

Análise dos fatores associados à demissão de funcionários em empresas privadas do Nordeste.

Fernando Moraes
Edleide de Brito

18 de novembro de 2024

Sumário

- 1** Introdução
- 2 Metodologia
- 3 Resultados
- 4 Considerações Finais
- 5 Referências

Introdução

- Professor Substituto do Departamento de Estatística.
- Bolsa de pesquisa FAPESB.
- Doutorando em Estatística - USP/UFSCar.

Introdução

- **Iniciação científica voluntária. (2017).**
- Ideia inicial: Trabalhar com o dados de empresas baianas para verificar o tempo de sobrevivência.
- Objetivo: Identificar os fatores de risco para o desligamento de funcionários nas empresas da Bahia.

Introdução

- Iniciação científica voluntária. (2017).
- Ideia inicial: Trabalhar com o dados de empresas baianas para verificar o tempo de sobrevivência.
- Objetivo: Identificar os fatores de risco para o desligamento de funcionários nas empresas da Bahia.

Introdução

- Iniciação científica voluntária. (2017).
- Ideia inicial: Trabalhar com o dados de empresas baianas para verificar o tempo de sobrevivência.
- Objetivo: Identificar os fatores de risco para o desligamento de funcionários nas empresas da Bahia.

Introdução



Introdução

- RAIS (2010) – Fonte de dados para controle e estatísticas do mercado de trabalho no Brasil.
- IBGE – Pesquisas sobre características demográficas e socioeconômicas.
 - Análises descritivas - tempos de emprego.

Introdução

- RAIS (2010) – Fonte de dados para controle e estatísticas do mercado de trabalho no Brasil.
- IBGE – Pesquisas sobre características demográficas e socioeconômicas.
 - Análises descritivas - tempos de emprego.

Introdução

- Santos e Nakano (2015) – Modelos de regressão lognormal e de Cox para tempo de emprego (RAIS 2002-2009).
- Arruda et al. (2016) – Modelo probit para desemprego no Nordeste (PNAD 2003 e 2013).

Introdução

Análise de Sobrevivência

- Estima o tempo até a ocorrência de eventos de interesse. Considera observações censuradas.
- Menezes e Cunha (2014), a análise de sobrevivência é empregada para avaliar a duração do desemprego no Brasil.

Introdução

Análise de Sobrevivência

- Estima o tempo até a ocorrência de eventos de interesse. Considera observações censuradas.
- Menezes e Cunha (2014), a análise de sobrevivência é empregada para avaliar a duração do desemprego no Brasil.

Introdução

Objetivos

- Avaliar fatores de risco associados às demissões em empresas privadas no Nordeste.
- Aplicação de Análise de Sobrevivência com ajuste de quatro modelos paramétricos.
- Contribuição para a tomada de decisões que reduzam taxas de demissão.

Sumário

- 1 Introdução
- 2 Metodologia**
- 3 Resultados
- 4 Considerações Finais
- 5 Referências

Base de Dados

- Foi retirada do site RAIS (Relação Anual de Informações Socioeconômicas).
 - Alagoas, Bahia, Ceará, Maranhão, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe.
- *Software R* (versão 3.4.2).

Base de Dados

- Foi retirada do site RAIS (Relação Anual de Informações Socioeconômicas).
 - Alagoas, Bahia, Ceará, Maranhão, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe.
- *Software R* (versão 3.4.2).

Base de Dados

- Trabalhadores do Nordeste registrados na RAIS em 2015 (12,6 milhões).
- Inclui desligamentos em empresas privadas (demissão, término de contrato, etc.).
- Limpeza dos dados: remoção de observações inconsistentes (ex.: idade e tempo de emprego igual a zero, maior que 35 anos).

Base de Dados

- Variáveis utilizadas: tempo de emprego, motivo de desligamento, idade, carga horária, remuneração, etc.
- Inclusão de variáveis adicionais: porte e localização do estabelecimento (capital ou interior)

Causas de Desligamento

- Demissão;
- Término de contrato;
- Transferência;
- Falecimento;
- Aposentadoria;
- Posse em outro cargo inacumulável, Redistribuição, Readaptação.

Amostragem

Amostragem estratificada

- Divisão da população em estratos.

