

**DESIGUALDADE SALARIAL ENTRE HOMENS E MULHERES NA  
TECNOLOGIA DA INFORMAÇÃO:  
UMA ANÁLISE DESCRITIVA DE DADOS COM PYTHON****WAGE INEQUALITY BETWEEN MEN AND WOMEN IN  
INFORMATION TECHNOLOGY:  
A DESCRIPTIVE DATA ANALYSIS WITH PYTHON****DESIGUALDAD SALARIAL ENTRE HOMBRES Y MUJERES EN  
TECNOLOGÍAS DE LA INFORMACIÓN:  
UN ANÁLISIS DESCRIPTIVO DE DATOS CON PYTHON**Yohann Gabriel Oliani Freddi<sup>1</sup>Cristiane Palomar Mercado<sup>2</sup>Maria das Graças Junqueira Machado Tomazela<sup>3</sup>

Artigo recebido em Março de 2024

Artigo aceito em Junho 2024

DOI: 10.26853/Refas\_ISSN-2359-182X\_v10n05\_07

**RESUMO**

Debates sobre a disparidade salarial entre homens e mulheres vem aumentando nos últimos anos e este é um tema bastante relevante atualmente. No contexto brasileiro essa temática tem sido objeto de discussão em inúmeros estudos, os quais abordam as discrepâncias salariais em relação ao sexo, à experiência profissional, à etnia, à região, entre outros fatores, evidenciando a importância fundamental deste debate. O setor da TI é significativo para o mundo, representando 5% do Produto Interno Bruto (PIB) em 2011 e 6,8% no Brasil em 2019. Visto esta importância, verificou-se a necessidade de trazer o debate das diferenças salariais para o setor de TI no Brasil. Assim, o objetivo deste trabalho foi analisar a disparidade salarial entre homens e mulheres na área de Tecnologia da Informação (TI). Para isso foi realizada uma pesquisa quantitativa descritiva, utilizando como base os dados provenientes da Relação Anual de Informações Sociais (RAIS), que passaram por um processo de pré-processamento para viabilizar a análise proposta. Os processos foram realizados por meio da linguagem de programação Python e seus pacotes. Como resultados verificou-se a existência de diferenças salariais em desvantagem para o sexo feminino. A diferença salarial nos dados analisados foi de -19,97% em 2009 e -22,86 em 2019, comportamento semelhante à desigualdade salarial em geral no país. Para os recortes conforme

---

<sup>1</sup> Graduado em Análise e Desenvolvimento de Sistemas pela Fatec Sorocaba. Graduado em Ciências Econômicas pela Universidade de São Carlos. E-mail: yohannmax\_10@hotmail.com. Lattes: <http://lattes.cnpq.br/2298689854202173>. OrcId: <https://orcid.org/0009-0005-5786-8139>.

<sup>2</sup> Mestranda em Estudos da Condição Humana na Universidade Federal de São Carlos. Docente da Fatec Sorocaba. E-mail: cristiane.mercado@fatec.sp.gov.br. Lattes: <http://lattes.cnpq.br/6546422898817208> OrcID: <https://orcid.org/0000-0002-3095-1849>.

<sup>3</sup> Doutora em engenharia de produção pela Universidade Metodista de Piracicaba. Docente da Fatec Sorocaba. E-mail: graca.tomazela@fatec.sp.gov.br. Lattes: <http://lattes.cnpq.br/9809874053943708>. OrcID: <https://orcid.org/0000-0002-5471-2658>.

as ocupações internas de TI (diretoria, gerência etc.) e raça, na maioria, desvantagem salarial para as mulheres, com significativa piora de cenário entre os anos.

**Palavras-chave:** Desigualdade Salarial; Gênero; Tecnologia da Informação; Python;

## ABSTRACT

Debates about the pay gap between men and women have been increasing in recent years and this is a very relevant topic today. In the Brazilian context, this topic has been the subject of discussion in numerous studies, which address salary discrepancies in relation to gender, professional experience, ethnicity, region, among other factors, highlighting the fundamental importance of this debate. The IT sector is significant for the world, representing 5% of the Gross Domestic Product (GDP) in 2011 and 6.8% in Brazil in 2019. Given this importance, there was a need to bring the debate on salary differences to the IT sector in Brazil. Therefore, the objective of this work was to analyze the salary disparity between men and women in the area of Information Technology (IT). To this end, descriptive quantitative research was conducted, using as a basis data from the Annual Social Information List (RAIS), which went through a pre-processing process to enable the proposed analysis. The processes were conducted using the Python programming language and its packages. As a result, it was verified the existence of salary differences to the disadvantage of females. The salary difference in the data analyzed was -19.97% in 2009 and -22.86 in 2019, a behavior similar to salary inequality in general in the country. For the selections according to internal IT occupations (board, management, etc.) and race, in the majority, there is a salary disadvantage for women, with a significant worsening of the situation between the years.

**Keywords:** Pay Inequality; Gender; Information Technology; Python;

## RESUMEN

Los debates sobre la brecha salarial entre hombres y mujeres han ido aumentando en los últimos años y este es un tema muy relevante hoy en día. En el contexto brasileño, este tema ha sido objeto de discusión en numerosos estudios, que abordan las discrepancias salariales en relación al género, experiencia profesional, etnia, región, entre otros factores, destacando la importancia fundamental de este debate. El sector de TI es importante para el mundo, representando el 5% del Producto Interno Bruto (PIB) en 2011 y el 6,8% en Brasil en 2019. Dada esta importancia, era necesario llevar el debate sobre las diferencias salariales al sector de TI en Brasil. Por ello, el objetivo de este trabajo fue analizar la disparidad salarial entre hombres y mujeres en el área de Tecnologías de la Información (TI). Para ello se realizó una investigación cuantitativa descriptiva, utilizando como base datos del Listado Anual de Información Social (RAIS), los cuales pasaron por un proceso de preprocesamiento para posibilitar el análisis propuesto. Los procesos se realizaron utilizando el lenguaje de programación Python y sus paquetes. Como resultado, se constató la existencia de diferencias salariales en desventaja para las mujeres. La diferencia salarial en los datos analizados fue de -19,97% en 2009 y -22,86 en 2019, comportamiento similar a la desigualdad salarial en general en el país. En las selecciones según las ocupaciones internas de TI (junta directiva, dirección, etc.) y la raza, en la mayoría de los casos se observa una desventaja salarial para las mujeres, con un empeoramiento significativo de la situación entre los años.

**Palabras clave:** Desigualdad salarial; Género; Tecnología de la información; Python;

## 1 INTRODUÇÃO

Ao se pensar sobre a sociedade, tem-se à nossa frente um universo de interações e possibilidades: o padrão de consumo capitalista, a cidadania de cada indivíduo, as relações comerciais globalizadas, o cuidado (ou não) da natureza, as relações trabalhistas, entre outros

diversos aspectos dos possíveis e existentes encadeamentos sociais. Compreender os mais variados fatores e relações que ocorrem em uma sociedade é importante para o entendimento de nós como seres sociais, assim como, para encontrar formas e caminhos de construir uma sociedade mais justa para todos os seus membros (Bauman e May, 2010).

Na sociedade brasileira as diferenças de remuneração entre homens e mulheres são muito altas, em comparação com outras sociedades em estágio de desenvolvimento equivalente (Coelho e Corseuil, 2002).

O setor de Tecnologia da Informação é de extrema importância para a economia mundial, pois é um setor capaz de transformar a sociedade, uma vez que está presente em grande parte das atividades desenvolvidas pelos seres humanos. Segundo documento da Associação Brasileira das Empresas de Tecnologia da Informação e Comunicação-BRASSCOM (2012), em 2011, a TIC (Tecnologia da Informação e Comunicação) representava cerca de US\$ 3,4 trilhões no mundo, aproximadamente 5% do Produto Interno Bruto (PIB) mundial da época.

O estudo das diferenças salariais tem um campo amplo de debate, dessa forma, é importante analisar estes diferenciais também na área da tecnologia da informação (TI). Ao fazê-lo, contribui-se para uma compreensão mais específica dessas diferenças em toda a área, considerando a relevância fundamental desse setor para o país e o mundo.

