



De 11 a 14 - NOVEMBRO DE 2024

**Inteligência Artificial
na Gestão de Operações:
Limitações e possibilidades**



UMA ABORDAGEM DE BIG DATA PARA PREDIÇÃO DA EMPREGABILIDADE DE MULHERES NA ÁREA DA ENGENHARIA POR REGIÕES DO BRASIL

RAVENNA LINS RODRIGUES - ravenna.lins@ufpe.br
UNIVERSIDADE FEDERAL DE PERNAMBUCO - UFPE

CECIR BARBOSA DE ALMEIDA FARIAS - cecir.barbosa@professor.ufcg.edu.br
UNIVERSIDADE FEDERAL DE CAMPINA GRANDE - UFCG

JAIRO RODRIGUES DA SILVA - jaiorodriguesdasilva@hotmail.com
UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO - UFRPE

ÁREA: 6. ENGENHARIA ORGANIZACIONAL

SUBÁREA: 6.1 - GESTÃO ESTRATÉGICA E ORGANIZACIONAL

RESUMO: O PRESENTE ESTUDO BUSCOU ANALISAR AS DISPARIDADES DE GÊNERO E SALARIAIS NO MERCADO DE TRABALHO DAS ENGENHARIAS NO BRASIL. PARA ISSO, FORAM UTILIZADOS DADOS DA RELAÇÃO ANUAL DE INFORMAÇÕES SOCIAIS (RAIS) E APLICADA UMA ABORDAGEM DE REGRESSÃO LOGÍSTICA BINÁRIA. OS RESULTADOS MOSTRARAM QUE, MESMO APÓS O CONTROLE DE VARIÁVEIS COMO IDADE, RAÇA E REGIÃO, AS MULHERES APRESENTARAM UMA MENOR PROBABILIDADE DE RECEBER SALÁRIOS ACIMA DA MÉDIA. ISSO EVIDENCIA A EXISTÊNCIA DE BARREIRAS SIGNIFICATIVAS NO RECONHECIMENTO SALARIAL DAS MULHERES NO SETOR DE ENGENHARIA, EVIDENCIANDO A PERSISTÊNCIA DE DESIGUALDADES DE GÊNERO. DESSA FORMA, CONCLUI-SE QUE POLÍTICAS IGUALITÁRIAS SÃO IMPRESCINDÍVEIS PARA PROMOVER A EQUIDADE DE GÊNERO NO MERCADO DE TRABALHO E SUPERAR ESSAS DISPARIDADES, GARANTINDO UM AMBIENTE MAIS JUSTO E INCLUSIVO PARA TODOS OS PROFISSIONAIS.

PALAVRAS-CHAVES: MERCADO DE TRABALHO; EMPREGABILIDADE FEMININA; DISPARIDADE SALARIAL; ENGENHARIA.

1. INTRODUÇÃO

São inúmeros os fatores que influenciam a empregabilidade de um país ou região, a empregabilidade está relacionada com o perfil da oferta de trabalho e da mão de obra disponível, sendo assim, o conceito de empregabilidade é um indicador crucial da estabilidade econômica e social de um país. Entre os oito Objetivos de Desenvolvimento do Milênio (ODM) apresentados pela Organização das Nações Unidas (ONU) está promover a igualdade de gênero e empoderamento feminino (ONU, 2017).

A ONU defende a participação equitativa das mulheres em todos os aspectos da vida e entre suas áreas prioritárias, estão: aumentar a liderança e a participação das mulheres; aprimorar o empoderamento econômico das mulheres e colocar a igualdade de gênero no centro do planejamento e dos orçamentos de desenvolvimento nacional (FRANCO, 2018). Nesse contexto, a ONU estabeleceu padrões globais para alcançar a igualdade de gênero. Os Estados Membros devem formular leis, políticas públicas, programas e serviços necessários à implementação desses padrões e consequentemente alcançar o objetivo (MEDEIROS e PINHEIRO, 2018; FRANCO, 2018).

De acordo com Felix e da Silva (2018), o crescimento da participação feminina no mercado de trabalho vem sofrendo mudanças significativas. Em um recorte por gênero, os dados mostraram que, no ano de 2018, o nível de emprego da mão de obra feminina cresceu 3,91% em comparação com 2,57% dos homens.

Entretanto, a taxa de desocupação em 2022 evidenciou altas taxas de desemprego. A taxa das mulheres (16,8%) está acima da média nacional de 14,6%, enquanto a taxa dos homens (12,8%) está abaixo da média (IPEA, 2022). Nesse contexto, outro aspecto da desigualdade de gênero é a disparidade salarial: Em média, homens ganham 23% a mais do que as mulheres nos mesmos postos de trabalho (IBGE 2018; ONU 2017).

