



UNIVERSIDADE DE BRASÍLIA (UnB)
FACULDADE DE ADMINISTRAÇÃO, CONTABILIDADE, ECONOMIA E GESTÃO
DEPARTAMENTO DE ECONOMIA

As Desigualdades no Rendimento Domiciliar per Capita do DF

CAIO ROSAL DE ÁVILA

ORIENTADOR: ROBERTO DE GÓES ELLERY

JANEIRO DE 2023

As Desigualdades no Rendimento Domiciliar per Capita do DF

Monografia apresentada ao Departamento de Economia da Universidade de Brasília, como requisito parcial à obtenção do grau de bacharelado em Ciências Econômicas.

Orientador: Dr. Roberto de Góes Ellery Júnior

Brasília

2023

LISTA DE ABREVIATURAS E SIGLAS

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
BPC	Benefício de Prestação Continuada
DF	Distrito Federal
IBGE	Instituto Brasileiro de Geografia e Estatística
IDH	Índice de Desenvolvimento Humano
POF	Pesquisa de Orçamentos Familiares

As Desigualdades no Rendimento Domiciliar per Capita do DF

Resumo

Este estudo tem como objetivo analisar as desigualdades existentes na distribuição da renda domiciliar per capita do DF [Distrito Federal], uma Unidade da Federação que inclui a capital, Brasília, conhecida por ter o maior produto interno bruto per capita do país e um alto nível de IDH segundo dados do IBGE (2021). No entanto, apesar desse contexto aparentemente positivo, este trabalho revelou a presença de disparidades significativas na distribuição de renda entre os domicílios do DF, abrangendo várias formas de desigualdade. Essas desigualdades se manifestam de diversas maneiras: racial, espacial, de gênero, de tipo familiar e de acesso a qualificação. Nesse contexto, o objetivo geral deste trabalho é fornecer uma análise detalhada e minuciosa dessas variáveis sociais no Distrito Federal, além de criar uma ferramenta de aprendizado de máquina que possa ser utilizada para auxiliar no desenvolvimento e avaliação de políticas públicas voltadas para a redução dessas disparidades. Para atingir esse propósito, foi desenvolvido um modelo de regressão multinível que se mostrou estatisticamente significativo a um nível de confiança de 95%. Esse modelo permite uma compreensão mais profunda das relações entre as variáveis sociais mencionadas e a distribuição da renda domiciliar per capita, o que pode ser valioso para a formulação de estratégias eficazes de intervenção e melhoria das condições de vida das populações afetadas por essas desigualdades no Distrito Federal.

Palavras-chave: distribuição de renda; racial; gênero; tipo familiar; acesso à qualificação

The Inequalities in per Capita Household Income in the Federal District.

Abstract

This study aims to analyze the existing inequalities in the distribution of per capita household income in the DF [Distrito Federal], a Federation Unit that includes the capital, Brasília, known for having the highest gross domestic product per capita in the country and a high level of HDI according to data from IBGE (2021). However, despite this apparently positive context, this work revealed the presence of significant disparities in the distribution of income among households in the Federal District, encompassing various forms of inequality. These inequalities manifest themselves in different ways: racial, spatial, gender, family type and access to qualifications. In this context, the general objective of this work is to provide a detailed and thorough analysis of these social variables in the Distrito Federal, in addition to creating a machine learning tool that can be used to assist in the development and evaluation of public policies aimed at reducing these disparities. To achieve this purpose, a multilevel regression model was developed and proved to be statistically significant at a 95% confidence level. This model allows for a deeper understanding of the relationships between the mentioned social variables and the distribution of per capita household income, which can be valuable for formulating effective intervention strategies and improving the living conditions of populations affected by these inequalities in the Distrito Federal.

Keywords: income distribution; racial; gender; family type; access to qualification

Introdução

Segundo Menezes et al. (2020), o mundo tem observado taxas crescentes no crescimento do produto nos mais diversos países do mundo. Entretanto, concomitantemente, vêm se observando o crescimento da desigualdade. Portanto esse crescimento não está sendo distribuído de forma equitativa entre os países e entre os indivíduos. Nesse contexto, o Brasil destaca-se como um dos países mais desiguais do mundo.

Relativamente ao Brasil, Barros et al. (2006) apontam para uma diminuição da desigualdade nos anos anteriores ao estudo e essa pode ser atribuída a uma série de fatores, dentro os quais, a expansão de programas de transferência de renda, como “Bolsa Família” ou BPC [Benefício de Prestação Continuada], o aumento do salário-mínimo, a melhora na distribuição de renda dos trabalhos e o aumento do nível de educação.

Posteriormente, entre 2006 e 2012, segundo Medeiros e Souza (2016), os indicadores de desigualdade e de resultados econômicos representaram uma evolução no combate à pobreza e desigualdade. Todavia, apesar dessa melhora, esse contraste econômico e social no país ainda representa um grande obstáculo na melhoria de qualidade de vida da população. Dessa forma, é necessário, além de uma política de continuidade, que haja aprimoramento a respeito da temática.

No entanto, Hoffmann (2020) afirma que, a partir do ano de 2016 houve uma inversão do quadro de melhora descrito anteriormente, causada majoritariamente pela crise de 2014 e por políticas de austeridade, que culminaram no aumento do desemprego.

Posteriormente, Quinzani (2020) aponta que a desigualdade foi mais ainda intensificada pela crise do coronavírus e suas consequências, mostrando a fragilidade das populações mais pobres, as quais tendem a sofrer maior impacto mediante crises políticas e econômicas.

Sendo este contexto histórico observado em todo Brasil, há ainda peculiaridades no âmbito regional de desigualdades. Barros (2011) afirma que diversos fatores contribuem para intensificar mais ainda o processo de desigualdade entre regiões, como a precariedade em infraestrutura, baixo investimento em educação, falta de estímulo ao mercado privado e enfraquecimento de instituições governamentais.

Outrossim, o DF [Distrito Federal], segundo dados do IBGE [Instituto Brasileiro de Geografia e Estatística] (2020), apesar de ter o maior rendimento per capita do Brasil, enquanto unidade da federação, apresenta o maior coeficiente de Gini do país em 2016. Dessa maneira, evidencia, dentro de seu espaço geográfico, as maiores disparidades de renda do país.

Ainda sobre o DF, Oliveira et al. (2016) destacam a grande diferença de indicadores sociais entre as regiões administrativas do DF, refletidas em diversos setores da economia das Regiões Administrativas - RA's, seja no analfabetismo, no acesso a saúde entre outros.

Martins (2019), afirma que a pobreza e a desigualdade estão associadas a diversos fatores, como: idade, gênero, função social, cultura, localização geográfica, infraestrutura entre outros. Para tanto, a desigualdade demonstra-se em diversos níveis de exclusão que acabam por perpetuar essa condição.

Este trabalho propôs a criação de modelos de regressão para estimar a renda de indivíduos no Distrito Federal, com o propósito de desenvolver uma ferramenta útil no contexto das políticas públicas. O objetivo é melhorar a compreensão da distribuição de renda na sociedade. Ao analisar as estimativas com base em diferentes características dos indivíduos, é possível identificar grupos ou áreas em situação de vulnerabilidade e pobreza. Isso facilita uma alocação mais eficiente dos recursos em programas sociais, como o bolsa-família.

Além disso, o modelo pode ser usado para avaliar o impacto das políticas em vigor, fornecendo dados essenciais para melhorar o bem-estar e a qualidade de vida dos cidadãos em diferentes estratos do Distrito Federal. Outro ponto importante é a aplicação do modelo na formulação de políticas de desenvolvimento econômico de combate a desigualdade. Com base nas estimativas de renda, os formuladores de políticas públicas podem tomar decisões estratégicas, direcionando investimentos em setores cruciais como infraestrutura, educação e saúde, com o objetivo de promover o crescimento econômico sustentável e mais equânime de forma a reduzir as disparidades de renda em diferentes grupos e estratos do DF.

Em resumo, este trabalho se destaca como uma ferramenta fundamental para a concepção, execução e avaliação de políticas públicas. As previsões feitas pelo modelo de aprendizado de máquina desenvolvido podem embasar as ações governamentais, tornando-as mais informadas e alinhadas com o objetivo de melhorar o bem-estar e a qualidade de vida da população de maneira abrangente, equitativa e eficaz.

