

# Plataforma para estudo de participação e desempenho dos gêneros no Centro de Informática

Rodrigo Augusto Vasconcelos Sarmento



CENTRO DE INFORMÁTICA  
UNIVERSIDADE FEDERAL DA PARAÍBA

João Pessoa, 2018

Rodrigo Augusto Vasconcelos Sarmiento

# Plataforma para estudo de participação e desempenho dos gêneros no Centro de Informática

Monografia apresentada ao curso Engenharia de Computação  
do Centro de Informática, da Universidade Federal da Paraíba,

Orientadora: Thaís Gaudêncio do Rêgo

Novembro de 2018

Ficha catalográfica: elaborada pela biblioteca do CI.

Será impressa no verso da folha de rosto e não deverá ser contada.

Se não houver biblioteca, deixar em branco.

**Catálogo na publicação**  
**Seção de Catalogação e Classificação**

S246p Sarmento, Rodrigo Augusto Vasconcelos.

Plataforma para estudo de participação e desempenho dos gêneros no Centro de Informática / Rodrigo Augusto Vasconcelos Sarmento. - João Pessoa, 2018.

63 f.

Orientação: Thaís Gaudêncio Rêgo.  
Monografia (Graduação) - UFPB/CI.

1. Gênero. 2. Computação. 3. Ciência de dados. 4. Visualização de dados. I. Rêgo, Thaís Gaudêncio. II. Título.

UFPB/BC



CENTRO DE INFORMÁTICA  
UNIVERSIDADE FEDERAL DA PARAÍBA

Trabalho de Conclusão de Curso de Engenharia de Computação intitulado  
***Plataforma para estudo de participação e desempenho dos gêneros no Centro de Informática***  
de autoria de Rodrigo Augusto Vasconcelos Sarmento, aprovada pela  
banca examinadora constituída pelos seguintes professores:

*Thaís Gaudêncio do Rêgo*

Prof. Dr. Thaís Gaudêncio do Rêgo  
Universidade Federal da Paraíba

*Ed Porto Bezerra*

Prof. Dr. Ed Porto Bezerra  
Universidade Federal da Paraíba

*Lincoln David Nery e Silva*

Prof. Dr. Lincoln David Nery e Silva  
Universidade Federal da Paraíba

João Pessoa, 8 de novembro de 2018

Centro de Informática, Universidade Federal da Paraíba  
Rua dos Escoteiros, Mangabeira VII, João Pessoa, Paraíba, Brasil CEP: 58058-600  
Fonc: +55 (83) 3216 7093 / Fax: +55 (83) 3216 7117

## **AGRADECIMENTOS**

À minha orientadora Thaís Gaudêncio pelos ensinamentos durante esses anos de graduação que foram cruciais para o meu desenvolvimento.

Agradeço aos meus avós e a toda minha família que me acompanharam na árdua criação e motivação do saber.

Por fim, agradeço à oportunidade de ter estudado em uma Universidade Federal com todo o apoio e suporte que o governo, por meio de suas bolsas de incentivo, me proporcionou.

## RESUMO

Problemas relacionados a gênero afetam instituições que são base para a sociedade, como universidades e empresas. Pensando nisso, e tendo em vista a importância da universidade como modelo social e do valor da entendimento de dados, é apresentado aqui um trabalho que expõe visualizações interativas quanto às questões de participação e desempenho dos discentes de acordo com o gênero, dentro de uma dessas instituições, a Universidade Federal da Paraíba (UFPB), mais especificamente nos cursos do Centro de Informática (CI), Engenharia de Computação, Matemática Computacional e Ciência da Computação, entre os períodos acadêmicos que vão de 1985.1 e 2018.2.

**Palavras-chave:** <Gênero>, <Computação>, <Ciência de dados> <Visualização de dados>

## ABSTRACT

Gender-related problems affect institutions that are the bases of our society, such as universities and companies, thinking about this, and considering the importance of the university as a social model and the value of understanding the data, is presented here a work that exposes interactive visualizations as to the issues of participation and performance of students according to gender, within one of these institutions, the Federal University of Paraíba (UFPB), more specifically in the Center of Informatics (CI), in the majors of Computer Engineering, Computational Mathematics and Computer Science, between the academic periods of 1985.1 and 2018.2.

**Key-words:** <Gender>, <Computer major>, <Data science> <Data visualization>

## LISTA DE FIGURAS

1	Rendimento habitual médio mensal de todos os trabalhos e razão de rendimentos por sexo. Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios Contínua 2012-2016 [14] . . . . .	17
2	Taxa de frequência escolar líquida, ajustada no ensino médio, por sexo e cor ou raça(%). Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios Contínua 2016 [14] . . . . .	18
3	O que mulheres e homens estudam? Dados globais da relação de participação por área de estudo. Adaptado do: Index Global de Desigualdade de Gênero 2017 (Global Gender Gap Index 2017), Fórum Econômico Mundial (World Economic Forum). 2017 [21] . . . . .	20
4	Exemplo da utilização de gráficos do tipo barra para mesurar o impacto individual de cada elemento ao longo de uma série temporal. Adaptado de Tableau [28]. . . . .	23
5	“ Pipeline de análise de big data mostrando o papel da visualização de dados e a colaboração do ser humano” [39]. Adaptado de Gandomi e Haider, 2015 . . . . .	25
6	Árvore representando os dados contidos nos arquivos JSON. Fonte: Autoria própria . . . . .	26
7	Seleção das cores. Adaptado de Tableau . . . . .	29
8	Visualização do menu de escolha das abas . . . . .	31
9	Página web com os resultados da pesquisa. . . . .	34
10	Visualização da <i>Dashboard</i> com o percentual de gênero da composição total dos cursos. Na figura, % do total de Contagem de Gênero para cada Curso. A cor mostra detalhes sobre Gênero. As marcas são rotuladas por % do total de Contagem de Gênero. Fonte: Autoria própria . . . . .	36
11	Visualização da <i>Dashboard</i> que contém a série temporal da quantidade de alunos e alunas que entram. Utilização do filtro curso. Fonte: Autoria própria . . . . .	37
12	Visualização da <i>Dashboard</i> com a série temporal do percentual de alunas que entram. Utilização do filtro curso. Fonte: Autoria própria . . . . .	38
13	Visualização da <i>Dashboard</i> com a composição dos cursos. Utilização do filtro período. Fonte: Autoria própria . . . . .	39
14	Visualização da <i>Dashboard</i> . Fonte: Autoria própria . . . . .	40

15	Visualização da <i>Dashboard</i> . Fonte: Autoria própria	41
16	Visualização da <i>Dashboard</i> com o percentual, por gênero, de conclusão do curso. Fonte: Autoria própria	42
17	Visualização da <i>Dashboard</i> com a série temporal do percentual de participação por disciplina. Utilização interativa dos filtros curso e disciplina. Fonte: Autoria própria	44
18	Visualização da <i>Dashboard</i> com a série temporal da média das notas considerando as situações aprovado e reprovado. Utilização interativa dos filtros curso e disciplina	45
19	Visualização da <i>Dashboard</i> com a série temporal do percentual de aprovação considerando as situações aprovado, reprovado, reprovado por falta e trancado. Utilização interativa dos filtros curso, disciplina e gênero. Fonte: Autoria própria	46
20	Visualização da <i>Dashboard</i> com o percentual de participação e percentuais de aprovação, dispensa, reprovação por falta, reprovação e trancamento. Utilização interativa dos filtros curso, período e disciplina. Fonte: Autoria própria	47
21	Legenda das cores para o Fluxograma da Computação. Fonte: Autoria própria	48
22	Visualização do recurso interativo fluxograma para o curso Engenharia de Computação. Fonte: Autoria própria	49
23	Visualização do recurso interativo fluxograma para o curso Ciência da Computação. Fonte: Autoria própria	50
24	Visualização do recurso interativo fluxograma para o curso de Matemática Computacional. Fonte: Autoria própria	51
25	Visualização da <i>Dashboard</i> com o local de origem dos discentes. Fonte: Autoria própria	53
26	Visualização da <i>Dashboard</i> com a comparação por gênero quanto à origem dos discentes. Fonte: Autoria própria	54
27	Visualização da <i>Dashboard</i> com o rank referente ao local de origem dos discentes. Fonte: Autoria própria	55
28	Visualização da <i>Dashboard</i> com informação referente ao local de origem dos discentes do Centro de Informática antes e depois do ENEM. Fonte: Autoria própria	56

29	Resultado para protonotopia utilizando [38]	62
30	Resultado para deuteranopia [38]	62
31	Resultado para tritanopia [38]	63
32	Resultado para achromatopsia [38]	63

## LISTA DE TABELAS

1	Respostas do questionário: “Como você se posiciona no problema relacionado ao azul-rosa? Se você fosse um designer escalado para definir cores para valores ligados a gênero em seu gráfico, qual seria a sua escolha?”.	
	Adaptada de [31] . . . . .	24

## LISTA DE ABREVIATURAS

- CI - Centro de Informática
- CNPq - Conselho nacional de Desenvolvimento Científico e Tecnológico
- ENADE - Exame Nacional de Desempenho dos Estudantes
- ENEM - Exame Nacional do Ensino Médio
- IME - Instituto Militar de Engenharia
- IDC - *International Data Corporation* (Corporação Internacional de Dados)
- ILO - *International Labor Organization* (Organização Internacional do Trabalho)
- INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
- ONU - Organização das Nações Unidas
- SISU - Sistema de Seleção Unificada
- SIGAA - Sistema Integrado de Gestão de Atividades Acadêmicas
- STEM - *Science, Technology, Engineering and Mathematics* (Ciência, Tecnologia, Engenharia e Matemática)
- TIC - Tecnologia da Informação e Comunicação
- UNESCO - *The United Nations Educational, Scientific and Cultural Organization* (Organização das Nações Unidas para a Educação, a Ciência e a Cultura)
- UFPB - Universidade Federal da Paraíba
- VISIO - Laboratório de Visão Computacional
- WEF - *World Economic Forum* (Fórum Econômico Mundial)

## Sumário

<b>1 INTRODUÇÃO</b>	<b>15</b>
1.1 Definição do problema	16
1.2 Objetivo geral	16
<b>2 CONCEITOS GERAIS E REVISÃO DA LITERATURA</b>	<b>17</b>
2.1 Discrepância de gênero	17
2.2 Discrepância de gênero no âmbito universitário	18
2.3 Computação e discriminação	21
2.4 Ciência de dados	21
2.5 Visualização de dados	22
2.6 As cores da visualização	24
<b>3 METODOLOGIA</b>	<b>25</b>
3.1 Aquisição de dados	25
3.2 Tratamento dos dados	27
3.3 Geração de infográficos	28
3.4 Página Web	30
3.5 Ambiente Computacional	30
<b>4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS</b>	<b>31</b>
4.1 Ambiente proposto	31
4.2 Visualizações Propostas	31
4.3 Pagina inicial	33
4.4 Participação	35
4.5 Desempenho	43
4.6 Mapa	52
<b>5 CONCLUSÕES E TRABALHOS FUTUROS</b>	<b>57</b>
<b>REFERÊNCIAS</b>	<b>57</b>



# 1 INTRODUÇÃO

A universidade é tida como um ambiente que assume compromissos de igualdade, devido ao seu papel de busca intensiva pelo conhecimento. O princípio da igualdade deve orientar as políticas das instituições de ensino superior [1], sabendo que a promoção da igualdade de gênero é um dos objetivos do milênio estabelecidos pela ONU, figurando em terceiro lugar entre suas oito metas [9] e também levando em consideração o seu papel emancipador, através da formação profissional [9]. Porém, o que diríamos ao observarmos que esses deveres para com o avanço e a igualdade nem sempre são confirmados?

É preciso voltar um pouco no tempo e entender a contextualização dos estudos de desigualdade de gênero e suas terminologias. Eles começam a surgir por volta da década de 1970, como consequência de uma nova era da revolução feminista, que ocorre quando o movimento começa a se preocupar com construções teóricas. Nesse momento histórico, temos o entendimento do conceito de “Gênero” tomando um caráter diferente do de “Sexo” [2], e isso tendo por objetivo “rejeitar um determinismo biológico implícito no uso de termos como sexo ou diferença sexual” [5].

A questão dessa rejeição é primordial para o debate da diferença na quantidade de homens e mulheres em diversos cursos de nossa instituição. Discordar do determinismo biológico, confronta as barreiras criadas pela sociedade, que limitam as possibilidades de participação feminina em certos âmbitos do conhecimento. Gênero é um marcador social de diferença que resulta em exclusão e influencia na posição de mulheres e homens em nossa sociedade [9].

Uma área que apresenta lacuna pertinente quanto a gênero é a da computação. No Instituto Militar de Engenharia (IME) a participação feminina na primeira turma, na década de 70, era de 70%, diminuindo ao longo dos anos em chegando a 15% em 2016 [16].

Essa queda de atuação feminina é fomentada pela construção de um imaginário que define áreas particulares de atuação para o gênero feminino e masculino. Esse cenário vem sendo contestado com programas do governo como o Programa Ensino Médio Inovador, que em 2013 buscou incentivar a participação feminina em cursos de Ciências Exatas [10]. Porém, o problema continua a existir, tendo em vista que a soma da participação feminina, no Brasil, no ano de 2017, nas áreas de graduação em **Engenharia, Manufatura e Construção 5,8%; Informação e Comunicação, Tecnologias 0,9%; Ciências Naturais, Matemática e Estatística 2,6%** que constituem esse imaginário, não ultrapassou 9,3% [21].

## 1.1 Definição do problema

Com a quantidade de dados provenientes do ambiente universitário, sabendo dos problemas de discrepância de gênero, do papel da Universidade como ambiente que proporciona o desenvolvimento social, das capacidades da área de Ciência de Dados e seus métodos atuais de análise e visualização, a pesquisa em questão visa expor dados de participação e desempenho de discentes, quanto ao gênero, dentro da instituição de ensino superior UFPB, nos cursos do centro de informática (CI). Isso tudo vem de acordo com a tendência global de utilização de dados para a compreensão do comportamento humano, especialmente em grupos [3], uma vez que os dados conseguem explicar as “estruturas sociais” que as pessoas tendem a formar [4].

## 1.2 Objetivo geral

Utilização de ferramentas de Ciência de Dados para a formação de base de dados organizadas e geração de conteúdo gráfico para a criação de uma página web com visualizações entrada, desistência e desempenho nos Cursos do Centro de Informática (CI) da Universidade Federal da Paraíba (UFPB).

## 2 CONCEITOS GERAIS E REVISÃO DA LITERATURA

### 2.1 Discrepância de gênero

As mulheres constituem quase metade da população mundial com 49,556% [11], porém, possuem uma participação menor no mercado de trabalho com uma taxa de desemprego 0,8% maior que a dos homens, que fica em torno dos 6% [12]. Além disso, as mulheres possuem uma taxa global de participação na força de trabalho de 48,5% em 2018, que ainda é 26,5% menor que a taxa dos homens. Para agravar a situação, na maior parte dos países, os homens tendem a ganhar mais para cargos semelhantes, que as mulheres. Segundo dados do IBGE [13], o salário médio pago às mulheres, em 2017, foi de apenas 77,5% do rendimento masculino. Homens recebem em média R\$ 2.410, enquanto mulheres R\$ 1.868. A Figura 1, traz dados sobre a diferença no rendimento entre os anos de 2012 e 2016.