É importante salientar que há uma certa escassez em pesquisas exploratórias e preditivas referentes à empregabilidade. Tal fato ocorre devido esta área de pesquisa apresentar obstáculos para obtenção de dados adequados, confiáveis e atualizados em razão da diversidade de contexto em que estes se inserem (GARCÍA-PEÑALVO et al., 2018).

No entanto, em 2013, a Comissão de Estatística das Nações Unidas organizou o Conjunto Mínimo de Indicadores de Gênero (CMIG), com 63 indicadores (52 quantitativos e 11 qualitativos) para harmonizar estatísticas sobre igualdade de gênero e empoderamento feminino (Boletim Geográfico, 2018). No Brasil, o IBGE, por meio do Sistema Nacional de

Informações de Gênero (SNIG), divulgou os resultados de muitos desses indicadores com base na Pesquisa Nacional por Amostra de Domicílios (PNAD) (BOLETIM GEOGRÁFICO, 2018).

O termo Big Data é utilizado para representar essa enorme quantidade de dados gerados diariamente. A exploração adequada desse volume pode fornecer informações rápidas, confiáveis e oportunas, essenciais para a sobrevivência e crescimento sustentável das organizações (DE ESPÍNDOLA, 2016). Big Data permite não apenas entender o passado, mas também prever o futuro com maior detalhamento, destacando a necessidade de encontrar sentido em grandes volumes de informação (MINELI, CHAMBERS & DHIRAJ, 2013).

Nesse sentido, a mineração de dados apresenta-se como uma importante técnica de extração de conhecimento em banco de dados. Sendo assim, o presente estudo teve como objetivo identificar e analisar indicadores de empregabilidade feminina na área da engenharia por regiões do Brasil a partir da base de dados do RAIS.

2. DADOS ABERTOS GOVERNAMENTAIS: UMA ABORDAGEM DE *BIG DATA*

O governo federal vem passando por um processo de transformação digital. Segundo a Cartilha Técnica para Publicação de Dados Abertos no Brasil, disponibilizada pelo Portal Brasileiro de Dados Aberto (2014), desde 2009 o governo brasileiro vem realizando ações para o desenvolvimento de uma política de disseminação de dados e informações governamentais para o livre uso pela sociedade, alinhadas com um crescente movimento global para democratização do acesso a dados e informações no paradigma de dados abertos (BRASIL, 2012).

O programa de desenvolvimento e implementação de estratégias de governo digital da OECD (Organização para a Cooperação e Desenvolvimento Econômico) apresenta 12 princípios fundamentais para o êxito, dentre os quais está a criação de uma cultura orientada a dados no setor público (OECD, 2018). A OECD, aponta que os governos devem desenvolver uma cultura de análise e uso de dados públicos que ajudem a prever novas necessidades e tendências e entender como melhorar os processos e dinâmicas existentes (OECD, 2018).

No Brasil, isso tem sido reforçado através de políticas de dados abertos no âmbito do *e-government* (*e-gov*), ou governo eletrônico. Segundo definição da *Open Knowledge Internacional*, organização global sem fins lucrativos focada em fomentar a cultura de dados abertos, dados são abertos quando qualquer pessoa pode livremente acessá-los, utilizá-los, modificá-los e compartilhá-los para qualquer finalidade, estando sujeito a, no máximo,

exigências que visem preservar sua proveniência e sua abertura (OPEN KNOWLEDGE INTERNACIONAL, 2019).

Na prática, esses critérios têm sido satisfeitos pelas entidades governamentais através da criação de portais que disponibilizam conjuntos de dados estruturados (linhas de dados dispostos em colunas pré-determinadas) e organizados sob uma licença aberta (SOUTO, ARRUDA e ARAÚJO, 2019). Tal fato, forma uma expressiva massa de dados, a qual pode ser compreendida como um fenômeno de *Big Data*.

O uso do termo *Big Data* representa uma grande quantidade de dados de alta variedade, velocidade, valor e veracidade (WALKER, 2015; AKOKA, 2017). Entretanto, a designação “Big” não se limita apenas ao tamanho dos dados, mas também ao quão inteligente esses dados são (GUNTER et al., 2017). É legítimo assumir, também, que mais importante do que o tamanho ou a inteligência dos dados, é a forma como estes são utilizados (GUNTER et al., 2017).

