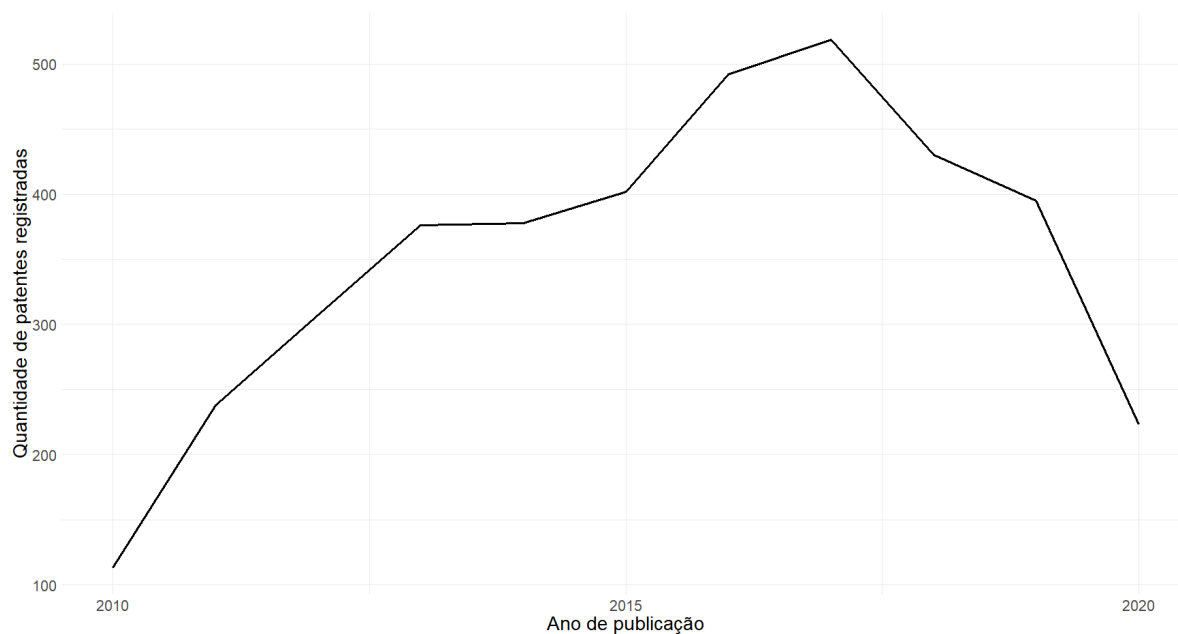
Fonte: Elaboração própria

Figura A.12 – Distribuição de *softwares* desenvolvidos e valor da bolsa (em reais) segundo tipo de bolsa



