

# Estudo sobre a Participação das Mulheres nos Cursos de Graduação na Área STEM

Luciane C. J. de Deus<sup>1</sup>, Lenise M. V. Rodrigues<sup>1</sup>,  
Lucas Lopes Felipe<sup>1</sup>, Alexandre N. Louzada<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal do Rio de Janeiro (UFRJ)

**Abstract.** *Although the presence of women in science has increased globally, there is still an underrepresentation of women in the scientific and technological system. The scarcity of research related to the presence of women in the STEM area in higher education, generated an initial motivation to expand the understanding of this scenario, from the acquisition, treatment and analysis of raw and open dataset from INEP - Instituto Nacional de Estudos e Pesquisas, regarding the Higher Education Census. To enable a temporal analysis, the second decade of the millennium was worked on, in the period of 2010 and 2019. Afterwards, the scope of courses was defined to meet the research in relation to the STEM area and, then, the courses of Biological Sciences and related subjects, Physical Sciences, Computing, Engineering, Mathematics and Statistics. The results contributed to the technical aspect, demonstrating an adequate computational architecture for the management of provenance metadata and supporting the understanding, reproducibility and reuse of these data. And also in the educational and social sphere, expanding the basis of the discussion about women in higher education in the STEM area. It was possible to verify the area of Engineering and related professions, as being where women are more present in undergraduate courses in the STEM area, in addition to other interesting investigations by region of the country, modality and teaching network.*

**Resumo.** *Apesar da presença de mulheres na ciência ter aumentado de forma global, ainda há uma sub-representação delas no sistema científico e tecnológico. A escassez de pesquisas relacionadas a presença das mulheres na área STEM no ensino superior, gerou uma motivação inicial em ampliar o entendimento desse cenário, a partir da aquisição, tratamento e análise de dataset bruto e aberto do INEP - Instituto Nacional de Estudos e Pesquisas, referente ao Censo de Educação Superior. Para possibilitar uma análise temporal, trabalhou-se a segunda década do milênio, no período de 2010 e 2019. Após, foi definido o escopo de cursos para atender a pesquisa em relação a área STEM e, então, selecionados os cursos de Ciências Biológicas e correlatas, Ciências Físicas, Computação, Engenharias, Matemática e Estatística. Os resultados contribuíram no aspecto técnico, demonstrando uma arquitetura computacional adequada para a gestão dos metadados de proveniência e apoiando a compreensão, reprodutibilidade e o reuso destes dados. E também no âmbito educacional e social, ampliando o embasamento da discussão sobre mulheres no ensino superior na área STEM. Foi possível averiguar a área de Engenharia e profissões correlatas, como sendo onde as mulheres estão mais presentes nos cursos de graduação da área STEM, além de outras investigações interessantes por região do país, modalidade e rede de ensino.*

## 1. Introdução

De acordo com os dados do último levantamento do IBGE, as mulheres representam 51,8% da população brasileira.<sup>1</sup> Contudo, apenas 51,56% das mulheres possuem colocação no mercado de trabalho, um índice 20% inferior ao masculino.<sup>2</sup> Ao ampliar o olhar para a área de *STEM* (Ciência, Tecnologia, Engenharia e Matemática, na sigla em inglês),<sup>3</sup> de direta atuação dos autores, é possível perceber que apenas 26% dessa força de trabalho é feminina.<sup>4</sup>

No Brasil, os dados de graduandos por sexo demonstram que as mulheres têm maior inserção no ensino superior. Entretanto, predominam em cursos de Ciências Humanas e Sociais Aplicadas, mas não nas áreas *STEM*, onde é maior a presença dos homens.<sup>5</sup>

Contudo, se é inegável a necessidade de mais mulheres em *STEM*,<sup>6</sup> por que as mulheres representam tão pouco nessa área? As mulheres fazem parte dos cursos de graduação de *STEM*? Quantas mulheres ocupam vagas de graduação nessa área? Quantas mulheres concluem essas graduações? Será que o prognóstico é positivo ou negativo? Essas perguntas serviram de base para a realização deste trabalho, que visa analisar a participação de mulheres nos cursos de graduação na área de *STEM* com base no censo de educação superior do INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.

**Estrutura do artigo.** Na Seção 2, a literatura relacionada é apresentada. Na Seção 3, além de ter a apresentação da metodologia, há a descrição do *dataset* bruto e sua origem e uma análise preliminar da mesma, assim como a especificação das ferramentas utilizadas e suas respectivas versões. Na Seção 4, é apresentada a nossa contribuição com base na metodologia de desenvolvimento curado, seguindo os princípios *FAIR*.<sup>7</sup> Por fim, na Seção 5, foram analisados o aprendizado e possíveis trabalhos futuros.

## 2. Trabalhos Relacionados

A participação feminina nos cursos das áreas de *STEM* é o objeto de estudo da pesquisa e encontramos alguns trabalhos correlatos. Apesar da presença de mulheres na ciência ter aumentado de forma global, ainda há uma sub-representação delas no sistema científico e tecnológico e um pequeno número de mulheres em determinadas áreas ou subáreas do conhecimento, por exemplo, nas ciências exatas e engenharias [de Araújo and Tonini 2020]. Essa também será nossa investigação.