Em 2016 foi publicada a Estratégia de Governança Digital (EGD) para o período de 2016 a 2019, alinhado com a Política de Governança Digital, instituída pelo Decreto nº 8.638, de 15 de janeiro de 2016 e com o Plano Plurianual (PPA) 2016-2019. Uma das iniciativas da EGD 2016-2019 (a IE.03.06) é promover cooperação nacional e internacional com setor produtivo e academia, visando contribuir para a promoção da consolidação do Governo Eletrônico (BRASIL, 2016).

2.1 Empregabilidade Feminina na Área da Engenharia no Brasil

Segundo Lombardi 2005, um dos setores profissionais que vem apresentando aumento da atividade feminina é a engenharia, que tem se tornado menos heterogênea, deixando de ser predominantemente ocupada pelo público masculino

É possível observar, também, a existência e prevalência de mulheres em áreas específicas da engenharia. Segundo as estatísticas do Conselho Federal de Engenharia e Agronomia (CONFEA, 2020), para os profissionais cadastrados com o título de “Engenheiro Civil”, totalizam-se 343.076 registros ativos, dos quais 269.627 (79%) são engenheiros e 73.449 (21%) são engenheiras. No estado da Paraíba, a quantidade de engenheiras civis registradas no Conselho Regional de Engenharia e Agronomia (Crea-PB) é um total de 2.499 mulheres, isto é, 13% dos profissionais inscritos (LIRA, 2018).

Com este quantitativo, é válido ressaltar que o registro como Engenheira Civil é o que mais concentra mulheres no Conselho Federal, seguido pelas engenheiras agrônomas

(20.252 profissionais), engenheiras de segurança do trabalho (13.256 profissionais) e engenheiras ambientais (10.739 profissionais) (CONFEA, 2020).

Outros obstáculos enfrentados pelo sexo feminino na área da engenharia é a questão salarial, visto que a remuneração é, em geral, inferior. De acordo com uma pesquisa realizada pelo IBGE, as mulheres receberam 77,7% do salário dos homens em 2019. Em cargos de maior rendimento, como diretores e gerentes, as mulheres ganharam apenas 61,9% do rendimento dos homens (IBGE, 2018).

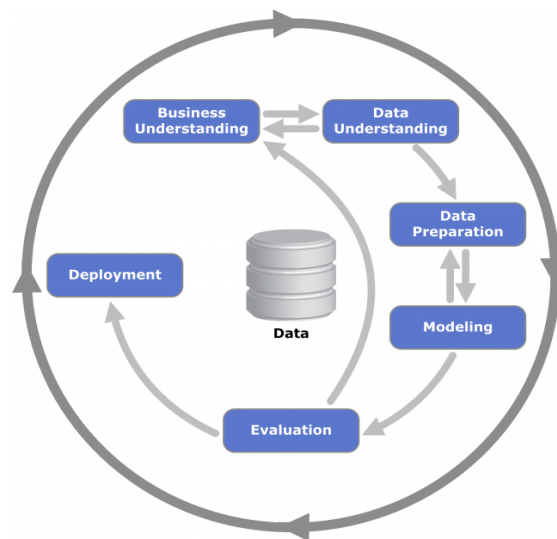
Mudanças no sentido da igualdade salarial e de direitos trarão benefícios para a sociedade como um todo, inclusive para engenheiras e engenheiros. Corrales (2016), aborda que a igualdade de gênero nas engenharias trará diversas vantagens ao Brasil, e quaisquer dificuldades ao longo deste caminho não devem ser vistas como impeditivas, mas sim como superáveis.

Ademais, a empregabilidade é um fator considerável na redução das desigualdade de gênero na estrutura econômica. Desse modo, o estudo do perfil da empregabilidade feminina pode auxiliar na redução dos índices de desempregos feminino e mitigação das disparidades salariais (Oliveira, Scorzafave e Pazello 2009).

3. PROCEDIMENTO METODOLOGICO

O método adotado para realização do estudo foi o *Crisp-DM (Cross Industry Standard Process for Data Mining)*, o qual se apresenta como um dos mais utilizados na área de Mineração de Dados devido fornecer uma abordagem estruturada para o planejamento de projetos.

Segundo Chapman (2000), a metodologia é composta por seis fases, conforme ilustrado na Figura 1, organizadas de maneira cíclica, cujo fluxo é unidirecional, possibilitando ir e voltar entre as suas fases e tarefas. São estas: Entendimento do Negócio (*Business Understanding*), Entendimento dos Dados (*Data Understanding*), Preparação dos Dados (*Data Preparation*), Modelagem (*Modeling*), Avaliação (*Evaluation*) e Aplicação (*Deployment*). A Figura 1, demonstra o ciclo de vida original do modelo CRISP-DM.