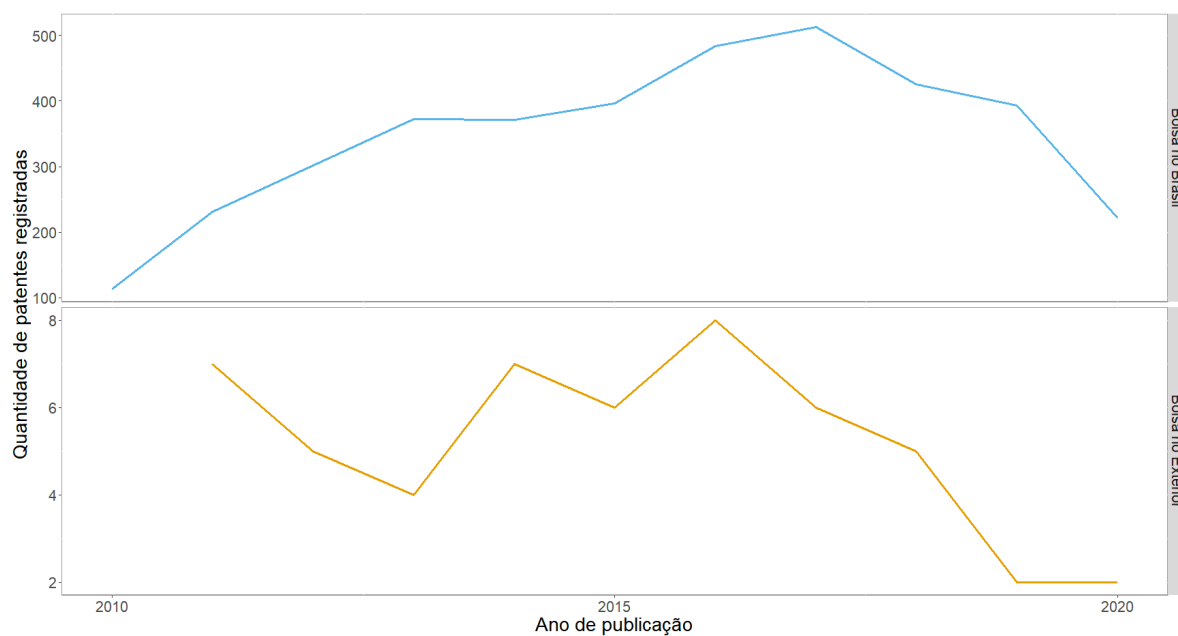
Fonte: Elaboração própria

Figura A.13 – Patentes registradas por ano



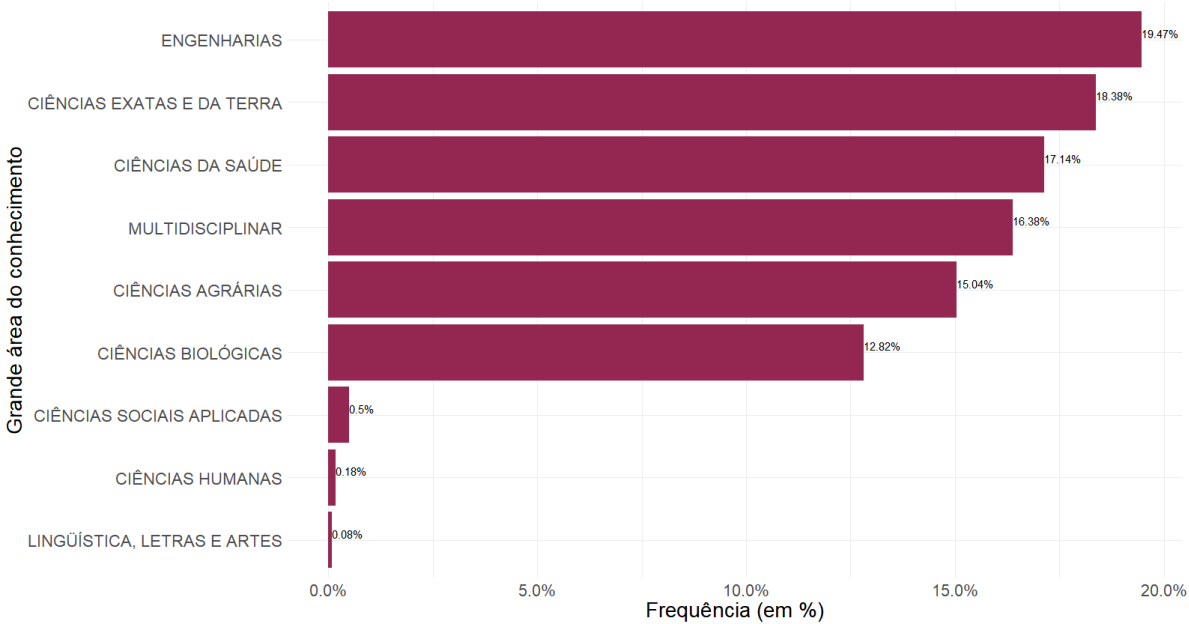
Fonte: Elaboração própria

Figura A.14 – Patentes registradas por ano e tipo de bolsa