Assim, o objetivo deste trabalho foi apresentar as relações salariais trabalhistas, considerando a possível diferenciação entre homens e mulheres, sendo tais profissionais situados nas ocupações de Tecnologia da Informação, tanto em setores diretamente ligados à área como entre outras áreas produtivas.

Propõe-se neste artigo abordar o tema de diferenças salariais voltadas ao sexo biológico dos trabalhadores, classificados como homens ou mulheres. Não foi possível utilizar a identidade de gênero visto adequação à classificação presente na base de dados que foi utilizada. A análise realizada tem um outro peso de importância quando considerada a desvalorização (menores salários e condições) da mulher no trabalho de forma geral e, em específico, na queda da representatividade no setor de TI. Por tanto, a partir dos resultados do trabalho em tela, foi possível compreender a existência de tais diferenças e o quão profundas são, considerando os recortes gerais específicos de dados.

A fim de avaliar estas diferenças, foram coletados, tratados e analisados dados da Relação Anual de Informações Sociais (RAIS) do Ministério da Economia, comparando a quantidade (frequência de pessoas) e salário mediano anual entre homens e mulheres para diversos cruzamentos de dados específicos, separando as diferenças em questões de ocupações (cargos) específicas de TI e raça. Para compreendê-los dentro da grande área da TI foram utilizados códigos referentes às diversas funções trabalhistas relacionadas a área, a partir da Classificação Brasileira de Ocupações (CBO), conforme quadro 2 apresentado no apêndice O recorte dos dados no tempo foi do ano de 2009 e 2019, com descrições e comparações em uma década.

Para a realização da análise dos dados extraídas da RAIS, foi utilizado como método a análise exploratória de dados. Como ferramenta facilitador destas análises, foi utilizada a linguagem de programação Python, juntamente dos softwares necessários para o controle de pacotes e bibliotecas, tanto quanto para a construção do código de forma mais interativa.

## 2 REFERENCIAL TEÓRICO

O referencial teórico aborda as disparidades de renda no Brasil, destacando a relevância do setor de Tecnologia da Informação e o papel das mulheres nesse campo específico.

### 2.1 Estudos sobre as desigualdades de renda no Brasil

No Brasil, o primeiro a abordar o tema de diferenças salariais foi Langoni em 1973, dando início ao debate do assunto. O autor propôs uma análise detalhada das desigualdades de renda demonstradas no mercado de trabalho, por meio dos diferenciais salariais conforme observa-se o nível educacional, a idade, o sexo, o setor e a região da qual um trabalhador possui ou está. Porém, por dificuldade de acesso a dados para estas análises, o debate não foi difundido e aprofundado em outros estudos nas duas décadas subseqüentes. A maior recorrência de estudos no tema foi a partir de 1990, visto liberação de acessos a dados do IBGE relativos a pesquisas domiciliares em formas digitais e ao avanço tecnológico de computadores utilizados para a análise dos dados (Coelho e Corseuil, 2002).

Coelho e Corseuil (2002) observam que são diversos os motivos para as desigualdades salariais e apontam cada estudo e seus detalhes a partir da primeira abordagem de Langoni. São três grupos de características trabalhistas que impactam essas diferenças: (1) pessoais produtivos, (2) não-produtivos e (3) dos postos de trabalho.

No grupo de características pessoais produtivas (1), temos que os salários podem ser impactados por meio: do nível educacional (fundamental, médio, técnico, superior e pós-graduação), que é o tópico mais estudado no Brasil; e do aumento da experiência profissional do trabalhador, geralmente medido pela idade e pelo tempo de trabalho.

Já no grupo de características pessoais não-produtivas (2) temos: as diferenças salariais pelo gênero do trabalhador; ou pela cor.

Por fim, temos as características referentes aos postos de trabalho (3): questões que envolvem a carteira de trabalho, ou a distribuição dos postos de trabalho não registrados e uma menor remuneração relacionada; o grau de sindicalização do trabalhador; o setor de atividade; a região da qual o posto de trabalho e a residência do trabalhador se encontra; e a natureza jurídica do posto, se é uma empresa pública ou privada.

Um dos primeiros trabalhos científicos de relevância na análise da discriminação por gênero nos salários no Brasil foi o de Camargo e Serrano (1983). Na análise foram utilizados dados da Relação Anual de Informações Sociais (RAIS) de 1976, estimando as diferenças conforme sexo dentro de características como escolaridade e do mercado de trabalho. A principal conclusão foi que há diferenças salariais de sexo, sem apontar a magnitude dessas diferenças (Coelho e Corseuil, 2002).

Um outro levantamento foi realizado com microdados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 1981 a 1989. Os autores, Barros, Ramos e Santos (1995) concluíram que além da existência de diferença salarial entre homens e mulheres, tal diferença era acima de 50%, sendo a desvantagem para as mulheres. Dentro do período analisado (oito anos) ocorreu a redução da desigualdade em 3% (Coelho e Corseuil, 2002).

Há também trabalhos que descrevem tais diferenças por sexo de forma geográfica, como em Ometto, Hoffmann e Alves (1997), que compararam dois estados, São Paulo e Pernambuco, demonstrando que a condição de nascer como mulher é mais desfavorável a seu salário quando

esta nasce em Pernambuco do que em São Paulo. Já Cavalieri e Fernandes (1998) com dados de 1989, indicam que as diferenças são menores nas regiões metropolitanas do Sul e Sudeste, comparadas às restantes. Também encontram que tais diferenças salariais são menores entre pessoas brancas, e a mesma tendência ocorre para os menos escolarizados, comparado a pessoas com mais escolaridade (tempo de estudo). Por fim, demonstram um formato de U invertido para a comparação entre idades, logo, nas faixas mais baixas e elevadas, as diferenças por sexo são menores do que nas intermediárias (Coelho e Corseuil, 2002).

Kassouf (1998) indica que as diferenças salariais entre homens e mulheres são ainda mais profundas quando os trabalhadores analisados estão no mercado informal, sendo significativa também para o formal (Coelho e Corseuil, 2002).

Em 2000, o trabalho de Leme e Wajnman demonstrou que existe tendência de redução das diferenças salariais por sexo conforme passam as gerações, ou seja, para cada geração mais jovem, a diferença salarial já observada em desfavor para as mulheres, tende a ser menor em todas as idades (Coelho e Corseuil, 2002).

## **2.2 A importância do setor da Tecnologia da Informação no Brasil**

As empresas que utilizam a TIC apresentam um melhor desempenho, aumento da produtividade de setores chaves e diminuição de gargalos no desenvolvimento econômico, principalmente voltado para as matrizes energéticas, os sistemas de transportes e as cidades como um todo. Esse conjunto de fatores ressaltam a importância do setor de Tecnologia da Informação para a economia mundial.

Também é um setor que transforma a sociedade, pois, pode trazer vantagens com muita rapidez. Dois exemplos são citados no relatório “Estratégia TIC Brasil 2022” da BRASSCOM: (1) a redução da pobreza em comunidades (aumento de 10% na renda) de pescadores na cidade de Kerala, na Índia, a partir do acesso a novos mercados facilitado pela utilização de telefones celulares; e (2) a variação de 22% para 90% no número de pessoas com tuberculose que conseguiram finalizar o tratamento no sistema de saúde da África do Sul.

A TIC auxilia em diversos outros setores como: segurança, educação, agronegócio, setor público em geral, meio ambiente, entre outros, além de permitir a automelhoria, utilizando de suas capacidades tecnológicas para aumentar a própria eficiência, produção e inovação (BRASSCOM, 2012).

No Brasil, para dados de 2019, o setor representava 6,8% do PIB, com forte importância para diversos setores, sejam estes públicos ou privados. Tal setor também empregou cerca de 1,56 milhão de pessoas no país todo. A remuneração média do setor era mais que o dobro (R\$ 4.170) da remuneração média nacional (R\$ 1.928). Tais dados demonstram não somente a sua relevância mundial, mas também no país (BRASSCOM, 2019).

## **2.3 A importância das mulheres na TI.**

A grande área de estudo da TI em seus primórdios era predominantemente preenchido por mulheres. O primeiro curso de Ciências da Computação do Instituto de Matemática e Estatística da Universidade de São Paulo (IME/USP), no ano de 1974, tinham as mulheres como maioria. No mercado de trabalho, entre o início da década de 1970 até a de 1980, houve um crescimento na quantidade de mulheres na computação, de 10% para 36% (Silveira, 2018). Nos Estados Unidos há dados semelhantes: no ano de 1986 a participação feminina era de 40% no mercado de trabalho da TI (Ramsey e Mccorduck, 2005).