FIGURA 1. Etapas do método *CRISP-DM*

Fonte: IBM Knowledge Center (2012)

3.1 Entendimento Do Negócio (*Business Understanding*)

A primeira fase contou com a definição e familiarização da ferramenta de captação dos dados. Dessa forma, foi escolhida a plataforma “Base dos Dados” (<https://basedosdados.org/>), a qual conta com um grande *datalake* público por meio do BigQuery, além da capacidade de busca e filtragem catalogada das bases de dados e seus metadados.

Para esse estudo, foram filtrados dados dos anos 2017 até 2021, pertencentes à base da Relação Anual de Informações Sociais (RAIS). A partir disso, foi utilizado um script em SQL para extração dos dados a partir da Classificação Brasileira de Ocupações (CBO), acerca de 5 engenharias (Engenharia Civil, Engenharia Elétrica, Engenharia Mecânica, Engenharia de Produção e Engenharia Química).

3.2 Entendimento Dos Dados (*Data Understanding*)

Após a extração dos dados e formulação da base a ser trabalhada no estudo deu-se início à segunda fase que compreendeu o entendimento e familiarização das amostras. Foram selecionadas 17 variáveis e construída 1 referente às regiões do país, a qual integrou todos os estados às suas respectivas regiões. Para a análise estatística por meio da Regressão Logística Binária foram utilizadas 5 variáveis (cor, sexo, faixa_etaria, região e cbo_2002), já para a construção dos modelos preditivos da análise de classificação, foram utilizadas 7 variáveis

(cor, sexo, faixa_etaria, região, cbo_2002, quantidade_horas_contratadas e faixa_tempo_emprego).

3.3 Preparação Dos Dados (*Data Preparation*)

Posteriormente, com o entendimento dos dados, foi possível organizá-los, isto é, renomear as amostras com suas respectivas referências dispostas no dicionários do RAIS, e realizar uma nova filtragem, tendo em vista um melhor direcionamento da pesquisa. Portanto, foram aplicados filtros a duas variáveis, sendo estas "tipo_vínculo", a qual considerou apenas amostras com vínculos CLT, "raca_cor" que extraiu amostras referentes às cores (branco, preto e pardo).

Em seguida deu-se início ao tratamento dos dados o qual buscou reduzir ou eliminar anomalias e/ou discrepâncias que poderiam comprometer a análise como um todo. Entretanto, não houve modificações expressivas visto que foi realizado o *download* da base tratada, disponível na plataforma Base dos Dados, citada anteriormente. Porém, como o novo filtro foi aplicado nessa fase da pesquisa, realizou-se um novo tratamento.

3.4 Modelagem (*Modeling*)

Nessa fase foi desenvolvido um modelo de regressão logística binária para analisar a influência de diferentes variáveis, como gênero, faixa etária, e região geográfica, na remuneração dos profissionais das engenharias. Utilizando a ferramenta *Jamovi*, que é um software estatístico de código aberto e interface gráfica, os dados foram preparados e os modelos foram construídos com base em critérios estatísticos rigorosos, visando identificar padrões e tendências significativas.

Assim, inicialmente a coluna da variável "valor_remuneracao_media_sm" foi duplicada e binarizada por meio de um filtro no próprio *software*. Esta foi identificada como "remuneracao_bin". e definida como a variável dependente do modelo. Nesse sentido, a variável binarizada classificou as amostras como "acima" ou "abaixo" da média salarial a depender do valor médio salarial, isto é, amostras com remuneração inferior a 8.47 salários mínimos foram classificadas como "abaixo" enquanto amostras com remunerações superiores a 8.47 salários mínimos foram classificadas como "acima".

Em seguida, foram selecionados os fatores e definidos os níveis de referência para as variáveis preditoras, tal como: remuneração = acima da média, cor = branca, região = sudeste,

faixa_etaria = 5 (30 a 39 anos), cbo_2002 = engenharia civil e sexo = feminino. Por fim, foi determinado um intervalo de confiança de 95%.

Posteriormente, foi gerado um filtro para a o quarto quartil da variável binarizada (remuneracao_bin) e um segundo modelo de regressão logística binária foi construído a partir do limiar inicial, isto é da média salarial (10.01 salários mínimos).

3.5 Avaliação (*Evaluation*)

Essa etapa buscou verificar quão bem o modelo de regressão logística binária foi capaz de prever a remuneração acima da média com base nos dados fornecidos. Isso envolveu a análise de métricas específicas, como a significância estatística dos coeficientes, o intervalo de confiança e a interpretação das odds ratios, para avaliar a eficácia do modelo em identificar os fatores que influenciam a remuneração.