Material e Métodos

Variáveis do Modelo

Inicialmente foi feita uma modelagem de regressão linear múltipla apenas para título de comparação com as mesmas variáveis e dados a seguir. Foi utilizada a base de dados do IBGE da última POF [Pesquisa de Orçamentos Familiares] de 2017-2018 a fim de se realizar uma análise de regressão multinível com as seguintes variáveis, dentre as quais formam assinaladas em vermelho as variáveis Dummies de referência que serão ocultadas no procedimento N-1 de Dummização e serviram para parâmetros de comparação

posteriormente entre as variáveis do modelo. A Tabela 1 abaixo demonstra essas informações:

Tabela 1 – Sumário variáveis dos modelos		(continua)
Variáveis	Descrição das variáveis	Expectativa
PC_RENDA_DISP = “RENDA” ao longo do trabalho	Variável Dependente APÓS procedimento de Box_Cox. Renda disponível familiar per capita. É a divisão do total da renda disponível da Unidade de Consumo pelo total de moradores da Unidade de Consumo. Considerou-se como renda disponível a soma dos rendimentos monetários e não monetários menos impostos diretos, contribuições sociais, e outras deduções compulsórias ou quase compulsórias. Moradores com condição na família “empregado doméstico” e “parente de empregado doméstico” são excluídos de todas as etapas de construção da variável.	Variável dependente do modelo APÓS procedimento de Box_Cox
PC_RENDA_DISP.1	Variável Dependente ANTES do procedimento de Box_Cox. Renda disponível familiar per capita. É a divisão do total da renda disponível da Unidade de Consumo pelo total de moradores da Unidade de Consumo. Considerou-se como renda disponível a soma dos rendimentos monetários e não monetários menos impostos diretos, contribuições sociais, e outras deduções compulsórias ou quase compulsórias. Moradores com condição na família “empregado doméstico” e “parente de empregado doméstico” são excluídos de todas as etapas de construção da variável.	Variável dependente do modelo ANTES do procedimento de BoxCox
Tabela 1 – Sumário variáveis dos modelos		(continua)
Variáveis	Descrição das variáveis	Expectativa

		Espera-se correlacionar positivamente com a variável dependente devido ao fato de características intrínsecas como experiência de vida, sabedoria e experiência profissional
IDADE	Idade (em anos)	
		Espera-se que haja uma forte correlação positiva com os anos de estudos e a variável dependente
ANOS_ESTUDO	Duração de Estudo (em anos)	
	Cor ou Raça é uma variável Dummy dividida em:	Aguarda-se que haja diferenças entre as raças abordadas devido a divergências culturais, históricas e sociais entre elas.
	Cor Branca	
	Cor Preta	
	Cor Parda	
	Cor Amarela	
	Cor Indígena	
	Cor NA (nenhuma alternativa)	
	Sexo é uma variável Dummy dividida em:	É esperado que exista uma diferença positiva entre Homens e Mulheres
	Sexo Feminino	
	Sexo Masculino	
SEXO		

Tabela 1 – Sumário variáveis dos modelos (conclusão)

Variáveis	Descrição das variáveis	Expectativa
Composição	Composição familiar da respectiva Unidade de Consumo da pessoa. Variável derivada, construída a partir do quesito idade, necessária para a produção das tabelas da publicação de perfil das despesas.	Espera-se que a variável composição apresente um coeficiente Beta maior que as demais.
	Considere como criança - pessoas até 14 anos, adultos - pessoas entre 15 e 64 anos e idosos – pessoas com mais de 65 anos.	
	Moradores com condição na família “empregado doméstico” e “parente de empregado doméstico” são excluídos de todas as etapas de construção da variável.	
	COMPOSICAO é uma variável dummy dividida em:	
	1 – Um adulto sem criança	
	2 – Um adulto com ao menos uma criança	
	3 – Mais de um adulto sem criança	
	4 – Mais de um adulto com ao menos uma criança	
	5 – Um ou mais idosos com ou sem crianças	
	6 – Um ou mais idosos, com ao menos um adulto, com ou sem crianças	

Fonte: Dados Originais da pesquisa POF (2018)

O nível será constituído pela variável “Estratos” do DF (indica divisões geográficas dentro do território conforme o Censo de 2010 do IBGE) cujas variáveis estão demonstradas na Tabela 2 abaixo:

Tabela 2 – Sumário dos estratos

Nível	Estratos do DF divididos em:	Expectativa
	5301, 5302, 5303, 5304, 5305 e 5306 – Área Urbana	

Fonte: Dados Originais da pesquisa POF (2018)

Formulação dos Modelos

A formulação dos modelos foi representada com base nas informações contidas nas tabelas 1 e 2.

Modelo de Regressão Linear Múltipla

PC_RENDA_DISP

$$\begin{aligned}
 &= \beta_0 + \beta_2 \times IDADE + \beta_3 \times ANOS_ESTUDO \\
 &+ \beta_4 \times TIPO_SITUACAO_REGRural \\
 &+ \beta_5 \times TIPO_SITUACAO_REGUrbano \\
 &+ \beta_6 \times CORAmarela + \beta_7 \times CORBranca \\
 &+ \beta_8 \times CORIndígena + \beta_9 \times CORNA \\
 &+ \beta_{10} \times CORParda + \beta_{11} \times CORPreta \\
 &+ \beta_{12} \times COMPOSICA01 \\
 &+ \beta_{13} \times COMPOSICA02 \\
 &+ \beta_{14} \times COMPOSICA03 \\
 &+ \beta_{15} \times COMPOSICA04 \\
 &+ \beta_{16} \times COMPOSICA05 \\
 &+ \beta_{17} \times COMPOSICA06 \\
 &+ \beta_{18} \times SEXOHomem + \beta_{19} \times SEXOMulher \\
 &+ \varepsilon
 \end{aligned} \tag{1}$$

Onde ε representa os erros aleatórios do modelo, β_0 representa o intercepto e os β s restantes indicam os coeficientes de inclinação.

Após o procedimento Step-Up a 5% de significância e do procedimento de N-1 Dummies (ocultar as variáveis a vermelho) resultou-se em:

$$\begin{aligned}
PC_RENDA_DISP &= \beta_0 + \beta_1 \times IDADE \\
&+ \beta_2 \times ANOS_ESTUDO \\
&+ \beta_4 \times CORAmarela \\
&+ \beta_5 \times CORParda \\
&+ \beta_6 \times CORPreta \\
&+ \beta_7 \times COMPOSICAO1 \\
&+ \beta_8 \times COMPOSICAO2 \\
&+ \beta_9 \times COMPOSICAO4 \\
&+ \beta_{10} \times COMPOSICAO5 \\
&+ \beta_{11} \times SEXOMulher + \varepsilon
\end{aligned} \tag{2}$$

Onde ε representa os erros aleatórios do modelo.

Modelo multinível

$$\begin{aligned}
PC_{RENDA_{DISP}} &= \beta_0 + \beta X \\
&+ v_{1j},
\end{aligned} \tag{3}$$

onde $\beta_0 = \gamma_{00} + \mu_{0j}$, $\beta = \gamma_{ij} + \mu_{ij}$ e X representa o vetor das variáveis IDADE, ANOS_ESTUDO, CORAmarela, CORParda, CORPreta, COMPOSICAO1, COMPOSICAO2, COMPOSICAO4, COMPOSICAO5, SEXOMulher. Nesse contexto, β_0 : representa o intercepto e β : indica o vetor dos β das variáveis respectivas.

Segundo Fávero e Belfiore (2017), “os modelos multinível encaixam-se em uma estrutura aninhada que é definida da seguinte forma pelos autores: “devem existir variáveis que apresentam dados que se alteram entre indivíduos que representam determinado nível, porém permanecem inalteradas para certos grupos de indivíduos, sendo que esses grupos representam um nível superior. “

Mais ainda, Fávero e Belfiore (2017) afirmam que “os modelos multinível permitem que sejam identificadas e analisadas as heterogeneidades individuais e entre grupos a que pertencem estes indivíduos, tornando possível a especificação de componentes aleatórios em cada nível da análise. “

Dessa forma, a combinação de ANOS_ESTUDO e IDADE nos efeitos aleatórios das variáveis implica que se está incorporando esses efeitos associados à variável ANOS_ESTUDO e à variável IDADE. Essencialmente, isso pressupõe que a relação entre a renda e essas duas variáveis não é uniforme para todos os indivíduos em diferentes áreas

geográficas. Nesse contexto, permite-se que os coeficientes e os interceptos dessas variáveis variem conforme os diferentes estratos geográficos do Distrito Federal, representados na tabela 2, onde existiriam vários modelos distintos pelos níveis dentro de um só modelo completo.

Eliminação de outliers

Os outliers foram removidos pelo procedimento Z-score 3+ a fim de que o modelo tivesse maior capacidade preditiva aderente à realidade e melhor observação gráfica dos dados, devido há presença de valores de outliers muito elevados. De acordo com Hodge e Austin (2004), o Z-score é uma maneira de identificar outliers por meio da média e do desvio padrão dos dados, sendo identificado como outlier o valor, cuja diferença com a média seja X vezes maior que os desvio padrão. Nesse caso, o valor de 3 é considerado uma medida conservadora.