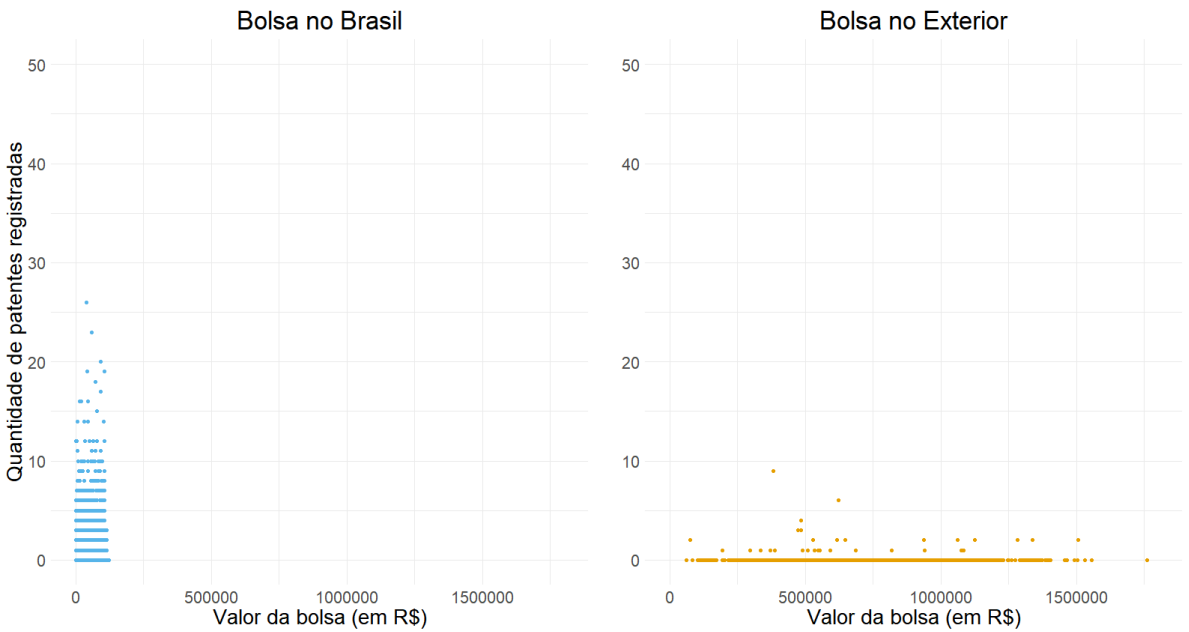
Fonte: Elaboração própria

Figura A.15 – Patentes registradas por grande área de conhecimento



Fonte: Elaboração própria

Figura A.16 – Distribuição de patentes registradas e valor da bolsa (em reais) segundo tipo de bolsa