Grandes criações e trabalhos foram realizados por mulheres na computação: Ada Lovelace foi quem desenvolveu o primeiro algoritmo da história; a freira Kennet Keller ajudou a criar a linguagem de programação BASIC; Grace Hopper por sua vez criou a linguagem Flow-Matic, base para o COBOL; tecnologias necessárias para as redes de telefone e wi-fi foram criadas por Hedy Lamarr. A lista de contribuições das mulheres para a sociedade a partir da TI é de alta importância e quantidade (Silveira, 2018).

Porém, a partir do final da década de 1990 começa a ocorrer uma diminuição da participação das mulheres nos cursos de computação e no mercado de trabalho. Nos Estados Unidos, em 1999, a proporção de mulheres era de 29%, em comparação com os 40% de 13 anos antes. No curso de Ciências da Computação do Instituto de Ciências Matemáticas e de Computação (ICMS) da USP, em 1997, 48% das pessoas formadas eram mulheres, já em 2003 (seis anos depois) há uma queda para 12% de formandas. Em 2016, para o mesmo curso, somente 3% dos formandos eram mulheres (Silveira, 2018).

Sobre esta diminuição na quantidade de mulheres na academia e no mercado de trabalho da TI, Ramsey e McCorduck (2005) levantam alguns pontos: no geral os estereótipos culturais e sociais moldam com o passar do tempo tais participações. As características estereotipadas da pessoa que trabalha com TI são: homens brancos, geralmente antissociais, individualistas, competitivos e com físico sedentário. Mulheres estão diretamente definidas para atividades mais humanas, como nutrição, comunicação e de expressões emocionais, e em contraponto, homens estão para tarefas mais orientadas e que utilizam ferramentas. Outro estereótipo é de que mulheres são mais voltadas a família e isto diminuiu a possibilidade de ter funções de trabalho que precisam de mais tempo e viagens. O resultado da pressão destes estereótipos sobre os sexos acaba por subvalorizar as mulheres na TI e diminuir a propensão de uma mulher escolher esta área, visto a possibilidade de julgamento, sofrimento e desvalorização.

### 3 MÉTODO

A metodologia utilizada neste trabalho foi quantitativa descritiva, chamada também de análise descritiva de dados, que segundo Gil (2002), objetiva primordialmente a descrição das características de determinada população ou fenômeno.

Os dados utilizados para o desenvolvimento do presente trabalho foram obtidos a partir da base de dados do Ministério do Trabalho e Emprego, a Relação Anual de Informações Sociais, na qual foi possível encontrar informações de cada indivíduo do sexo feminino e masculino, em combinação com as informações das ocupações pelas Classificação Brasileira de Ocupações (Apêndice A), as quais estão também presentes na RAIS.

A coleta de dados foi realizada por meio de software de transferência de dados via <sup>4</sup>File Transfer Protocol (FTP), capturados para cada ano necessário para a realização da pesquisa, 2009 e 2019. Os dados são disponibilizados no formato de arquivo de texto (.txt), separados por ponto e vírgula (;).

---

<sup>4</sup> File Transfer Protocol ou Protocolo de Transferência de Arquivos permite a transferência de arquivos de e para um computador remoto.

Foi utilizada a ferramenta <sup>5</sup>Filezilla para realizar o download dos dados desagregados (microdados) disponibilizados em FTP pelo próprio Ministério citado.

Para o ano de 2009, os dados coletados estavam separados para cada UF do Brasil, totalizando 18,9 GB de dados brutos. Já para 2019 os dados estavam agregados conforme as macrorregiões do país, totalizando 30,2 GB de dados brutos.

Visto a disponibilidade dos dados de cada ano em agregações diferentes e a necessidade de agrupar e filtrar somente os dados referentes a ocupações de TI, foi necessário realizar a agregação, limpeza e filtragem destes dados para cada ano. O primeiro tratamento foi a eliminação de diversas colunas dos arquivos não úteis à análise e foram filtrados os dados (separados por trabalhador) que estavam com vínculo ativo em 31/12. Na sequência foram descartados os trabalhadores que estavam com remuneração nominal em dezembro zerados. Então foram capturados os dados referentes a trabalhadores dentro das famílias consideradas por este trabalho como ocupações vinculadas à TI. As famílias consideradas foram: Diretores de tecnologia da informação, Gerentes de Tecnologia da Informação, Engenheiros de Computação, Analistas de Tecnologia da Informação, Técnicos de desenvolvimento de sistemas e aplicações, Técnicos de Suporte e monitoração ao usuário de tecnologia da informação e administradores de tecnologia da informação (quadro 1).

Um outro filtro necessário foi pelos códigos setoriais. Para isto foram selecionados os códigos referentes aos setores de TI pela Classificação Nacional de Atividades Econômicas (CNAE 2.0), sendo estes: 620 - Atividades dos serviços de tecnologia da informação e 631 – Tratamento de dados, hospedagem na internet e outras atividades. Por fim foi criada uma coluna definindo a união federativa do grupo de dados do arquivo lido. Após isto, todos os arquivos foram agrupados e então finalizada a coleta e filtragem dos dados.

Para análise dos dados foi utilizada a linguagem <sup>6</sup>Python e seus <sup>7</sup>pacotes: Time<sup>8</sup>, para anotação de tempo de execução; <sup>9</sup>Pandas, para manipulação de tabelas; <sup>10</sup>Glob, para manipulação de arquivos e pastas; <sup>11</sup>Matplotlib, para criação de gráficos; <sup>12</sup>Seaborn, para criação de gráficos; <sup>13</sup>Scipy, para testes estatísticos em geral; <sup>14</sup>Statsmodels, para outros testes

<sup>5</sup> É um programa de código aberto e de distribuição gratuita, usado como cliente e servidor FTP. <https://filezilla-project.org/>.

<sup>6</sup> É uma linguagem de programação e que pode ser aplicada a diversos tipos de problema, desde desenvolvendo de sistemas web a ciência de dados. <https://docs.python.org/>.

<sup>7</sup> Pacotes Python são uma maneira de estruturar o “espaço de nomes” dos módulos Python. Módulos executam funções específicas. <https://docs.python.org/pt-br/3/tutorial/modules.html>.

<sup>8</sup> Time: Módulo do Python que provê várias funções relacionadas à tempo (datas e horários). <https://docs.python.org/pt-br/3.9/library/time.html>.

<sup>9</sup> Pandas é uma ferramenta de análise e manipulação de dados de código aberto. <https://pandas.pydata.org/>.

<sup>10</sup> Glob é um módulo do Python usado para procurar nomes de caminhos no computador que correspondam a um padrão especificado. <https://docs.python.org/3/library/glob.html>.

<sup>11</sup> Matplotlib é uma biblioteca usada para criar visualizações estáticas, animadas e interativas em Python. <https://matplotlib.org/>.

<sup>12</sup> Seaborn é uma biblioteca de visualização de dados Python baseada em matplotlib. Fornece uma interface de alto nível para desenhar gráficos estatísticos. <https://seaborn.pydata.org/>.

<sup>13</sup> Scipy é um software de código aberto para resolver algoritmos de matemática, ciência e engenharia. <https://docs.scipy.org/doc/scipy/>.

<sup>14</sup> StatsModels é um módulo Python que fornece funções para a estimativa de modelos estatísticos. <https://www.statsmodels.org/stable/index.html>.

estatísticos; <sup>15</sup>Numpy, para manipulação matemática de dados; <sup>16</sup>Geopandas, para manipulação de dados geográficos e criação de mapas.

Já para criar a estrutura dos <sup>17</sup>workbooks foi utilizada a ferramenta <sup>18</sup>anaconda; o pacote e ferramenta jupyter, para criação e manipulação de arquivos workbooks e a ferramenta <sup>19</sup>visual studio code, para edição de <sup>20</sup>scripts e workbooks, ver Quadro 1.