Procedimentos e Testes Auxiliares

Outrossim, foi feita primariamente uma breve análise descritiva dos dados do modelo, apresentando-se medidas estatísticas por meio de boxplots. Neto et al (2017) afirmam que o boxplot é uma forma visual de representar a distribuição de um conjunto de dados, destacando a mediana, a dispersão e a presença de valores extremos. Ele permite uma rápida identificação da posição central dos dados (mediana) e a variação em torno dessa medida. Além disso, é possível visualizar se os dados estão simetricamente distribuídos, ou se há assimetria, ou ainda se existe a presença de outliers.

A transformação de Box & Cox (1964) foi aplicada para otimizar o modelo em relação à normalidade e à homoscedasticidade. Ela consiste em encontrar o valor ideal do parâmetro lambda (λ) que maximiza a seguinte função:

$$Y(\text{transformado}) = \frac{(Y^\lambda - 1)}{\lambda} \quad (4)$$

Onde, $Y(\text{transformado})$ é o novo conjunto de dados que se obtém após aplicar a transformação. Y é a variável que se está tentando explicar ou prever com seu modelo de regressão. λ (lambda): O parâmetro lambda (λ) é o valor que se está tentando encontrar para maximizar a função.

Adicionalmente, cabe ressaltar que a modelagem hierárquica abordada funciona intrinsecamente na presença de heteroscedasticidade devido à interação entre as variáveis explicativas e o erro do modelo, nesse sentido, não foi realizado o procedimento de verificação de homoscedasticidade de Breusch- Pagan, que verifica a hipótese de presença de homoscedasticidade em relação à variância dos resíduos do modelo.

O procedimento Step-UP foi adotado para remover variáveis não significativas a 5% de maneira iterativa, de forma a selecionar o conjunto de variáveis mais relevantes manualmente. O processo consiste em calcular diferentes combinações de variáveis do modelo de maneira a encontrar um arranjo mais adequado e, assim, remover as variáveis não significativamente relevantes ao modelo a 5% de significância estatística.

Analisou-se o coeficiente de determinação (R^2), que indica a proporção da variabilidade da variável dependente explicada pelo modelo e que pode variar de 0 a 1. Um valor de R^2 próximo a 1 significa que o modelo consegue explicar uma grande parte da variabilidade da variável dependente, o que indica um ajuste muito bom dos dados. Por outro lado, um valor de R^2 próximo a 0 significa que o modelo não consegue explicar adequadamente a variabilidade da variável dependente, sugerindo que o ajuste dos dados não é tão bom. (Martins e Rodrigues, 2014)

O coeficiente é calculado pela seguinte equação:

$$R^2 = \left[\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \right]^2 \quad (5)$$

Onde n: Número de observações nos dados, x_i : O valor da variável independente (ou preditora) para a i -ésima observação, \bar{x} (x_i barra): A média dos valores da variável independente (x_i) para todas as observações, y_i : O valor da variável dependente (ou resposta) para a i -ésima observação e \bar{y} (y_i barra): A média dos valores da variável dependente (y_i) para todas as observações.

Mais ainda, utilizaram-se o log-verossimilhança (LOGlik), o AIC [Akaike Information Criterion] (AIC) e o BIC [Bayesian Information Criterion] que são métricas amplamente utilizadas na seleção de modelos estatísticos. O LOGlik mede a probabilidade de observarmos os dados com base nos parâmetros do modelo ajustado. O AIC e o BIC, por outro lado, levam em conta tanto o ajuste do modelo quanto a complexidade dele, penalizando modelos com um número excessivo de parâmetros.

Por fim, foram calculados os coeficientes de Gini e a curva de Lorenz do DF com os dados da pesquisa POF. Segundo Soares et al. (2006), o coeficiente de Gini é uma ferramenta estatística que mede a desigualdade e varia de 0 a 1, sendo 1 o critério mais desigual e 0

completa igualdade. A curva de Lorenz é uma representação gráfica em que se observa de forma gradativa a concentração de riqueza dos 1% mais pobres até os 1% mais ricos.

Foram analisados 4 tipos de modelos e selecionado o modelo de regressão multinível com efeito de intercepto e efeitos aleatórios a fim de garantir maior flexibilidade aos pressupostos do modelo e aderência à realidade.

Todos os procedimentos acima serão realizados por meio do Software R versão 4.2.1. Essas etapas de teste e análise são fundamentais para garantir a robustez e a validade dos resultados obtidos pelo modelo de regressão multinível.

Resultados e Discussão

Análise Descritiva dos Dados

A tabela a seguir foi retirada do software R 4.2.1 por meio do código base do R “summary()”

Tabela 3 – Sumário das variáveis quantitativas		(continua)
Variáveis Quantitativas	Indicadores Estatísticos	Resultado
PC_RENDA_DISP = “RENDA” ao longo do trabalho	Valor Mínimo	250.10
	Primeiro Quartil	1096.00
	Mediana	1838.00
	Média	2991.50
	Terceiro Quartil	3411.80
	Valor Máximo	26610.80
PC_RENDA_DISP.1	Valor Mínimo	4.14
	Primeiro Quartil	4.88
	Mediana	5.11
	Média	5.13
	Terceiro Quartil	5.38
	Valor Máximo	6.13
Tabela 3 – Sumário das variáveis quantitativas		(conclusão)
Variáveis Quantitativas	Indicadores Estatísticos	Resultado
IDADE	Valor Mínimo	19.00
	Primeiro Quartil	31.00
	Mediana	41.00

ANOS_ESTUDO	Média	43.23
	Terceiro Quartil	54.00
	Valor Máximo	95.00
	Valor Mínimo	0.00
	Primeiro Quartil	8.00
	Mediana	12.00
	Média	10.89
	Terceiro Quartil	15.00
	Valor Máximo	16.00

Fonte: Resultados originais da pesquisa

Tabela 4- Sumário das variáveis categóricas

Variáveis Categóricas	Variáveis Categóricas	Número de Unidades
COR ou Raça	Cor Branca	1081
	Cor Preta	371
	Cor Parda	1411
	Cor Amarela	37
	Cor Indígena	19
	Cor NA (nenhuma alternativa)	45
Sexo ou Gênero	Sexo Feminino	1603
	Sexo Masculino	1361
Composição ou Tipo Familiar	1 – Um adulto sem criança	148
	2 – Um adulto com ao menos uma criança	47
	3 – Mais de um adulto sem criança	1023
	4 – Mais de um adulto com ao menos uma criança	1026
	5 – Um ou mais idosos com ou sem crianças	84
	6 – Um ou mais idosos, com ao menos um adulto, com ou sem crianças	636

Fonte: Resultados originais da pesquisa

Por meio dos dados acima já é possível aferir uma grande desigualdade na variável PC_RENDA_DISP.1 que representa a variável “RENDA” sem procedimento BoxCox.

Quando a mediana de um conjunto de dados está abaixo da média, isso indica que a distribuição dos valores é assimétrica à direita, ou seja, há valores maiores que puxam a média para cima.

A mediana é o valor central quando os dados estão ordenados em ordem decrescente ou crescente, como é o caso. Ela representa o ponto em que metade dos valores está acima e metade está abaixo. Já a média é a soma de todos os valores dividida pelo número de observações.

Se a mediana está abaixo da média, significa que há valores relativamente maiores na parte superior da distribuição de dados. Esses valores mais altos influenciam a média, puxando-a para cima. É um indicativo de uma assimetria positiva, em que existem alguns valores atípicos ou extremos no lado direito da distribuição.

Essa análise pode ser melhor vista por meio do processo de boxplot da variável em questão, por isso foram realizados dois desses procedimentos por meio do código "boxplot()" do pacote "MASS" do R, a fim de analisar os dados .