3.6 Aplicação (*Deployment*)

Por fim, a validação do modelo de regressão logística binária ocorreu por meio da corroboração dos resultados com a literatura existente sobre disparidades de gênero e salariais no mercado de trabalho das engenharias. Isso ajudou a confirmar a consistência e a relevância dos resultados obtidos no estudo.

4. RESULTADO E DISCUSSÃO

O modelo de regressão logística binária, apresentado na Tabela 1, evidenciou a princípio que indivíduos do sexo feminino apresentaram uma perspectiva salarial inferior ao dos indivíduos do sexo masculino com um $odds = 0.793$ (95% IC = 0.764, 0.823, $p < 0.001$), isto é, aproximadamente 0,20 menos chances de receber uma remuneração acima do valor salarial médio (8.47 salários mínimos).

TABELA 1. Regressão logística binária a partir da variável dependente “remuneracao_bin”

Preditor	Estimativas	Erro Padrão	Z	p	Rácio de Chances	Intervalo de Confiança a 95%	
						Lim. Inferior	Superior
Intercepto	0.0730	0.0212	3.45	< .001	1.076	1.032	1.121
Cor:							
Amarelo – Branco	-0.2243	0.0659	-3.40	< .001	0.799	0.702	0.909
Pardo – Branco	0.0939	0.0164	5.73	< .001	1.098	1.064	1.134
Preto – Branco	0.2796	0.0471	5.94	< .001	1.323	1.206	1.450
Região:							
Centro Oeste – Sudeste	0.1693	0.0281	6.03	< .001	1.185	1.121	1.252
Nordeste – Sudeste	0.2425	0.0215	11.27	< .001	1.274	1.222	1.329
Norte – Sudeste	0.4724	0.0349	13.52	< .001	1.604	1.498	1.718
Sul – Sudeste	0.5629	0.0226	24.94	< .001	1.756	1.680	1.835
Faixa Etária:							
3 – 5	2.5742	0.0522	49.28	< .001	13.121	11.844	14.536
4 – 5	1.1303	0.0187	60.35	< .001	3.097	2.985	3.212
6 – 5	-0.3525	0.0196	-17.97	< .001	0.703	0.676	0.731
7 – 5	-0.3298	0.0245	-13.44	< .001	0.719	0.685	0.754
8 – 5	-0.1032	0.0581	-1.78	0.076	0.902	0.805	1.011
Cbo_2002:							
Eng Elétrica – Eng Civil	-0.4573	0.0226	-20.19	< .001	0.633	0.606	0.662
Eng Mecânica – Eng Civil	-0.1871	0.0230	-8.15	< .001	0.829	0.793	0.868
Eng Produção – Eng Civil	-0.3767	0.0200	-18.86	< .001	0.686	0.660	0.714
Eng Química – Eng Civil	-0.5195	0.0459	-11.31	< .001	0.595	0.544	0.651
Sexo:							
Masculino – Feminino	-0.2321	0.0190	-12.20	< .001	0.793	0.764	0.823

Fonte: Autoria Própria (2023)

A respeito da cor, apesar dos indivíduos de cor amarela terem menor participação (1%) no mercado de trabalho, conforme demonstrado no apêndice 1, estes apresentam mais chances de ganhar acima do valor salarial médio em comparação aos indivíduos da cor branca que evidenciaram uma razão de chances de 0.799 (95% IC = 0.702, 0.909, $p < 0.001$). Em contrapartida, indivíduos da cor branca possuem maiores chances em comparação à

empregados pardos odds = 1.098 (95% IC = 1.064, 1.134, $p < 0.001$) e principalmente pretos odds = 1.323 (95% IC = 1.206, 1.450, $p < 0.001$).

Em relação às regiões do país, destaca-se que a região Sudeste é conhecida por possuir o maior polo industrial do Brasil. Está concentra 60% (Gráfico 4) dos empregados do *dataset* em estudo e apresenta valores salariais superiores às demais regiões do país, conforme exposto no apêndice 2. Além disso, a região Sudeste evidenciou uma razão de chances de 1.756 (95% IC = 1.680, 1.835, $p < 0.001$) e 1.604 (95% IC = 1.498, 1.718, $p < 0.001$) em comparação às regiões Sul e Norte. É válido ressaltar que a região Sul apresenta a maior disparidade salarial dos país, com as mulheres recebendo um salário médio 24,18% inferior ao dos homens.