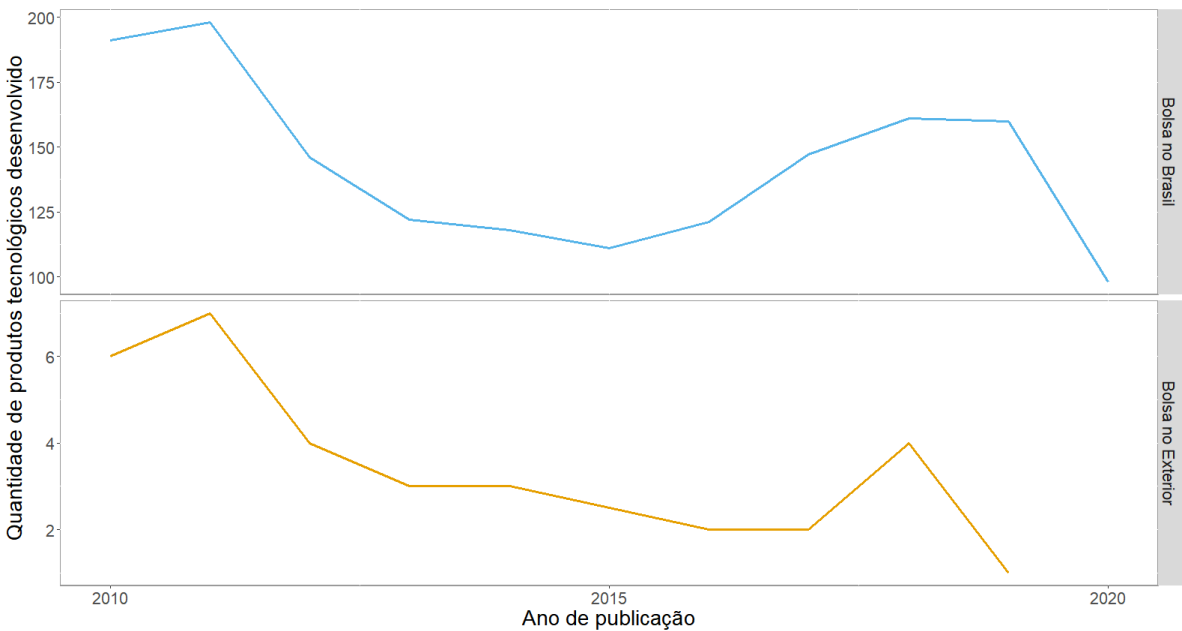
Fonte: Elaboração própria

Figura A.17 – Produtos tecnológicos desenvolvidos por ano



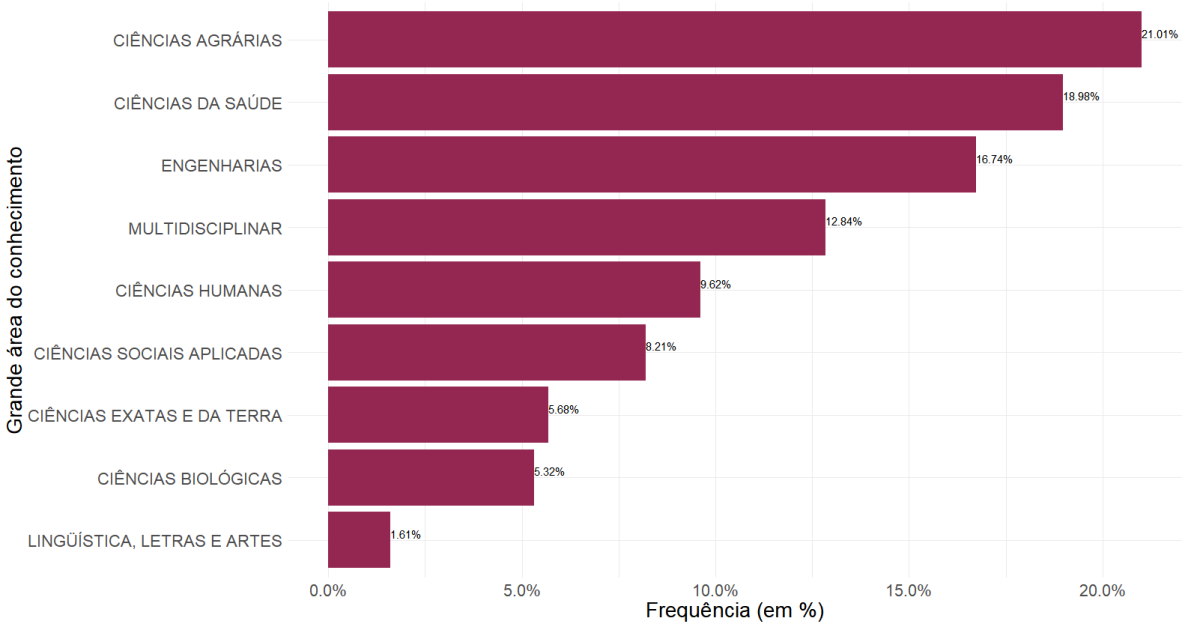
Fonte: Elaboração própria

Figura A.18 – Produtos tecnológicos desenvolvidos por ano e tipo de bolsa



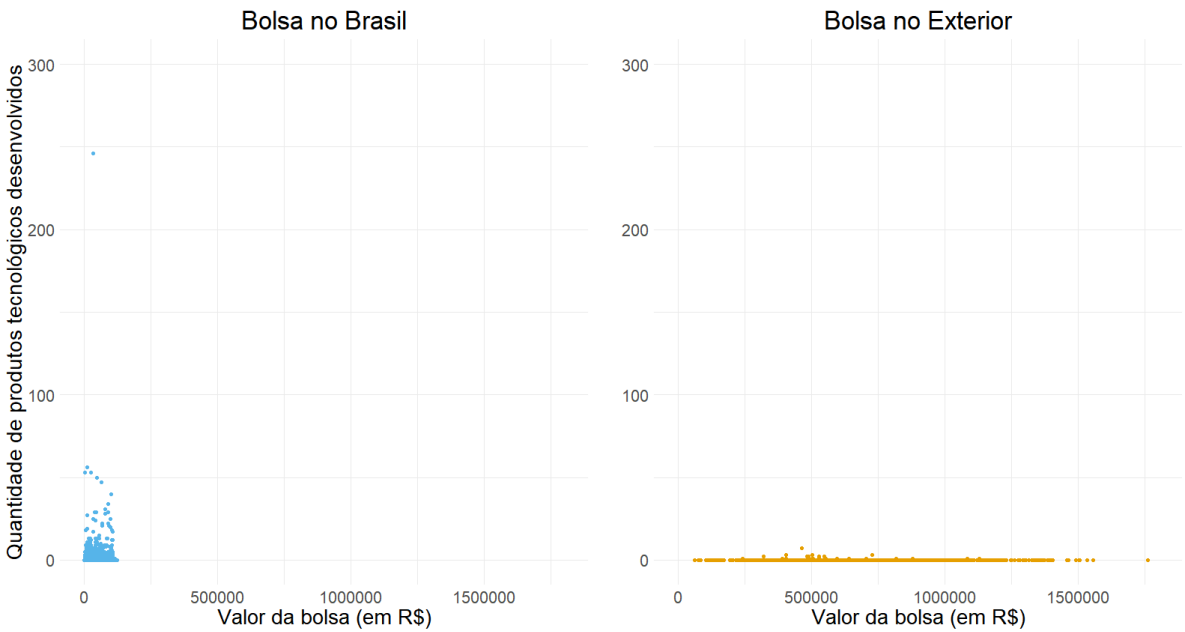
Fonte: Elaboração própria

Figura A.19 – Produtos tecnológicos desenvolvidos por grande área de conhecimento