Quadro 1 - CBOs utilizadas no filtro dos dados

FAMÍLIA	TÍTULO
1236	Diretores de tecnologia da informação
1425	Gerentes de tecnologia da informação
2122	Engenheiros da computação
2124	Analistas de tecnologia da informação
3171	Técnicos de desenvolvimento de sistemas e aplicações
3172	Técnicos de suporte e monitoração ao usuário de tecnologia da informação
2123	Administradores de tecnologia da informação

Fonte: Classificação Brasileira de Ocupações - Ministério da Economia

#### 4 RESULTADOS E DISCUSSÃO

Com os dados devidamente agregados e filtrados, foi possível realizar correlações entre sexo e remuneração em diversas perspectivas, abrangendo ocupação, UF, raça e escolaridade. Na próxima seção, são apresentadas informações referentes a todos os indivíduos e remunerações analisadas, seguida por uma análise mais específica, envolvendo as ocupações no setor de TI e recortes raciais.

<sup>15</sup> NumPy é um projeto de código aberto que permite computação numérica com Python.

<sup>16</sup> Geopandas é um projeto de código aberto que facilita o trabalho com dados espaciais (dados sobre locais na superfície terrestre) em python. <https://geopandas.org/en/stable/>.

<sup>17</sup> Workbooks são documentos para escrever comandos para serem executados.

<sup>18</sup> Anaconda é uma plataforma usada para construir e implantar iniciativas de ciências de dados e IA. <https://www.anaconda.com/>.

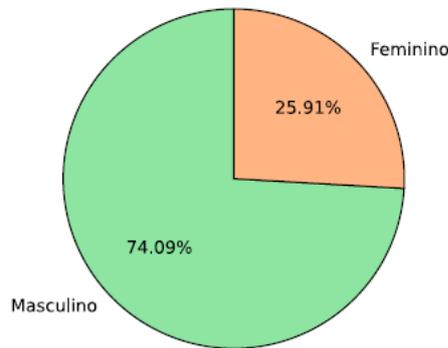
<sup>19</sup> É um editor de código-fonte (comandos de programação) desenvolvido pela Microsoft. <https://code.visualstudio.com/>.

<sup>20</sup> Script é um conjunto de comandos a serem executados pelo computador.

#### 4.1 Diferença salarial entre homens e mulheres em 2009 para ocupações de TI em todos os setores do mercado de emprego formal.

O arquivo de 2009 somou um total de 393 mil e 937 trabalhadores, destes são 102 mil e 86 mulheres e 291 mil e 851 homens. Do conjunto, 25,91% são mulheres e 74,09% são homens. Logo, são 2,86 vezes mais homens que mulheres, como pode ser observado na Figura 1, o que corrobora com o trabalho de Silveira (2018), no qual a quantidade de mulheres na área de TI é bem menor que a de homens.

Figura 1 - Participação de homens e mulheres como trabalhadores de TI em 2009



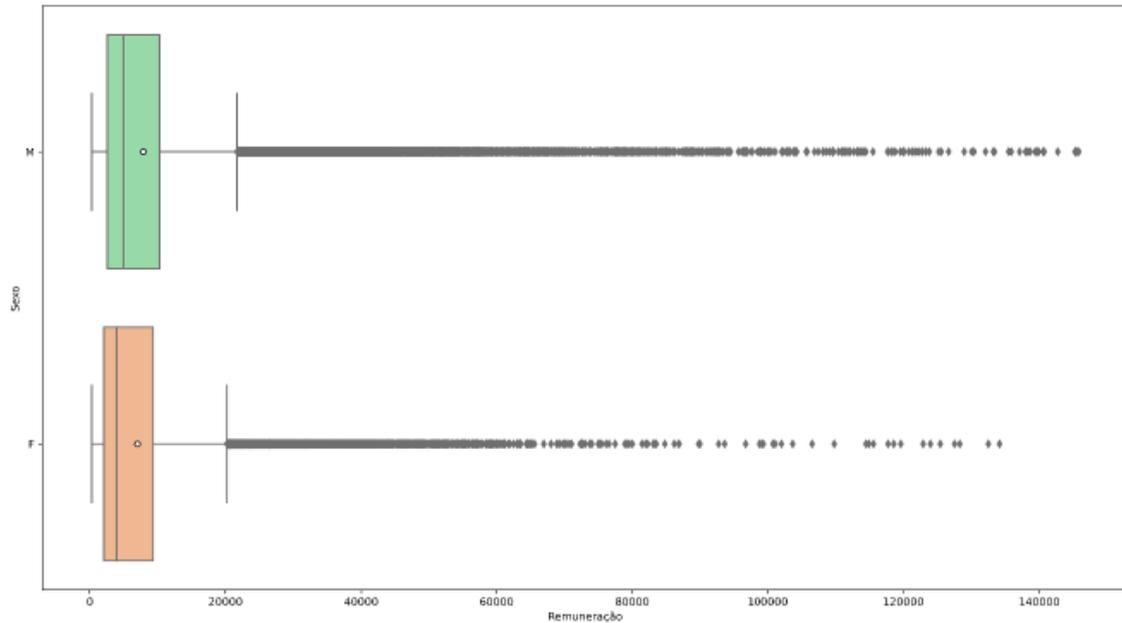
Fonte: Elaborado pelo autor, com dados da RAIS 2009

Com os dados referentes à massa salarial, temos para o sexo masculino um valor de aproximadamente dois bilhões e trezentos milhões de reais. Já para a massa do sexo feminino tem-se o valor aproximado de 718 milhões de reais. A proporção entre os valores é de 3,22 vezes mais massa salarial para os homens que para as mulheres. Em termos percentuais, temos 76,32% representando a parcela do montante para homens e 23,68% para as mulheres. O gráfico da Figura 2, que mostra a distribuição da massa salarial entre os sexos feminino e masculino, revela algumas observações a) para remunerações acima de R\$ 60.000,00, há mais pontos na linha dos homens do que das mulheres, indicando uma predominância masculina em faixas salariais mais altas; b) as barras à esquerda representando as mulheres estão visualmente deslocadas mais para a esquerda em comparação com as dos homens, sugerindo uma possível disparidade salarial desfavorável às mulheres, assim como a posição da mediana.

Ao verificar as medianas tem-se R\$ 4.950,41 para os homens e R\$ 3.962,06 para as mulheres, representando uma variação mediana percentual de -19,97%, ou seja, as trabalhadoras do sexo feminino recebem 20% menos que os trabalhadores do outro sexo.

Os testes estatísticos para validação das medianas deste grupo analisado resultaram também em rejeição da hipótese nula, validando que são diferentes entre si e confirmando estatisticamente que a mediana do grupo do sexo feminino é menor que a do sexo masculino, como observado.

Figura 2 - Distribuição das remunerações entre o sexo masculino (M) e feminino (F) para trabalhadores de TI em 2009

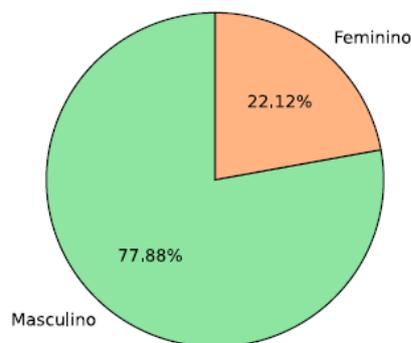


Fonte: Elaborado pelo autor, com dados da RAIS 2009

#### 4.2 Diferença salarial entre homens e mulheres em 2019 para ocupações de TI em todos os setores de emprego formal.

Para o ano de 2019, considerando o grupo de trabalhadores em ocupações de TI independente do setor de atuação, podemos verificar as seguintes quantidades entre os sexos: são 455 mil e 327 homens e 129 mil e 352 mulheres, um total de 584 mil e 679 pessoas, um aumento aproximadamente de 31% na população total de trabalhadores comparado a 2009. As representações percentuais de cada sexo são de 22,12% mulheres e 77,88% homens. No gráfico da Figura 3 pode-se visualizar esta proporção.

Figura 3 - Participação de homens e mulheres como trabalhadores de TI em 2019



Fonte: Elaborado pelo autor, com dados da RAIS 2019

Com os dados referentes à massa salarial, tem-se para o sexo masculino um valor de aproximadamente três bilhões e 600 milhões de reais. Já ao verificar a massa do sexo feminino

tem-se o valor de 890 milhões de reais, aproximadamente. A proporção entre os valores é de 4,10 vezes mais massa salarial para os homens que para as mulheres. Em termos percentuais, tem-se 80,38% representando a parcela da massa para homens e 19,62% para as mulheres.