Amostragem Estratificada

$$n = \frac{z_{1-\alpha}^2 \sum_{H=1}^L \frac{N_H^2((1-P_H)P_H)}{W_H}}{N^2 d^2},$$

em que:

- $z_{1-\alpha}^2$: Quantil da distribuição normal padrão ($\alpha = 0,05$);
- d : Precisão (0,01);
- L : Número de estratos;
- P_H : Proporção de indivíduos no estrato H com a característica;
- N_H : Número total de elementos do estrato H;
- N : Número total de elementos;
- $W_H = \frac{N_H}{N}$.

Amostragem Estratificada

$$n = \frac{z_{1-\alpha}^2 \sum_{H=1}^L \frac{N_H^2((1-P_H)P_H)}{W_H}}{N^2 d^2},$$

em que:

- $z_{1-\alpha}^2$: Quantil da distribuição normal padrão ($\alpha = 0,05$);
- d : Precisão (0,01);
- L : Número de estratos;
- P_H : Proporção de indivíduos no estrato H com a característica;
- N_H : Número total de elementos do estrato H;
- N : Número total de elementos;
- $W_H = \frac{N_H}{N}$.

Estados do nordeste



Estados do nordeste

Tabela 1: Tamanho dos estratos.

Estratos	Número de indivíduos
Alagoas	500720
Bahia	2495585
Ceará	1649567
Maranhão	669703
Paraíba	558304
Pernambuco	1833927
Piauí	414658
Rio Grande do Norte	614759
Sergipe	403655

Análise de Sobrevivência

- Foi estimado o tempo até o funcionário ser desligado da empresa.
- Os modelos de análise de sobrevivência são caracterizados pela presença de censura, que é definida como a ausência do evento no tempo da análise.
- Censura ocorre quando o evento de interesse (demissão) não é observado.

Análise de Sobrevivência

- Foi estimado o tempo até o funcionário ser desligado da empresa.
- Os modelos de análise de sobrevivência são caracterizados pela presença de censura, que é definida como a ausência do evento no tempo da análise.
- Censura ocorre quando o evento de interesse (demissão) não é observado.

Análise de Sobrevivência

- Foi estimado o tempo até o funcionário ser desligado da empresa.
- Os modelos de análise de sobrevivência são caracterizados pela presença de censura, que é definida como a ausência do evento no tempo da análise.
- Censura ocorre quando o evento de interesse (demissão) não é observado.

Função de Sobrevivência

Função de Sobrevivência

- É uma das principais funções para descrever estudos de sobrevivência.
 - A probabilidade de uma observação não falhar até um tempo t .

$$S(t) = P(T \geq t).$$

Kaplan-Meier

- Função de sobrevivência empírica na presença de censura:

$$S(t) = \prod_{(j:t_j < t)} \left(1 - \frac{d_j}{n_j} \right),$$

em que:

- $t_1 < t_2 < \dots < t_k$: k tempos distintos e ordenados em falhas;
- d_j : número de falhas em t_j , $j = 1, \dots, k$;
- n_j : número de indivíduos sob risco em t_j .

Kaplan-Meier

- Função de sobrevivência empírica na presença de censura:

$$S(t) = \prod_{(j:t_j < t)} \left(1 - \frac{d_j}{n_j} \right),$$

em que:

- $t_1 < t_2 < \dots < t_k$: k tempos distintos e ordenados em falhas;
- d_j : número de falhas em t_j , $j = 1, \dots, k$;
- n_j : número de indivíduos sob risco em t_j .

- Os modelos mais utilizados em Análise de Sobrevida são o exponencial, o Weibull e o log-normal.

$$S_E(t) = e^{-\frac{t}{\alpha}}; S_W(t) = e^{-\left(\frac{t}{\alpha}\right)^\gamma}; S_{ln}(t) = \Phi\left(\frac{-\log(t) + \mu}{\sigma}\right),$$

em que:

- α : Tempo médio de vida na distribuição exponencial e o parâmetro de escala na distribuição Weibull;
- μ : Média do logaritmo do tempo de falha;
- σ : Desvio-padrão;
- γ : Parâmetro de forma;
- Φ : Função de distribuição acumulada da distribuição normal padrão.