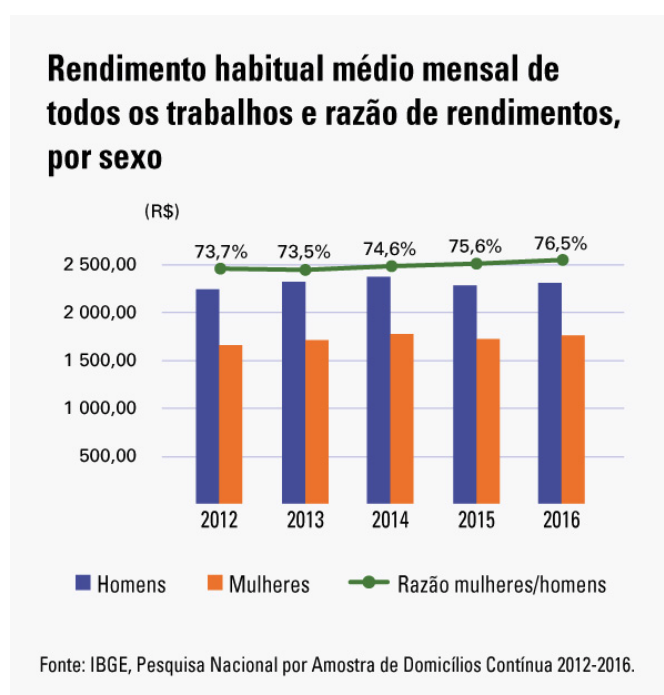
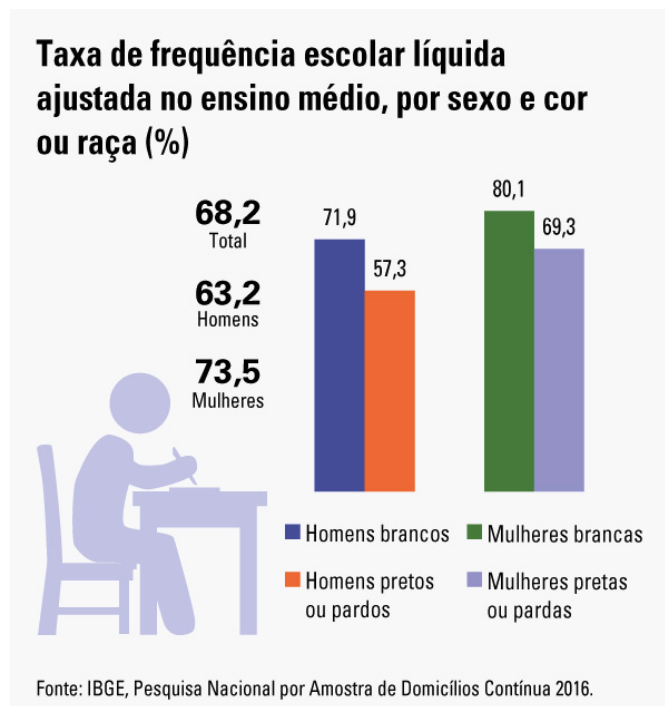


Figura 1: Rendimento habitual médio mensal de todos os trabalhos e razão de rendimentos por sexo. Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios Contínua 2012-2016 [14]

É importante observar que essa diferença salarial não condiz com a capacitação. Estudos do próprio IBGE demonstram que mulheres estudam mais, trabalham mais e ganham menos do que o homens, mesmo com um nível educacional superior [14], como pode ser observado na Figura 2.

Mesmo estudo da Organização Internacional do Trabalho (OIT) [12] reforça a importância do equilíbrio entre os gêneros no mercado de trabalho, o documento aponta



**Figura 2:** Taxa de frequência escolar líquida, ajustada no ensino médio, por sexo e cor ou raça(%). Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios Contínua 2016 [14]

um crescimento do PIB, produto interno bruto que representa a soma de todos os bens e serviços finais produzidos numa região, durante um período, nacional em até R\$ 382 bilhões se, por exemplo, o Brasil conseguir reduzir em 25% o desequilíbrio no mercado, incentivando a participação feminina nas mais diversas áreas de atuação.

Ainda com relação ao mercado de trabalho, a visão estereotipada de algumas áreas da ciência tende a prejudicar a participação feminina nos chamados “empregos do futuro”, profissões que derivam de Ciência (Science), Tecnologia (Technology), Engenharia (Engineering) e Matemática (Mathematics), resumidas pela sigla STEM [19], principalmente, as que exigem habilidades das áreas de Tecnologia da Informação e Comunicação (TIC), que dominarão grande parte do mercado. Estima-se que, até 2020, 98% dos empregos relacionados a STEM exigirão habilidades em TIC, e haverá por volta de 1 milhão de postos vazios na área de computação [19]. As mulheres representam apenas 3% dos estudantes graduados em áreas ligadas a TIC em todo o mundo [19].

## 2.2 Discrepância de gênero no âmbito universitário

Quando falamos de sociedade e especificamente de mercado de trabalho é importante refletir sobre o ambiente de onde os membros que o constituem derivam e esse lugar é comumente a universidade que, como foi dito na introdução, é tido como um espaço que busca compromissos de igualdade e que reflete as tendências de uma sociedade.

Atualmente, no ambiente universitário, dados oficiais divulgados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), apontam um número crescente de mulheres no ensino superior, com um aumento de 81,14% no número total de matrículas. Entre 2000 e 2007 houve um aumento de 76,92% no número de mulheres [6], e em números brutos de participação, temos as mulheres liderando em quantidade de matrículas e concluintes desde 2009, de acordo com o Censo da Educação Superior [7]. Porém, esse processo de avanço na participação feminina é definido como “lento” e pontual, devido à implementação de políticas universalistas de inclusão nos últimos anos dentro do Brasil [9].

Também é importante observar que as diferenças de gênero não podem ser superadas apenas pelo crescimento de participação feminina, em números brutos e sem contexto de atuação, lembrando que permanece, dentro do âmbito educacional, a perpetuação de preconceitos que definem áreas de atuação específicas para mulheres e homens. Isso se deve a um imaginário sexista, que classifica as mulheres como secundárias, as afasta de determinadas áreas e dificulta a inserção feminina em cargos efetivos de poder dentro das instituições de ensino superior [22].

Por isso, é relevante observar o quanto a relação de pares, influência na entrada e permanência das mulheres, nas áreas de STEM, tendo em vista as oportunidades de emprego do futuro, por serem os ramos que mais crescem com a automatização do mercado de trabalho. Estudos predizem que aproximadamente 800 milhões de trabalhos serão perdidos, até 2030, com a automatização [20].

Desde os primeiros níveis as meninas internalizam estereótipos, pois observam a baixa quantidade de mulheres representando profissões ligadas às áreas de STEM, e quando conseguem ingressar em algum curso da área se deparam com uma baixa quantidade de mulheres docentes gerando um intenso sentimento de não pertencimento. Em [19] se descreve a necessidade de pertencimento com o campo de estudo. A falta de representatividade explica o fato das diferenças de gênero por área.

A Figura 3 demonstra a divisão de gênero por áreas de atuação, tanto no Brasil quanto no mundo. Refletindo a permanência de um desequilíbrio, dentro dos campos da ciência, tendo maior participação feminina as áreas de educação e saúde, que são sinônimos de um antigo, e ainda presente, estereótipo da mulher na sociedade.

### O que Mulheres e Homens Estudam?

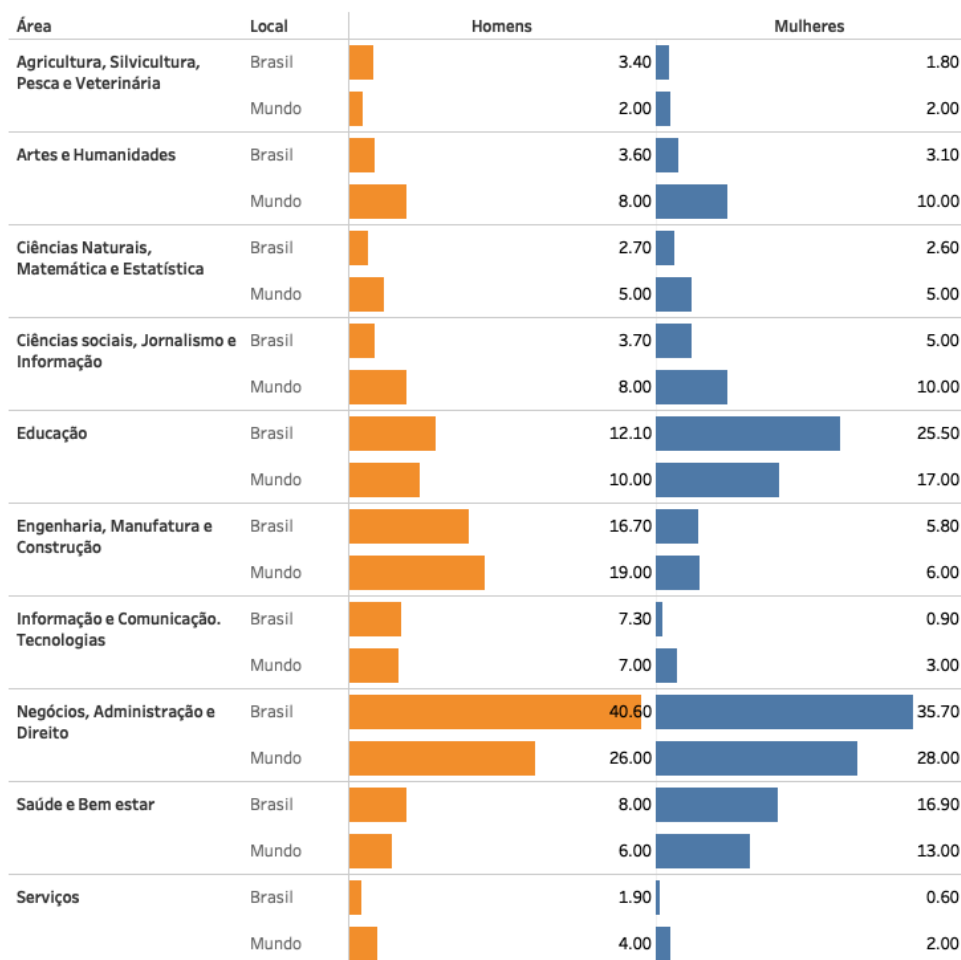


Figura 3: O que mulheres e homens estudam? Dados globais da relação de participação por área de estudo. Adaptado do: Index Global de Desigualdade de Gênero 2017 (Global Gender Gap Index 2017), Fórum Econômico Mundial (World Economic Forum). 2017 [21]

### 2.3 Computação e discriminação

Particularmente, na área da computação, os números estão piorando. Tomando, por exemplo, a ciência da computação na América Latina e Caribe, temos uma queda de 2 a 13%, desde 2000 [15], no número de mulheres graduadas. Mais especificamente no Brasil, desde meados da década de 90, é observável uma queda na participação feminina. Na computação a atuação caiu de 30% para 5 ou 10% nos últimos anos [8]. A mesma coisa acontece em países de outras regiões, internacionalmente a representação feminina em ciência da computação varia entre 9% e 25%, especificamente nos Estados Unidos da América (EUA), o número de mulheres na graduação de computação, em 1985, era de 37% [8], em 2010-2011, esse número caiu para 17,6% [24].

Importante relembrar, para o caso brasileiro, que a computação nasce nas universidades, na década de 70, com turmas majoritariamente compostas por mulheres, exemplo das primeiras turmas da Universidade Estadual de São Paulo (USP) [16]. Nesse momento, existia uma ligação da computação com a área da matemática (licenciatura), e era presente a relação da licenciatura com o gênero feminino [22]. Assim, nessa circunstância inicial de correlação entre computação e licenciatura existiam mais mulheres participando da área. Porém, com o crescimento e popularização da tecnologia, surge a ideia da computação “pop” que se conecta com o universo masculino [17] [23], por meio dos primeiros computadores pessoais que eram em sua essência vendidos como sistemas de vídeo game que atendiam às necessidades dos meninos, dessa forma, o setor rapidamente “expurga” as mulheres, criando um estigma com relação à sua inserção [16].

### 2.4 Ciência de dados

Sabendo da problemática da desigualdade de gênero e tendo em vista a questão de essa ser uma das metas do milênio [9] surge a necessidade da utilização de formas novas de análise do comportamento humano para a área, e uma delas é a ciência de dados [3].

Ao longo dos últimos anos, essa ciência, que envolve as habilidades de obtenção, organização, previsão e visualização dos dados, começa a sua expansão. Surge assim, devido ao crescimento exorbitante de dados, a necessidade de um cientista de dados, alguém que possui a capacidade de tratar toda essa informação e gerar resultados filtrados da mesma. Por exemplo, a Corporação Internacional de Dados ( *International Data Corporation*, IDC) [18] estima que em 2020 teremos aproximadamente 44 zetabytes de informação, sendo o custo de manter essa informação lucrativo. Porém, ao mesmo tempo, a quantidade de pessoas capazes de lidar com esse montante continua com baixo crescimento, no presente ano temos a estimativa da falta de 1,5 milhões de analista e gestores de dados, somente nos Estados Unidos [26], por isso as possibilidades da área ainda são pouco aproveitadas e a tendência é de uma expansão ainda maior pro setor.

Beneficiados, dessa expansão, serão países em crescimento que devido às suas próprias deficiências, quanto à criação de dados concretos sobre suas populações, sofrem no combate de suas mazelas. Dados são extremamente importantes nas mãos daqueles que tomam as decisões de uma população, por retratarem o ambiente em que são coletados, ou seja, o crescimento dos dados digitais com a “revolução de dados”, surge para baratear o custo ao acesso à informação, para as instituições que lidam com a sociedade, especialmente nesses países em ascensão [3], que antes gastavam tempo e dinheiro com a necessidade de pesquisa em campo para tomadas de ações.

Um exemplo de utilização foi a análise de informação política. Durante a campanha de presidenciais, nos Estados Unidos, Obama, em 2008, já possuía equipes para confrontar dados eleitorais. Sua equipe teve que lidar com cinco terabytes de dados desorganizados de registro eleitoral, de múltiplas fontes e formatos, e dessa forma foram capazes de limpar mais de 1,6 milhões de eleitores da lista de ligações à serem realizadas, diminuindo em 25% o trabalho do setor [25], demonstrando o poder da área.

## 2.5 Visualização de dados

Tendo noção do que foi tratado na seção anterior é possível compreender o papel fundamental dos dados em nossa sociedade. A grande dificuldade quanto aos dados reside no seu entendimento, muitos se apresentam em formatos pouco palatáveis para o público geral. Partindo do princípio, citado em [27], que aponta que uma imagem pode conter uma vasta quantidade de informação, que pode ser processada muito mais rapidamente que um conjunto de palavras, a questão é: “Como transformar nossos dados em visualizações que gerem impacto e disseminem a informação de forma concisa e clara?”.

A base para construção de uma boa visualização de dados parte da cautela para não encher o usuário de informação, criando formas para que ele consiga desenhar conclusões e interagir de forma direta [27].

Para isso, umas das necessidades é se ater à forma como o dado se distribui. A maior parte dos dados apresentados neste artigo, por exemplo, são orientados ao tempo, relação período a período da participação e desempenho dos alunos, por gênero, seguindo o que é sugerido em [27], que defende a necessidade de uma abordagem padronizada de dados desse tipo, são utilizados neste projeto, preferencialmente, gráficos do mesmo tipo (barra e linha).