Fonte: Elaboração própria

Figura A.20 – Distribuição de produtos tecnológicos desenvolvidos e valor da bolsa (em reais) segundo tipo de bolsa



Fonte: Elaboração própria

APÊNDICE B – *SCRIPTS* UTILIZADOS NO PROCESSO ETL

B.1 *Script em Bash - Captura de nome e ID das telas paginadas da plataforma Lattes*

```
#!/bin/bash

for (( c=0; c<=700; c++ ))
do
    echo "Welcome $c times"

    i=$((c*10000))

    echo $i

    curl http://buscatextual.cnpq.br/buscatextual/busca.do/?metodo
    \=forwardPaginaResultados\&registros\=$i\;10000\&query
    \=%28+%2Bidx_nacionalidade%3Ae%29+or+%28+%2Bidx_nacionalidade
    %3Ab+%5E500%29\&analise\=cv\&tipoOrdenacao\=null\&paginaOrigem
    \=index.do\&mostrarScore\=false\&mostrarBandeira\=true\
    &modoIndAdhoc\=null | grep -E "abreDetalhe" >> tmp.dmp

done
```

B.2 *Script em Python - Relacionar Nome e ID 1*

```
import re
from unicode import unicode
import io

accent_map = {
    u'\u00c0': u'A',
    u'\u00c1': u'A',
    u'\u00c2': u'A',
    u'\u00c3': u'A',
    u'\u00c4': u'A',
    u'\u00c5': u'A',
    u'\u00c6': u'A',
    u'\u00c7': u'C',
```

u'\u00c8': u'E',
u'\u00c9': u'E',
u'\u00ca': u'E',
u'\u00cb': u'E',
u'\u00cc': u'I',
u'\u00cd': u'I',
u'\u00ce': u'I',
u'\u00cf': u'I',
u'\u00d0': u'D',
u'\u00d1': u'N',
u'\u00d2': u'O',
u'\u00d3': u'O',
u'\u00d4': u'O',
u'\u00d5': u'O',
u'\u00d6': u'O',
u'\u00d7': u'x',
u'\u00d8': u'0',
u'\u00d9': u'U',
u'\u00da': u'U',
u'\u00db': u'U',
u'\u00dc': u'U',
u'\u00dd': u'Y',
u'\u00df': u'B',
u'\u00e0': u'a',
u'\u00e1': u'a',
u'\u00e2': u'a',
u'\u00e3': u'a',
u'\u00e4': u'a',
u'\u00e5': u'a',
u'\u00e6': u'a',
u'\u00e7': u'c',
u'\u00e8': u'e',
u'\u00e9': u'e',
u'\u00ea': u'e',
u'\u00eb': u'e',
u'\u00ec': u'i',
u'\u00ed': u'i',
u'\u00ee': u'i',
u'\u00ef': u'i',
u'\u00f1': u'n',
u'\u00f2': u'o',
u'\u00f3': u'o',

```

        u'\u00f4': u'o',
        u'\u00f5': u'o',
        u'\u00f6': u'o',
        u'\u00f8': u'o',
        u'\u00f9': u'u',
        u'\u00fa': u'u',
        u'\u00fb': u'u',
        u'\u00fc': u'u'
    }

def accent_remove (m):
    return accent_map[m.group(0)]

with open('temporarioBolsitas_dout_pleno_2010_2014_Manoel.csv') as f:
    lines = [line.rstrip('\n\r') for line in f]
    #print(lines)

with io.open('relacao_nome_id.csv', encoding="utf-8") as f:
    lines_arquivo_full = [line.rstrip('\n\r') for line in f]

#print(len(lines_arquivo_full))

f_destino = open("encontrados.csv", "a")

for linha in lines:

    nome_arr = linha.split(';')
    nome = nome_arr[0]
    nome = nome.upper()

    print(nome)

for l_arquivo_full in lines_arquivo_full:

    nome_a_comparar = l_arquivo_full.split('|')[2]
    nome_a_comparar = nome_a_comparar.upper()
    nome_a_comparar = nome_a_comparar

    #print(nome_a_comparar)
    nome_a_comparar = unicode(nome_a_comparar)
    #nome_a_comparar = re.sub(u'([\u00C0-\u00FC])',
    #accent_remove, nome_a_comparar.encode().decode('utf-8'))