---

<sup>1</sup><https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/18320-quantidade-de-homens-e-mulheres.html#:~:text=Segundo\%20dados\%20da\%20PNAD\%20Cont\%C3\%ADnu,51\%2C8\%25\%20de\%20mulheres.>

<sup>2</sup><https://www.cnnbrasil.com.br/business/participacao-de-mulheres-no-mercado-de-trabalho-e-20-inferior-a-dos-homens/>

<sup>3</sup><https://pt.wikipedia.org/wiki/STEM>

<sup>4</sup><https://movimentomulher360.com.br/mulheres-ocupam-apenas-26-das-vagas-em-areas-stem/>

<sup>5</sup><http://revista.educacao.ws/revista/index.php/abenge/article/viewFile/1693/905/>

<sup>6</sup><http://portal.sbpnet.org.br/noticias/o-mundo-precisa-de-mais-mulheres-nas-carreiras-stem/>

<sup>7</sup><https://www.gov.br/ibict/pt-br/central-de-conteudos/noticias/2022/marco-2022/principios-go-fair>

As autoras ainda complementam que é preciso incentivar crianças e jovens para as carreiras científicas, e que o Estado Brasileiro tem papel fundamental de fomentar ações para que as meninas possam ter contato com a ciência, tecnologia e inovação desde os ciclos mais básicos da educação, o que corrobora com nossa pesquisa que pretende servir de embasamento para discussões desse cunho.

No caso do Brasil, conforme o trabalho de [Iwamoto 2022] a legislação federal que aborda a inclusão dos cidadãos em geral nas *STEM* é escassa, e, no caso de inclusão de mulheres nas áreas, raríssima. O que demonstra o quanto ainda temos que avançar, a nível de legislação e também na permanência de mulheres na área.

Para [Silva et al. 2022] o ambiente acadêmico nas áreas de *STEM* apresenta diversos problemas que prejudicam o desempenho acadêmico das mulheres. E mesmo que o viés dessa pesquisa não seja explorar os motivos da desistência, foram considerados estudos que relatam tanto a atuação de mulheres na ciência e tecnologia, como a inclusão e permanência de mulheres em cursos relacionados à *STEM*. Podendo auxiliar na análise dos dados extraídos em nossa pesquisa quanto a estatísticas de ingresso, evasão e outras análises.

Não foi encontrada literatura científica relacionada a metodologia de desenvolvimento de *datasets* curados, com proveniência, seguindo os princípios *FAIR* sobre a presença de mulheres nos cursos de graduação nas áreas de *STEM*. Portanto, este trabalho se baseia nesta metodologia de pesquisa para apresentar um estudo sobre a participação feminina nos cursos de graduação na área de *STEM*.

### 3. Materiais e Métodos

O percurso metodológico envolveu os seguintes passos:

- (i) Pesquisa bibliográfica sobre a inserção das mulheres nos cursos *STEM* do Brasil, conforme descrito na seção anterior;
- (ii) Seleção da base de dados, conforme descrito na subseção 3.1;
- (iii) Aquisição dos dados no *Dataset* Bruto 3.1;
- (iv) Tratamento e criação de dicionário de dados 3.2
- (v) Verificação da proveniência 4.2 e investigações diversas 4.1

#### 3.1. Seleção da Base de Dados

Buscamos uma base de dados de procedência confiável, que possuisse também dados históricos para possibilitar uma análise temporal mais apurada e não apenas um olhar limitado na última amostra.

#### 3.2. Aquisição dos dados no *Dataset* Bruto

As bases selecionadas foram provenientes do Censo de Educação Superior,<sup>8</sup> fornecido pelo INEP, usando a tabela *Microdados — Censo da Educação Superior*, disponível em [https://download.inep.gov.br/microdados/microdados\\_censo\\_da\\_educacao\\_superior\\_{ano}.zip](https://download.inep.gov.br/microdados/microdados_censo_da_educacao_superior_{ano}.zip), em que o *ano* corresponde ao ano alvo para cada coleta.

---

<sup>8</sup><https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-da-educacao-superior/resultados>

Levamos em consideração dados coletados no período entre 2010 e 2019 e, a partir disso, fizemos uma agregação de todas as tabelas em um único arquivo que chamamos de *cursos*.

A tabela bruta para trabalho ficou com 246.081 linhas e 196 colunas.

### 3.3. Tratamento de dados

Para atender o escopo da pesquisa em relação ao *STEM*, foram selecionados cursos dessas áreas, observando os dados do *dataset* onde encontrou-se para:

**S**(Science): Ciências Biológicas e correlatas e Ciências Físicas;

**T**(Technology): Computação;

**E**(Engineering): Engenharias e correlatas;

**M**(Mathematics): Matemática e Estatística.

A Tabela 1 a seguir, apresenta o código e áreas correspondentes utilizadas para a filtragem do *dataset*.

código	nome da área
51	ciências biológicas e correlatas
54	matemática e estatística
53	ciências físicas
61	computação
71	engenharias e correlatas

**Tabela 1. Áreas da tabela correspondentes a cursos *STEM***

Após essa filtragem a tabela gerada finalizou com 17.979 linhas e 18 colunas.

#### 3.3.1. Dicionário de dados

Para facilitar o entendimento e manipulação dos dados, criou-se um dicionário de dados, onde centraliza informações sobre o dataset, apresentando a descrição de cada Nome da Coluna, conforme Tabela 2 a seguir:

### 3.4. Reprodutibilidade

O ambiente computacional preparado utilizou Python<sup>9</sup> v3.9.12, com a interface jupyter<sup>10</sup> v.2.3.0, usando anaconda<sup>11</sup> v.6.4.12 como gerenciador de pacotes.

As bibliotecas utilizadas e suas respectivas versões foram: requests<sup>12</sup> v2.28.1, pandas<sup>13</sup> v1.4.2, datatable<sup>14</sup>, tqdm<sup>15</sup> v4.64.