Os preditores mais significativos foram referentes a faixa etária 5 (30 a 39 anos) com um odds = 13.121 (95% IC = 11.844, 14.536, $p < 0.001$) em relação à faixa etária 3 (18 a 24 anos) e um odds = 3.097 (95% IC = 2.985, 3.212, $p < 0.001$) em relação à faixa etária 4 (25 a 29 anos). Isso indica que indivíduos nesta faixa etária (30 a 39 anos) são cerca de 13.121 e 3.097 vezes mais chances de obterem salários superiores à média geral (8.47 salários) em relação a sujeitos pertencentes às faixas etárias 3 e 4. As demais faixas etárias (40 a 49 anos) e 7 (50 a 64 anos) apresentaram uma perspectiva de salários superiores a faixa etária 5. Tal fato, possivelmente está relacionado à experiência profissional, tempo de trabalho, dentre outros. Ademais, destaca-se que a faixa etária 8 (65 anos ou mais) apresentou não apresentou correlação estatisticamente significativa ($p < 0.076$).

No que se refere às áreas da engenharia, verificou-se que apesar da engenharia civil deter o maior percentual de empregados (63%), esta apresenta uma perspectiva salarial inferior às demais engenharias, especialmente em relação às engenharias química e elétrica com aproximadamente 0,40 (odds = 0.595, 95% IC = 0.544, 0.651, $p < 0.001$) e 0,37 (odds = 0.633, 95% IC = 0.606, 0.662, $p < 0.001$) menos chances.

Tendo em vista identificar a existência de diferenças significativas por meio dos níveis de referência das variáveis preditoras, foi construído um modelo de regressão logística binária a partir do limiar inicial do quarto quartil (10.01 salários mínimos) da variável binarizada (remuneracao_bin), representado na Tabela 2.

TABELA 2. Regressão logística binária do quarto quartil da variável dependente

Preditor	Estimativas	Erro Padrão	Z	p	Rácio de Chances	Intervalo de Confiança a 95%	
						Lim. Inferior	Superior
Intercepto	1.31012	0.0257	50.929	< .001	3.707	3.524	3.898
Cor:							
Amarelo – Branco	-0.34579	0.0691	-5.006	< .001	0.708	0.618	0.810
Pardo – Branco	0.23399	0.0188	12.450	< .001	1.264	1.218	1.311
Preto – Branco	0.35921	0.0548	6.554	< .001	1.432	1.286	1.595
Região:							
Centro Oeste – Sudeste	-0.00403	0.0321	-0.126	0.900	0.996	0.935	1.061
Nordeste – Sudeste	0.27371	0.0254	10.766	< .001	1.315	1.251	1.382
Norte – Sudeste	0.30933	0.0411	7.518	< .001	1.363	1.257	1.477
Faixa Etária:							
3 – 5	3.29257	0.1141	28.867	< .001	26.912	21.521	33.654
4 – 5	1.40406	0.0264	53.202	< .001	4.072	3.866	4.288
6 – 5	-0.46754	0.0204	-22.912	< .001	0.627	0.602	0.652
7 – 5	-0.55469	0.0251	-22.069	< .001	0.574	0.547	0.603
8 – 5	-0.35106	0.0610	-5.753	< .001	0.704	0.625	0.793
Cbo_2002:							
Eng Elétrica – Eng Civil	-0.57378	0.0244	-23.483	< .001	0.563	0.537	0.591
Eng Mecânica – Eng Civil	-0.25676	0.0258	-9.970	< .001	0.774	0.735	0.814
Eng Produção – Eng Civil	-0.66576	0.0222	-30.036	0.001	0.514	0.492	0.537
Eng Química – Eng Civil	-1.03484	0.0479	-21.588	< .001	0.355	0.323	0.390
Sexo:							
Masculino – Feminino	-0.39701	0.0236	-16.806	< .001	0.672	0.642	0.704

Fonte: Autoria Própria (2023)

Desse modo, verificou-se que as chances de sujeitos do sexo feminino ganharem acima do valor salarial médio (10.01) em relação a indivíduos do sexo masculino é ainda menor (odds = 0.672, 95% IC = 0.642, 0.704, $p < 0.001$) no que se refere às maiores remunerações do país, ou seja, aproximadamente 0,33 menos chances.

Quanto às regiões do país, não houve correlação significativa entre as chances das regiões Sudeste e Centro-Oeste ($odds = 0.996$, 95% IC = 0.935, 1.061, $p < 0.900$). Entretanto, a região Sudeste apresentou maiores chances em comparação às demais regiões.

Conforme esperado, constatou-se uma razão de chances ainda maior da faixa etária 5 (30 a 39 anos) em relação a 3 (18 a 24 anos) e 4 (25 a 29 anos) na modelagem realizada a partir do quarto quartil, visto que tais amostras reportam os maiores salários. Desse modo, observou-se que indivíduos pertencentes a faixa etária 5 (30 a 39 anos) apresentam $odds = 26.912$ (95% IC = 21.521, 33.654, $p < 0.001$) e $odds = 4.072$ (95% IC = 3.866, 4.288, $p < 0.001$) mais chances, de obter remunerações acima do valor médio do quarto quartil, se comparado aos indivíduos do grupo 3 (18 a 24 anos). Em contrapartida, os sujeitos da faixa etária 5 apresentam menores chances em comparação às demais faixas etárias (6,7 e 8).