```

```

        if (nome == nome_a_comparar):
            print("Achei: " + nome + "=>" + nome_a_comparar)
            print(l_arquivo_full.split('|')[0])

            f_destino.write(l_arquivo_full.split('|')[0] + "\n")

print(len(lines))

f_destino.close()

```

B.3 Script em Python - Relacionar Nome e ID 2

```

from unicode import unicode
import io
import pandas as pd

with open('temporarioBolsitas_dout_pleno_2010_2014_Manoel.csv') as f:
    lines = [line.split(';')[0].upper() for line in f]
    #print(lines)

with io.open('relacao_nome_id.csv', encoding="utf-8") as f:
    lines_arquivo_full =
    [line.rstrip('\n\r').upper().split('|') for line in f]

for l in lines_arquivo_full:
    l[2] = unicode(l[2])

#print(lines_arquivo_full)

nomes_list = pd.DataFrame( {'Nome': lines }, columns = ['Nome'])
print(nomes_list)

nomes_full = pd.DataFrame(lines_arquivo_full ,
columns =['Id', 'NomeUnderline', 'Nome'])
print(nomes_full)

merge = pd.merge(nomes_list , nomes_full , on=['Nome'], how='left')
print(merge)

```

```
f_destino = open("encontrados.csv", "a")

for index, row in merge.iterrows():
    f_destino.write(row['Nome'] + "|" + str(row['Id']) + "\n")

f_destino.close()

#merge.to_csv(path_or_buf='arquivo_tratado.csv', sep='|')

exit(0)

f_destino = open("encontrados.csv", "a")
f_destino.write("NOME_ORIGINAL|ID\n")

for linha in lines:

    nome_arr = linha.split(';')
    nome = nome_arr[0]
    nome = nome.upper()

    print(nome)
    f_destino.write(nome + "|")

for l_arquivo_full in lines_arquivo_full:

    nome_a_comparar = l_arquivo_full.split('|')[2]
    nome_a_comparar = nome_a_comparar.upper()
    nome_a_comparar = nome_a_comparar

    #print(nome_a_comparar)
    nome_a_comparar = unicode(nome_a_comparar)
    #nome_a_comparar = re.sub(u'([\u00C0-\u00FC])', ,
    #accent_remove, nome_a_comparar.encode().decode('utf-8'))

    if (nome == nome_a_comparar):
        print("Achei: " + nome + "=>" + nome_a_comparar)
        print(l_arquivo_full.split('|')[0])

        f_destino.write(l_arquivo_full.split('|')[0] + "\n")

print(len(lines))
```

```
f_destino.close()
```

B.4 Script em Python – Download do XML 1

```
from lattes.pages import Curriculum, Xml
import io

with io.open('aprocessar.csv', encoding="utf-8") as f:
    lines = [line.rstrip('\n\r') for line in f]

#print(lines)

for l in lines:
    print(l)

    cod = l.split('|')[1]

    if(cod[0:1] == 'K'):

        short_id = cod
        curriculum = Curriculum(short_id)
        #print(short_id)
        xml = Xml(curriculum.long_id, 'xmls/')
    else:
        print("Pulei ..... "+ l)
```

B.5 Script em Python – Download do XML 2

```
#!/usr/bin/env python
# -*- coding: utf-8 -*-
from multiprocessing import Process
from lattes.pages import Curriculum, Xml
from lattes.config import BaseLogger
import io
import time
import os.path
```

```

logger = BaseLogger.from_file('simple_client', file_name='simple_client.log')
ZIP_PATH = 'xmls'

with open('processado.csv', 'a') as out_file:
    out_file.write('')
with open('erro.csv', 'a') as out_file:
    out_file.write('')

with io.open('aprocessar.csv', encoding="utf-8") as f:
    lines = [line.rstrip('\n\r') for line in f]

l_aprocessar = []
for l in lines:

    cod = l.split('|')[1]

    if(cod[0:1] == 'K'):
        #if not os.path.isfile(ZIP_PATH + '/' + cod + '.zip'):
            l_aprocessar.append(cod)

def single_cored_example(short_ids):
    worker(short_ids)

def multi_cored_example(short_ids):
    """Simple enough multicore example."""
    chunk = 10000
    #print(l_aprocessar)

    short_ids = [short_ids[x:x+chunk]
                  for x in range(0, len(short_ids), chunk)]
    #logger.info('Spawning processes')
    for split_list in short_ids:
        p = Process(target=worker, args=(split_list,))
        p.start()

def worker(short_ids):
    """Run through a list of short_ids downloading it's respective xmls."""