---

<sup>9</sup><https://pypi.org/>

<sup>10</sup><https://jupyter.org/>

<sup>11</sup><https://www.anaconda.com/>

<sup>12</sup><https://pypi.org/project/requests/>

<sup>13</sup><https://pandas.pydata.org/>

<sup>14</sup><https://pypi.org/project/datatable/>

<sup>15</sup><https://pypi.org/project/tqdm/>

Nome da Coluna	Coluna Renomeada	Descrição
NU_ANO_CENSO	ano	Ano do Censo
NO_REGIAO	regiao	Nome da região geográfica do local de oferta do curso
SG_UF	uf	Sigla da Unidade da Federação do local de oferta do curso
CO_MUNICIPIO	cod_municipio	Código do IBGE do município do local de oferta do curso
NO_MUNICIPIO	nome_municipio	Nome do Município do local de oferta do curso
IN_CAPITAL	capital	Informa se o local de oferta do curso está localizado em capital de Estado 0 - Não 1 - Sim (.) Não aplicável (Curso a distância)
CO_CINE_AREA_ESPECIFICA	cod_area	Código de identificação da área específica, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco
NO_CINE_AREA_ESPECIFICA	area	Nome da área específica, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco
NO_CINE_ROTULO	curso	Nome do curso, conforme adaptação da Classificação Internacional Normalizada da Educação Cine/Unesco
TP_REDE	rede_publica	Rede de Ensino 1 - Pública 2 - Privada
TP_MODALIDADE_ENSINO	presencial	Tipo da modalidade de ensino do curso 1 - Presencial 2 - Curso a distância
TP_GRAU_ACADEMICO	bacharel	Tipo do grau acadêmico conferido ao aluno pela conclusão dos requisitos exigidos pelo curso 1 - Bacharelado 2 - Licenciatura 3 - Tecnológico 4 - Bacharelado e Licenciatura (.) Não aplicável (cursos com nível acadêmico igual a sequencial de formação específica ou cursos de área básica de curso identificada pela variável TP_ATRIBUTO_INGRESSO)
QT_ING	ingressantes	Quantidade de ingressante no curso
QT_ING_FEM	ingressantes_fem	Quantidade de ingressante do sexo feminino no curso
QT_CONC	concluintes	Quantidade de concluinte no curso
QT_CONC_FEM	concluintes_fem	Quantidade de concluinte no curso do sexo feminino
-	frac_ingressantes	Fração de ingressantes mulheres no total de ingressantes
-	frac_concluintes	Fração de concluintes mulheres no total de concluintes

**Tabela 2. Colunas selecionadas e suas descrições**

Após isso, para entender mais sobre as principais características e atributos da base, que nos seriam interessantes, realizou-se uma pré-análise do arquivo, utilizando o software Microsoft Excel,<sup>16</sup> verificando as colunas e interpretando-as, elencando as que relacionam ao escopo da pesquisa, deixando um total de 246.081 linhas e 18 colunas.

## 4. Análise dos Dados e Resultados

### 4.1. Análise e Investigações

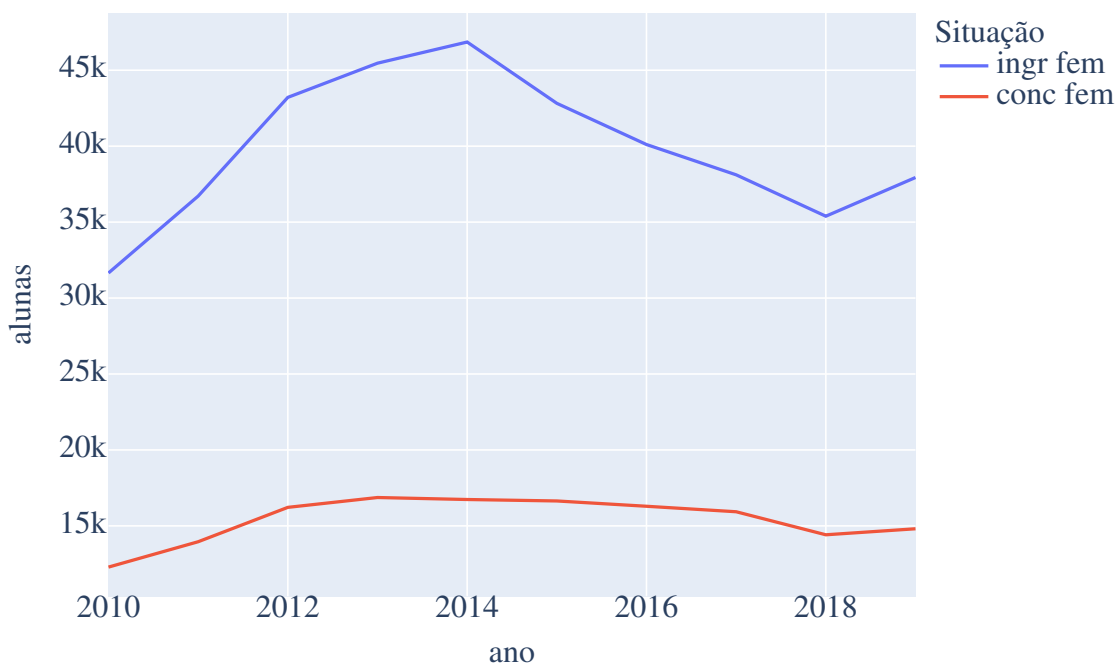
A investigação embasou-se nas perguntas norteadoras da pesquisa conforme a seguir.

#### 4.1.1. Qual a relação das mulheres que ingressam com as que concluem a graduação na área STEM de 2010 à 2019 ?

Sabemos que ao considerar o número de mulheres concluintes em um ano, não se refere às mulheres ingressantes daquele mesmo ano, devido ao tempo de cada curso de graduação. Por isso, decidimos por uma série histórica para possibilitar uma melhor análise dessa comparação, conforme apresenta a Figura 1.

Ao observar a série histórica referente às mulheres ingressantes e concluintes de cursos de graduação na área STEM, de 2010 à 2019, verifica-se de imediato a diferença entre os valores, sendo que as concluintes de um ano tem um número menor que metade das mulheres ingressantes do mesmo ano. Ou seja, as mulheres ingressam, porém muitas delas não concluem os cursos.