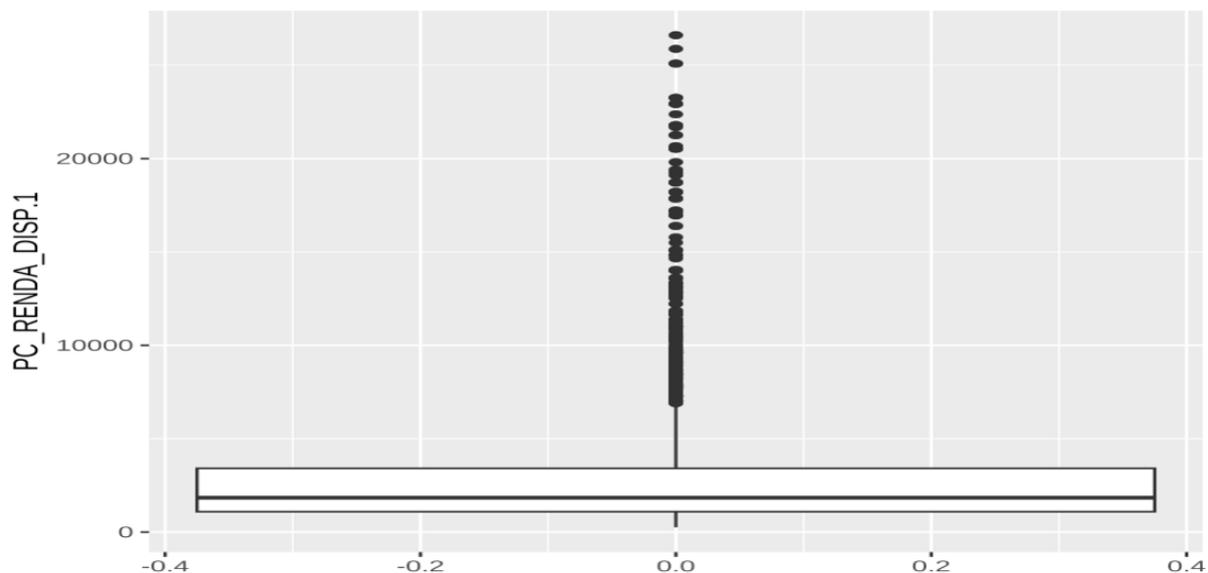


Figura 1- BoxCox da variável "RENDA"
Fonte: Resultados originais da pesquisa

Nesses gráficos, é possível observar melhor os fatos narrados no parágrafo anterior. Há uma quantidade grande de dados concentrada entre os valores 0 e 3000 reais e a mesma quantidade de valores é distribuída de forma bem mais diluída acima deste valor, com um valor médio de 2991.5 reais, o que aponta novamente a uma distribuição assimétrica positiva.

Resultados dos Modelos

Comparação dos Modelos

Tabela 5 – Indicadores de qualidade dos modelos dispostos de forma comparada

	Modelo Linear sem transformação BoxCox (modelo 1)	Modelo Linear com transformação BoxCox\ (modelo 2)	Modelagem Multinível apenas com intercepto aleatório (modelo 3)	Modelagem Multinível com intercepto e efeito aleatórios (modelo 4)
Coefficiente de Determinação	0.2703	0.34	-	-
BIC	55653.13	1335.956	1196.219	1156.275
AIC	55575.20	1258.031	1124.328	1054.429
LogLik	-27774.60	-616.0153	-550.1639	-510.2145

Fonte: Resultados originais da pesquisa

Conforme demonstrado na Tabela 5, há ganhos de eficiência e qualidade dos modelos baseado nos indicadores acima. O Modelo Linear com transformação BoxCox tem o R^2 mais alto (0.34) em relação ao que não apresenta o procedimento, indicando que ele ajusta os dados melhor em comparação com o Modelo Linear sem transformação BoxCox (0.2703).

Relativamente ao BIC e AIC, a presença de valores de BIC e AIC indicam um melhor ajuste. Portanto, os modelos de Modelagem Multinível com intercepto e efeito aleatórios têm os menores valores de BIC e AIC, o que sugere que eles se ajustam melhor aos dados em comparação com os modelos lineares.

Mais ainda, Valores mais altos indicam um melhor ajuste do modelo aos dados. Novamente, os modelos de Modelagem Multinível têm valores de LogLik mais altos, o que sugere um melhor ajuste.

Por fim, valores mais altos de LogLik indicam um melhor ajuste do modelo aos dados. Novamente, os modelos de Modelagem Multinível têm valores de LogLik mais altos (menos negativos), o que também corrobora para o ganho de qualidade do modelo na forma hierárquica e conclui-se que todos os indicadores apontam para o modelo 4 com melhor ajuste. Nesse sentido, optou-se por utilizar o modelo 4.

Aferição da Significância Estatística do uso do Modelo Multinível

Tabela 6 - Análise estatística da utilização da modelagem multinível

	Estimativas da Variância	Erro Padrão	Variável z	P-Value
Var(v0j)	0.009895859	0.005478963	1.806155	0.071
Var(e)	0.0822272673	0.002143622	38.380215	0.0

Fonte: Resultados originais da pesquisa

"Var(v0j)": Variância entre os grupos (unidades superiores).

"Var(e)": Variância residual dentro dos grupos (unidades inferiores).

A primeira linha corresponde à variância entre grupos, cujo valor de cerca de 0.009 aponta que exista variabilidade nos resultados entre grupos que não seja inicialmente explicada pelas variáveis adotadas. Para além disso, a um nível de significância de 10% conclui-se pela que esse valor é significativo estatisticamente.

A segunda linha refere-se à variância intragrupos. O valor de cerca de 0.082 aponta para a existência dessa variância dentro dos grupos de estratos diferentes significativo a 0% de nível de significância.

A fim de calcular o valor da correção intraclasse (ICC - Intraclass Correlation Coefficient) em um modelo multinível, utilizou-se a seguinte fórmula:

$$ICC = \frac{Var(v0j)}{(Var(v0j) + Var(e))} \quad (6)$$

Onde:

Var(v0j) é a variância entre os grupos (variância aleatória no nível superior).

Var(e) é a variância residual dentro dos grupos (variância aleatória no nível inferior).

Os valores de Var(v0j) e Var(e) podem ser obtidos a partir das estimativas fornecidas na saída da fórmula utilizada no R acima. Usando os valores da sua saída:

$$Var(v0j) = 0.009895859$$

$$Var(e) = 0.0822272673$$

Agora podemos calcular o ICC:

$$ICC = \frac{0.009895859}{(0.009895859 + 0.082272673)}$$

$$ICC = \frac{0.009895859}{0.092168532}$$

$$ICC \approx 0.1074$$

O valor da correlação intraclasse é aproximadamente 0.1074, o que indica que cerca de 10.74% da variação total nos dados ocorre entre os grupos (nível superior), enquanto os 89.26% restantes são devido à variação dentro dos grupos (nível inferior). Esse valor é importante para entender a proporção da variação explicada pela estrutura hierárquica dos dados no modelo multinível. Quanto maior o ICC, maior a variação entre os grupos em relação à variação dentro dos grupos, e vice-versa. Nesse sentido, optou-se, contudo, por utilizar-se modelagem multinível com efeito aleatório de intercepto e efeito aleatório de inclinação a nível de 10% de significância estatística e considerando o valor do ICC apontado.

Resultados do Procedimento Step-Up

No Procedimento Step-Up do modelo 4 foram retiradas as seguintes variáveis não estatisticamente significantes a 95% de nível de confiança:

Tabela 7 - Variáveis Retiradas no Procedimento Step-UP

Cor Amarela
 Cor Indígena
 Composição 6
 COR NA

Fonte: Resultados originais da pesquisa

Resultados Gerais do Modelo Final (modelo 4)

Tabela 8 - Output do Modelo 4

Fixed Effects	Resultados da Tabela 8	Transformação inversa do resultado do procedimento	Erro Padrão	T- Valor	P-valor
---------------	------------------------	--	-------------	----------	---------

	BoxCox com Lambda= - 0.11				
(Intercept)	4.608178	618.723500	0.046277	99.58713	0.00
ANOS_ESTUDO	0.035687	1.036400	0.005199	6.86360	0.00
IDADE	0.004909	1.004922	0.000485	10.13067	0.00
COR_Preta	-0.081173	0.922366	0.017158	-4.73009	0.00
COR_Parda	-0.072957	0.929911	0.011470	-6.36052	0.00
SEXO_Mulher	-0.028764	0.971689	0.105010	-2.73926	0.0062
Composicao_1	0.170599	1.187939	0.024382	6.99664	0.00
Composicao_2	-0.194618	0.824841	0.042173	-4.61476	0.00
Composicao_4	-0.101553	0.891542	0.011690	-8.74056	0.00
Composicao_5	0.179524	1.198880	0.033290	5.39268	0.00

Fonte: Resultados originais da pesquisa

Na tabela 8, foram dispostos alguns dos mesmos dados presentes na Figura 2, mas de maneira a aplicar a transformação inversa de Boxcox a fim de que se possa analisar de maneira mais clara.

De acordo com Chein (2019) a interpretação dos Betas, como o lambda é próximo de 0, assemelha-se a interpretação de uma função Log-Linear. Nesse tipo de Função, os β 's representam a elasticidade da variável dependente em relação a independente. Desse modo, a variação de 1 unidade adicional de X induz a variação percentual de β (x)% da variável Y, todo o resto constante.

Ou seja, no caso em específico, para cada variação de 1 unidade adicional da variável "ANOS_ESTUDO" gera-se cerca 3,57% de acréscimo na variável "REND", *ceteris paribus*.

Ainda segundo Chein(2019), quando se trata de uma variável Dummy os Betas relativos referem-se às diferenças médias entre as variáveis adjacentes e as ocultados no procedimento N-1 Dummies.