```

```

for short_id in short_ids:
    #logger.info('Getting curriculum for {}'.format(short_id))

    print(short_id)

    with open('processado.csv', 'r') as out_file:
        lines = out_file.read()
    with open('erro.csv', 'r') as out_file:
        lines_erro = out_file.read()

    if (lines.find(short_id)==-1 and lines_erro.find(short_id)==-1):
        try:
            #time.sleep(10)
            curriculum = Curriculum(short_id)
            Xml(curriculum.long_id, ZIP_PATH)

            with open('processado.csv', 'a') as out_file:
                out_file.write(short_id + '\n')

            #logger.info('Curriculum for {} has been
            #downloaded'.format(short_id))
        except Exception as e:
            with open('erro.csv', 'a') as out_file:
                out_file.write(short_id + '\n')
            print(str(e))

if __name__ == "__main__":
    short_ids = l_aprocessar
    #long_ids = ['6380212729787758', '7639569152487589', '1024601314143406']
    #logger.info('Starting single_core...')
    #single_cored_example(short_ids)
    logger.info('Starting multi_core...')
    multi_cored_example(short_ids)

```

B.6 Script em Python – Descompactar XML

```

import zipfile
import os
import xml.etree.ElementTree as ET
import time

```

```
def getldfromXml(f_full, f_single):
    tree = ET.parse(f_full)
    xml = tree.getroot()

    if not 'NUMERO-IDENTIFICADOR' in xml.attrib:
        f_single = f_single.replace('.zip', '')

        print('Numero identificador nao encontrado no xml,
adicionando o nome do arquivo ao xml: ' + f_single)
        xml.set('NUMERO-IDENTIFICADOR', f_single)
        tree.write(f_full)

    return xml.attrib['NUMERO-IDENTIFICADOR']

def unzipfile(arquivo, arquivo_full, destino):

    if(os.path.isfile(destino + "/curriculo.xml")):
        os.remove(destino + "/curriculo.xml")

    with zipfile.ZipFile(arquivo_full, 'r') as zip_ref:
        zip_ref.extractall(destino)

    if(not os.path.isfile(destino + "/curriculo.xml")):
        raise Exception("Arquivo nao encontrado apos dezip.
Verifique: " + destino + "/curriculo.xml")

    try:
        novonome = getldfromXml(destino + "/curriculo.xml", arquivo)
        os.rename(destino + "/curriculo.xml",
destino + "/" + novonome + ".xml")
    except Exception as e:
        print('Erro. Removendo arquivo ' + arquivo_full)
        print('Erro encontrado: ')
        print(e)
        os.remove(arquivo_full)
        time.sleep(10)

files = os.listdir('xmlszipados/')
for f in files:
    print f
```

```

if '.zip' in f:
    unzipfile(f, 'xmlszipados/' + f, 'xmlsdezipados')

```

B.7 Script em Python – Carga no banco MySQL

```

import xml.etree.ElementTree as ET
import os
import json
import time
import mysql.connector

"""
verificar
- mais de um doutorando – vem em sequencia a tag doutorado
- nao tem area de atuacao mas tem empresa
- pais de nacionalidade dados gerais nao tem alguns
- area de atuacao SEQUENCIA-AREA-DE-ATUACA 4073572814178465.xml
"""

def getXmlFull(f):
    with open(f, 'rb') as source:
        parser = ET.XMLParser(encoding = "iso-8859-1")
        #print f
        tree = ET.parse(source, parser)
        xml = tree.getroot()

    return xml

def getAttrs(el, lAttrs):

    r= {}

    for a in lAttrs:
        valor = el.attrib[a] if el.attrib.has_key(a) else None
        r[a] = valor

    return r

def getElements(xml, tag):

```

```

r = xml.findall(tag)
return r if len(r) else None

def getInfo(xml, lTags, lAttrs ):

    global retAttr
    for t in lTags:
        els = [xml] if (xml.tag == 'CURRICULO-VITAE' and
            t == 'CURRICULO-VITAE' and len(lTags)==1) else
            getElements(xml, t)
        if (els is None): break
        newLtags = lTags
        newLtags.remove(t)

        if(els is None): print 'chegou'

        for e in els:
            if(len(newLtags)==0):
                d = getAttrs(e, lAttrs)
                retAttr.append(d)

            getInfo(e, newLtags, lAttrs)

def buildSql(table_name, dAttrs, nid = None):

    lista_campos = ''
    lista_valores = ''

    if(nid is not None):
        dAttrs['numero_identificador'] = nid

    for c,v in dAttrs.items():
        if(v is not None):
            lista_campos += c.lower().replace('-', '_') + ', '
            lista_valores += "'" + v.replace("'", "'") + "', "

    lista_campos = lista_campos[:-2]
    lista_valores = lista_valores[:-2]

    s = "insert into " + table_name +
        " (" + lista_campos + ") values (" + lista_valores + ")"