- Os modelos mais utilizados em Análise de Sobrevida são o exponencial, o Weibull e o log-normal.

$$S_E(t) = e^{-\frac{t}{\alpha}}; S_W(t) = e^{-\left(\frac{t}{\alpha}\right)^\gamma}; S_{ln}(t) = \Phi\left(\frac{-\log(t) + \mu}{\sigma}\right),$$

em que:

- α : Tempo médio de vida na distribuição exponencial e o parâmetro de escala na distribuição Weibull;
- μ : Média do logaritmo do tempo de falha;
- σ : Desvio-padrão;
- γ : Parâmetro de forma;
- Φ : Função de distribuição acumulada da distribuição normal padrão.

- Os modelos mais utilizados em Análise de Sobrevida são o exponencial, o Weibull e o log-normal.

$$S_E(t) = e^{-\frac{t}{\alpha}}; S_W(t) = e^{-\left(\frac{t}{\alpha}\right)^\gamma}; S_{ln}(t) = \Phi\left(\frac{-\log(t) + \mu}{\sigma}\right),$$

em que:

- α : Tempo médio de vida na distribuição exponencial e o parâmetro de escala na distribuição Weibull;
- μ : Média do logaritmo do tempo de falha;
- σ : Desvio-padrão;
- γ : Parâmetro de forma;
- Φ : Função de distribuição acumulada da distribuição normal padrão.

Gamma Generalizada

$$f_{GG}(t) = \frac{\gamma}{\Gamma(\kappa)\alpha^{\gamma\kappa}} t^{\gamma\kappa-1} e^{-\frac{t}{\alpha}^{\gamma}} \quad \text{e} \quad S_{GG}(t) = \int_t^{\infty} \frac{\gamma}{\Gamma(\kappa)\alpha^{\gamma\kappa}} u^{\gamma\kappa-1} e^{-\frac{u}{\alpha}^{\gamma}} du$$

- α é o parâmetro de escala;
- κ e γ os parâmetros de forma;
- $\Gamma(\kappa) = \int_0^{\infty} x^{\kappa-1} e^{-x} dx$.

Método de Máxima Verossimilhança

$$L(\theta) = \prod_{(i=1)}^n [f(t_i; \theta)]^{\delta_i} [S(t_i; \theta)]^{1-\delta_i},$$

em que:

- δ_i : Variável indicadora de falha (censura);
- t_i : Tempo até o evento de interesse;
- θ : Vetor de parâmetros a serem estimados;
- $f(t_i; \theta)$: Função densidade de probabilidade.

Método de Máxima Verossimilhança

$$L(\theta) = \prod_{(i=1)}^n [f(t_i; \theta)]^{\delta_i} [S(t_i; \theta)]^{1-\delta_i},$$

em que:

- δ_i : Variável indicadora de falha (censura);
- t_i : Tempo até o evento de interesse;
- θ : Vetor de parâmetros a serem estimados;
- $f(t_i; \theta)$: Função densidade de probabilidade.

Seleção de Modelo

- Critério de Informação de Akaike (AIC):

$$AIC = -2\log[L(\theta)] - 2p$$

em que:

- $L(\theta)$: Função de verossimilhança;
- p : Número de parâmetros do modelo.

Seleção de Modelo

- Critério de Informação de Akaike (AIC):

$$AIC = -2\log[L(\theta)] - 2p$$

em que:

- $L(\theta)$: Função de verossimilhança;
- p : Número de parâmetros do modelo.

Seleção de Modelo

- Critério de Informação de Akaike (AIC):

$$AIC = -2\log[L(\theta)] - 2p$$

em que:

- $L(\theta)$: Função de verossimilhança;
- p : Número de parâmetros do modelo.

Análise de Resíduo

Worm Plots

- Avaliar a distribuição empírica via quantis teóricos.
- Buuren e Fredriks (2001)
- Identificar regiões com possíveis violações no modelo.
- **Interpretação:**
 - Eixo vertical: diferença entre distribuições teórica e empírica
 - Pontos dentro das faixas de confiança indicam adequação

Sumário

- 1 Introdução
- 2 Metodologia
- 3 Resultados**
- 4 Considerações Finais
- 5 Referências

Resultados

- Base de dados inicial com mais de 9 milhões de observações e 21 variáveis
- Análise descritiva para identificar padrões e comportamentos.

Resultados

Tabela 2: Características dos empregados de empresas com vínculo não público da Região Nordeste em 2015.

	Funcionários	%
Situação		
Desligado	2939611	32.2
Não Desligado	6201267	67.8
Situação		
Admitido	3093230	33.8
Não Admitido	6047648	66.2
Total	9140878	100.0

Resultados

Tabela 3: Características dos empregados de empresas com vínculo não público da Região Nordeste em 2015.