Portanto, o Beta da variável "Composicao_1" indica que os indivíduos pertencentes a essa classe têm em média uma "REND" cerca de 18% maior se comparados aos indivíduos pertencentes a "Composicao_3", ocultada anteriormente.

Por fim, os resultados da Figura 2, demonstram a significância estatística de todas as variáveis do modelo a 5% de significância. Mais ainda, os resultados contidos na tabela 8 serão debatidos e contextualizados no tópico a seguir acerca da realidade do DF.

Discussões sobre o Tema

Curva de Lorenz do DF

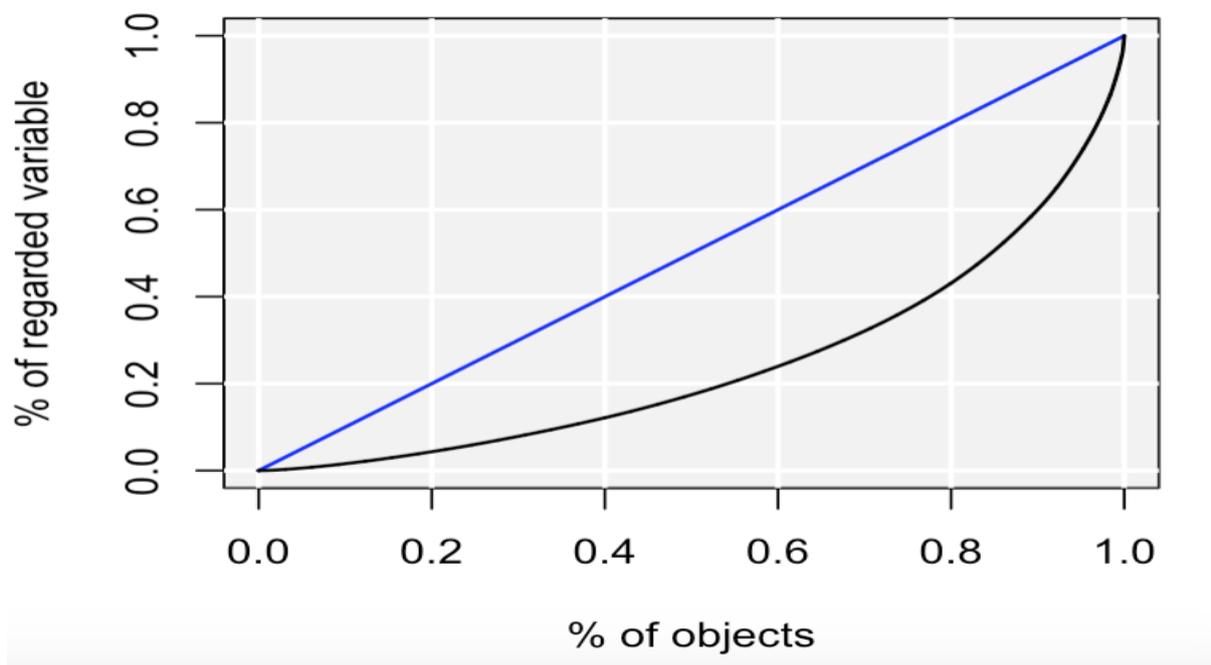


Figura 2 – Curva de Lorenz do DF

Fonte: Resultados originais da pesquisa

A figura 2 representa a curva de Lorenz do Distrito Federal obtida por meio da função Lorenz() do pacote “REAT” do R. Foram utilizados os dados da pesquisa POF do IBGE e optou-se por utilizar-se os dados antes do procedimento de eliminação de outliers a fim de captar de forma mais condizente as disparidades da variável “RENDIA”.

Nesse contexto, é possível inferir que os 80% mais pobres detêm pouco mais de 40% dos rendimentos per capita em questão. Por conseguinte, é possível notar que os 20% mais ricos auferem quase 60% dos rendimentos per capita totais.

Coeficiente de GINI por variáveis

Tabela 9 - Coeficiente de Gini

Variáveis	Coeficiente de Gini
GINI do DF	0.5089754
COR_Preta	0.4533566
COR_Parda	0.4698907
COR_Branca	0.5195823
COR_Amarela	0.5741215
SEXO_Mulher	0.5050519

SEXO_Homem	0.5130016
Composicao_1	0.5189186
Composicao_2	0.4563596
Composicao_3	0.5233293
Composicao_4	0.480379
Composicao_5	0.4629021
Composicao_6	0.4338285
Estrato 5301	0.4530201
Estrato 5302	0.5152931
Estrato 5303	0.4781762
Estrato 5304	0.3863336
Estrato 5305	0.3278318
Estrato 5306	0.6802469
Estrato 5307	0.3815777
Estrato 5308	0.3699076

Fonte: Resultados originais da pesquisa

O coeficiente foi obtido por meio da função `gini()` do pacote "REAT" do R aplicada à variável "RENDA". Por meio dos valores da Tabela 9, observou-se que, dentre as cores ou raças, a variável "COR_Amarela" tem o maior coeficiente de Gini (0.5741215), indicando maior desigualdade na distribuição de renda entre pessoas com essa cor. Por outro lado, a variável "COR_Branca" também apresenta alta desigualdade (0.5195823) um pouco acima da total (0.5089754). Para além disso, as variáveis "COR_Parda" (0.4698907) e "COR_Preta" (0.4533566) têm coeficientes de Gini relativamente mais baixos, indicando menor desigualdade em comparação com as outras cores dentro de seus próprios grupos.

Relativamente ao sexo, não há uma grande diferença nos coeficientes de Gini entre "SEXO_Mulher" (0.5050519) e "SEXO_Homem" (0.5130016). Ambos têm níveis semelhantes de desigualdade.

Sobre a composição familiar, a variável "Composicao_3" tem o maior coeficiente de Gini (0.5233293), indicando maior desigualdade na distribuição de renda entre os grupos de composição familiar. A variável "Composicao_6" tem o coeficiente de Gini mais baixo (0.4338285), indicando menor desigualdade.

Por fim, acerca dos estratos, o coeficiente de Gini varia significativamente entre os estratos sociais, com "Estrato 5306" (0.6802469) tendo o coeficiente mais alto, indicando alta desigualdade. Não obstante, o "Estrato 5304" (0.3863336) tem o coeficiente de Gini mais baixo, indicando menor desigualdade nesse estrato.

Disparidade da “RENDÁ” Relativa ao Sexo

As mulheres desde a infância no seio familiar sofrem tratamento diferenciado. A cultura da sociedade tende a sobrecarregar as mulheres. Jornadas duplas ou até triplas são comuns. Mais ainda, as mulheres apesar de possuírem maior qualificação se comparado aos homens recebem um salário menor. (Ortiz et al. 2018).

Com base na pesquisa realizada pela Secretaria de Estado de Planejamento e Orçamento do Distrito Federal (CODEPLAN) em 2014, no setor público do DF, apesar dos avanços como a criação da Secretaria de Estado da Mulher e a implementação de cotas eleitorais, os homens ainda ocupam predominantemente as funções públicas. Mais ainda, as mulheres são mais comumente direcionadas às áreas de cuidados, como saúde, educação e assistência social, fortalecendo o estereótipo de gênero.

Para além disso, os vieses cognitivos representam uma barreira invisível na ascensão das mulheres dentro das empresas e até mesmo as mulheres não se apercebem de sua exclusão, mesmo que os dados sejam claros, revelando-se um preconceito implícito. (Oliveira, 2022).

O modelo desenvolvido neste trabalho corrobora os fatos narrados acima. O Beta relativo à variável Dummy "SEXO_mulher" representa que no modelo as mulheres obtiveram na sua “RENDÁ” em média menos cerca de 3% em comparação com os homens, *ceteris paribus*.

Diferenças da “RENDÁ” referentes à Cor ou à Raça

É importante ressaltar que, apesar da redução da discriminação racial nas elites ocupacionais brasileiras, esse processo é marcado por uma evolução lenta e duradoura ao longo dos séculos. (de Almeida e da Silva, 2021).

Conforme destacado por Martins (2014), o autor resalta a presença de determinações do racismo que influenciam negativamente as oportunidades de emprego e ascensão profissional para a população negra. Apesar de grande evolução a respeito da política de cotas e combate ao racismo estrutural, a exclusão dos negros remonta desde a época da escravidão e perpetua-se no mercado de trabalho ao longo do tempo, nos cargos eletivos e nos cargos de direção.

O modelo aponta que os negros recebem menos em média cerca de 8% em relação aos brancos acerca da “RENDÁ”, todo o resto constante. No mesmo sentido, o modelo apontou uma “RENDÁ” menor em média de cerca de 7,3% se comparada aos dos brancos, *ceteris paribus*. Nesse sentido, observa-se uma desigualdade de rendimentos entre raças no DF, assim como descrito no meio acadêmico nas mais diversas regiões do Brasil.