Utiliza-se, então, para todas visualizações de séries históricas, por exemplo, gráficos do tipo barra. Esse tipo de visualização busca, o que é descrito em [29], a percepção de tendências, tão importantes, em séries do tipo temporal. Em [28], guia da empresa de software Tableau, também é defendida a utilização dos mesmos tipos, existe ainda a explicação da utilização específica de gráficos do tipo barra com o objetivo de mensurar

o impacto individual de cada elemento ao longo do tempo [28], como pode ser observado na Figura 4, ideal para os comparativos de desempenho e participação de gênero ao longo dos períodos.

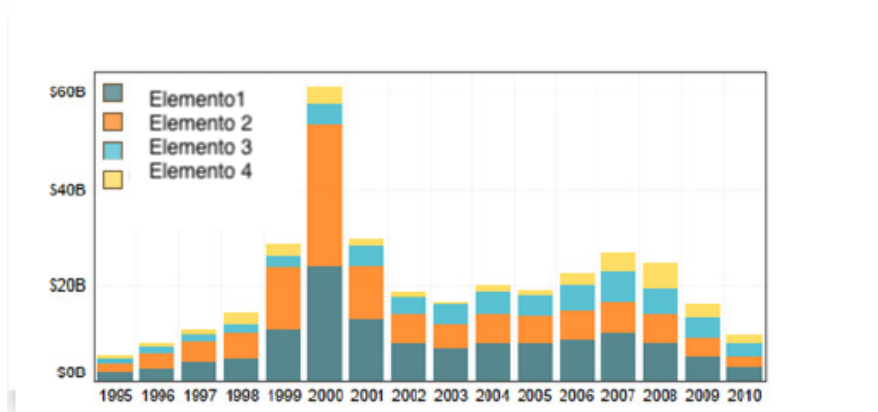


Figura 4: Exemplo da utilização de gráficos do tipo barra para mesurar o impacto individual de cada elemento ao longo de uma série temporal. Adaptado de Tableau [28].

## 2.6 As cores da visualização

Por uma questão de percepção, sobre o tópico apontado em [30] que demonstra que o cerne do esteriótipo está representado na associação de cor e gênero, neste trabalho leva-se em consideração as cores advindas das discussões e visualizações presentes em [31, 19], como pode observar-se na Tabela 1. Para evitar generalizações relacionadas a esteriótipos e também considera-se o problema do daltonismo, tendo em vista, que afeta 8% dos homens e 0,5% das mulheres [32].

Assim, as cores laranja e azul são escolhidas, de forma tal que laranja representa os homens e azul as mulheres, pelo respeito quantos às limitações visuais e à quebra de padrões relacionados a gênero.

**Tabela 1: Respostas do questionário: “Como você se posiciona no problema relacionado ao azul-rosa? Se você fosse um designer escalado para definir cores para valores ligados a gênero em seu gráfico, qual seria a sua escolha?”.** Adaptada de [31]

	Mulher	Homem
% que utilizaria azul (H) e rosa (M)	14	41
% que utilizaria azul (H) e outra cor (M)	36	29
% que utilizaria dois tipos pares de cores	50	30

### 3 METODOLOGIA

Aqui é apresentada a forma de aquisição e tratamento dos dados, geração de infográficos, criação da página web e ambiente computacional.

A metodologia segue uma parte do pipeline de análise de dados apresentado na Figura 5. Porém, esse projeto não visa realizar a etapa final de interpretação a ideia principal, mais uma vez, é a implementação de um ambiente para que interessados possam analisar o tema.

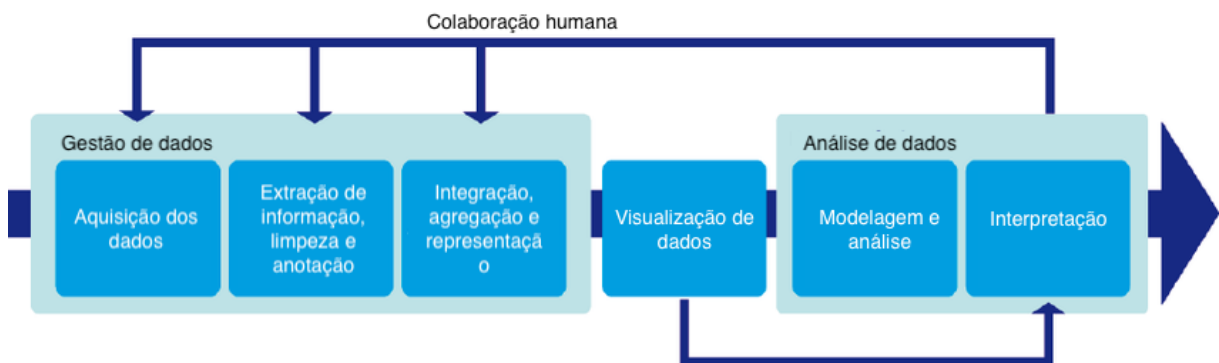


Figura 5: “ Pipeline de análise de big data mostrando o papel da visualização de dados e a colaboração do ser humano” [39]. Adaptado de Gandomi e Haider, 2015

#### 3.1 Aquisição de dados

Foi implementado em [34] um mecanismo para absorção dos dados presentes na plataforma SIGAA (Sistema Integrado de Gestão das Atividades Acadêmicas) da UFPB.

Os dados obtidos vieram em formato **JSON** com um único arquivo para cada aluno dos cursos de Engenharia de Computação Matemática Computacional e Ciência da Computação. Cada arquivo contendo os atributos apresentados na Figura 6

Cada **JSON** aluno continha uma sequência de matérias com todos os atributos listados em **Disciplinas**.

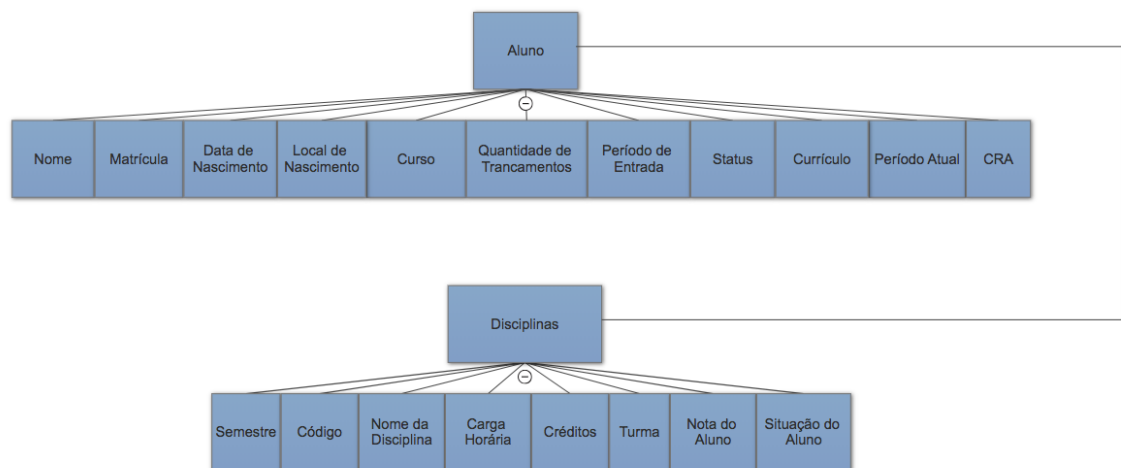


Figura 6: Árvore representando os dados contidos nos arquivos JSON. Fonte: Autoria própria

## 3.2 Tratamento dos dados

Tendo em vista o formato **JSON**, apresentado na seção anterior, fez-se necessária a utilização de um *script* na linguagem *Python* para a construção dos dois seguintes arquivos *Comma-Separated Values* (CSV):

- **Dados alunos**

- **Gênero do Discente**
- Data de Nascimento
- Local de Nascimento
- Curso
- Currículo
- Status
- Quantidade de Trancamentos
- Período de Entrada
- Período Atual
- CRA

- **Dados disciplinas**

- **Gênero do Discente**
- Semestre
- Código
- Nome da Disciplina
- Ch
- Cr
- Turma
- Nota do Aluno
- Situação do Aluno
- Curso

Cada arquivo JSON produziu uma linha no novo arquivo “Dados Alunos” e múltiplas linhas para o arquivo “Dados Disciplinas”, variando de acordo com a quantidade de disciplinas cursadas pelo aluno. Assim, temos a seguinte composição dos arquivos:

- “Dados Alunos” contém 3480 linhas referente aos 3480 alunos que já foram matriculados
- “Dados Disciplinas” contém 58054 disciplinas relacionadas à algum discente matriculado

Em outras palavras cada arquivo JSON cria uma linha em “Dados Alunos” e em média 16,68 linhas “Dados Disciplinas”.

## Pandas

A biblioteca **Pandas** [36] foi uma das utilizadas durante esse processo de organização, por utilizar uma estrutura de dados denominada **pandas.DataFrame**, que facilita a manipulação dos mesmos. Além disso, possui uma estrutura bidimensional de tamanho variável, que funciona como um dicionário para outro tipo de objeto **Pandas**, o **pandas.Series** que é um *array numpy.ndarray*, unidimensional. Sendo assim possível a criação dos arquivos.

## Atributo gênero

O atributo **Gênero do Discente** apresentado nos dois arquivos finais, foi criado através da utilização da biblioteca **Genderize** [37]. Essa biblioteca fornece, de acordo com o nome, a probabilidade de dado nome ser do gênero masculino ou feminino, essa foi a possibilidade encontrada para definição do gênero.

## Problemas

Indivíduos que foram definidos com gênero com probabilidade menor que 80% passaram por uma segunda análise para uma possível modificação. De um total de 3480 nomes analisados, a média da probabilidade de certeza do gênero foi de **98,2338%**, foram retirados dessa contagem os **389** casos de incerteza onde o gênero foi considerado indefinido, ocorrendo também a necessidade de uma análise particular.

Com relação aos casos com porcentagem menor que 80%, foi necessário reanalisar o gênero de **94** alunos. **Pandas** também foi aproveitada para todas as checagens e cálculos feitos.

### 3.3 Geração de infográficos

Após a construção da base de dados, foi escolhida a ferramenta Tableau, para visualização da informação, devido à sua praticidade e opções de inserção de tabelas interativas em páginas web [35].

Em Tableau existem as seguinte formas de organização de elementos gráficos:

1. *Worksheet* (planilha): Uma planilha contém uma única exibição e é onde você cria visualizações de seus dados arrastando e soltando elementos.
2. *Dashboard* (painel): Um painel é uma coleção de várias planilhas e informações de suporte exibidas em um único local para que você possa comparar e monitorar uma variedade de dados simultaneamente.
3. *Stories* (história): Uma história é uma planilha que contém uma sequência de planilhas e painéis que trabalham juntos para transmitir informações.

### Teste das cores

Utilização da ferramenta “Colorblind Web Page Filter ” [38], para visualização do efeito causado pelas cores definidas.

As cores foram definidas dentre as opções determinadas em Tableau para daltonismo, observar-se na Figura 7, seguindo também o debate citado em [31].



Figura 7: Seleção das cores. Adaptado de Tableau

### 3.4 Página Web

Para a produção do ambiente interativo foram utilizadas as linguagens **HTML**, **CSS** e **Javascript**, além do framework **Semantic UI** para acabamento do “Front-end”, que define a aparência da página, pelos elementos de interface reusáveis para páginas em **HTML**.

A página é hospedada em **GitHub Pages**, que fornece todo um ambiente gratuito.

### 3.5 Ambiente Computacional

Toda construção do trabalho foi realizada em um macOS Mojave versão 10.14.1 com processador 2.7 GHz Intel Core i5 com 8 GB 1867 MHz DDR3 de memória. A aplicação Tableau utilizada foi a de versão 20183.18.1018.1932.

## 4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

### 4.1 Ambiente proposto

Seguem os resultados obtidos com a produção do ambiente, página web , interativo, fruto dos dados retirados da plataforma SIGAA. A página contém as abas:

- Página inicial
- Participação
- Desempenho
- Mapa
- Envolvidos

A página pode ser acessada no seguinte link:

<https://7rodrigoaugusto.github.io/generoUFPB/index.html>

A Figura 8 apresenta o *layout* do menu de escolha das abas desenvolvidas. Cada aba da plataforma representa uma categoria de resultados que estão expostos de forma tal qual o usuário possa interagir e refletir sobre a questão de gênero no espaço universitário do centro de informática.

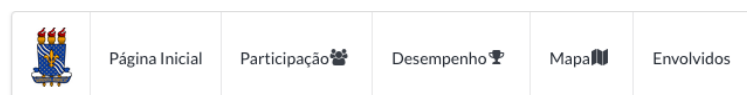


Figura 8: Visualização do menu de escolha das abas

### 4.2 Visualizações Propostas

Para as abas comentadas na seção anterior, serão apresentadas as propostas de visualização para cada.

Com Tableau foram idealizadas 6 *Stories* interativas, para popular as respectivas abas de **participação, desempenho e mapa**. Segue a distribuição dos *Stories*:

- Participação - 1 *Story* e 7 *Dashboards*
- Desempenho - 4 *Stories* e 26 *Dashboards*
- Mapa - 1 *Story* e 3 *Dashboards*

Cada aba é apresentada, a seguir, com uma descrição de seus *Stories* e *Dashboards* contendo a informação de cada visualização presente e de qual filtro interativo é disponibilizado para que o usuário possa aprimorar sua visualização e conseqüente análise:

## 1. Participação

- **Story única**

- **1° Dashboard:** Percentual da composição total dos cursos. Sem filtros
- **2° Dashboard:** Série temporal, visão da quantidade de alunos e alunas que entram . Utilização interativa do filtro curso, para definir qual curso irá aparecer na visualização
- **3° Dashboard:** Série temporal, visão do percentual de alunas que entram. Utilização do filtro curso para definir qual curso irá aparecer na visualização
- **4° Dashboard:** Composição dos cursos. Utilização do filtro período para definir qual período irá aparecer na visualização
- **5° Dashboard:** Situação dos discentes por gênero no Centro de Informática com relação a: ativos, cadastrados, cancelados, concluintes, formandos e trancados. Sem a opção de filtros
- **6° Dashboard:** Situação dos discentes por gênero e curso com relação a: ativos, cadastrados, cancelados, concluintes, formandos e trancados. Filtro curso
- **7° Dashboard:** Percentual, por gênero, de conclusão do curso. Sem a opção de filtros

## 2. Desempenho

- **1° Story:** Disciplinas

- **1° Dashboard:** Série histórica do percentual de participação por disciplina. Utilização interativa dos filtros curso e disciplina, para definir qual disciplina e de qual curso irá aparecer na visualização
- **2° Dashboard:** Série histórica da média das notas considerando as situações aprovado e reprovado. Utilização interativa dos filtros curso e disciplina para definir qual disciplina e de qual curso irá aparecer na visualização
- **3° Dashboard:** Percentual das situações aprovado, reprovado, reprovado por falta e trancado por gênero. Além da série temporal com o percentual de aprovações. Utilização interativa dos filtros curso, disciplina e gênero.