	Funcionários	%
Sexo		
Feminino	3224328	35.3
Masculino	5916550	64.7
Região		
Capitais	4270901	46.7
Região metropolitana	1556239	17.0
Interior	3313738	36.3
Total	9140878	100.0

Resultados

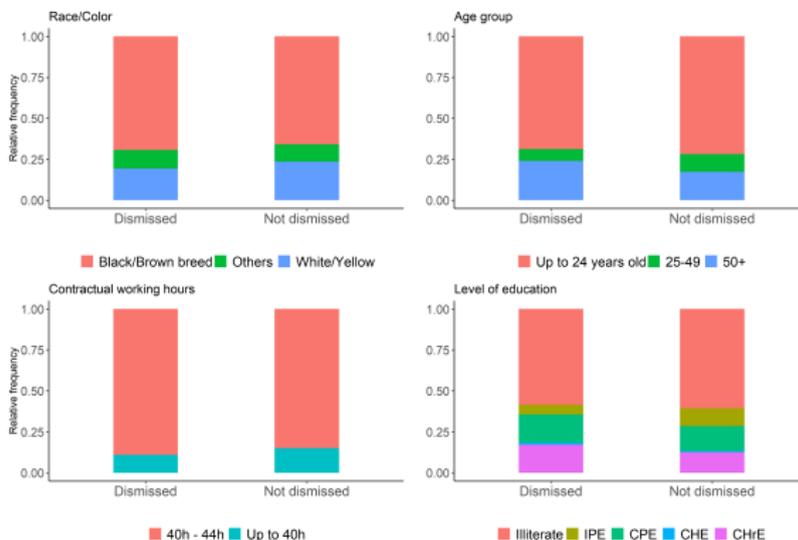


Figura 1: Situação dos empregados por categorias de variáveis qualitativas.

Resultados

- Para comparar o efeito das categorias de variáveis no tempo de sobrevivência dos funcionários de empresas não públicas no Nordeste, foram gerados gráficos com as curvas de sobrevivência para as covariáveis.

Resultados

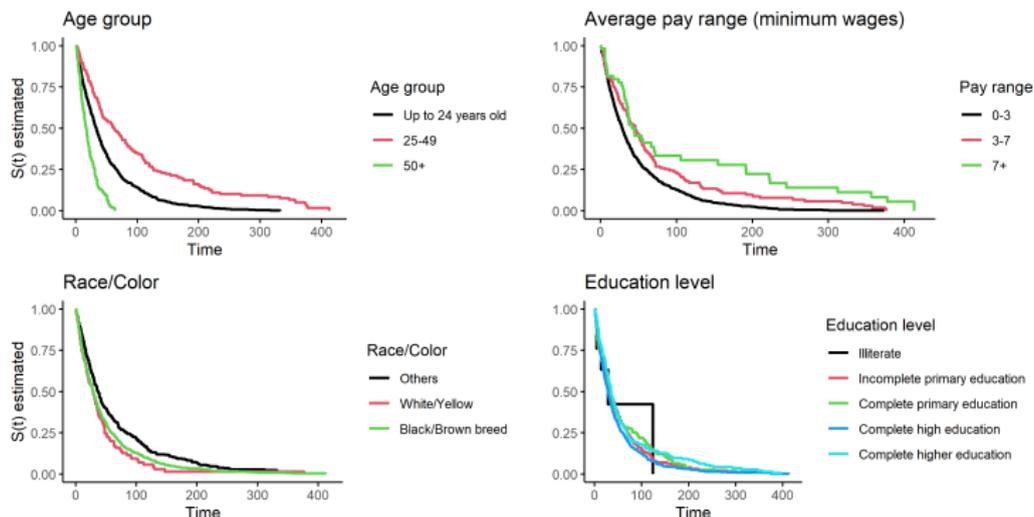


Figura 2: Curvas de sobrevivência para covariáveis estimadas por Kaplan-Meier.

Resultados

Tabela 4: Size of strata.