Relativamente à população indígena, Gomes e da Silva (2016) remontam às diversas formas de abuso cometidos aos povos indígenas, desde escravidão, exploração, invasão de terras. Posteriormente, percebe-se que os indígenas, apesar de também terem políticas de cotas de acesso à universidade e concursos, ainda sofrem exclusão e preconceito, inclusive no que diz respeito à participação no mercado de trabalho e no acesso aos bens e serviços disponíveis.

O modelo não apontou, com significância estatística a 5%, que exista diferença entre salários de brancos e indígenas no DF. Entretanto é necessário destacar que isso possa se dever ao fato das variáveis do censo do IGBE não considerarem os indígenas que vivam em regiões afastadas ou em aldeias em variáveis relevantes ao modelo, o que pode ter enviesado a amostra.

Relativamente aos povos amarelos, Segundo Yanagida (2022), os povos amarelos também sofreram com preconceito e tipos de racismo no Brasil, inclusive com o apagamento de sua história e suas origens, resumidas ao termo “amarelo”.

Entretanto, de acordo com Dezem (2005, p. 53) é necessário evidenciar que, diferentemente dos povos indígenas e negros, grande parte dos povos amarelos, assim como imigrantes europeus vieram subsidiados pelo governo do Brasil e, posteriormente, foram-lhe doadas terras a fim de serem trabalhadores livres, muito embora, as condições de trabalho desses imigrantes fossem bastante precárias.

Portanto, a origem desses povos de certa forma diverge dos povos citados anteriormente, pois tiveram vantagens em sua chegada ao País que se refletem nos rendimentos atuais de seus descendentes.

Desigualdades Geográficas na “RENDA”

De acordo com Nunes e Costa (2016), em seu estudo sobre a dinâmica urbana, violência e heterogeneidade social no Distrito Federal e em Brasília, destacam-se as complexidades e desigualdades presentes nessa região. Eles argumentam que a cidade reflete as contradições e diversidades socioeconômicas do país, evidenciando uma realidade marcada por diferentes dinâmicas e desafios. Por um lado, há um controle do crescimento do mercado imobiliário no Plano Piloto de Brasília devido ao seu tombamento histórico e sua arquitetura peculiar, por outro lado há um crescimento seguido da lógica de mercado nas regiões administrativas (RA's).

Nesse contexto, há uma exclusão geográfica de acesso à moradia por parte das populações de menor renda a áreas centrais da cidade e há um crescimento desordenado nas RA's da cidade acompanhado pela lógica de mercado sem grandes restrições e planejamento.

De acordo com o estudo de Ratima (2003), é evidente a exclusão espacial vivenciada pela população residente nas RA's do Distrito Federal, apesar de constituírem a maior parte do território. Essa exclusão é caracterizada pela falta de acesso a serviços de qualidade, como postos de saúde, escolas, teatros e estabelecimentos comerciais, os quais estão concentrados de forma desproporcionada na área central conhecida como Plano Piloto. Essa disparidade na distribuição de serviços revela uma clara desigualdade na provisão de recursos e oportunidades para a população residente nas RAs em comparação com aquela que vive no Plano Piloto que somados às desigualdades de renda tendem a acentuar a disparidade de qualidade de vida entre os habitantes do DF.

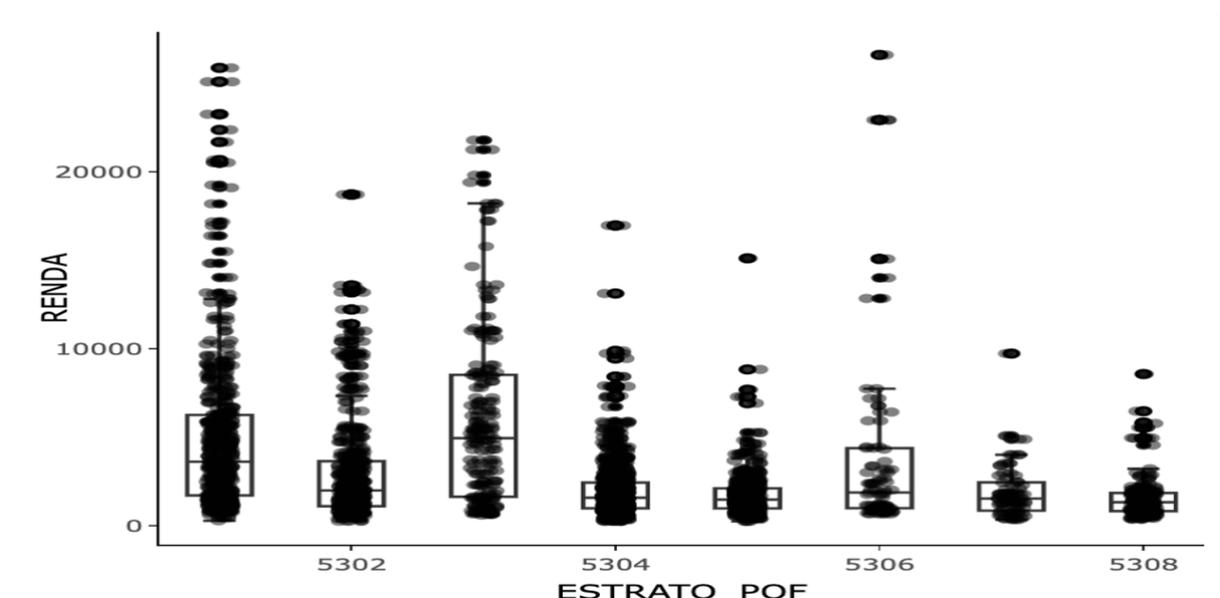


Figura 3 - Boxplots da "REND" por estratos
Fonte: Resultados originais da pesquisa

Com base no gráfico obtido por meio do pacote "Ggplotly" e "MASS" por meio da função `boxplot()` acima, que representa a variável dos estratos do Distrito Federal (DF) no contexto do modelo de regressão hierárquica, pode-se inferir que há discrepâncias tanto entre quanto dentro dos grupos, corroborando resultados obtidos anteriormente neste estudo. Os estratos 5301 e 5303 exibem as maiores médias, o que também se reflete em seus valores mais elevados. Em contraste, os estratos 5307 e 5308, caracterizados como estratos rurais, apresentam médias e valores menores, conforme esperado.

Adicionalmente, observa-se que os valores dos estratos 5304 e 5305 estão próximos dos encontrados no meio rural, o que sugere que esses estratos possuem rendimentos inferiores quando comparados com o meio urbano, apesar de exibirem alguns valores mais altos do que os observados no meio rural.

Por fim, destaca-se a presença de estratos localizados dentro da cidade com rendimentos menores em relação aos demais, comparáveis aos encontrados nos meios rurais, possivelmente relativos às cidades satélites que apesar de densamente povoadas, algumas delas apresentam uma precariedade de diversos aspectos sociais, inclusive a renda.

Desigualdade do Acesso ao Ensino no DF e conseqüente impacto na "REND"

O estudo conduzido por Pinto et al. (2018) aborda a temática do acesso à Educação Infantil no Distrito Federal. Os pesquisadores investigaram as disparidades existentes entre diferentes grupos socioeconômicos no que diz respeito à participação das crianças na Educação Infantil.

A pesquisa buscou compreender como o acesso à Educação Infantil tem sido influenciado por questões históricas e sociais, destacando as diferenças entre o passado e o presente. Apesar de os autores identificarem no DF como uma ideia inicial baseada em sua estrutura arquitetônica de uma cidade em que pobres e ricos estudariam nos mesmos ambientes.

Posteriormente, observou-se o contrário com a gentrificação da cidade e exclusão de parcela da população. Por meio do surgimento das cidades satélites e expulsão dos pobres do centro da cidade.

Por outro lado, é possível identificar melhorias no que diz respeito ao acesso ao ensino superior. Melo (2009) observa que há maior democratização ao ensino superior devido a ações afirmativas por parte do governo como as políticas de quotas.

Disparidade da “RENDIA” devido à Composição Familiar

De acordo com as conclusões do estudo de Wajnman, Turra e Agostinho (2006), a diferença nos rendimentos per capita das famílias, relativamente à sua estrutura, dão-se devido ao fato de os indivíduos da família estarem ou não a trabalhar, ou ainda se estão ou não aposentados, se recebem ou não benefícios e, mais ainda, se esses indivíduos pertencerem ou não a grupos historicamente marginalizados, como mulheres e negros.

Nesse contexto ainda sobre o estudo de Wajnman et al. (2006), observou-se que houve uma pequena mudança no rendimento per capita, entre os anos de 2000 a 2005 no Brasil, motivada pela reengenharia de algumas composições familiares, como envelhecimento da população, o aumento de pessoas do sexo feminino a sustentarem seus lares. Não obstante, identificou-se que as pessoas de idade mais avançada apresentam rendimentos mais elevados na medida em que se beneficiaram ao longo do tempo de sistema de previdência mais favorável, como em relação aos funcionários públicos e ainda por priorização em benefícios como BPC [Benefício de Prestação Continuada].