```

```

    return s

def farmAndInsert(xml, table_name, ITags, IAttrs, nid=None):

    global retAttr

    retAttr = []
    getInfo(xml, ITags, IAttrs)

    for e in retAttr:
        s = buildSql(table_name, e, nid)
        #print s
        mycursor.execute(s)

    """
    ElsAttrs = { 'DADOS-BASICOS-DO-ARTIGO': [ 'ANO-DO-ARTIGO', 'IDIOMA' ],
                 'DETALHAMENTO-DO-ARTIGO': [ 'TITULO-DO-PERIODICO-OU-REVISTA' ]
    }
    """

def getSliced(xml, ITags, dSliceAttrs, table_name, nid):

    global retAttr

    #verificar se existe pelo menos um elemento
    for t in ITags:

        els = xml.findall(t)

        #ver se eh ultimo elemento
        if(t == ITags[-1]):
            break

        if(els is None or len(els)==0 ):
            return
        else:
            xml = els[0]

    lItems = []
    for e in els:

        dItem = {}

```

```

    for eak,eav in dSliceAttrs.items():

        retAttr = []

        getInfo(e, [eak], eav )

        if (len(retAttr)>0):
            dItem.update( retAttr[0])

    lItems.append(dItem)

for a in lItems:
    s = buildSql(table_name, a, nid)
    #print s
    mycursor.execute(s)

```

B.8 Script em Python – Gerar CSV

```

f = open("../tmp_01.dmp", "r")

f_destino = open("relacao_nome_id.csv", "a")
f_destino.write("id|nome_normalized|nome_full\n")

for l in f:
    #print l
    key_nome = l
    key_nome = l[l.find("(")+1:l.find(")")]
    key_nome = key_nome.replace("'", "")
    key_nome = key_nome.split(',')
    #print key_nome

    nome_full = l[l.find('">')+2:l.find("</a>")]
    #print nome_full

    f_destino.write(key_nome[0] + "|" + key_nome[1] + "|" + nome_full + "\n")

f_destino.close()

f = open("../tmp_02.dmp", "r")
f_destino = open("relacao_nome_id.csv", "a")
for l in f:

```

```

    #print l
    key_nome = l
    key_nome = l[l.find("(")+1:l.find(")")]
    key_nome = key_nome.replace("'", "")
    key_nome = key_nome.split(',')
    #print key_nome

    nome_full = l[l.find('">')+2:l.find("</a>")]
    #print nome_full

    f_destino.write(key_nome[0] + "|" + key_nome[1] + "|" + nome_full + "\n")

f_destino.close()

f = open("../tmp_03.dmp", "r")
f_destino = open("relacao_nome_id.csv", "a")
for l in f:

    #print l
    key_nome = l
    key_nome = l[l.find("(")+1:l.find(")")]
    key_nome = key_nome.replace("'", "")
    key_nome = key_nome.split(',')
    #print key_nome

    nome_full = l[l.find('">')+2:l.find("</a>")]
    #print nome_full

    f_destino.write(key_nome[0] + "|" + key_nome[1] + "|" + nome_full + "\n")

f_destino.close()

f = open("../tmp_04.dmp", "r")
f_destino = open("relacao_nome_id.csv", "a")
for l in f:

    #print l
    key_nome = l
    key_nome = l[l.find("(")+1:l.find(")")]
    key_nome = key_nome.replace("'", "")

```

```
key_nome = key_nome.split(',')
#print key_nome

nome_full = l[l.find('>')+2:l.find("</a>")]
#print nome_full

f_destino.write(key_nome[0] + "|" + key_nome[1] + "|" + nome_full + "\n")

f_destino.close()
```

B.9 Script em Python – Verificação do arquivo zip

```
import zipfile
import os

def testZip(f):
    print "Testing zip file: %s" % f

    try:

        the_zip_file = zipfile.ZipFile(f)
        ret = the_zip_file.testzip()

    except:
        print "File %s is corrupt..." % f
        ret = "algo"

    if ret is not None:
        print "First bad file in zip: %s" % ret
        os.remove(f)

    else:
        print "Zip file is good."

files = os.listdir('xmlszipados/')
for f in files:
    print f
    if '.zip' in f:
        testZip('xmlszipados/' + f)
```