<sup>16</sup><https://www.microsoft.com/pt-br/microsoft-365/excel>



**Figura 1. Série histórica de mulheres ingressantes e concluintes em STEM**

Ampliando as percepções sobre a série histórica é possível perceber que em relação aos ingressantes, teve-se uma crescente de 2010 a 2014, sendo que nesse último ano, podemos destacar que foi quando mais mulheres entraram nos cursos de graduação na área STEM. Após o número foi diminuindo até 2018, onde voltou a ter crescimento.

Quanto as mulheres concluintes pode-se dizer nos 4 primeiros anos (2010 a 2013) teve-se uma crescente, que logo após se estabilizou até 2017, onde o valor diminui para 2018 e que permanece em 2019.

#### **4.1.2. Quais são as áreas STEM que concentram maior e menor número de mulheres ingressantes e concluintes?**

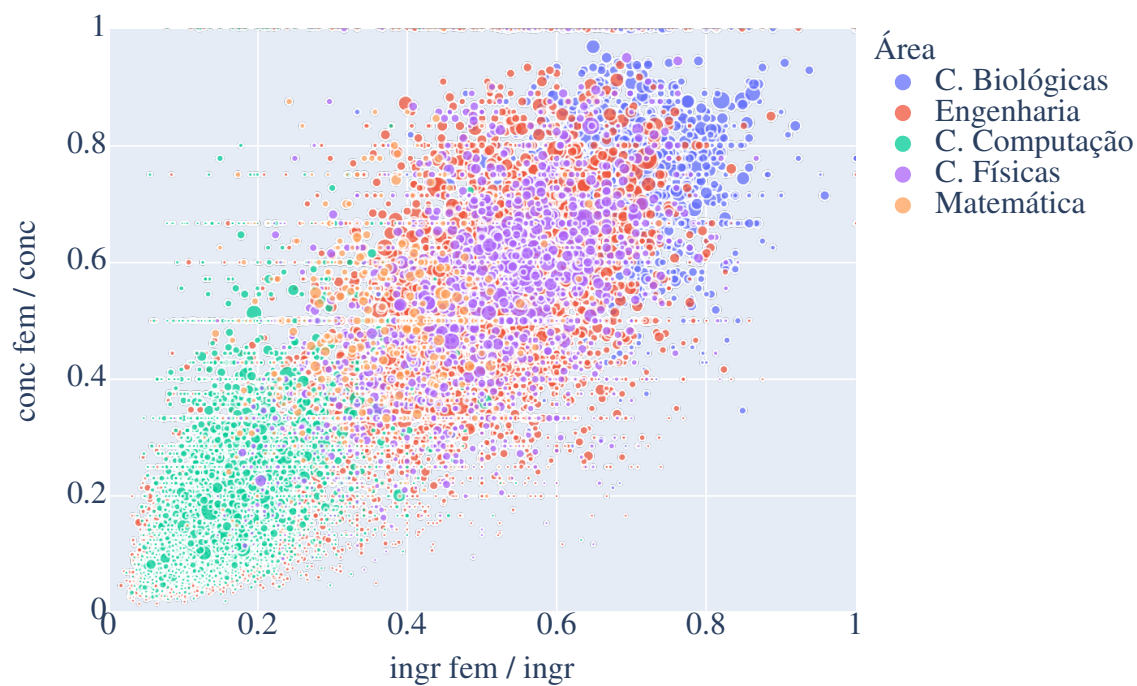
Os gráficos de dispersão foram criados utilizando Python, com Plotly Express <sup>17</sup>, que é uma interface que opera em uma variedade de tipos de dados e produz figuras fáceis de estilizar. O `px.scatter`, possibilita que cada ponto de dados seja representado como um ponto marcador, cuja localização é dada pelas colunas x e y, no caso a `frac_ingressantes` e a `frac_concluintes`, respectivamente. Sendo que representam as seguintes expressões:

$$\text{frac\_ingressantes} = \frac{\text{ingressantes mulheres}}{\text{total de ingressantes}} \quad \text{frac\_concluintes} = \frac{\text{concluintes mulheres}}{\text{total de concluintes}}$$

$$\text{tamanho do ponto} = \text{número de concluintes femininas}$$

<sup>17</sup><https://plotly.com/python/line-and-scatter/>

Considerando o apresentado na Figura 2, é possível observar que se tem mais concentração de mulheres concluintes na área de Engenharia e profissões correlatas.



**Figura 2. Gráfico de dispersão de área**

Porém, ao selecionarmos as áreas que se referem diretamente a Ciência, que são as áreas 51 - ciências biológicas e correlatas e 53 - ciências físicas, teremos também uma grande concentração de mulheres por área STEM. Porém, ainda menor que a Engenharia.

Em relação a menor concentração de mulheres é na área de Matemática e estatística.