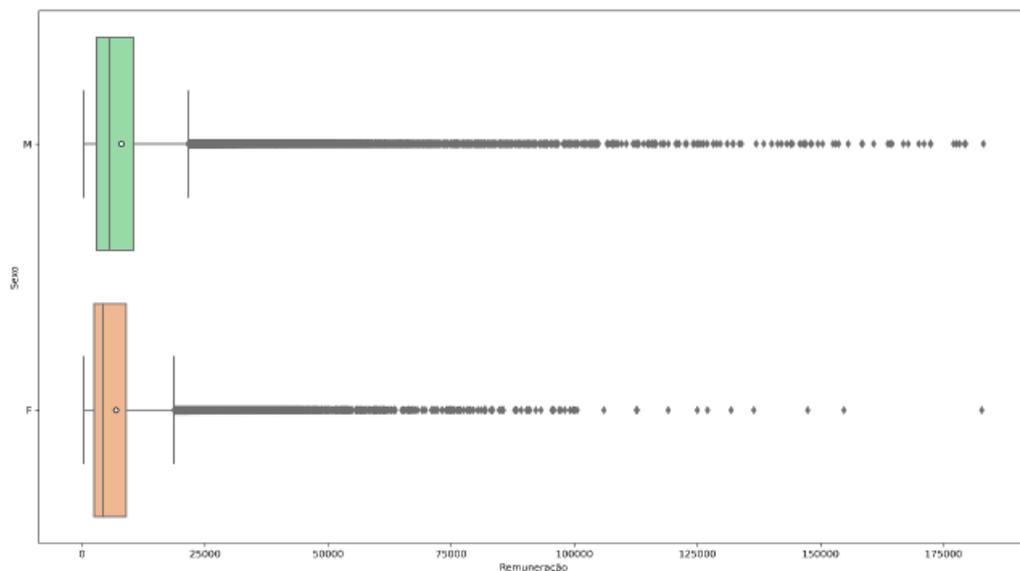
Então ao comparar a proporção de homens e mulheres e a proporção da massa salarial de cada, tem-se 0,58 a mais para a massa salarial referente à quantidade de cada trabalhador, ou seja, possivelmente existe diferença salarial entre os dois grupos.

O gráfico da Figura 4 que mostra a distribuição da massa salarial entre os sexos feminino e masculino ficou semelhante ao gráfico do mesmo tipo com dados de 2009. Algumas observações: a) para remunerações maiores que R\$ 75.000,00 existem muito mais pontos na linha dos homens que no das mulheres, b) as barras à esquerda das mulheres estão mais à esquerda que a dos homens e mais achatadas, indicando uma possível diferenciação salarial em desvantagem das mulheres, assim como a linha que divide a caixa (mediana).

Ao visualizar as medianas tem-se R\$ 5.528,63 para os homens e R\$ 4.264,72 para as mulheres, representando uma variação mediana percentual de -22,86%, ou seja, as trabalhadoras do sexo feminino recebem 22,86% a menos que os trabalhadores do sexo masculino.

Os testes estatísticos para validação das medianas deste grupo analisado resultaram, também, em rejeição da hipótese nula, validando, então, que são diferentes entre si e confirmando estatisticamente que a mediana do grupo do sexo feminino é menor que a do sexo masculino.

Figura 4 - Distribuição das remunerações entre o sexo masculino (M) e feminino (F) para trabalhadores de TI em 2019



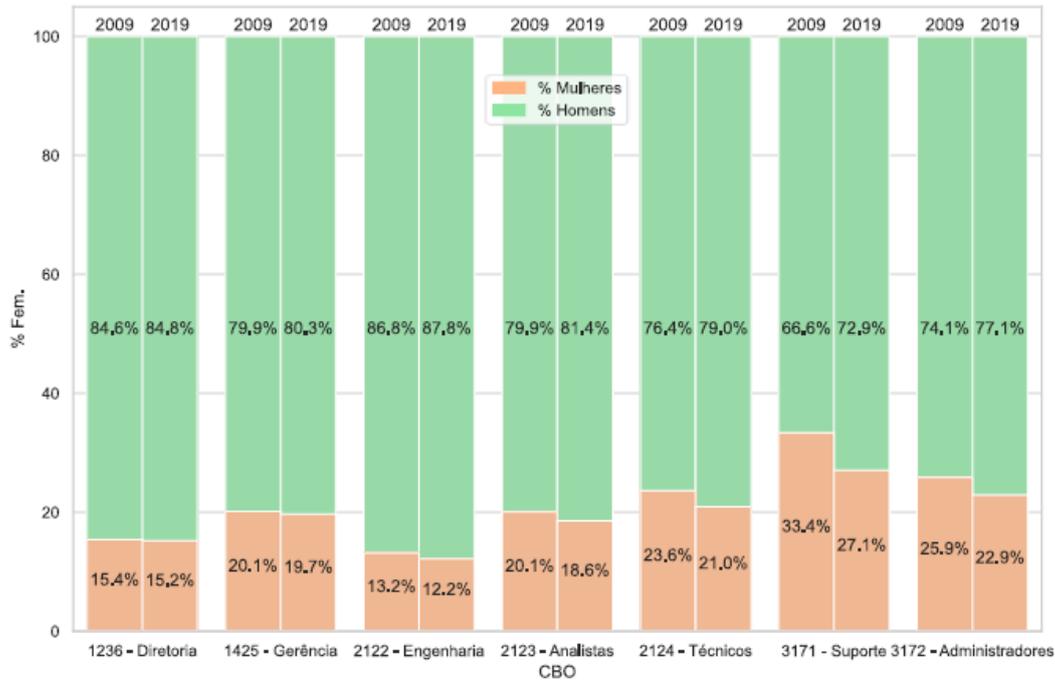
Fonte: Elaborado pelo autor, com dados da RAIS 2019

### 4.3. Diferença salarial entre homens e mulheres em 2009 e 2019 conforme suas ocupações específicas.

Ao considerar a seleção dos dados referentes a trabalhadores com ocupações de TI e distribuídos por todos os setores existentes, foi possível cruzar os dados para cada ocupação específica existente e o sexo dos trabalhadores. Para isso, foi calculada a mediana, visto que é

mais ajustada para remuneração de forma geral. Também é mostrada a quantidade de trabalhadores de acordo com sua ocupação para os anos de 2009 e 2019, ver Figura 5.

Figura 5 - Quantidade percentual de trabalhadores conforme sexo e ocupação para os anos de 2009 e 2019



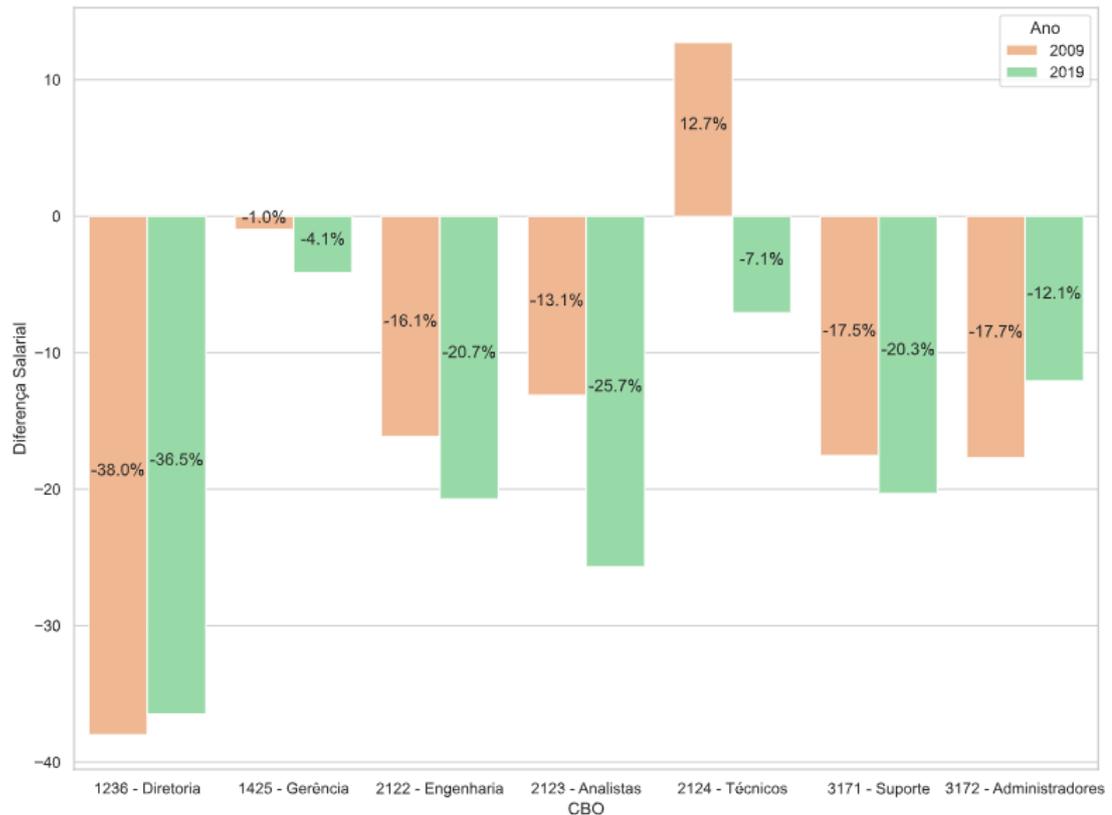
Fonte:

Elaborado pelo autor, com dados da RAIS 2009/2019

Conforme o gráfico da Figura 5, que mostra a quantidade percentual de trabalhadores conforme sexo e ocupação, comenta-se as seguintes mudanças em cada ocupação conforme os anos: para o grupo de ocupações de diretoria não houve mudanças significativas na composição da participação de cada sexo, isto também é observado na ocupação de gerência, de engenharia e de analistas; já a partir dos técnicos observa-se uma redução de 2,6% na participação de mulheres na ocupação, seguindo de 3% de redução para a categoria de administradores de TI e 6,3% para os técnicos. Logo, é possível observar que em todas as ocupações houve queda da participação feminina, demonstrando um cenário pior em 10 anos.

Para a análise de salários, observa-se na figura 6, que somente para o ano de 2009 e para a ocupação de técnicos tem-se uma vantagem salarial em favorecimento às mulheres, o restante das ocupações em ambos os anos existe diferença salarial mediana negativa. Para as ocupações de gerência, engenharia, analistas, técnicos e suporte, ocorre a piora do cenário de desvantagem. Já para diretoria e administração de TI, ocorre uma diminuição das diferenças em mediana. Porém, conclui-se que em mediana as mulheres recebem consideravelmente menos que os homens, independente da ocupação, sendo a maior diferença observada para os cargos de diretoria e a menor para de gerência, sendo um grande contraponto, visto serem dois cargos executivos.

Figura 6 - Diferenças salariais medianas conforme variação percentual entre homens e mulheres para cada ano entre as ocupações



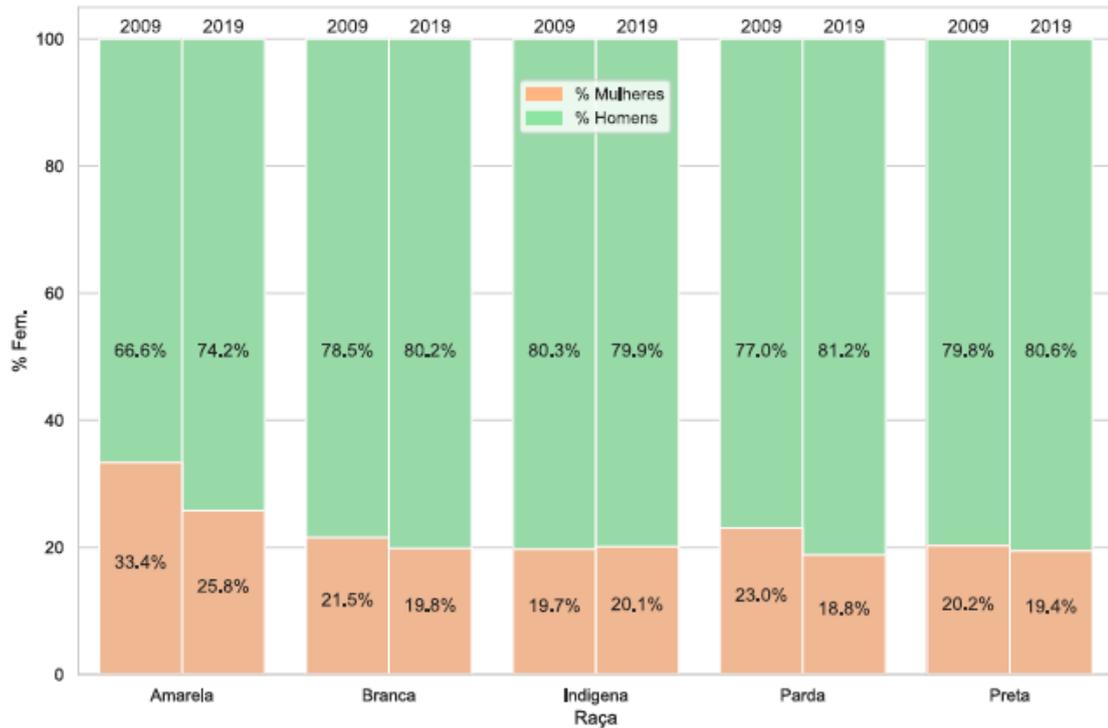
Fonte: Elaborado pelo autor, com dados da RAIS 2009/2019

#### 4.4. Diferença salarial entre homens e mulheres em 2009 e 2019 conforme a raça.

Nesta etapa foi realizada a análise das possíveis diferenças salariais entre homens e mulheres em ocupações de TI conforme a raça de cada indivíduo. As raças classificadas nos dados da RAIS são definidas como: indígena, branca, preta, amarela e parda, sendo essas nomenclaturas mantidas durante o texto aqui apresentado. É importante ressaltar que a quantidade total de pessoas aqui é diferente em cada ano do que para as análises anteriores, visto que nos dados foram retiradas as pessoas que não declararam raça.

O gráfico da Figura 7 mostra a participação de homens e mulheres no mercado de TI em cada ano, por raça. A participação feminina nas raças branca, indígena e preta tem pequenas mudanças de um período para o outro, seguindo da parda com uma diminuição de 4,2% e, por fim, na amarela com 7,6%. Novamente observa-se a baixa participação das mulheres, com a maioria das taxas próximas de 20%.

Figura 7 - Quantidade percentual de trabalhadores conforme sexo e raça para os anos de 2009 e 2019