Destarte, no contexto do modelo pode observar-se que as composições com idosos apresentam rendimentos maiores em média, todo o resto constante, se comparados às outras, possivelmente pelo fato explicitado acima. Por outro lado, é relevante destacar que a composição 1 apresentou em média, todo o resto constante, rendimentos maiores que a composição número 3. Desse modo os adultos que moram sozinhos apresentaram rendimentos per capita em média, *ceteris paribus*, superiores aos adultos sem filhos que moram sozinhos. Pode-se dever ao fato de apenas um dos adultos trabalhar e outro cuidar das atividades domésticas, como no caso de muitos casais.

Considerações Finais

O objetivo deste trabalho foi analisar a relação entre a renda familiar per capita no DF e algumas variáveis sociodemográficas. Para tanto, foram utilizados modelos de regressão linear múltipla, modelos de regressão multinível, a curva de Lorenz e os coeficientes de Gini a fim de analisar a disparidade no rendimento per capita de renda no DF em relação às mais diversas variáveis sociodemográficas.

Os resultados obtidos mostram que existem diferenças significativas entre os indivíduos por diversas razões. Constatou-se que as mulheres tendem a ganhar rendimentos mais baixos do que os homens.

Além disso, foram observadas disparidades de rendimento entre diferentes grupos étnicos. Os negros e pardos obtiveram rendimentos mais baixos que os brancos, confirmando a existência de persistentes desigualdades raciais na sociedade. Não foram encontradas

diferenças significativas de renda entre nativos indígenas e amarelos em relação aos brancos, mas cabe ressaltar que a amostra pode não ter sido suficientemente representativa para capturar essas especificidades.

Outra razão importante encontrada foi a desigualdade de renda geográfica. As classes urbanas tinham rendimentos mais elevados do que as classes rurais, e alguns estratos tinham rendimentos mais baixos e outros bem mais elevados. Essa desigualdade regional reflete as desigualdades econômicas do Distrito Federal.

Além disso, observou-se que o acesso à educação impacta diretamente a renda familiar per capita. A existência de políticas de cotas e ações afirmativas para o ensino superior têm contribuído para uma maior democratização do acesso ao ensino superior, mas ainda existem desafios quanto à qualidade da Educação Infantil e do ensino.

Discutiu-se como o tipo de composição familiar influencia a renda per capita da população, de forma que os adultos que moram sozinhos e as estruturas familiares que possuem idosos tenham rendimentos per capita mais elevados se comparados aos restantes.

Em resumo, este estudo destacou as implicações das desigualdades sociais e econômicas existentes no Distrito Federal. As diferenças de raça, gênero, região geográfica e acesso à educação e estrutura familiar são fatores que influenciam diretamente o rendimento familiar. A compreensão dessas disparidades é necessária para desenvolver políticas públicas mais eficazes para promover a igualdade e o desenvolvimento sustentável no Distrito Federal.

Vale ressaltar que os múltiplos modelos utilizados neste estudo permitiram uma análise mais aprofundada da estrutura hierárquica dos dados, trazendo uma maior compreensão das relações entre as variáveis e sua influência sobre os chefes de família, onde o modelo é possível. e é tanto mais relevante compreender os diferentes DF de acordo com os diferentes contextos já mencionados, e tanto mais efetivamente nos resultados das atividades sociais.

Mais ainda, a utilização de dados atualizados e de uma amostra mais representativa poderia contribuir para resultados mais sólidos e relevantes. Com isso, poderá apoiar uma administração pública mais eficaz, voltada à redução das desigualdades e à promoção de uma sociedade mais justa e equitativa para todos os cidadãos do Distrito Federal.

Agradecimento

Agradeço a Belaura Rosal de Ávila (minha mãe), Milce Lemos Rosal de Ávila (minha avó), Paulo Lemos Rosal de Ávila (meu tio) e Belisa Moura de Ávila (minha prima) pelo apoio na minha jornada acadêmica e conclusão da tese. Sou grato por tê-los em minha vida.

Referências

Banco Central do Brasil [BACEN]. 2023. Calculadora do Cidadão.

Disponível em:

<https://www3.bcb.gov.br/CALCIDADA0/publico/corrigirPorIndice.do?method=corrigirPorIndice>. Acesso em: 09 de Jun. 2023.

[Barros, A. R. \(2011\). Desigualdades regionais no Brasil: natureza, causas, origens e soluções. Disponível em: https://www.redalyc.org/pdf/4017/401746386001.pdf](#). Acesso em: 29 jun. 2023.

Barros, R. P. D., Carvalho, M. D., Franco, S., & Mendonça, R. S. P. D. 2006. Uma análise das principais causas da queda recente na desigualdade de renda brasileira. Disponível em: [Uma análise das principais causas da queda recente na desigualdade de renda brasileira](#). Acesso em: 29 jun. 2023.

Carmona, P. A. C.; Oliveira, T. R.; Coelho, J. M.; Amaral, R. 2022. Habitação de interesse social no Plano Piloto: operações urbanas consorciadas como resposta à segregação espacial no DF. Revista de Direito da Cidade, 14(3), 1707-1729. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/rdc/article/viewFile/57145/44420>. Acesso em: 29 jun. 2023.

Chein, F. 2019. Introdução aos modelos de regressão linear: um passo inicial para compreensão da econometria como uma ferramenta de avaliação de políticas públicas. Acesso em: <https://repositorio.enap.gov.br/handle/1/4788>. Acessado em 25 jul. 2023.

de Almeida Vieira, C.; da Silva, R. M. 2021. Desigualdade racial na elite ocupacional brasileira: considerações acerca da discriminação no Mercado de trabalho 2005-2015. *Revista da ABET*, 20(1). Disponível em: <https://scholar.archive.org/work/c6mjlrxtqhblhfezwuozz564/access/wayback/https://periodicos.ufpb.br/index.php/abet/article/download/41504/33883>. Acesso em: 29 jun. 2023.

de Matos Pires, L. 2018. PLANO BRASIL SEM MISÉRIA: GERAÇÃO AUTÔNOMA DE RENDA PARA REDUÇÃO DA EXTREMA POBREZA NA ÁREA RURAL DO DISTRITO FEDERAL. Disponível em: <https://www.ipe.df.gov.br/wp-content/uploads/2018/02/TD-47-Plano-Brasil-Sem-Miséria-Geração-autônoma-de-renda-para-redução-da-extrema-pobreza-na-área-rural-do-DF.pdf>. Acesso em: 29 jun. 2023.

de Ratima Eactilhos, M. 2003. Segregação socioespacial no Distrito Federal. *Revista Katálysis*, 6(2), 237-248. Disponível em: <https://www.redalyc.org/pdf/1796/179617959009.pdf>. Acesso em: 29 jun. 2023.

Dezem, R. 2005. *Matizes do "Amarelo": a Gênese dos Discursos sobre os Orientais no Brasil (1878-1908)* (Vol. 4). Editora Humanitas. Disponível em:

https://books.google.com/books?hl=pt-PT&lr=&id=p9UTH4f3bVIC&oi=fnd&pg=PA15&dq=povos+amarelos+no+mercado+de+trabalho+no+brasil&ots=re1gM0ofak&sig=2qm_kXNmT2bHrI0oWxl4uq3SIGc. Acesso em: 29 jun. 2023.

Dias, V. H. C. 2017. Relações entre índices de criminalidade e indicadores socioeconômicos nas regiões administrativas do Distrito Federal. *Acta de Ciências e Saúde*, 1(1), 1-5. Disponível em: <http://www2.ls.edu.br/actacs/index.php/ACTA/article/download/154/142>. Acesso em: 29 jun. 2023.

Fávero L. F.; Belfiore P. (2017). *Manual de Análise de Dados - Estatística e Modelagem Multivariada com Excel, SPSS e Stata*. Gen LTC. 855- 864.

Gomes, E. B.; da Silva, A. R. L. 2016. O DIREITO DO INDÍGENA AO TRABALHO: “INTEGRAÇÃO”, SOCIEDADE E A CONVENÇÃO 169 DA OIT – (ESTUDO DE CASO). *Revista Direitos Culturais*, 10(21), 15-30. Disponível em: <https://core.ac.uk/download/pdf/322639809.pdf>. Acesso em: 29 jun. 2023.

Hoffmann, R. (2020). Desigualdade de renda no Brasil, 1995-2019: diversas distribuições e o impacto do desemprego. *RBEST Revista Brasileira de Economia Social e do Trabalho*, 2, e020007-e020007. Disponível em: <https://econtents.bc.unicamp.br/inpec/index.php/rbest/article/download/14205/9771>. Acesso em: 29 jun. 2023.