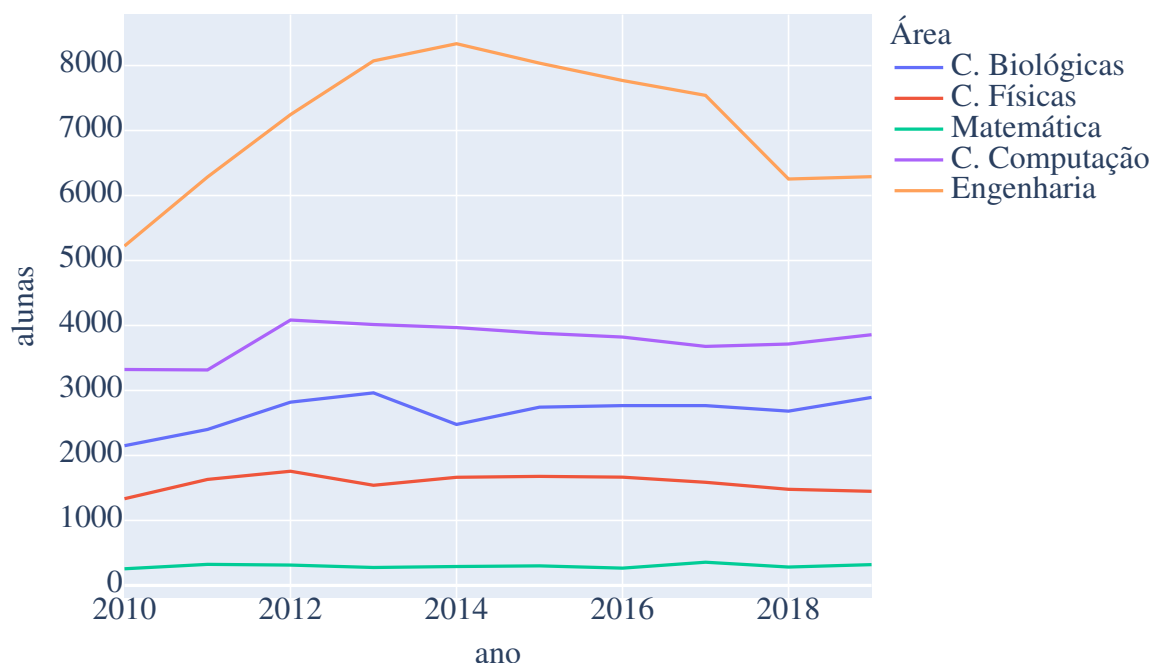
#### **4.1.3. Quantas mulheres ocupam vagas de graduação por área STEM de 2010 a 2019?**

Mais uma vez a série histórica foi utilizada para responder as indagações e nesse caso apresentar a quantidade de mulheres por área STEM durante os anos de 2010 a 2019.

Conforme a Figura 3, a área de Engenharia nessa última década teve destaque das mulheres na área STEM, demonstrando uma crescente de concluintes mulheres de 2010 à 2014, que comparando com os demais cursos entende-se que se mantiveram estáveis ou abaixo em 2014 e que foi a área de Engenharia que alavancou a presença da mulher na área STEM esse ano, que se demonstrou também na Figura 2, o ano de maior número de concluinte mulher.

Corroborando com esse destaque em 2014, para o IPEA comparações internacionais indicaram que a formação em áreas de Ciências, Matemática, Tecnologia e Engenharia ainda era pouco numerosa em

meados da década de 2000, mas foi justamente a partir desse momento que se observou uma mais forte expansão da oferta. E que em 2010 a 2012, foi a demanda por cursos dessa área, revelada pelo número de candidatos por vaga, o indicador que mais cresceu, apresentando, nesses três anos, salto ainda mais significativo do que os indicadores de oferta ao longo de todo o período compreendido entre 2000 e 2012



**Figura 3. Mulheres que ocupam vagas de graduação por área**

A área de Matemática e Estatística se manteve na série histórica como a que as mulheres ocupam as menores vagas em relação as demais áreas STEM.

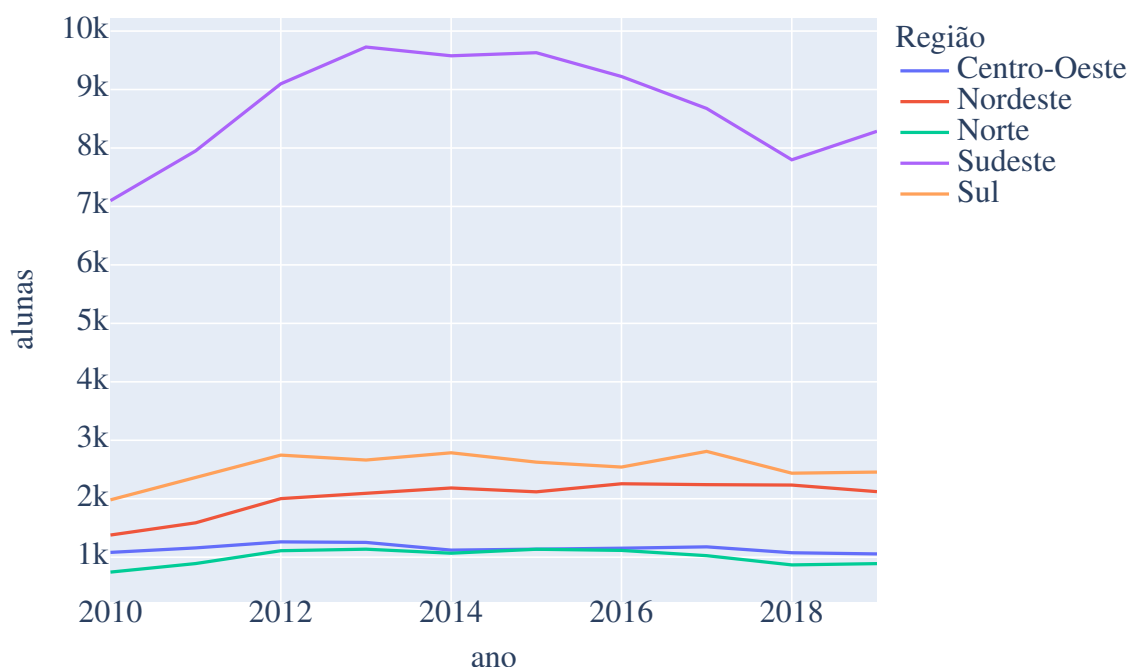
#### **4.1.4. Onde no Brasil se tem mais mulheres na graduação na área STEM?**

Em relação a região do Brasil, podemos identificar, através da Figura 4, que na série histórica apresentada a região Sudeste aparece em toda a década como a região onde se tem mais mulheres nos cursos de graduação da área STEM. Sendo mais que triplo a quantidade das mulheres em relação as demais regiões.