Acerca das áreas da engenharia, nota-se que a engenharia civil apresenta menos chances em comparação com as demais, especialmente em relação à engenharia química com um $odds = 0.355$ (95% IC = 0.323, 0.390, $p < 0.001$), isto é, 0.65 menos chances de receber uma remuneração superior.

Ademais destaca-se que os indicadores de empregabilidade feminina nas áreas das engenharias, analisados neste estudo, incluíram: (1) Representatividade de gênero nas especialidades de engenharia; (2) Diferença salarial de gênero; (3) Impacto da raça na empregabilidade e na diferença salarial; (4) Variações salariais por região geográfica; (5) Influência da faixa etária na remuneração. Esses indicadores foram utilizados para compreender e quantificar as disparidades de gênero e raciais no mercado de trabalho das engenharias.

5. CONCLUSÃO

O estudo abordou as disparidades de gênero e salariais no mercado de trabalho das engenharias no Brasil, concentrando-se nas especialidades de Engenharia Civil, Engenharia Mecânica, Engenharia Elétrica, Engenharia de Produção e Engenharia Química. Utilizando dados da Relação Anual de Informações Sociais (RAIS) dos anos 2017 a 2021 e uma abordagem metodológica abrangente seguindo o modelo CRISP-DM, a análise destacou a sub-representação das mulheres, que ocupam apenas 20% das posições, apesar de comporem 51,8% da população brasileira. Além disso, as mulheres enfrentam uma diferença salarial média de 15% em relação aos homens, mesmo desempenhando funções similares com cargas horárias semelhantes.

A análise estatística, utilizando a regressão logística binária, ratificou essas constatações, indicando que as mulheres têm menor chance de receber salários acima da média em comparação aos homens, mesmo após o controle de fatores como idade, raça, região e especialização em engenharia. A odds ratio de 0.793 (95% IC = 0.764, 0.823) para mulheres receberem salários acima da média confirma a existência de barreiras significativas no reconhecimento salarial.

No que tange às disparidades raciais, os resultados revelaram que indivíduos de cor amarela, embora representem apenas 1% do mercado, têm mais chances de ganhar acima do valor salarial médio em comparação aos indivíduos brancos. Indivíduos pardos e pretos enfrentam maiores desafios, com odds ratios de 1.098 (95% IC = 1.064, 1.134) e 1.323 (95% IC = 1.206, 1.450), respectivamente, em comparação aos brancos.

O estudo limitou-se a cinco especialidades de engenharia devido a restrições computacionais. Assim, para trabalhos futuros, sugere-se a exploração de outras especialidades de engenharia para obter uma visão mais abrangente do mercado de trabalho. Além disso, a inclusão de variáveis adicionais, como nível de escolaridade, tempo de experiência, tipo de vínculo empregatício e outras características demográficas, poderá enriquecer as análises e oferecer insights mais detalhados sobre os fatores que influenciam a empregabilidade e as disparidades salariais no setor de engenharia.

REFERÊNCIAS

AKOKA, Jacky; COMYN-WATTIAU, Isabelle; LAOUFI, Nabil. **Pesquisa sobre Big Data - Um estudo sistemático de mapeamento**. Padrões e interfaces de computador , v. 54, p. 105-115, 2017.

BOLETIM GEOGRÁFICO. **Estatísticas de gênero indicadores sociais das mulheres no Brasil**. IBGE, 2018. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101551_informativo.pdf>. Acesso em: 19 de Maio de 2024.

BRASIL. Decreto nº 8.638, de 15 de janeiro de 2016. **Institui a Política de Governança Digital**. Brasília. 2016. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2016/Decreto/D8638.htm. Acesso em: 18 de Maio de 2024.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Logística e Tecnologia da Informação. **Cartilha Técnica para Publicação de Dados Abertos no Brasil**. v. 1.0. Brasília, 2012. Disponível em: <http://dados.gov.br/pagina/cartilha-publicacao-dadosabertos>. Acesso em: 18 de Maio de 2019.

CONSELHO FEDERAL DE ENGENHARIA E AGRONOMIA (CONFEA). **Estatísticas**. 2020. Disponível em: <http://estatistica.confea.org.br:8080/EstatisticaSic/ModEstatistica/Pesquisa.jsp?vw=TitCadGenero>. Acesso em: 22 de Maio de 2024.