Estratos	Region Total (N_h)	(n_h)
Alagoas	500720	115
Bahia	2495585	572
Ceará	1649567	379
Maranhão	669703	154
Paraíba	558304	128
Pernambuco	1833927	421
Piauí	414658	96
Rio Grande do Norte	614759	141
Sergipe	403655	93

Resultados

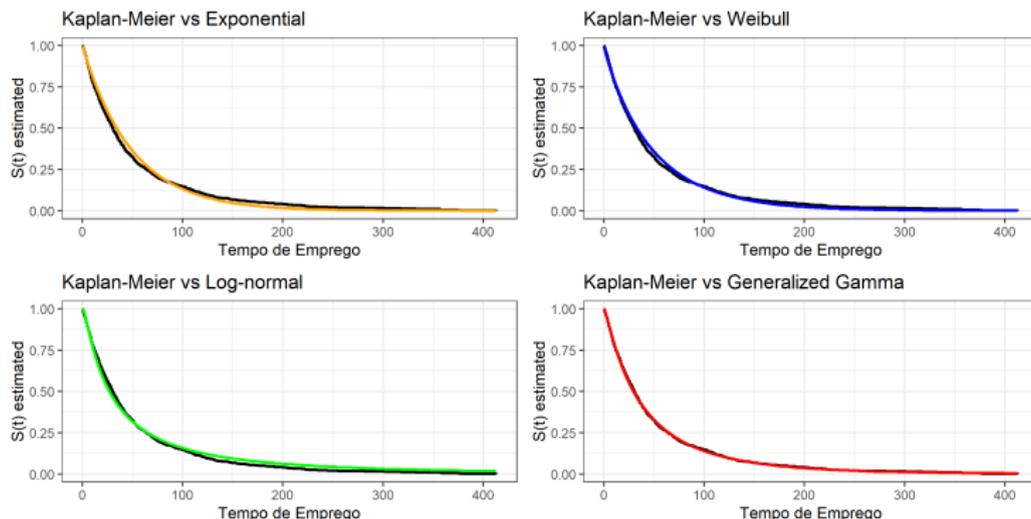


Figura 3: Curvas de sobrevivência estimadas pelos modelos versus curva de sobrevivência estimada por Kaplan-Meier, sem adição de covariáveis.

Resultados

Tabela 5: Comparação dos modelos pelo critério de informação de Akaike.

Modelos	AIC
Exponencial	2288.37
Weibull	2288,93
Log-normal	2286,53
Gama Generalizada	2279.50

Resultados

- O modelo Gama Generalizada com quatro covariáveis é definido por:

$$\mu(\mathbf{X}) = \mathbf{X}'\boldsymbol{\beta} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4,$$

Estimativas dos parâmetros

Tabela 6: Estimativas dos parâmetros do modelo gama generalizado.

Variáveis		Estimativas
β_0		2, 133
Sexo	Masculino	-
	Feminino	0, 048
Raça/cor	Outros	-
	Preto/Parda	0, 061
	Branca/Amarela	0, 214
Idade		0, 035
Remun. média (SM)	0-3	-
	3-7	0, 292
	7+	0, 424

Estimativas dos parâmetros

Tabela 7: Estimativas dos parâmetros de forma e escala.

Estimador	Estimativa
$\hat{\alpha}$	4,765
$\hat{\gamma}$	0,758
$\hat{\kappa}$	1,765

Exemplo

- Supondo que queremos analisar o tempo de permanência no emprego de uma mulher branca, com 25 anos, e salário médio de 3 salários mínimos. Precisamos:

Resultados

$$\mu(\mathbf{X}) = \mathbf{X}'\beta = 2.133 + 0.048 + 0.214 + 0.035 * 25 + 0.292 = 3.562$$

$$\hat{S}_{GG}(t|x) = 1 - \frac{0,759}{\Gamma(1,765)} \int_0^t z^{0,34} e^{-(\frac{z}{4,765})^{0,758}} dz = 71.10\%$$

Resultados

$$\mu(\mathbf{X}) = \mathbf{X}'\beta = 2.133 + 0.048 + 0.214 + 0.035 * 25 + 0.292 = 3.562$$

$$\hat{S}_{GG}(t|x) = 1 - \frac{0,759}{\Gamma(1,765)} \int_0^t z^{0,34} e^{-(\frac{z}{4,765})^{0,758}} dz = 71.10\%$$