- 4° **Dashboard**: Percentual de participação e percentuais de aprovação, dispensa, reprovação por falta, reprovação, trancamento e cancelamento. Utilização interativa dos filtros curso, período e disciplina para definir qual disciplina, de qual período e de qual curso irá aparecer na visualização
- 2° **Story**: Fluxograma da Computação com a apresentação de 26 *Dashboards* mapa de calor, representando um período de cada curso, com o percentual de aprovação em cada uma das disciplinas. São apresentados aqui são apresentados o:
  - Fluxograma Engenharia de Computação
  - Fluxograma Ciência da Computação
  - Fluxograma Matemática Computacional

### 3. Mapa

- **Story única**
  - 1° **Dashboard**: Apresentação do local de nascimento dos discentes do Centro. Utilização do filtro curso
  - 2° **Dashboard**: Mapa com a relação de gênero quanto ao local de origem dos discentes
  - 3° **Dashboard**: Rank de participação por estado. Utilização do filtro curso
  - 4° **Dashboard**: Comparação interativa do mapa da computação quanto à chegada dos alunos antes e depois da implementação do Exame Nacional do Ensino Médio (ENEM)

Em cada aba existe uma explicação para o que é proposto, quais perguntas a regem e para cada um dos recursos interativos uma análise parcial dos resultados com o intuito de demonstrar o potencial da plataforma.

### 4.3 Pagina inicial

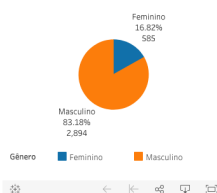
A Figura 9 apresenta o *layout* da página inicial que contém informação sobre:

- O problema de gênero
- Objetivo da plataforma
- Dados
- Visualização da composição do centro de informática

### Gênero

Problemas relacionados a gênero afetam instituições que são base para a sociedade, como universidades e empresas, pensando nisso, e tendo em vista a importância da universidade como modelo social, é mostrado aqui um trabalho que apresenta visualizações gráficas sobre as discrepâncias de gênero dentro de uma dessas instituições, a Universidade Federal da Paraíba (UFPB), mais especificamente nos cursos do centro de informática (CI), Engenharia de Computação, Matemática Computacional e Ciência da Computação.

### Participação por Gênero nos Cursos do Centro de Informática



### Objetivo

Ferramenta cujo o propósito é apresentar uma visão temporal quanto à participação, desistência e desempenho dos gêneros no centro de informática (CI). Assim, será possível, para interessados no tema, a análise do perfil observável dos alunos quanto ao gênero.

### Dados

Os dados foram coletados da plataforma SIGAA UFPB

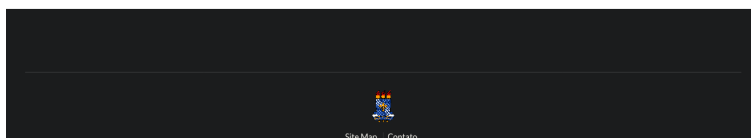


Figura 9: Página web com os resultados da pesquisa.

## 4.4 Participação

A aba participação, que contém 1 *Story* e 7 *Dashboards*, é representada nas Figuras [10](#), [11](#), [12](#), [13](#), [14](#), [15](#) e [16](#). O intuito da aba é representar o perfil da composição dos discentes do Centro de Informática.

### Perguntas que regem a aba

- Qual gênero foi mais representado ao longo dos anos?
- Qual a situação (ativo, cadastrado, cancelado, concluído, formando e trancado) mais comum de cada gênero?
- Qual gênero termina o curso?

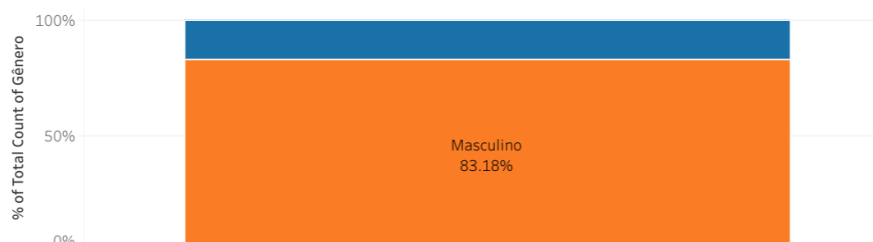
### Análise parcial

A composição do Centro de Informática da *Dashboard* representada na Figura [10](#) mostra que ao longo de todos os seus períodos, por gênero, a participação feminina é de 16,82%, esse percentual é semelhante quando observamos cada um dos cursos de forma independente, o percentual também se assemelha a padrões encontrados em outras Universidades, como por exemplo, Unicamp, UFRJ e UFMG onde o percentual fica em torno de 10% a 12% [40](#). O percentual específico do curso de Ciência da Computação, que é o mais antigo da instituição, também segue o padrão citado em [8](#), que observou que desde meados da década de 90 houve uma queda da participação feminina de 30% para 5 ou 10% nos últimos anos.

## Participação



### Composição Centro de Informática



### Composição por Curso

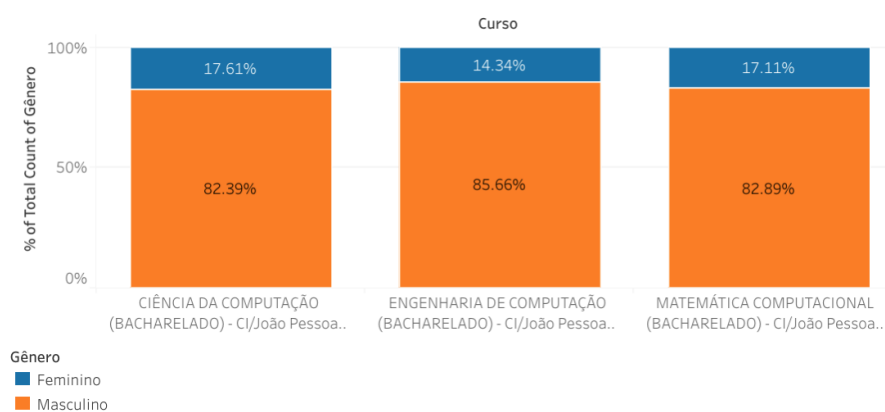


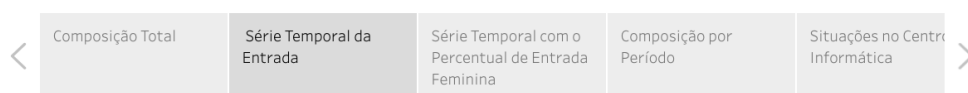
Figura 10: Visualização da *Dashboard* com o percentual de gênero da composição total dos cursos. Na figura, % do total de Contagem de Gênero para cada Curso. A cor mostra detalhes sobre Gênero. As marcas são rotuladas por % do total de Contagem de Gênero. Fonte: Autoria própria

A *Dashboard* da Figura 11 oferece uma visualização complementar à apresentada na 10 por possibilitar a observação dos seguintes pontos:

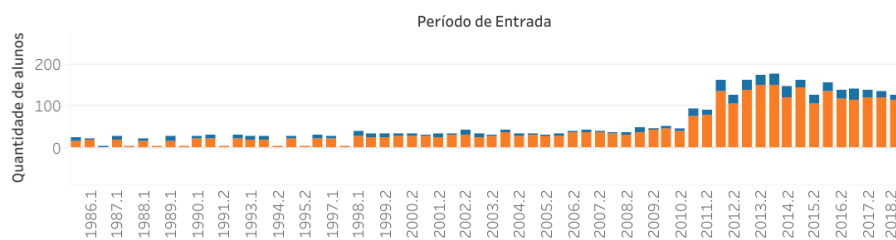
- Quantos discentes entram por período no Centro, e qual seu gênero?
- Quantos discentes entram por período em um dos cursos do Centro, e qual seu gênero?

A observação confirma a baixa entrada feminina no Centro e revela que a maior inserção ocorreu no período de 2014.1 com a entrada de 27 mulheres nos cursos de computação, esse valor correspondeu a 18,49% dos novos discentes. No exemplo, também é possível observar que, com a escolha do curso Engenharia de Computação, 2014.2 foi o período que mais se recebeu mulheres, com um total de 12.

## Participação



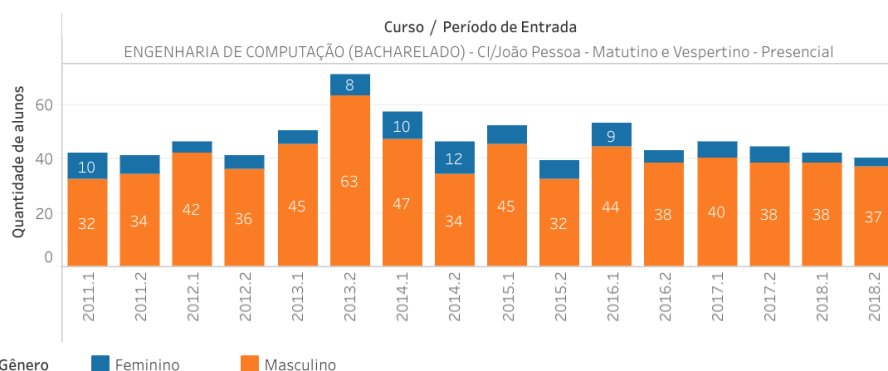
### Série Temporal da Entrada no Centro de Informática



Curso

ENGENHARIA DE COMPUTAÇÃO (BACHARELADO) - CI/João Pessoa - Matutino e Vespertino - Presencial

### Série Temporal da Entrada por Curso



Gênero ■ Feminino ■ Masculino

+ a b l e a u

← → ↶ 🔍 📄

Figura 11: Visualização da *Dashboard* que contém a série temporal da quantidade de alunos e alunas que entram. Utilização do filtro curso. Fonte: Autoria própria

Assim como a *Dashboard* apresentada na Figura 11 a *Dashboard* representada na Figura 12 oferece informação de entrada por período, porém, aqui temos o percentual de participação feminina. O período com a maior percentual, por acaso, é o mesmo apresentado na *Dashboard* anterior, 2014.2 com um percentual de 26,09%.

## Participação

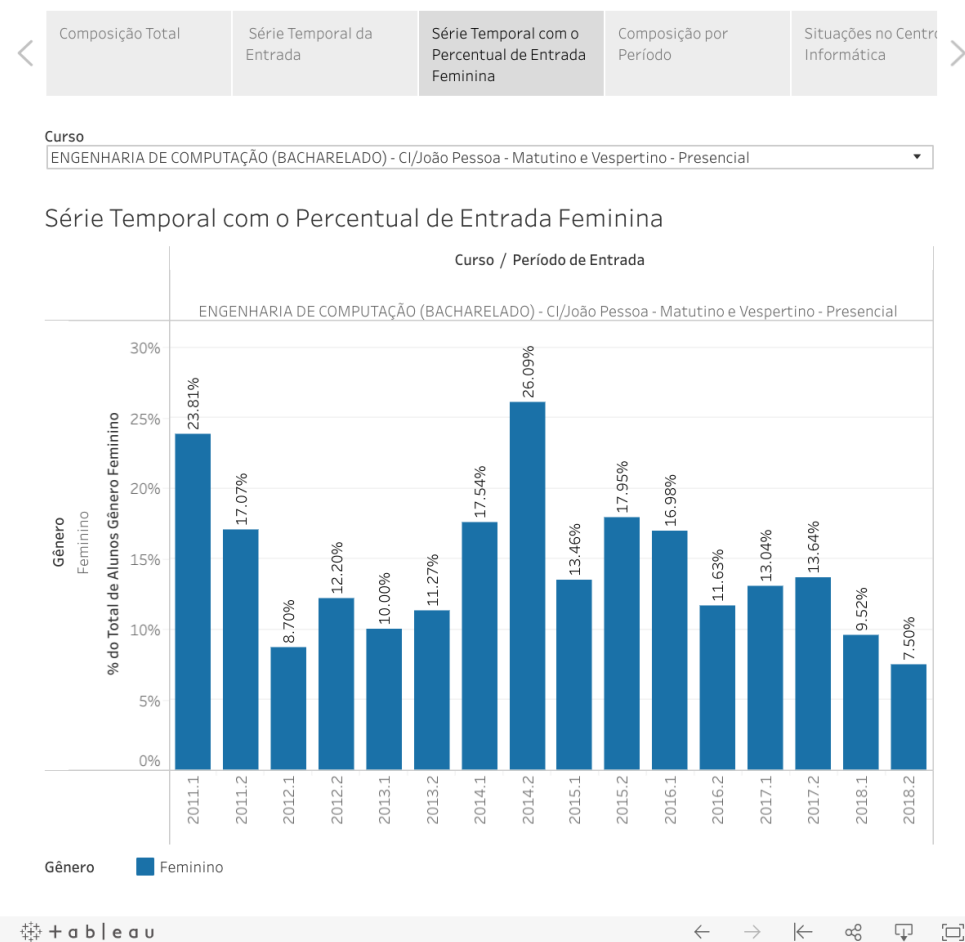


Figura 12: Visualização da *Dashboard* com a série temporal do percentual de alunas que entram. Utilização do filtro curso. Fonte: Autoria própria

Já com a *Dashboard* da Figura 13, aproveitando-se do filtro interativo de período, é possível ter uma visão ampliada dos dados demonstrados nas 3 *Dashboards* anteriores. No exemplo, é possível observar que o curso com percentual de entrada mais equilibrado é o de Matemática Computacional com 28,57%.

## Participação



### Participação por período

Período  ◀ ▶

### Percentual de Entrada por Período

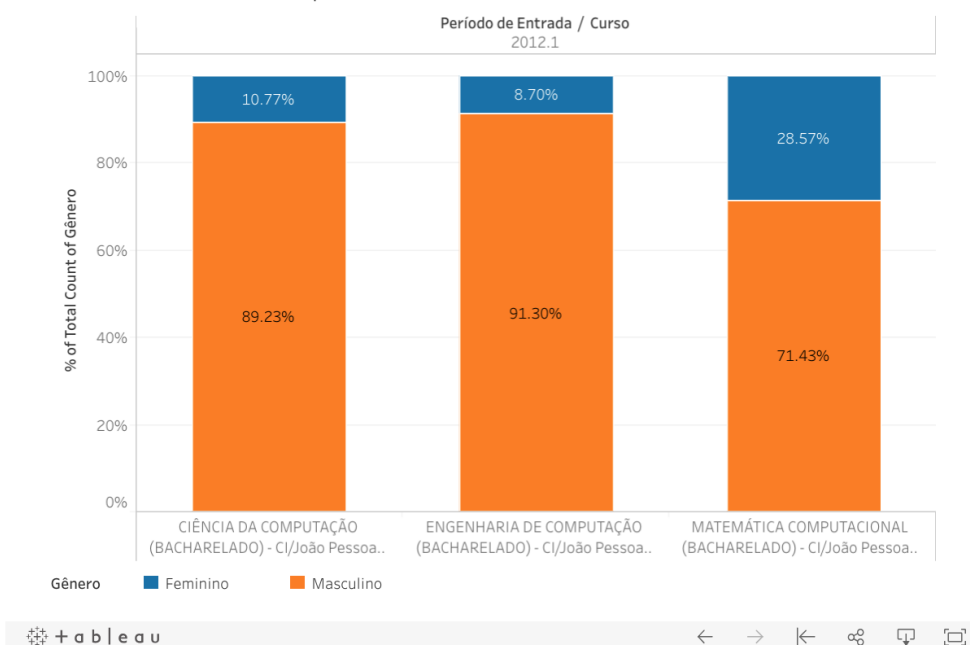


Figura 13: Visualização da *Dashboard* com a composição dos cursos. Utilização do filtro período. Fonte: Autoria própria

A *Dashboard* representada pela Figura 14 apresenta o percentual de situações (ativo, cadastrado, cancelado, concluído, formando e trancado) por gênero. Esse dado demonstra que o percentual de mulheres que cancelaram cursos, do Centro de Informática, é menor que o dos homens e o percentual que concluí é maior. Temos que das mulheres que se propõe a cursar computação 26,67% terminam o curso, esse valor é superior aos 18,69% de homens.