As regiões Norte e Centro-oeste são as que menos possuem mulheres na área STEM, tendo uma variação discreta entre os anos. Enquanto o Sul e o Nordeste nos anos de 2010 até 2012 tiveram ascensão das mulheres na graduação, mas após mantiveram linear com discretas oscilações.

Para ampliar essa percepção quanto a região, analisou-se através do gráfico de dispersão as mulheres ingressantes e concluintes dos estados da região Sudeste. São Paulo se destacou como o estado de maior número de mulheres que ocupam as vagas de graduação na área STEM, conforme Figura 5 .





**Figura 4. Gráfico comparativo entre mulheres concluintes por região**

#### **4.1.5. Tem mais mulheres concluintes na graduação da área STEM da Rede de Ensino Pública ou da Privada?**

De acordo com a Figura 6 as mulheres na rede de Ensino Pública continua em ascensão desde o início da década, enquanto na rede de Ensino Privada teve um aumento nos anos de 2010 a 2013, porém vem desacelerando.

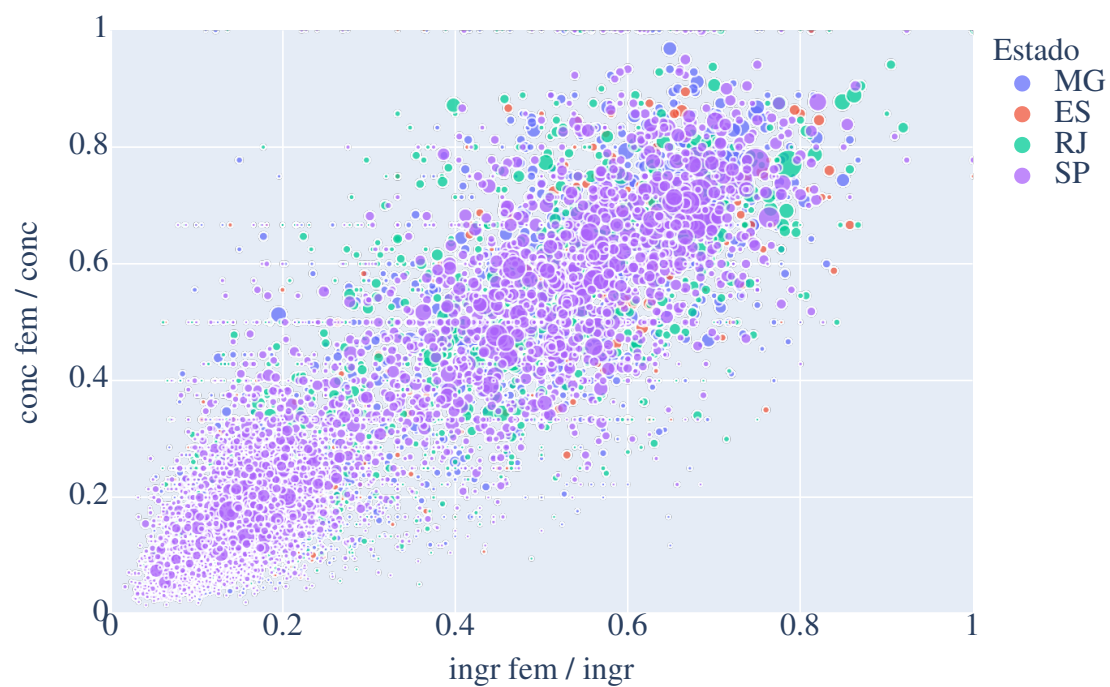
#### **4.1.6. Quantas mulheres concluintes na graduação da área STEM cursaram na modalidade presencial ou à distância ?**

Afim de contextualizar antes de discutir os resultados, o dataset utilizado apresentou microdados liberados para consulta até 2019 e assim, antes do momento da pandemia no mundo e no Brasil. A análise do gráfico de dispersão a seguir não tem influência do aspecto educacional durante a pandemia, onde o ensino remoto, híbrido e à distância tiveram um aumento devido a necessidade de isolamento e a impossibilidade do ensino presencial.

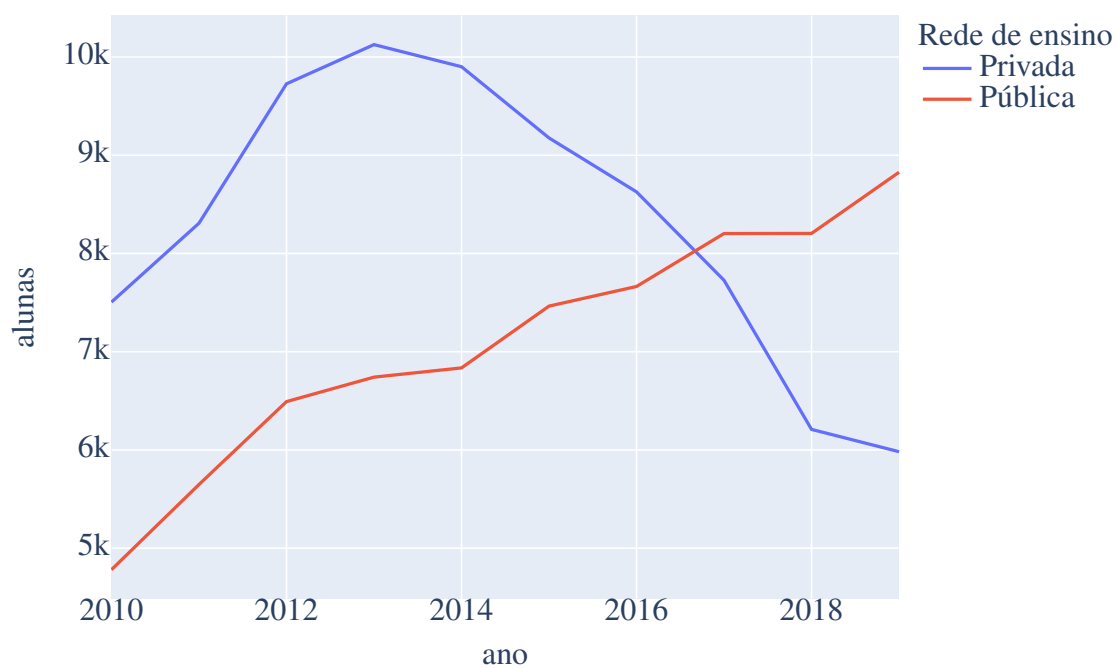
Então, em relação aos anos de 2010 à 2019, as mulheres concluintes cursaram o ensino presencial na maior proporção, conforme Figura 7.