CORRALES, Bruna Rossi. **Igualdade De Gênero Na Engenharia: Desafios E Benefícios**. 2016. 35 f. TCC (Graduação) - Curso de Economia, Instituto de Economia, Universidade Estadual de Campinas, Campinas, 2016.

DE ESPÍNDOLA, André Mauro Santos et al. **Big Data e Inteligência Estratégica: Um Estudo de Caso Sobre a Mineração de Dados como Alternativa de Análise**. Revista ESPACIOS| Vol. 37 (Nº 04) Año 2016.

FELIX, Carla Roberta Marques; DA SILVA, Sidney Vergínio. **A presença das mulheres e sua representatividade em cargos de liderança**. Mythos (Interdisciplinary), 2019.

FRANCO, S. A. **Mulheres que abandonam a carreira profissional: Uma análise da ocorrência do fenômeno opt-out entre brasileiras**. Universidade de Brasília - UnB, 2018.

GARCÍA-PEÑALVO, F. J.; CRUZ-BENITO, J.; MARTÍN-GONZÁLEZ, M.; VÁZQUEZ-INGELMO, A.; SÁNCHEZ-PRIETO, J. C.; THERÓN, R. **Proposing a machine learning approach to analyze and predict employment and its factors**. International Journal of Interactive Multimedia and Artificial Intelligence, v. 5, n. 7, p. 39–45, 2018.

GUNTER, W., MEHRIZI, M., HUYSMAN, M., & FELDBERG, F. (2017). **Debating big data: A literature review on realizing value from big data**. Journal of Strategic Information Systems 26 (3), 191-209.

IBGE. Estatísticas de gênero: indicadores sociais das mulheres no Brasil. n. 38, IBGE, **de População e Indicadores Sociais: Rio de Janeiro**, 2018. Disponível em:

<https://biblioteca.ibge.gov.br/index.php/bibliotecacatalogo?view=detalhes&id=2101551>.
Acesso em: 22 Maio 2024.

IPEA – INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **Desemprego já é menor que o observado antes da pandemia em diversos segmentos**. Brasília: Ipea, 2022. Disponível em:
<https://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=39076>
Acesso em: 25 Maio 2024.

LIRA, M. **Cada Vez Mais Mulheres Ocupam Os Canteiros de Obras**. A União - Jornal, Editora E Gráfica. 25 de março de 2018. Disponível em:
<auniao.pb.gov.br/noticias/caderno_paraiba/cada-vez-mais-mulheres-ocupam-os-canteiros-de-obras> Acessado em: 13 de Abril de 2024.

LOMBARDI, Maria Rosa. **Perseverança e resistência: a Engenharia como profissão feminina**. 2005. 292 f. Tese (Doutorado em Educação) – Universidade Estadual de Campinas, Faculdade de Educação. Campinas, 2005.

MEDEIROS, M.; PINHEIRO, L. S. **Desigualdades de gênero em tempo de trabalho pago e não pago no Brasil**. Sociedade e Estado, v. 33, p. 159 – 185, 04 2018.

MINELLI, M.; CHAMBERS, M.; DHIRAJ, A. **Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses**. Wiley CIO Series. 2013.

OECD, **Digital Government Toolkit: 12 principles**. 2018. Disponível em:
<http://www.oecd.org/governance/digitalgovernment/toolkit/12principles/>. Acesso em: 20 de Maio de 2024.

OLIVEIRA, P. R. de; SCORZAFAVE, L. G.; PAZELLO, E. T. **Desemprego e inatividade nas metrópoles brasileiras: as diferenças entre homens e mulheres**. Nova Economia, v. 19, p. 291 – 324, 09 2009.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS, ONU, 2017. **Glossário de termos do objetivo de desenvolvimento sustentável 5: Alcançar a igualdade de gênero e empoderar todas as mulheres e meninas**. Organização da Unidas Unidas para as Mulheres- ONU Mulher, 2017. Acesso em: 18 de Maio de 2024. Disponível em:<<http://www.onumulheres.org.br/onu-mulheres/documentos-de-referencia/#>> .

OPEN KNOWLEDGE INTERNACIONAL. **What is Open?**, 2019. Disponível em:
<https://okfn.org/about/>. Acesso em: 14 de Maio de 2024.

SOUTO, Hugo Medeiros; ARRUDA, Eduardo Martins de; ARAÚJO, Wagner Junqueira de. **Mineração de dados no contexto dos pregões eletrônicos**. 2019.

WALKER, R. **From Big Data to big profits: success with data and analytics**. New York: Oxford University Press, 2015.