## Participação

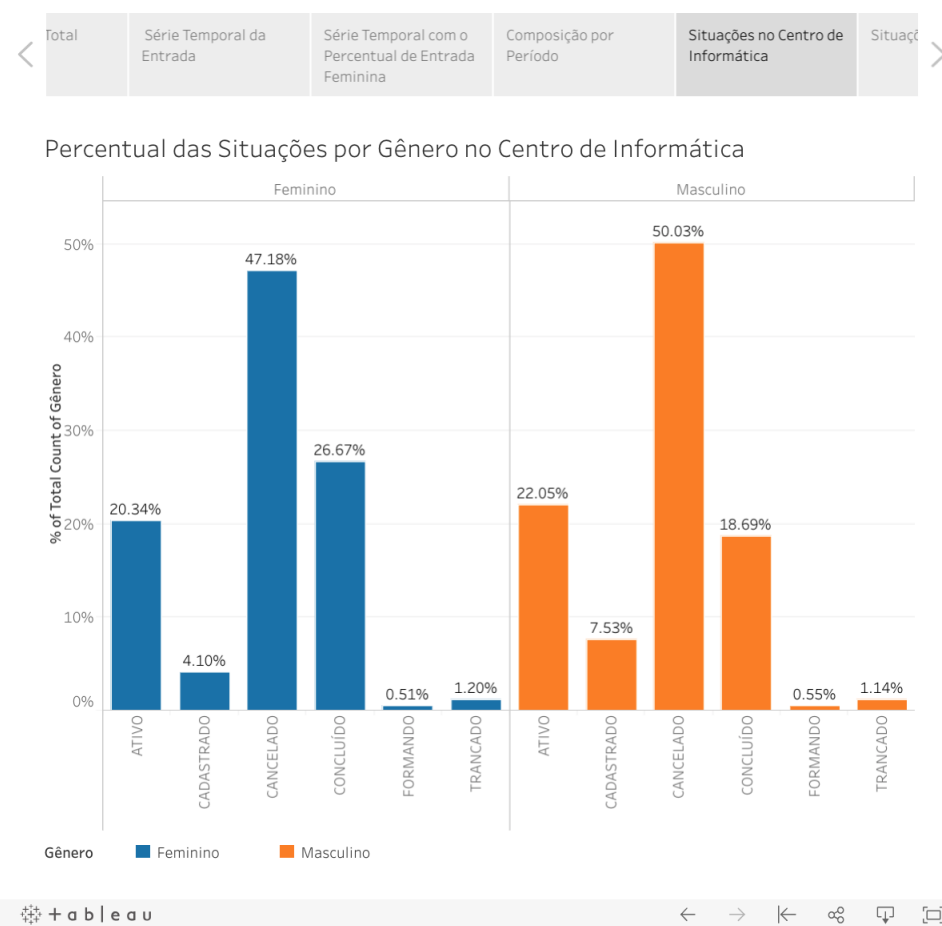


Figura 14: Visualização da *Dashboard*. Fonte: Autoria própria

A *Dashboard* da Figura 15 apresenta o percentual das mesmas situações da Figura anterior por curso e gênero. Como exemplo, foi escolhido o curso de Engenharia de Computação e ele demonstra análise similar à apresentada para o Centro de Informática. As mulheres concluem mais o curso com um percentual de 5,56% contra 2,48% dos homens.

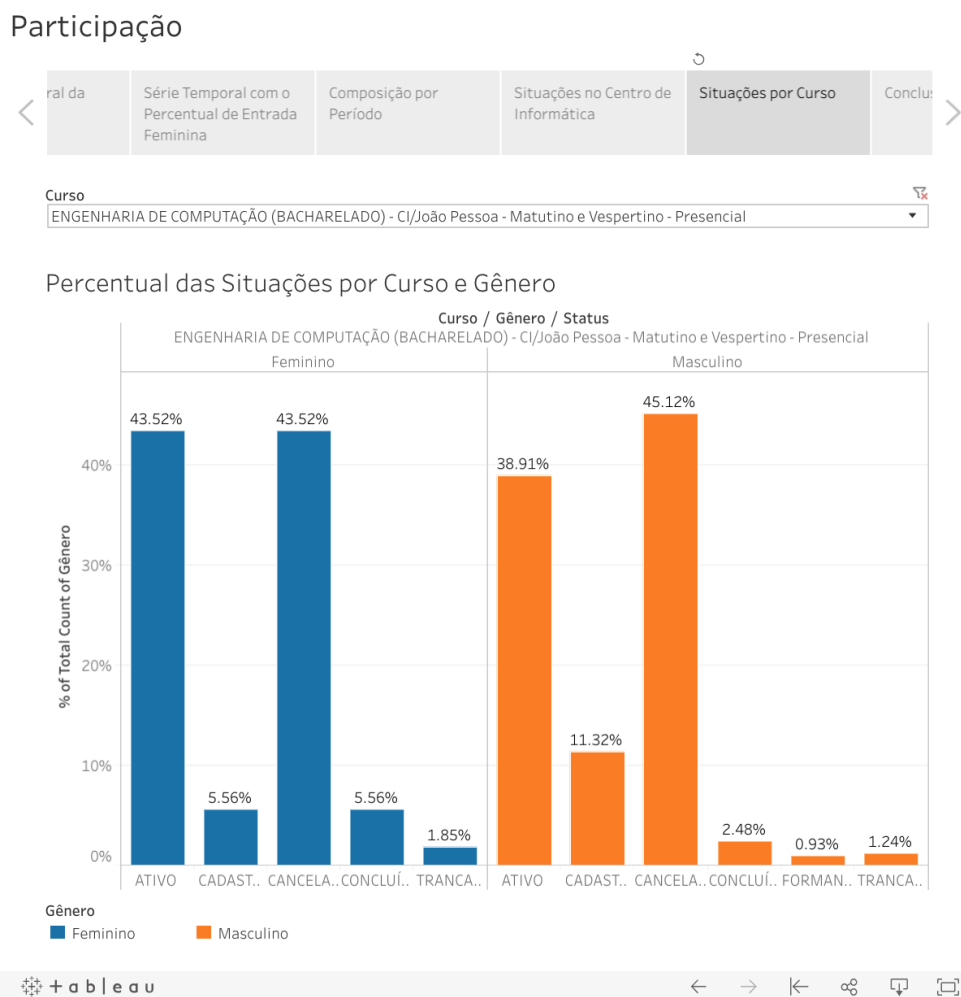


Figura 15: Visualização da *Dashboard*. Fonte: Autoria própria

Na *Dashboard* apresentada na Figura 16 é apresentado o percentual de concluintes por gênero. Percentualmente a quantidade de mulheres que concluem é menos discrepante que o percentual de participação.

### Participação

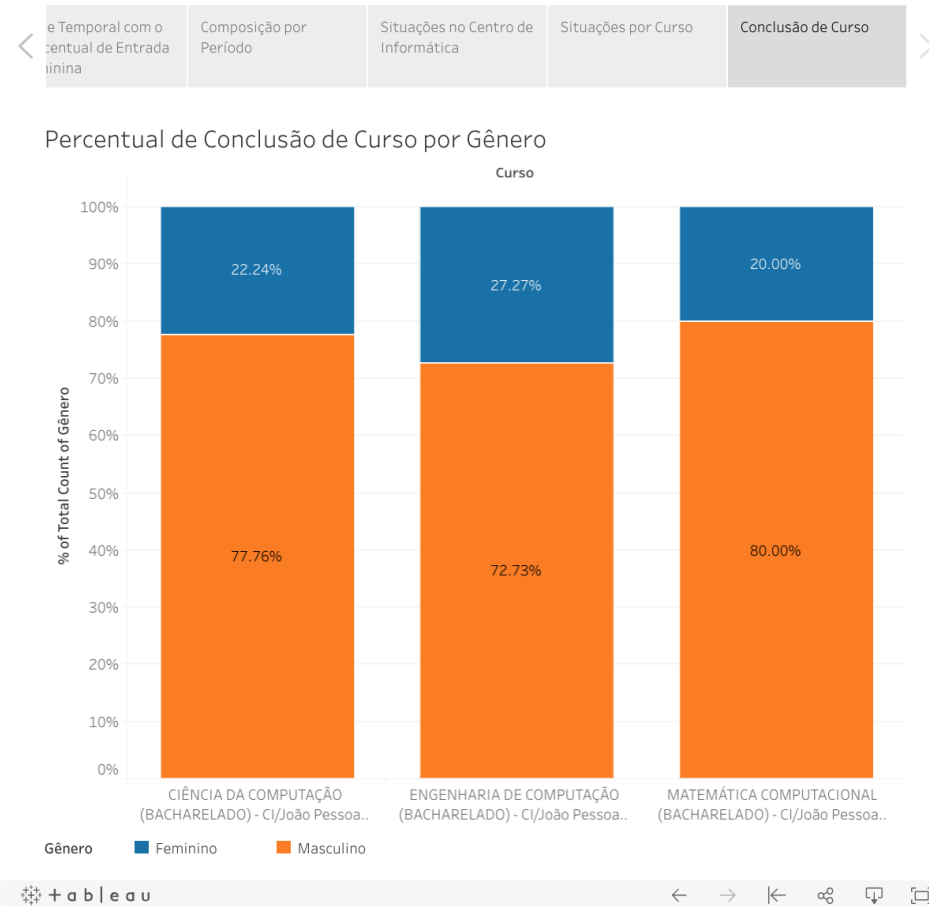


Figura 16: Visualização da *Dashboard* com o percentual, por gênero, de conclusão do curso. Fonte: Autoria própria

## 4.5 Desempenho

A aba contém 4 *Stories* e 30 *Dashboards* e contém dois recursos interativos **Disciplinas** e **Fluxograma da Computação**. O perfil dos discentes por gênero também deve ser analisado pelo comportamento e participação nas disciplinas do Centro.

### Perguntas que regem a aba

- Qual a participação por gênero nas disciplina?
- Qual o desempenho por gênero?
- Quais disciplinas são problemáticas?

### Análise parcial

A *Dashboard* da Figura [17](#) oferece ao usuário três visualizações quanto ao curso e disciplina escolhidos:

- Percentual de participação por gênero
- Série temporal com a quantidade de discentes por gênero
- Série temporal com o percentual de participação por gênero

No exemplo é possível avaliar a baixa participação feminina na disciplina “Cálculo Diferencial e Integral I”, consequência da também baixa participação no curso “Engenharia de Computação”, tendo em vista que é disciplina obrigatória do primeiro período do curso.

## Disciplinas

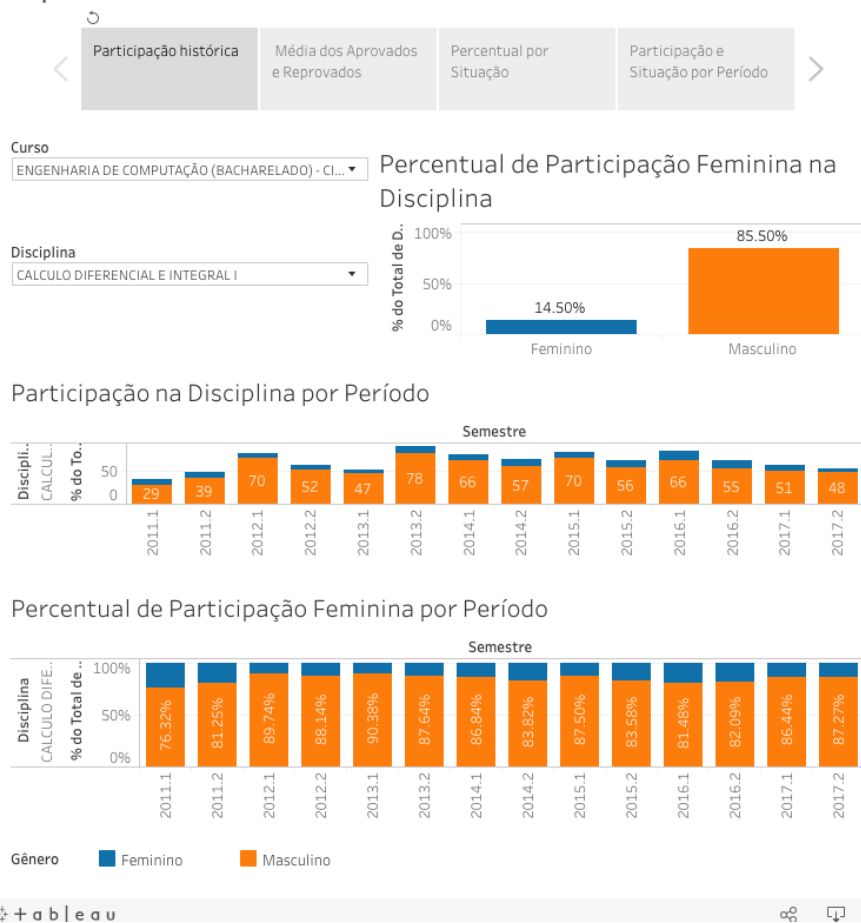


Figura 17: Visualização da *Dashboard* com a série temporal do percentual de participação por disciplina. Utilização interativa dos filtros curso e disciplina.  
 Fonte: Autoria própria

Segue a análise com a utilização da *Dashboard* presente na Figura 18, que apresenta a média dos alunos e alunas, filtrada por disciplina e curso. É possível observar, com a utilização das mesmas opções de filtros, que a média quanto às situações aprovado e reprovado, que representam os alunos que foram até o fim da disciplina, são semelhantes.