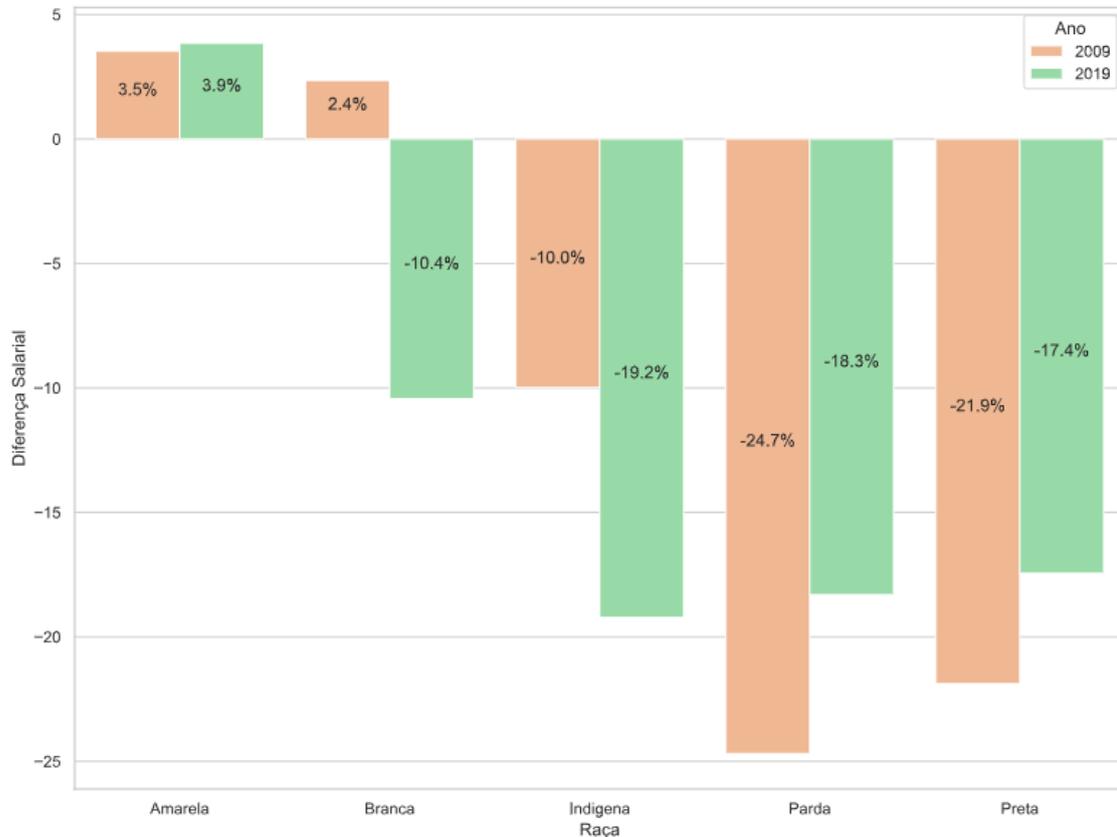


Fonte: Elaborado pelo autor, com dados da RAIS 2009/2019

A Figura 8 mostra as diferenças salariais medianas conforme variação percentual entre homens e mulheres para cada ano entre as raças. Realizando os testes estatísticos necessários para comparação das medianas, apontou-se algumas medianas como não diferentes estatisticamente, ou seja, não foi possível rejeitar a hipótese nula de igualdade entre as medianas. No ano de 2009 para as medianas entre as raças: amarela, branca e indígena, o teste apontou que estas medianas não são diferentes entre si ( $p > 0,01$ ). Isto ocorre também para as raças amarela e indígena para 2019.

Logo, as conclusões sobre a raça amarela do qual as remunerações medianas das mulheres são levemente maiores (3,5% e 3,9%) que a dos homens. Até que faz sentido, visto pequena diferença percentual em vantagem, então o salário mediano das mulheres e dos homens são iguais estatisticamente. Para a diferença salarial de 2,4% em vantagem das mulheres na raça branca, faz sentido as medianas serem iguais, exatamente pelo mesmo critério de serem medianas bem próximas. Por fim, na raça indígena temos percentuais de -10% e -19,2%, sendo o p-valor para 2009 de 0,335, muito maior que o nível de confiança considerado de 0,01, logo, não podemos afirmar que esta diferença salarial em desvantagem de 10% para as mulheres é estatisticamente possível, visto que pelo teste as medianas são consideradas iguais. Porém, para o ano de 2019 temos um p-valor de 0,027 que é muito próximo do nível de confiança de 0,01, então, não rejeitamos a hipótese nula de igualdade das medianas, porém, por muito pouco. Estas não constatações de medianas diferentes na raça indígena, provavelmente deve-se pela baixa quantidade de observações para o grupo, menos de mil pessoas no total dos sexos, em cada ano.

Figura 8 - Diferenças salariais medianas conforme variação percentual entre homens e mulheres para cada ano entre as raças



Fonte: Elaborado pelo autor, com dados da RAIS 2009/2019

Os resultados das análises corroboram com o trabalho de Coelho e Corseuil (2002) que usaram a mesma base de dados do presente trabalho, com dados de 1976 e chegaram à conclusão de que há diferenças salariais, mas não apontaram a extensão dessas diferenças. Outro trabalho, utilizando dados do PNAD entre os anos de 1981 e 1989, chegou à mesma conclusão sobre a existência de uma diferença salarial entre homens e mulheres acima de 50% ocorrendo redução da desigualdade em 3%.

O trabalho de Leme e Wajnman, em 2000, que demonstrou que existe tendência de redução das diferenças salariais por sexo conforme passam as gerações (COELHO e CORSEUIL, 2002), difere dos dados obtidos na presente pesquisa, que analisou os trabalhadores de TI em empregos formais e observou que as trabalhadoras do sexo feminino recebem 22,86% a menos que os trabalhadores do sexo masculino.

## 5 CONSIDERAÇÕES FINAIS

Diante do exposto, ao analisar os anos de 2009 e 2019, observa-se, de maneira geral, uma profunda diferença salarial mediana em desvantagem para o sexo feminino no Brasil. O valor da desigualdade salarial em 2009 foi de -19,97% em desvantagem para as mulheres. Já

para o ano de 2019 tem-se o valor percentual de -22,86% em mediana, demonstrando que em 10 anos, para trabalhadoras de TI em empregos formais distribuídas em todos os setores de empresas, a diferença salarial é ainda maior.

Para o recorte de dados conforme as ocupações de cada indivíduo, sendo estas de diretoria, gerência, engenharia, analistas, técnicos, suporte e administradores, conclui-se que há existência de diferença salarial em desvantagem às mulheres nos seguintes valores observados para as medianas (todos estatisticamente válidos): Para o grupo de diretoras, em 2009 o valor foi de -38%, pouco mudando o cenário em 2019 para -36,5%; Para as profissionais em ocupações de gerência, em 2009 observa-se o valor de -1,0%, aumentando a diferença em 10 anos para -4,1%; já para engenharia, de -16,1% em 2009 passa para -20,7%; para as analistas mulheres ocorre piora significativa: em 2009 a diferença era de -13,1% em desvantagem, do qual passa para -25,7% em 2019; o grupo de técnicos começa 2009 com uma alta vantagem salarial de 12,7% de maior remuneração para as mulheres, porém, ao passar uma década, ocorre a piora do cenário e as mulheres passam a ganhar em mediana menos 7,1% que os homens; para trabalhadores de suporte, as mulheres recebem -17,5% que os homens em mediana no ano de 2009, passando para -20,3% em 2019; Enfim, para o grupo de administradores, ocorre uma pequena melhora de um cenário de desvantagem para o sexo feminino: de -17,7% em 2009 para -12,1% em 2019. Logo, conclui-se que, na maioria dos percentuais (13 dos 14 anotados) existe diferenças salariais significativas em desvantagem as mulheres. Somente em um grupo ocorre uma melhora razoável da diferença, sendo este para administradores, no restante ocorre a piora ou a proximidade.

Para as diferenças salariais entre os sexos e a raça de cada indivíduo, novamente observa-se na maioria desvantagens para as mulheres: Na raça amarela, as mulheres recebem mais que os homens em mediana, 3,5% em 2009 e 3,9% em 2019; Já para o grupo de raça branca, em 2009 as mulheres recebiam mais, por volta de 2,4%, mas, em 10 anos, ocorre a desvantagem salarial no valor de -10,4%; para o grupo indígena em 2009 já existia diferença salarial para o sexo feminino em desvantagem: -10%, passando para -19,2% em 2019; tanto o grupo de raça parda e preta observa-se diferenças salariais mais profundas: -24,7% (2009) e -18,3% (2019) para a raça parda, ocorrendo uma melhora de cenário e -21,9%, em 2009, para -17,4% em 2019, para a raça preta. Então, para os grupos de raças temos somente para a raça amarela mulheres recebendo mais que os homens. Isto em 2009 para a raça branca foi observado também, porém ao passar da década acaba por se juntar aos outros grupos das outras raças recebendo muito menos que os homens conforme suas medianas. Porém, ao realizar o teste de diferença das medianas, pode-se concluir que as medianas de cada sexo para a raça amarela são iguais, sendo no valor absoluto realmente próximas. Já para o grupo da raça indígena que demonstrou diferenças salariais medianas em desvantagem das mulheres, não se pode afirmar estatisticamente estas diferenças, visto que segundo o teste apontou igualdade das medianas, provavelmente pelo baixo número de indivíduos em cada sexo para este grupo.

Esses resultados evidenciam uma semelhança com o contexto descrito na revisão bibliográfica referente aos trabalhadores em geral no país.

Com base nessa consideração, afirma-se que o debate sobre disparidade salarial é crucial. Os dados analisados, ao longo de uma década, indicam uma profunda diferença salarial em desfavor das mulheres no Brasil, ressaltando a importância de abordar e enfrentar essa questão no âmbito do debate público e das políticas de igualdade de gênero.