Resultados

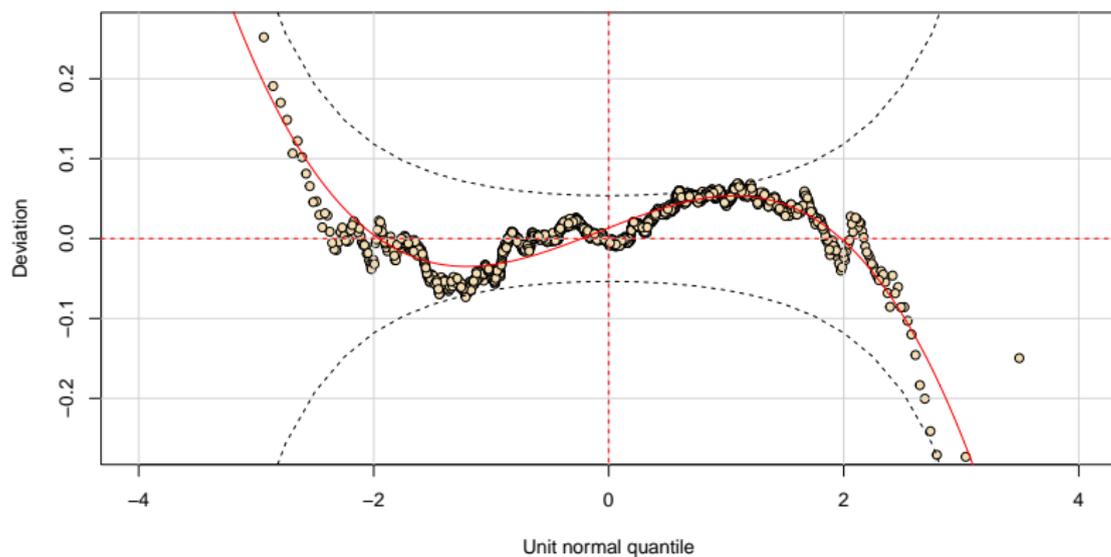


Figura 4: Warm Plot para o modelo gama generalizado.

Sumário

- 1 Introdução
- 2 Metodologia
- 3 Resultados
- 4 Considerações Finais**
- 5 Referências

Considerações Finais

- Comparação de modelos de Análise de Sobrevivência para avaliar fatores de risco associados à demissão de funcionários em empresas não públicas do Nordeste.
- **Resultados:**
 - Trabalhadores majoritariamente homens, pardos, brasileiros, de 18 a 39 anos, com salário até 3 salários mínimos.
 - Trabalham em capitais, com carga horária entre 41 e 44 horas, permanecendo em média 3 meses.
 - Terminações por iniciativa do empregador, principalmente sem justa causa.

Considerações Finais

- Compreensão dos fatores que influenciam a saída de funcionários, auxiliando empresas e formuladores de políticas públicas.
- Efeito positivo da educação e experiência na duração do emprego.
- Políticas que incentivem jovens a concluir estudos e a buscar formação alinhada ao mercado.

Considerações Finais

Limitações

- Grande base de dados.

Considerações Finais

Publicações

- Análise dos fatores associados à demissão de funcionários em empresas privadas do Nordeste - Revista Brasileira de Economia de Empresas (2025).

Considerações Finais

Riscos Competitivos

- **Eventos competitivos:** situações onde múltiplos eventos podem ocorrer, mas apenas um será observado
 - Demissão sem justa causa e desligamento sem justa causa como eventos mutuamente exclusivos

Sumário

- 1 Introdução
- 2 Metodologia
- 3 Resultados
- 4 Considerações Finais
- 5 Referências**

Referências

- AKAIKE, H. A new look at the statistical model identification. IEEE transactions on Automatic Control, v.19, n.6, p716-723,1974.
- BOLFARINE, H.; BUSSAB, W. O. Elementos de Amostragem. São Paulo: ABE - Projeto Fisher, Edgard Blücher, 2005.
- IBGE, Instituto Brasileiro de Geografia e Estatística. Disponível em: <https://www.ibge.gov.br/>. Acesso em: 19 fevereiro de 2018.
- R development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Disponível em: <https://www.r-project.org/>. Acesso em: 19 fevereiro de 2018.
- RAIS, Relação Anual de Informações Sociais. Disponível

Obrigado!

Dúvidas ou perguntas?

Contato:

- Email: fernando.humberto@ufba.br
- LinkedIn: <https://www.linkedin.com/in/fernandomoraesds/>

Sinta-se à vontade para entrar em contato!