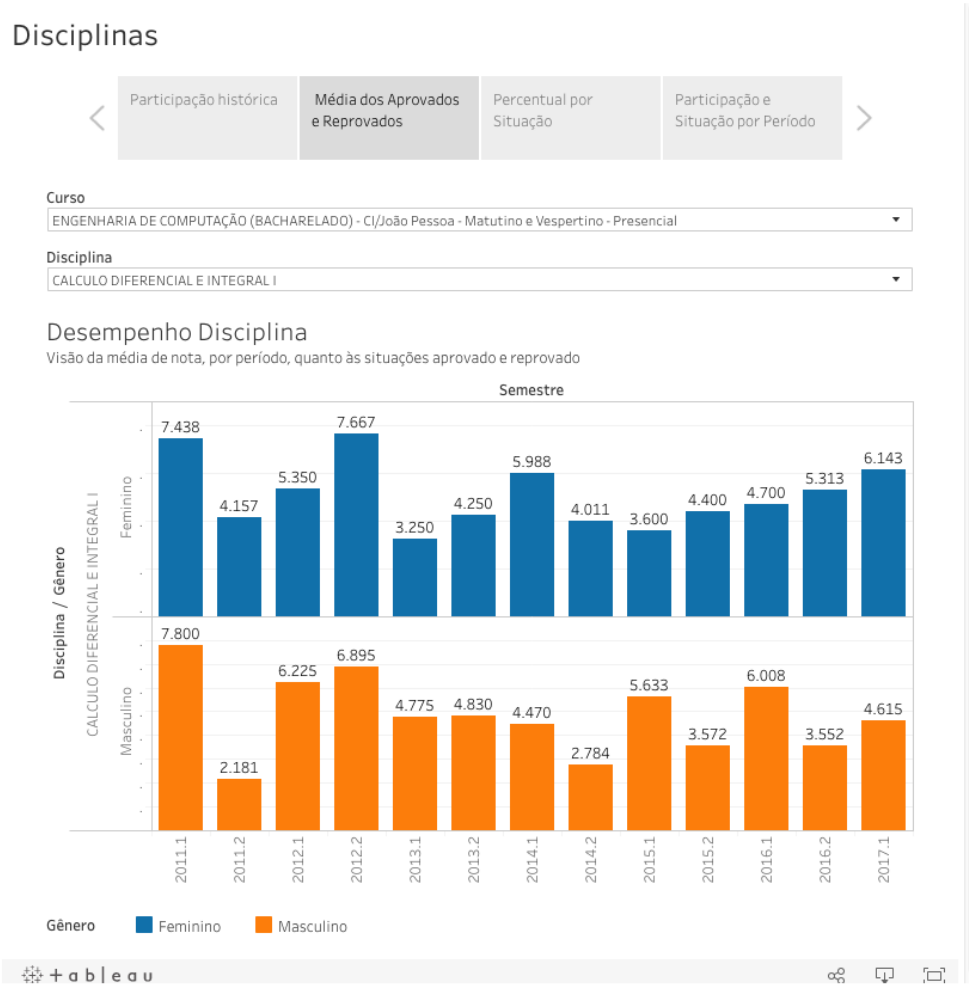


Figura 18: Visualização da *Dashboard* com a série temporal da média das notas considerando as situações aprovado e reprovado. Utilização interativa dos filtros curso e disciplina

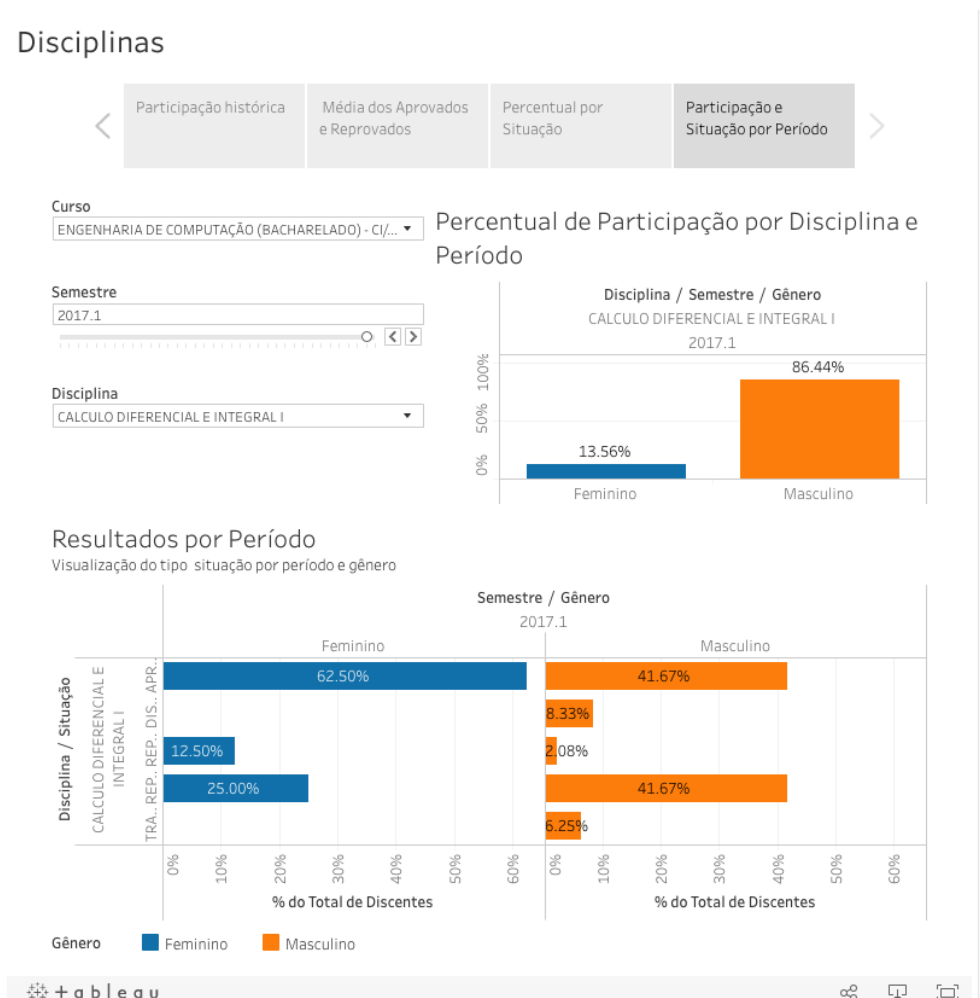
A média por gênero talvez não seja o melhor parâmetro de comparação, por isso contamos com a *Dashboard* da Figura 19, que por meio dos filtros de curso e disciplina garante a observação do percentual das situações aprovado, dispensado, reprovado por falta, reprovado e trancado, além de uma visão da série temporal, período a período, do mesmo percentual de aprovação, por gênero escolhido. No exemplo é perceptível a maior aprovação da disciplina “Cálculo Diferencial e Integral I” por mulheres. Essa *Dashboard* apresenta informação similar à contida nos fluxogramas apresentados nesta mesma aba.

## Disciplinas



**Figura 19: Visualização da *Dashboard* com a série temporal do percentual de aprovação considerando as situações aprovado, reprovado, reprovado por falta e trancado. Utilização interativa dos filtros curso, disciplina e gênero. Fonte: Autoria própria**

Por fim, a *Dashboard* da Figura 20 apresenta o percentual de participação nas disciplinas e o percentual da situação obtida (aprovado, reprovado, reprovado por falta, dispensado e trancado), diferente da primeira e terceira *Dashboard*, aqui é apresentada uma informação filtrada quanto ao período escolhido. O exemplo mostra a baixa participação das alunas de apenas 13,56% na disciplina “Cálculo Diferencial e Integral I” e um desempenho superior com as mulheres tendo um percentual de aprovação de 62,50% no período escolhido.



**Figura 20: Visualização da *Dashboard* com o percentual de participação e percentuais de aprovação, dispensa, reprovação por falta, reprovação e trancamento. Utilização interativa dos filtros curso, período e disciplina. Fonte: Autoria própria**

Nos exemplos apresentados, utilizando-se da disciplina de “Cálculo Diferencial e Integral I”, as mulheres possuem um percentual de aprovação maior que os homens. Tudo isso demonstra que elas ainda conseguem ter um rendimento superior mesmo em um ambiente com baixa representatividade. A questão das disciplinas problemáticas é abordada com clareza no recurso interativo apresentado a seguir.

## Fluxograma da Computação

O recurso interativo fluxograma, representado nas Figuras [22](#), [23](#) e [24](#), advém da ideia apresentada na *Dashboard Percentual por Situação*, onde é possível observar o percentual de aprovação, reprovação, reprovação por falta, trancamento e dispensa por disciplina. O recurso é um refinamento da ideia e utiliza-se do fluxograma dos cursos para representação das disciplinas importantes, de forma tal qual seja possível analisar o desempenho feminino e masculino por disciplina ao longo do curso e consequente entendimento das matérias problemáticas dos cursos.

As cores utilizadas para gênero seguem o mesmo padrão das outras visualizações e é representada na Figura [21](#).



**Figura 21: Legenda das cores para o Fluxograma da Computação. Fonte: Autoria própria**

### Análise parcial

A observação da Figura [22](#) demonstra o percentual de aprovação superior feminino nas disciplinas do primeiro período do curso de Engenharia de Computação. Merece destaque a disciplina de “Cálculo Diferencial e Integral I”, que anteriormente foi utilizada como exemplo, com percentual de aprovação de 45,24% de aprovação das mulheres que se propõem a cursar contra 36,11% de aprovação dos homens.

## Fluxograma Engenharia de Computação

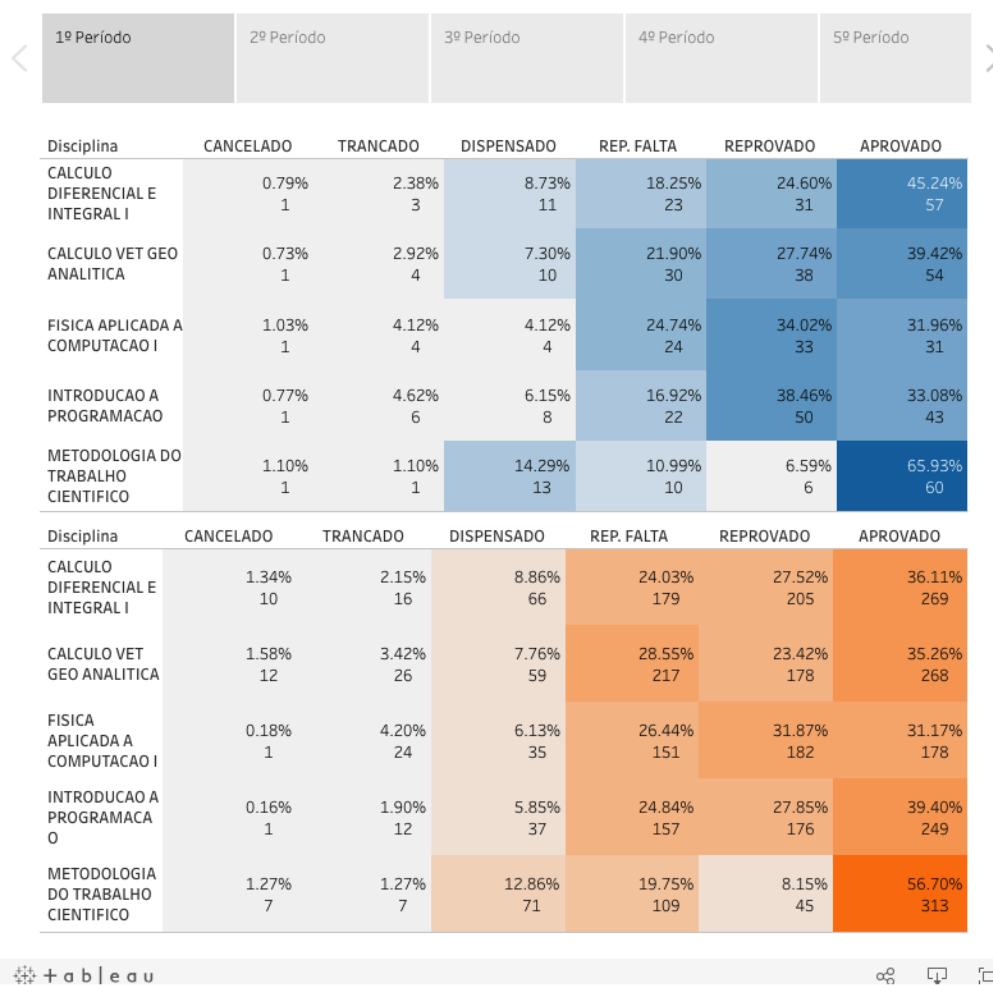


Figura 22: Visualização do recurso interativo fluxograma para o curso Engenharia de Computação. Fonte: Autoria própria

A análise da Figura 23 também apresenta um percentual de aprovação superior feminino em todas as disciplinas do primeiro período do curso de Ciência da Computação com exceção de “Introdução ao Computador”.

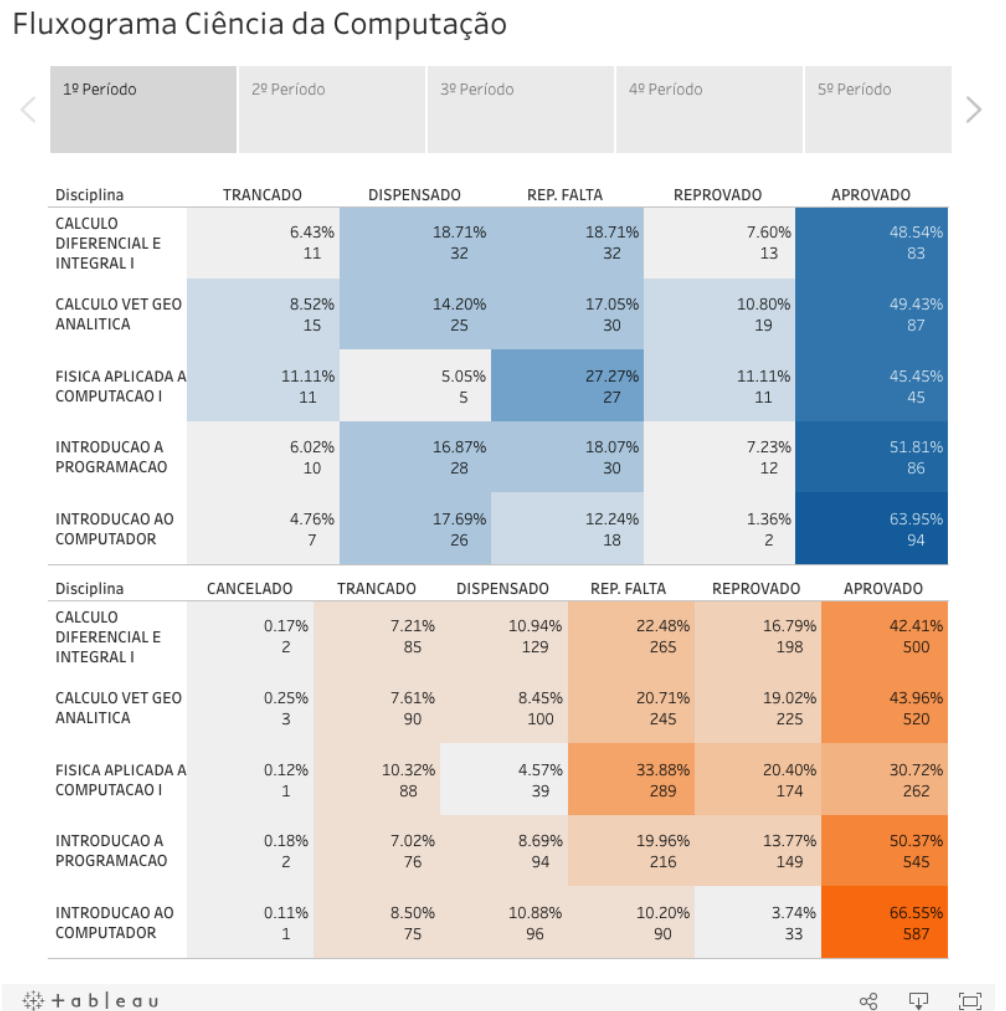


Figura 23: Visualização do recurso interativo fluxograma para o curso Ciência da Computação. Fonte: Autoria própria

A visualização da Figura 24 apresenta, assim como a primeira do curso de Engenharia de Computação, um percentual de aprovação superior feminino em todas as disciplinas do primeiro período do curso de Matemática Computacional.