#### **4.2. Proveniência do Dataset**

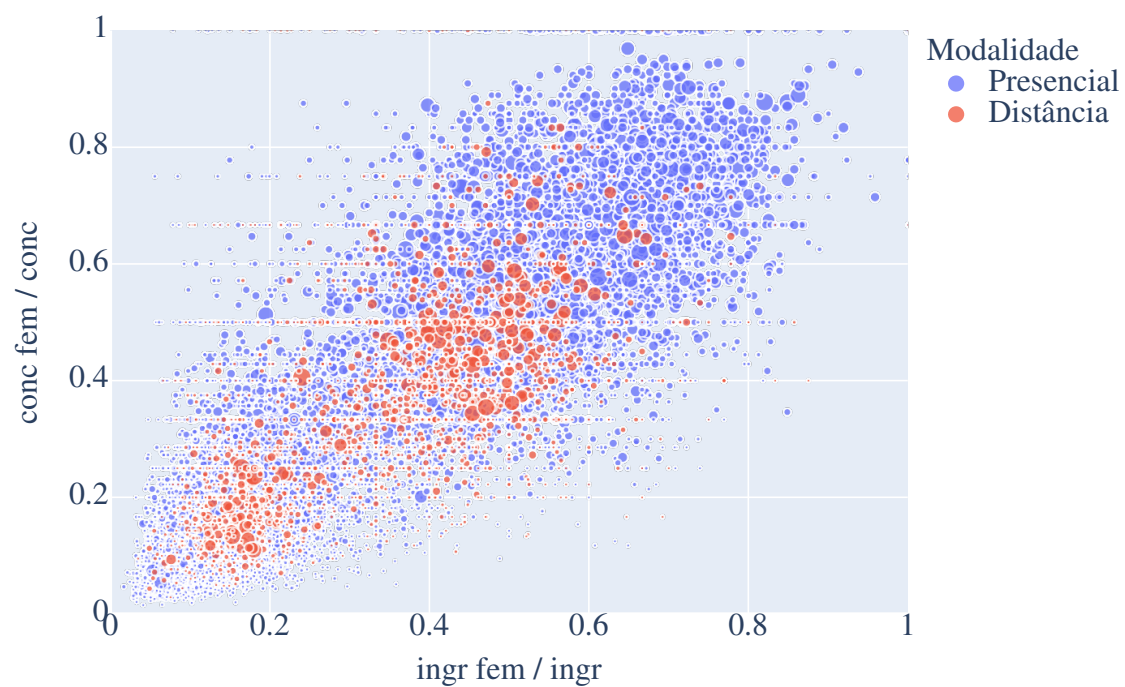
Proveniência é definida como um registro que descreve as pessoas, instituições, entidades e atividades envolvidas na produção, influência ou entrega de um dado ou coisa. Em



**Figura 5. Mulheres que ocupam vagas de graduação na região Sudeste na área STEM**



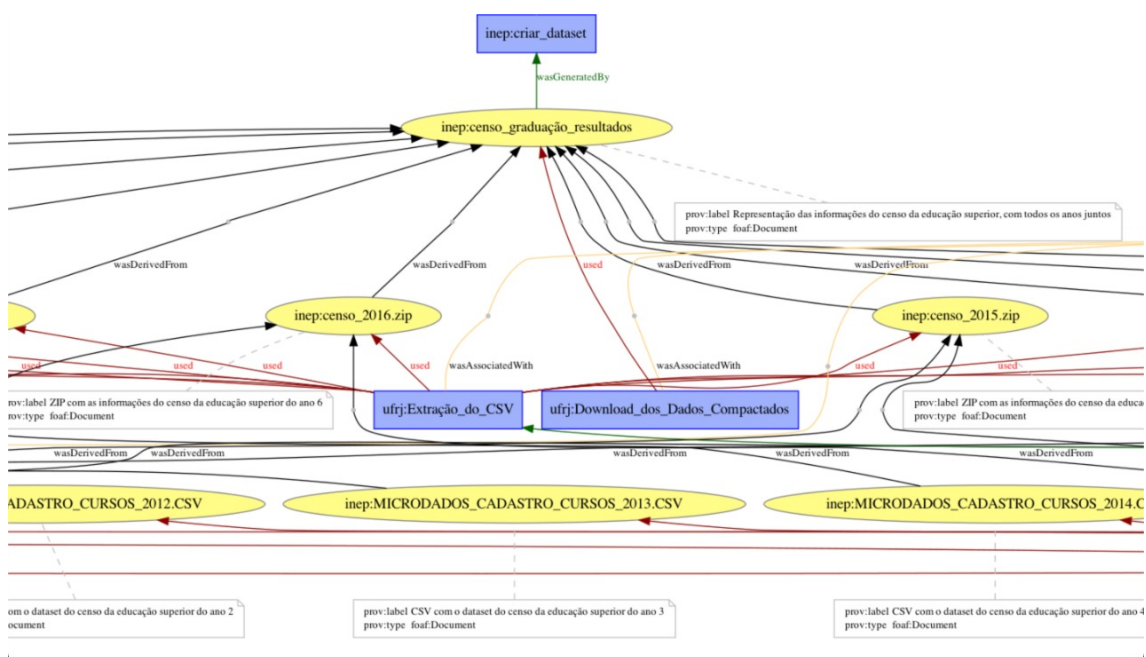
**Figura 6. Mulheres que ocupam vagas de graduação na rede pública e privada na área STEM**