## 6 REFERÊNCIAS

- BARROS, R. P. de, RAMOS, L., SANTOS, E. Gender differences in Brazilian labor market. In: SCHULTZ, P. **Investments in women human capital**. Chicago University Press, 1995.
- BAUMAN, Z., MAY, T. **Aprendendo a pensar com a sociologia**. Rio de Janeiro: Zahar, 2010.
- BLACK et al. **A Dictionary of Economics**. 3 ed. Oxford University Press. 2009. ISBN 9780199237043.
- BRASSCOM (Brasil). **Estratégia TIC Brasil 2022**. 2012. Disponível em: [https://brasscom.org.br/wp-content/uploads/2017/08/estrategia\\_tic\\_brasil\\_2022\\_forum\\_nacional\\_portugues.pdf](https://brasscom.org.br/wp-content/uploads/2017/08/estrategia_tic_brasil_2022_forum_nacional_portugues.pdf) . Acesso em: nov. 2020.
- BRASSCOM (Brasil). **Relatório Setorial 2019**. São Paulo, Abril de 2020. Disponível em: <https://brasscom.org.br/wp-content/uploads/2020/04/P-2020-04-09-Coletiva-de-Imprensa-Relat%C3%B3rio-Setorial-2019-v15.pdf>. Acesso em: nov. 2020.
- CAMARGO, J. M., SERRANO, F. L. P. **Os dois mercados: homens e mulheres na indústria brasileira**. Rio de Janeiro: PUC/Departamento de Economia, Texto Para Discussão, 46, 1983.
- CAVALIERI, C., FERNANDES, R. Diferenciais de salários por gênero e cor: uma comparação entre as regiões metropolitanas brasileiras. **Revista de Economia Política**, v. 18, n. 1, 1998.
- COELHO, Allexandro Mori; CORSEUIL, Carlos Henrique. **Diferenciais Salariais no Brasil: Um Breve Panorama**, Rio de Janeiro: IPEA, Texto Para Discussão, n. 898, 2002. Disponível em: [https://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td\\_0898.pdf](https://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_0898.pdf). Acesso em: nov. 2020.
- DIEGUES, A. C.; ROSELINO, J. E. Uma proposta metodológica para a mensuração das atividades de software realizadas fora da indústria de software. **Revista Brasileira de Inovação**. V. 10, n. 2. UNICAMP, Campinas, jul/dez, 2011. Disponível em: <http://goo.gl/PE8dEX>. Acesso em: out. 2020.
- KASSOUF, A. L. The wage rate estimation using the Heckman procedure. **Revista de Econometria**, n. 1, 1994.
- LANGONI, C. G. **Distribuição da renda e desenvolvimento econômico do Brasil**. Rio de Janeiro: Expressão e Cultura, 1973. 100
- LEME, C., WAJNMAN, S. Tendências de corte nos diferenciais de rendimento por sexo. In: HENRIQUES, R. M. (org.). **Desigualdade e pobreza no Brasil**. Rio de Janeiro: IPEA, 2000.
- LEY, Michelle Bellas Romariz Gaudie; LIPPI, Marcelo; NASCENTES, Alexandre Lioi. Análise de Dados do SNIS Sobre Resíduos Sólidos Utilizando Linguagem Python. **Encontro de Engenharia Sanitária e Ambiental**, Porto, ed. 18, 2018. Disponível em: [https://www.researchgate.net/publication/339121413\\_ANALISE\\_DE\\_DADOS\\_DO\\_SNIS\\_SOBRE\\_RESIDUOS\\_SOLIDOS\\_UTILIZANDO\\_LINGUAGEM\\_PYTHON](https://www.researchgate.net/publication/339121413_ANALISE_DE_DADOS_DO_SNIS_SOBRE_RESIDUOS_SOLIDOS_UTILIZANDO_LINGUAGEM_PYTHON). Acesso em: mai. 2021.
- LIMA, Guilherme. **Python: A origem do nome**. Alura. Abr. 2021. Disponível em: <https://www.alura.com.br/artigos/python-origem-do-nome>. Acesso em: mai. 2021.
- NOLETO. **O que é meta charset="utf-8" no HTML e para que serve?** Disponível em: <https://blog.betrybe.com/tecnologia/meta-charset-utf/>. Acesso em: mai. 2022.

OMETTO, A. M. H., HOFFMANN, R., ALVES, M. C. A segregação por gênero no mercado de trabalho nos estados de São Paulo e Pernambuco. **Economia Aplicada**, v. 1, n. 3, 1997.

RAMSEY, Nancy; MCCORDUCK, Pamela. **Where are the Women in Information Technology?**. Boulder, Colorado, n. 0413538, Maio, 2005. Disponível em: [https://www.alejandrobarrros.com/wp-content/uploads/old\\_old/Where\\_are\\_the\\_Women\\_in\\_Information\\_Technology.pdf](https://www.alejandrobarrros.com/wp-content/uploads/old_old/Where_are_the_Women_in_Information_Technology.pdf). Acesso em: jan. 2021.

ROSA, Daniel. **Linguagens de programação interpretadas x compiladas: qual é a diferença?** Disponível em: <https://www.freecodecamp.org/portuguese/news/linguagens-de-programacao-int-erpretadas-x-comp-iladas-qual-e-a-diferenca/>. Acesso em: jun. 2022.

SILVEIRA, Evanildo da. **Como as mulheres passaram de maioria a raridade nos cursos de informática.** BBC Brasil, 13 de abr. 2018. Disponível em: <https://www.bbc.com/portuguese/geral-43592581>. Acesso em: jan. 2021.

VICENTAINER, Denis; MATTEDI, Marcos; MELLO, Bruno. Aplicação das Bibliotecas Python para tratamento de dados em tempo real: A análise dos dados de isolamento social em Santa Catarina. **Metodologias e Aprendizado**, Santa Catarina, ed. 3, 2020. Disponível em: <https://publicacoes.ifc.edu.br/index.php/metapre/article/view/1392/1061>. Acesso em: mai. 2021.

## 7 ELEMENTOS PÓS-TEXTUAIS

### 7.1 Apêndice A

Quadro 2 - Famílias e ocupações utilizadas.

FAMÍLIA	TÍTULO	OCUPAÇÃO	TÍTULO OCUPAÇÃO
1236	Diretores de tecnologia da informação	1236-05	Diretor de tecnologia da informação
1425	Gerentes de tecnologia da informação	1425-05	Gerente de infraestrutura de tecnologia da informação
		1425-10	Gerente de desenvolvimento de sistemas
		1425-15	Gerente de operação de tecnologia da informação
		1425-20	Gerente de projetos de tecnologia da informação
		1425-25	Gerente de segurança da informação
		1425-30	Gerente de suporte técnico de tecnologia da informação
		1425-35	Tecnólogo em gestão da tecnologia da informação
2122	Engenheiros em computação	2122-05	Engenheiro de aplicativos em computação
		2122-10	Engenheiro de equipamentos em computação
		2122-15	Engenheiros de sistemas operacionais em computação
2124	Analistas de tecnologia da informação	2124-05	Analista de desenvolvimento de sistemas
		2124-10	Analista de redes e de comunicação de dados
		2124-15	Analista de sistemas de automação
		2124-20	Analista de suporte computacional
		2124-25	Arquiteto de soluções de tecnologia da informação
		2124-30	Analista de testes de tecnologia da informação
3171	Técnicos de desenvolvimento de sistemas e aplicações	3171-05	Desenvolvedor web (técnico)
		3171-10	Desenvolvedor de sistemas de tecnologia da informação (técnico)

		3171-15	Programador de máquinas - ferramenta com comando numérico
		3171-20	Desenvolvedor de multimídia
<b>3172</b>	Técnicos de suporte e monitoração ao usuário de tecnologia da informação	3172-05	Operador de computador
		3172-10	Técnico de suporte ao usuário de tecnologia da informação
<b>2123</b>	Administradores de tecnologia da informação	2123-05	Administrador de banco de dados
		2123-10	Administrador de redes
		2123-15	Administrador de sistemas operacionais
		2123-20	Administrador em segurança da informação

Fonte: Classificação Brasileira de Ocupações 5.1.7 – Ministério da Economia, 2002