### Fluxograma Matemática Computacional

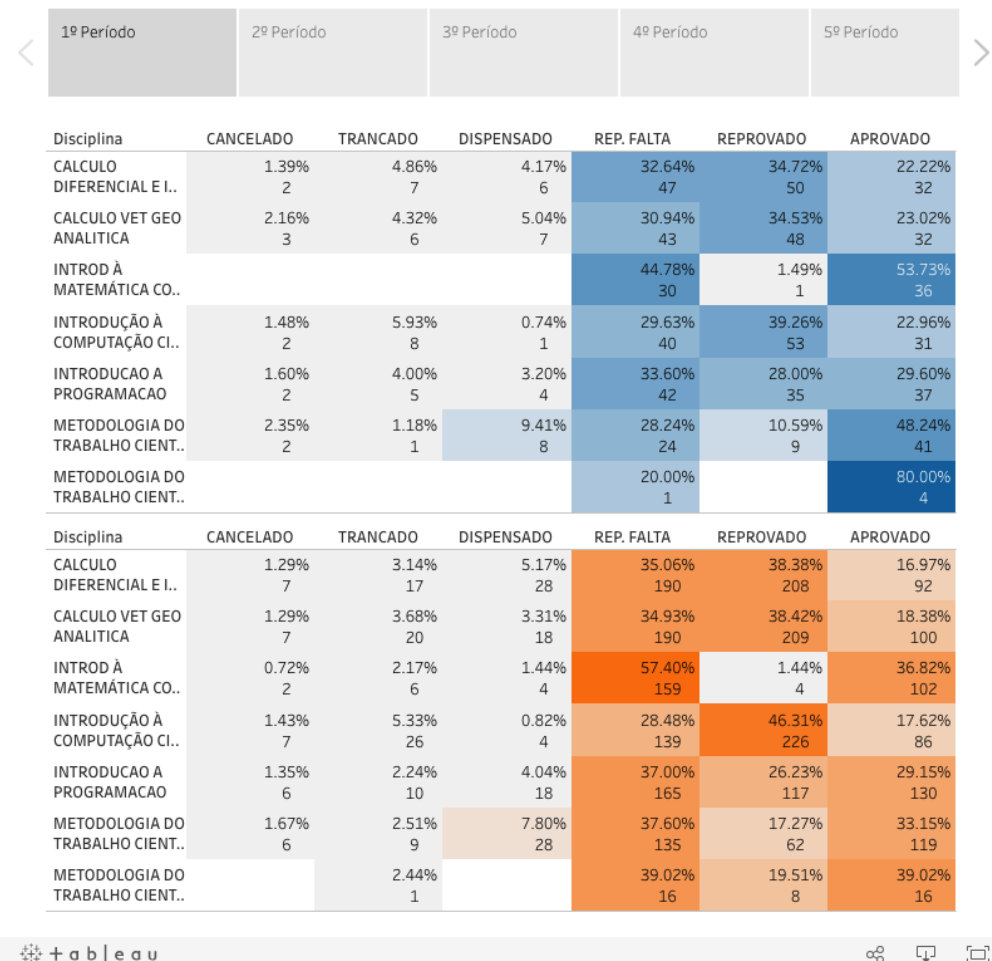


Figura 24: Visualização do recurso interativo fluxograma para o curso de Matemática Computacional. Fonte: Autoria própria

## 4.6 Mapa

A aba contém 1 *Story* com 4 *Dashboards* apresentando informação sobre o local de origem dos alunos do CI, as planilhas podem ser observadas nas Figuras [25](#), [26](#), [27](#) e [28](#). A *Dashboard* do mapa serve para que interessados possam compreender a relação da origem dos discentes e o gênero.

### Perguntas que regem a aba

- Qual o local de origem dos discentes do Centro?
- Qual a relação do gênero com o local de origem?
- Qual a relação do ENEM com o local de origem?

### Análise parcial

A *Dashboard* inicial, representada na Figura [25](#), demonstra a participação majoritária de discentes do próprio estado participando dos cursos do Centro.

## Mapa da Computação



Figura 25: Visualização da *Dashboard* com o local de origem dos discentes.  
Fonte: Autoria própria

A *Dashboard* representada na Figura 26 demonstra o problema de gênero quanto à participação feminina de discentes de outros estados, existe um desequilíbrio representativo, fica o questionamento quanto à independência das mulheres em buscar cursos fora de seu estado de origem.

### Mapa da Computação

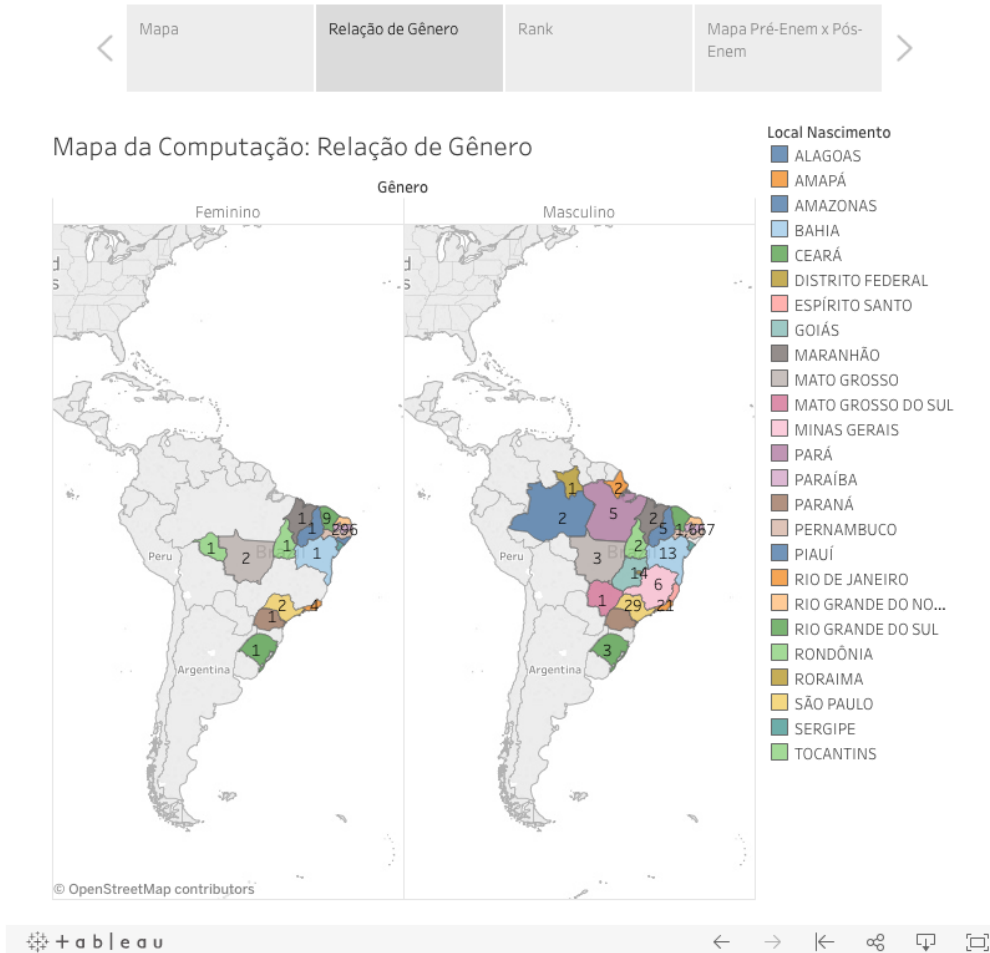


Figura 26: Visualização da *Dashboard* com a comparação por gênero quanto à origem dos discentes. Fonte: Autoria própria

A *Dashboard* representada na Figura 27, assim como o mapa inicial, demonstra que a maior parte dos discentes que participam do curso são do próprio estado.

### Mapa da Computação

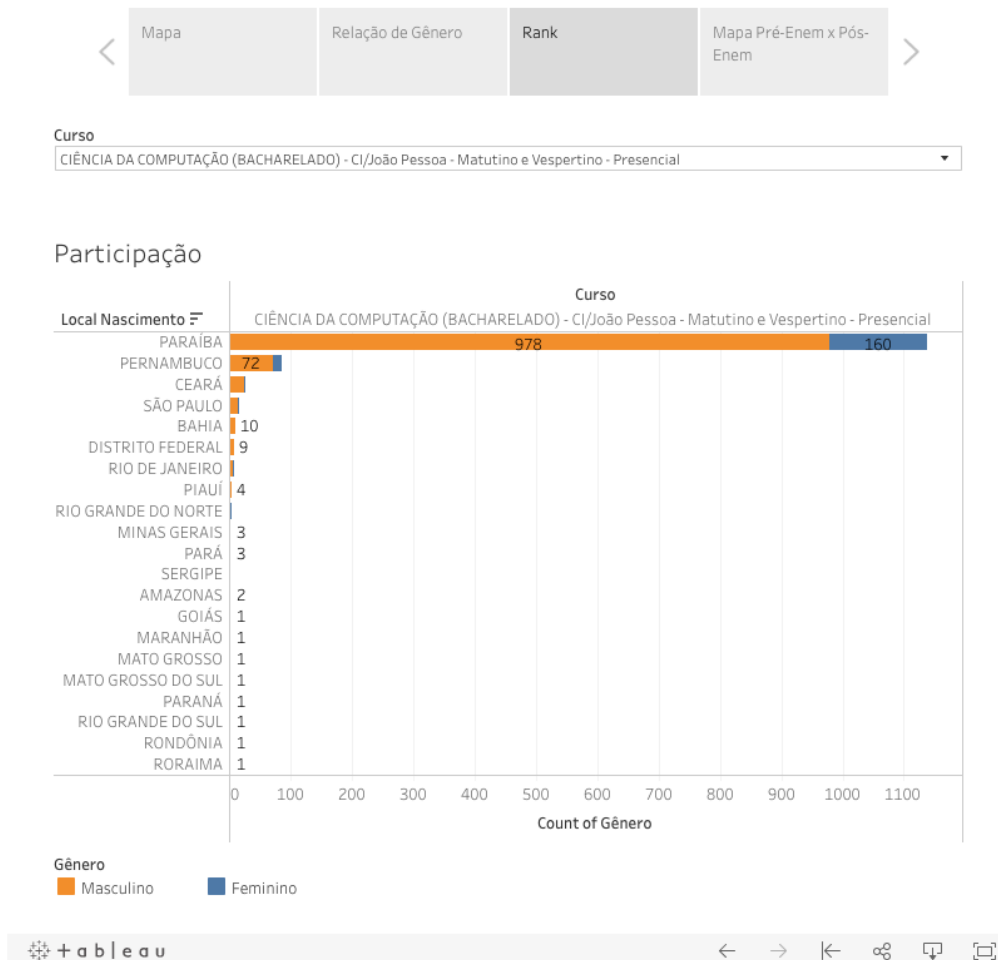


Figura 27: Visualização da *Dashboard* com o rank referente ao local de origem dos discentes. Fonte: Autoria própria

Já a *Dashboard* representada na Figura 28 demonstra o impacto que o ENEM causa quanto à participação de alunos de outros estados no CI. Ocorre um claro incremento nos números advindos de outros estados e na quantidade de estados representados.

### Mapa da Computação

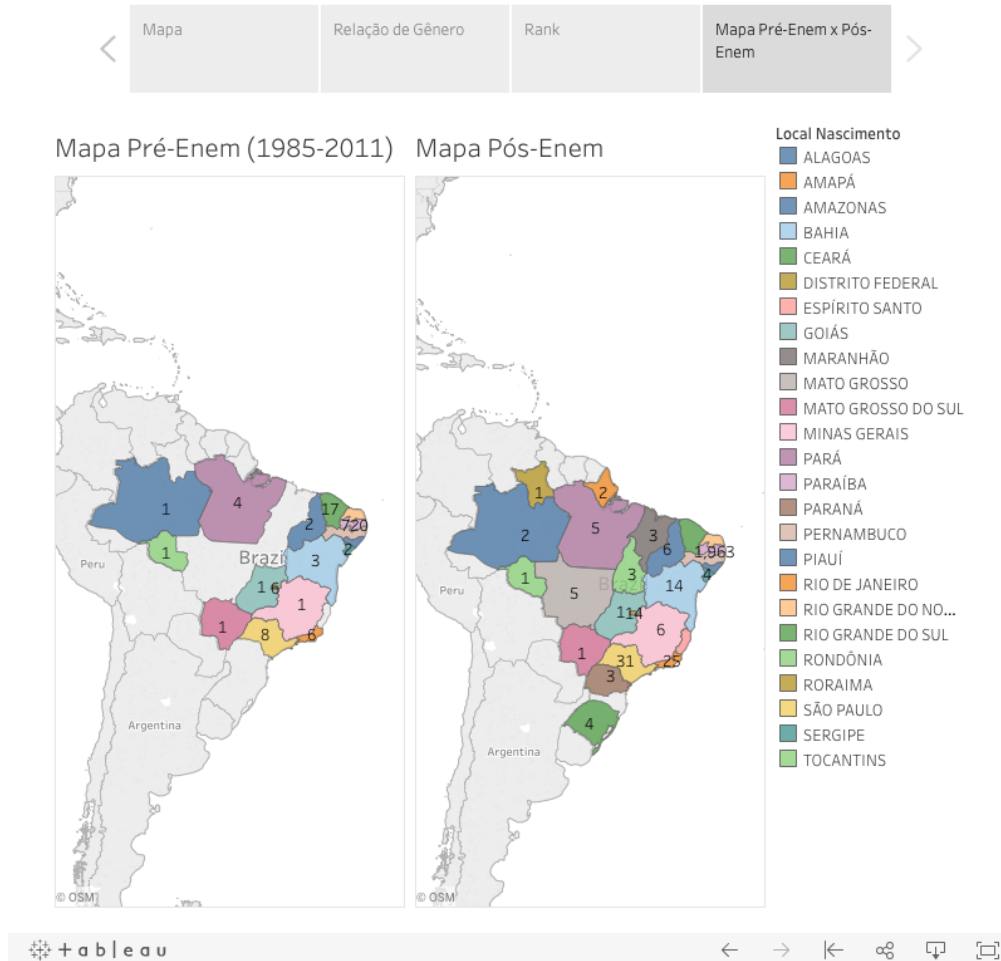


Figura 28: Visualização da *Dashboard* com informação referente ao local de origem dos discentes do Centro de Informática antes e depois do ENEM. Fonte: Autoria própria

## 5 CONCLUSÕES E TRABALHOS FUTUROS

A discrepância de gênero é real dentro do nosso próprio campus, demonstrando assim o quão importante é a temática para a criação de projetos que incentivem a participação feminina em cursos tradicionalmente masculinos, tendo em vista a cobrança cada vez maior por novos especialistas em áreas ligadas à tecnologia.

A universidade como espaço formador tem a obrigação de equilibrar a participação de gênero por área, já pensando no novo mercado de trabalho que aflora. A ferramenta de visualização funciona como uma importante metodologia de exposição destes dados, e mesmo sem capacidade de uma análise profunda de tudo que foi criado, por não ser o objetivo fim deste projeto, já é possível entender, com este método facilitador, discrepâncias estruturais com a simples observação dos elementos visuais propostos.

Com uma análise detalhada, por exemplo, da seção extra de fluxogramas, seria possível derrubar o mito da competitividade dentro das exatas, pois as mulheres conseguem ter uma aprovação tão boa ou melhor que a dos homens na maioria das disciplinas, similar com resultados de desempenho, conforme a literatura.