Hodge, V.; Austin, J. 2004. A survey of outlier, detection methodologies. *Artificial intelligence review*, 22(2), 85–126. Disponível em: <https://link.springer.com/article/10.1007/s10462-004-4304-y>. Acesso em: 29 jul. 2023.

Instituto Brasileiro de Geografia e Estatística [IBGE]. Descrição das variáveis - Microdados da amostra do Censo Demográfico 2010. Rio de Janeiro, Brasil: IBGE, 2010. Acesso em: 29 jun. 2023.

Instituto Brasileiro de Geografia e Estatística [IBGE]. 2016. Rendimentos de Todas as Fontes e Trabalho Infantil – Dados da PNAD contínua 2016. Recuperado de: https://agenciadenoticias.ibge.gov.br/media/com_mediaibge/arquivos/bd466f98f27dac67181148ebe5d960de.pdf. Acessado em jun. 2023.

Instituto Brasileiro de Geografia e Estatística [IBGE]. 2021. Pesquisa Nacional por Amostra de Domicílios Contínua: PNAD Contínua. Recuperado de <https://www.ibge.gov.br/estatisticas/sociais/trabalho/17270-pnad-continua.html>. Acesso em: 29 jun. 2023.

Instituto Brasileiro de Geografia e Estatística [IBGE]. 2018. Pesquisa de Orçamentos Familiares, Recuperado de: <https://www.ibge.gov.br/estatisticas/sociais/saude/24786-pesquisa-de-orcamentos-familiares-2.html>. Acesso em: 19 jul. 2023.

Instituto Brasileiro de Geografia e Estatística [IBGE]. 2021. Pesquisa Nacional por Amostra de Domicílios Contínua: PNAD Contínua. Recuperado de <https://www.ibge.gov.br/estatisticas/sociais/trabalho/17270-pnad-continua.html>. Acesso em: 29 jun. 2023.

Instituto Brasileiro de Geografia e Estatística [IBGE]. 2022. Contas Regionais do Brasil 2010-2019. Recuperado de <https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/9054-contas-regionais-do-brasil.html>. Acesso em: 29 jun. 2023.

Instituto Brasileiro de Geografia e Estatística [IBGE] . 2022. Produto Interno Bruto dos Municípios 2010-2019. Recuperado de <https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/26359-produto-interno-bruto-dos-municipios-2010-2019.html?=&t=o-que-e>. Acesso em: 29 jun. 2023.

Martins, J. D. S. 2019. Análise da desigualdade dos últimos 15 anos no Distrito Federal. Disponível em: [https://repositorio.idp.edu.br/bitstream/123456789/2709/1/Jean da Silva Martins.pdf](https://repositorio.idp.edu.br/bitstream/123456789/2709/1/Jean%20da%20Silva%20Martins.pdf). Acesso em: 29 jun. 2023.

Martins, M. E. G.; Rodrigues, J. F. 2014. Coeficiente de correlação amostral. *Revista de Ciência Elementar*, 2(2), 34-36. Disponível em: <https://rce.casadasciencias.org/rceapp/static/docs/artigos/2014-042.pdf>. Acesso em: 29 jun. 2023.

Martins, T. C. S. 2014. Determinações do racismo no mercado de trabalho: implicações na “questão social” brasileira. *Temporalis*, 14(28), 113-132. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=5010758> Acesso em: 29 jun. 2023.

Medeiros, M.; Souza, P. H. G. F. D. 2016. A estabilidade da desigualdade no Brasil entre 2006 e 2012: resultados adicionais. Disponível em: https://repositorio.ipea.gov.br/bitstream/11058/7497/1/PPE_v46_n03_Estabilidade.pdf. Acesso em: 29 jun. 2023.

Menezes, R. G.; Kraychete, E. S. 2022. Desigualdade Global E Desenvolvimento. *Caderno CRH*, 35, e022001. Disponível em: <https://www.scielo.br/j/ccrh/a/vgpn8Qcn5JMkkmkyFXS35VD/?lang=pt>. Acesso em: 29 jul. 2023.

Melo, L. V. D. S. 2009. Democratização do acesso à educação superior pública no Distrito Federal-Universidade de Brasília/Faculdade UnB Planaltina. Disponível em: https://repositorio.unb.br/bitstream/10482/4094/1/2009_LiviaVeledaSousaMelo.pdf . Acessado em: 21 jun. 2023.

Müller, F., Monasterio, L. M.; Dutra, C. P. R. 2018. “Por que tão longe?”. Mobilidade de crianças e estrutura urbana no Distrito Federal. *Cadernos Metrópole*, 20, 577-598. Disponível em: <https://www.scielo.br/j/cm/a/5T7RrCWptJhGrZfnrmpFMp/abstract/?lang=pt>. Acesso em: 29 jun. 2023.

Neto, J. V., dos Santos, C. B., Torres, É. M., & Estrela, C. 2017. Boxplot: um recurso gráfico para a análise e interpretação de dados quantitativos. *Revista Odontológica do Brasil Central*, 26(76). Disponível em: <https://robrac.org.br/seer/index.php/ROBRAC/article/view/1132>. Acesso em: 29 jun. 2023.

Nunes, B. F., Costa, A. 2020. Distrito Federal e Brasília: dinâmica urbana, violência e heterogeneidade social. *Cadernos Metrópole*, 17, 35-57. Disponível em: <https://www.redalyc.org/pdf/4028/402837795002.pdf>. Acesso em: 29 jun. 2023.

Oliveira, G. A. S. D. 2022. A servidora pública e a desigualdade de gênero nos cargos de chefia: uma análise sob a ótica do viés do status quo (Doctoral dissertation). Disponível em: <https://bibliotecadigital.fgv.br/dspace/handle/10438/33001>. Acesso em: 29 jun. 2023.

Ortiz, R. M.; Lobato, A. O. C. 2018. ANÁLISE DA IGUALDADE DE GÊNERO NO BRASIL: IMPACTOS DOS INDICADORES SOCIAIS DIVULGADOS PELO IBGE EM 2018 NA PROMOÇÃO DA JUSTIÇA SOCIAL E DOS DIREITOS FUNDAMENTAIS DAS MULHERES. Disponível em:

http://www.editorarealize.com.br/editora/ebooks/senacorporus/2018/TRABALHO_EV103_MD4_SA10_ID663_15032018182801.pdf. Acesso em: 29 jun. 2023.

Pereira, R. H. M.; Schwanen, T. 2013. Tempo de deslocamento casa-trabalho no Brasil (1992-2009): diferenças entre regiões metropolitanas, níveis de renda e sexo (No. 1813). Texto para Discussão. Disponível em:

<https://www.econstor.eu/bitstream/10419/91337/1/744913446.pdf>. Acesso em: 29 jun. 2023.

Pinto, V. F. F.; Müller, F.; Anjos, J. J. T. D. 2018. Entre o passado e o presente: contrastes de acesso à Educação Infantil no Distrito Federal. Educação em Revista, 34. Disponível em:

<https://www.scielo.br/i/edur/a/JR7zfT5B3dzJvLbM5XZbWBm/?lang=pt&format=html>. Acesso em: 29 jun. 2023.

Quinzani, M. A. D. 2020. O avanço da pobreza e da desigualdade social como efeitos da crise da covid-19 e o estado de bem-estar social. Boletim de Conjuntura (BOCA), 2(6), 43-47. Disponível em:

<https://revista.ioles.com.br/boca/index.php/revista/article/download/121/119>. Acesso em: 29 jun. 2023.

Samarco Neto, W. R. 2013. Educação e renda: um estudo comparativo sobre os retornos de escolaridade no Estado de Goiás e no Distrito Federal. Disponível em:

https://repositorio.unb.br/bitstream/10482/15147/1/2013_WilsonRamosSamarcoNeto.pdf. Acessado em: 20 de set. 2023.

Soares, S. S. D. 2006. Análise de bem-estar e decomposição por fatores da queda na desigualdade entre 1995 e 2004. *Econômica*, 8(1). Acesso em:

<https://periodicos.uff.br/revistaeconomica/article/download/34915/20170>. Acessado em: 20 de set. 2023.

Wajnman, S.; Turra, C. M.; Agostinho, C. S. 2006. Estrutura domiciliar e distribuição da renda familiar no Brasil. *Desigualdade de Renda no Brasil: uma análise da queda recente*, 1, 423-442. Disponível em:

https://www.academia.edu/download/69761423/Estrutura_Domiciliar_e_Distribuio_da_Ren20210916-18171-10317qq.pdf. Acesso em: 20 de set. 2023.

Yanagida, E. 2022. Racismo Amarelo: o retrato de uma violência social no Brasil. Disponível em: <https://repositorio.animaeducacao.com.br/handle/ANIMA/24243>. Acesso em: 29 jun. 2023.