**Figura 7. Mulheres concluintes que cursaram na modalidade presencial ou à distância**

particular, a proveniência da informação é crucial para decidir se a informação é confiável, como deve ser integrada com outras fontes de informação diversas e como dar crédito aos seus originadores ao reutilizá-la [Moreau and Missier 2013].

A proveniência dos dados é importante para agregar metadados sobre a origem, sua obtenção e mais informações sobre os dados, o que enriquece a pesquisa científica e a identificação da autoria [Buneman and Tan 2007]. Na pesquisa utilizamos a biblioteca *prov*, que é a alternativa em *Python* para o modelo fornecido em [Missier et al. 2013]. A Figura 8 traz um recorte do grafo de proveniência gerado a partir disso, cujo código e imagem completa se encontram no repositório<sup>18</sup>.



**Figura 8. Parte do Grafo de Proveniência**

A figura retrata:

1. Nome do agente responsável pela execução de cada atividade;
2. Informações adicionais sobre o agente que descrevem sua identificação e seu papel;
3. Os nomes das atividades executadas;
4. Atribuição da execução das atividades ao agente;
5. Especificação do dataset utilizado, assim como em que lugar pode ser encontrado;
6. Data e hora das atribuições;

## 5. Considerações Finais

No Brasil, apesar das mulheres terem maior inserção no ensino superior, predominam em cursos de Ciências Humanas e Sociais Aplicadas. Com escassez de pesquisas relacionadas a presença das mulheres na área STEM no ensino superior, nossa motivação inicial foi ampliar o entendimento desse cenário, a partir de dataset bruto e aberto do INEP.

<sup>18</sup><https://github.com/lucaslopes/alunas>

O desenvolvimento deste trabalho envolveu pesquisa bibliográfica sobre a inserção das mulheres nos cursos STEM do Brasil, seleção, tratamento e aquisição dos dados resultando em dataset curados e abertos, criação de dicionário de dados, verificação da proveniência e outras investigações.

Os resultados demonstraram que a arquitetura computacional proposta é adequada para a gestão dos metadados de proveniência, contendo informações de como, quando, onde e por que os dados foram obtidos e quem os produziu. A pesquisa possui disponibilização na Web dos metadados de proveniência, para possibilitar pesquisadores consultar a proveniência de recursos de dados e experimentos científicos, apoiando a compreensão, reprodutibilidade e o reuso destes dados, através do notebook que fornece as características do ambiente em que executamos e como fizemos.

A pesquisa também engloba uma visão educacional e social, explorando uma série histórica de dados de mulheres no ensino superior, equivalente a uma década, na área STEM. Uma contribuição foi a averiguação da área de Engenharia e profissões correlatas como sendo onde as mulheres estão mais presentes nos cursos de graduação da área STEM, seja como ingressante e também concluinte. O gráfico apresentado na Figura 4, que sinaliza as vagas por área, se assemelha a série histórica por área apresentado na Figura 2 e também o da Figura 5, por região. Assim, leva-nos a entender que a área da Engenharia e profissões correlatas tem uma grande influência no cenário das mulheres na área STEM, com destaque para a região Sudeste, especificadamente em São Paulo.

Outras informações também foram possíveis e trouxeram clareza para o estudo, como a presença crescente nos cursos da Rede de Ensino Pública, contrastando com o declínio da participação na Rede de Ensino Privada. A participação nesse período ser predominantemente presencial. Além de informações por regiões e estados.

Como trabalhos futuros pretende-se aplicar os princípios FAIR que visam tornar os datasets mais localizáveis, acessíveis, interoperáveis e reutilizáveis [Wilkinson et al. 2016]. Após prover os itens no Zenodo <sup>19</sup>, que atende aos padrões relevantes da comunidade de interesse além de fornecer informações sobre a proveniência, com os dados liberados pela licença de uso clara e acessível, a Creative Commons CC-BY, que permite distribuir, alterar ou reinventar o dataset, desde que se cite a fonte inicial.

## Referências

- Buneman, P. and Tan, W.-C. (2007). Provenance in databases. In *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*, pages 1171–1173.
- de Araújo, M. T. and Tonini, A. M. (2020). A participação das mulheres nas áreas de stem (science, technology engineering and mathematics). *Revista de Ensino de Engenharia*, 38(3).
- Iwamoto, H. M. (2022). Mulheres nas stem: um estudo brasileiro no diário oficial da união. *Cadernos de Pesquisa*, 52:e09301–e09301.
- Missier, P., Belhajjame, K., and Cheney, J. (2013). The w3c prov family of specifications for modelling provenance metadata. In *Proceedings of the 16th International Conference on Extending Database Technology*, pages 773–776.

---

<sup>19</sup><https://zenodo.org/>

- Moreau, L. and Missier, P. (2013). Prov-dm: The prov data model. <https://www.w3.org/TR/2013/REC-prov-dm-20130430/>. (Accessed on 09/25/2022).
- Silva, U. F., Ferreira, D. J., Ambrósio, A. P. L., and Oliveira, J. L. d. S. (2022). Problemas enfrentados por alunas de graduação em ciência da computação: uma revisão sistemática. *Educação e Pesquisa*, 48.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., et al. (2016). The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9.