A maior dificuldade do trabalho residiu na aquisição de uma base de dados completa, do nosso sistema SIGAA, desta forma o trabalho consistiu na criação exclusiva de visualizações para os cursos do Centro de Informática, com os dados provenientes do sistema de coordenações SIGAA.

Por fim, mesmo com as limitações, os resultados obtidos apresentam dados potencialmente importantes e este trabalho espera ser ampliado com a criação de um ambiente de visualização que abranja todos os cursos da UFPB, com uma análise e estudo particular para cada área do conhecimento, validação e levantamento de novos requisitos, para assim, dar mais transparência aos dados universitários, seguindo padrões internacionais e abrindo espaço para debates sobre o futuro da participação feminina.

## REFERÊNCIAS

- [1] DELORS, Jacques. **Educação: um tesouro a descobrir; Relatório para a UNESCO da Comissão Internacional sobre educação para o século XXI.** São Paulo: UNESCO, Ed. Cortez, 1999.
- [2] SCOTT, J. W. **Gender: A Useful Category of Historical Analysis.** The American Historical Review, vol. 91, nº 5. (Dec., 1986), pp. 1053-1075. Disponível em: <<http://www.jstor.org/stable/1864376>>. Acesso em: 4 de Maio de 2018.
- [3] UN Global Pulse. **Big Data for Development: Challenges and Opportunities.** ONU, Maio 2012. Disponível em: <<https://www.unglobalpulse.org/projects/BigDataforDevelopment>>. Acesso em 10 de Maio de 2018.
- [4] Onella, Jukka- Pekka. **Social Networks and Collective Human Behavior.** UN Global Pulse. 10 Nov. 2011. <<http://www.unglobalpulse.org/node/14539>>. Acesso em: 15 de Julho de 2018.
- [5] SANTO, Patrícia **Os estudos de gênero na Ciência da Informação.** Em Questão, v. 14, n. 2 (2008). Disponível em: <<http://seer.ufrgs.br/index.php/EmQuestao/article/view/6389>>. Acesso em: 15 de Julho de 2018.
- [6] Assessoria de Imprensa do Inep. **Mulheres estão em maior número na educação superior.** INEP, 2009. Disponível em: <<https://goo.gl/apkwqb>>. Acesso em: 4 de Maio de 2018.
- [7] INEP **Resumo Técnico Censo da Educação Superior 2009.** INEP, 2018. Disponível em: <<https://goo.gl/axrjSQ>>. Acesso em: 8 de Maio de 2018.
- [8] LIMA, Michelle Pinto. **As mulheres na Ciência da Computação.** Rev. Estud. Fem., Florianópolis , v. 21, n. 3, p. 793-816, Dec. 2013 . Disponível em: <<https://periodicos.ufsc.br/index.php/ref/article/view/S0104-026X2013000300003>>. Acesso em 1 de Junho de 2018. <http://dx.doi.org/10.1590/S0104-026X2013000300003>.
- [9] BARRETO, Andreia. **A mulher no ensino superior: distribuição e representatividade.** Cadernos do GEA. FLACSO/LPP/UERJ 2016. Disponível em: <[http://flacso.org.br/files/2016/04/caderno\\_gea\\_n6\\_digitalfinal.pdf](http://flacso.org.br/files/2016/04/caderno_gea_n6_digitalfinal.pdf)> Acesso em: 10 de Maio de 2018.

- [10] Oliveira dos Santos, Wilk. **Mulheres na Computação: Uma Análise da Participação Feminina nos Cursos de Licenciatura em Computação**. 2017. 814. 10.5753/cbie.wcbie.2017.814. Disponível em: Acesso em: 10 de Maio de 2018.
- [11] The World Bank. **Data Bank. Population, female (% of total)**. Disponível em: <<https://goo.gl/XGqgHC>>. Acesso em: 10 de Maio de 2018.
- [12] ILO. **International Labour Organization (ILO). WORLD EMPLOYMENT SOCIAL OUTLOOK TRENDS FOR WOMEN 2018 Global snapshot**. Disponível em: <[https://www.ilo.org/wcmsp5/groups/public/—dgreports/—dcomm/—publ/documents/publication/wcms\]\\_619577.pdf](https://www.ilo.org/wcmsp5/groups/public/—dgreports/—dcomm/—publ/documents/publication/wcms]_619577.pdf)>. Acesso em: 10 de Maio de 2018.
- [13] IBGE. **Estatísticas de Gênero - Indicadores sociais das mulheres no Brasil**. Disponível em: <<https://www.ibge.gov.br/estatisticas-novoportal/multidominio/genero/20163-estatisticas-de-genero-indicadores-sociais-das-mulheres-no-brasil.html?=&t=o-que-e>>. Acesso em: 10 de Maio de 2018.
- [14] IBGE Estatísticas Sociais. **Mulher estuda mais, trabalha mais e ganha menos do que o homem**. Disponível em: <<https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/20234-mulher-estuda-mais-trabalha-mais-e-ganha-menos-do-que-o-homem.html>>. Acesso em: 10 de Maio de 2018.
- [15] UNESCO. **Women still a minority in engineering and computer science - Mulheres continuam uma minória em engenharia e ciência da computação**. Disponível em: <[http://www.unesco.org/new/en/unesco/themes/gender-equality/resources/single-view-gender/news/women\\_still\\_a\\_minority\\_in\\_engineering\\_and\\_computer\\_science/](http://www.unesco.org/new/en/unesco/themes/gender-equality/resources/single-view-gender/news/women_still_a_minority_in_engineering_and_computer_science/)>. Acesso em: 10 de Maio de 2018.
- [16] SANTOS, Carolina. **Por que as mulheres “desapareceram” dos cursos de computação?** São Paulo: Jornal da USP, 2018. Disponível em: <<http://jornal.usp.br/universidade/por-que-as-mulheres-desapareceram-dos-cursos-de-computacao/>>. Acesso em: 20 de Agosto de 2018.
- [17] KENDALL, Graham. **Here’s why there’s a gender gap in computer science**. The World Economic Forum, 2017. Disponível em: <<https://www.weforum.org/agenda/2017/01/heres-why-theres-a-gender-gap-in-computer-science>>. Acesso em: 21 de Agosto de 2018.
- [18] REINSEL, David; GANTZ, John; RYDNING, John (13 de abril de 2017). **Data Age 2025: The Evolution of Data to Life-Critical (PDF)**. seagate.com. Framingham, MA, US: International Data Corporation. Acesso em: 4 de Maio de 2018.

- [19] UNESCO. **Decifrar o código: educação de meninas e mulheres em ciências, tecnologia, engenharia e matemática (STEM)**. Brasília: UNESCO, 2018. 86 p., il. Acesso em: 10 de Agosto de 2018.
- [20] MCKINSEY. **JOBS LOST, JOBS GAINED:WORKFORCE TRANSITIONS IN A TIME OF AUTOMATION**. McKinsey Global Institute, 2017.
- [21] World Economic Forum (2017). **Global Gender Gap Report 2017**. World Economic Forum, pp.99. Disponível em: <[http://www3.weforum.org/docs/WEF\\_GGGR\\_2017.pdf](http://www3.weforum.org/docs/WEF_GGGR_2017.pdf)>. Acesso em 10 de Maio de 2018.
- [22] LOPES, Guacira. **A emergência do gênero**. Tomaz Tadeu da Silva. Gênero sexualidade e educação. Petrópolis: Editora Vozes Ltda. 1997.
- [23] HACHÉ, Alex; CRUELS, Eva; VERGÉS, Nuria. **Mujeres programadoras y Mujeres hackers. Una aproximación des de Lela Coders**. 2011. Disponível em: <<http://www.rebelion.org/docs/141550.pdf>> . Acesso em: 17/03/2018.
- [24] Computer Science Org, 2018. **The Current State of Women in Computer Science**. Disponível em: <<https://www.computerscience.org/resources/women-in-computer-science/>>. Acesso em: 28 Out. 2018.
- [25] MAGOULAS, Roger; LORICA, Ben. **Introduction to Big Data**. Fevereiro 2009. Release 2.0. Sebastopol CA: O'Reilly Media.
- [26] Manyika, J., Chui, M., Brown, B., et al. **Big Data: The Next Frontier for Innovation, Competition, and Productivity**. 2011. McKinsey Global Institute.
- [27] MATTHEW O. W.; GRINSTEIN, G.; KEIM, D. **Interactive Data Visualization Foundations Techniques and Applications**. 2nd Edition, CRC Press, 2010.
- [28] TABLEAU. **Visual Analysis Best Practices Simple Techniques for Making Every Data Visualization Useful and Beautiful**. Tableau.
- [29] YAU, N. **Visualize This: The FlowingData Guide to Design, Visualization, and Statistics**. Ed. Willey. 384p. 2011.
- [30] Cunningham, S. J., & Macrae, C. N. **The colour of gender stereotyping**. *British Journal of Psychology*, 102, 598–614. doi: 10.1111/j.2044-8295.2011.02023.x. 2011.
- [31] KIRK, Andy. **Poll results: Let's have a think about blue-pink**. Visualising Data, 2015. Disponível em: <<http://www.visualisingdata.com/2015/02/poll-results-lets-think-blue-pink/>>. Acesso em: 10 Out 2018.

- [32] TABLEAU. **5 tips on designing colorblind-friendly visualizations**. Tableau, 2016. Disponível em: <<https://www.tableau.com/about/blog/2016/4/examining-data-viz-rules-dont-use-red-green-together-53463>>. Acesso em: 20 Out 2018
- [33] AMARAL, Marília Abrahão et al . **Investigando questões de gênero em um curso da área de Computação**. Rev. Estud. Fem., Florianópolis , v. 25, n. 2, p. 857-874, Agosto 2017 . Disponível em: <<https://goo.gl/uUaraV>>. Acesso em: 10 de Maio de 2018.
- [34] ISMAEL, LUCAS. **Estudo de usabilidade do SIGAA-UFPB como ferramenta de identificação e prevenção dos fenômenos de retenção e evasão**, Trabalho de Conclusão de Curso, 06/2018
- [35] TABLEAU. Compartilhamento das pastas de trabalho com pessoas que não têm o Tableau Desktop. Disponível em: <<http://kb.tableau.com/articles/howto/sharing-workbooks-without-tableau-desktop?lang=pt-br>>, Acesso em: 20 de Julho de 2018.
- [36] MCKINNEY, Wes. **Data Structures for Statistical Computing in Python**. Proceedings of the 9th Python in Science Conference, 51-56 (2010).
- [37] EHRHARDT, Erica **Genderize MIT License (MIT)**. Disponível em: <<https://pypi.org/project/Genderize/>>. Acesso em: 21 de Maio de 2018.
- [38] Color Laboratory.**Colorblind Web Page Filter**. Disponível em: <<http://colorlab.wickline.org/colorblind/colorlab/>>. Acesso em: 10 Agosto 2018.
- [39] GANDOMI, Amir. Haider, Murtaza. Beyond the hype: Big data concepts, methods, and analytics. **International Journal of Information Management**, Volume 35, Issue 2, Pages 137-144, April, 2015.
- [40] Schmidt, A. E., Britto, H., and Koehler, M. **Levantamento e análise dos dados da presença feminina discente no curso de bacharelado em sistemas de informação do instituto federal catarinense, campus camboriú**. Em 10<sup>o</sup> Mulheres na Tecnologia da Informação (Women in Information Technology, WIT), 2016, Porto Alegre, RJ, Em: Anais do XXXVI Congresso da Sociedade Brasileira de Computação (CSBC 2016).

# ANEXO A – ANEXOS E APÊNDICES 1

## Cores da visualização

As cores escolhidas, laranja para homens e azul para mulheres, respeitam os questionamentos levantados em [30, 31], sobre estereótipos relacionados a cor e gênero, e também as recomendações quanto à questão do daltonismo apontadas em [32].

Observa-se nas Figuras 29, 30, 31 e 32 os resultados relacionados a cada tipo de daltonismo:

- Protonotopia: Pessoas menos sensíveis à luz vermelha
- Deuteranopia: Pessoas com problemas relacionados à luz verde e vermelha
- Tritanopia: Pessoas que confundem azul com verde e amarelo com violeta.
- Achromatopsia: Pessoas que não conseguem perceber cores.

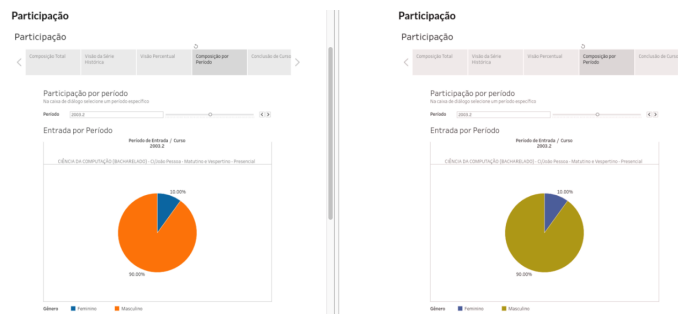


Figura 29: Resultado para protonotopia utilizando [38]

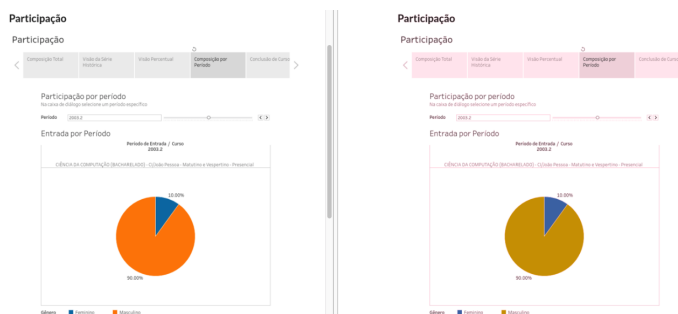


Figura 30: Resultado para deuteranopia [38]

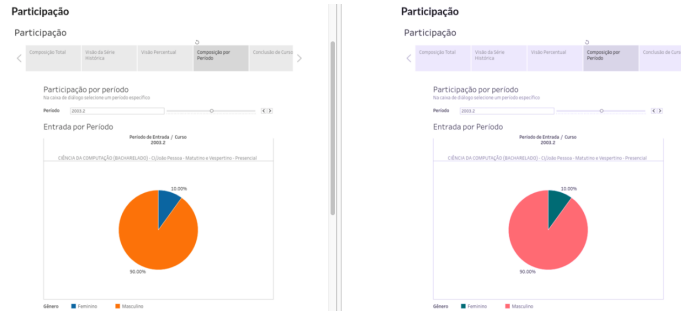


Figura 31: Resultado para tritanopia [38]

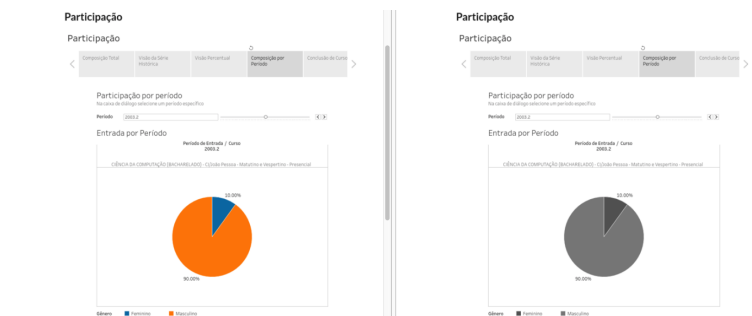


Figura 32: Resultado para achromatopsia [38]