

## ➔ Introdução

O Atlas do Desenvolvimento Humano no Brasil é um órgão que mede o IDH (Índice de Desenvolvimento Humano) nas regiões do país, com intuito de conhecer a realidade deste tema no território nacional. Conforme definições do próprio órgão, desenvolvimento humano é definido como “o processo de ampliação das escolhas e liberdades das pessoas para que elas tenham capacidades e oportunidades para serem aquilo que desejam ser.”. Para medir o índice, são considerados indicadores das principais áreas que impactam a qualidade de vida do ser humano: saúde, educação e renda.

O intuito da proposta deste estudo é verificar se determinadas variáveis pré-dispostas na base de dados coletada, tem influência e explicam as variações sobre a renda percapita atribuída a cada município.

Premissas: escolher e explorar 6 variáveis INDEPENDENTES dipostas na base de dados que possam explicar as variações na renda percapita.

As ferramentas para desenvolver o estudo serão a interface de desenvolvimento (IDE) RSTUDIO junto com o interpretador de comandos do R.

## ➔ Procedimentos de coleta

Foi efetuado download da base de dados do “Atlas do Município” (link: [http://www.atlasbrasil.org.br/2013/data/rawData/atlas2013\\_dadosbrutos\\_pt.xlsx](http://www.atlasbrasil.org.br/2013/data/rawData/atlas2013_dadosbrutos_pt.xlsx)), e considerado as variáveis para todos municípios disponíveis.

A base de dados em estudo contém a seguinte estrutura:

```
> cat("Quantidade de observacoes (LINHAS):", dim(atlas2013_dadosbrutos_pt_somente_municipio)[1])
Quantidade de observacoes (LINHAS): 16695
> cat("Quantidade de variaveis (COLUNAS):", dim(atlas2013_dadosbrutos_pt_somente_municipio)[2])
Quantidade de variaveis (COLUNAS): 237
```

Dessas 237 variáveis, foram selecionadas aleatoriamente 6, conforme premissas do estudos. São elas:

```
ESPVIDA (Esperanca de vida ao nascer)> cat("GINI (Indice de Gini)")
GINI (Indice de Gini)> cat("FECTOT (Taxa de fecundidade total)")
FECTOT (Taxa de fecundidade total)> cat("T_FBSUPER (Taxa de frequencia bruta ao ensino superior)")
T_FBSUPER (Taxa de frequencia bruta ao ensino superior)> cat("PMPOB (Proporcao de pobres)")
PMPOB (Proporcao de pobres)> cat("IDHM (Indice de Desenvolvimento Humano Municipal)")
IDHM (Indice de Desenvolvimento Humano Municipal)
```

Além de considerar a variável que queremos explicar a sua variação:

```
RDPC (Renda per capita media)
```

Logo, será criado um “novo” dataset de estudo que conterá apenas as variáveis citadas acima: ESPVIDA, GINI, FECTOT, T\_FBSUPER, PMPOB, IDHM, RDPC.

## ➔ Cálculos e Análise Estatística

Análise exploratória nas medidas de posição:

```
--->> MEDIDAS DE POSICAO <<---  
> summary(df_01)
```

ESPVIDA	GINI	FECTOT	T_FBSUPER	PMPOB	IDHM	RDPC
Min. :50.97	Min. :0.2700	Min. :1.21	Min. : 0.000	Min. : 0.00	Min. :0.1200	Min. : 33.24
1st Qu.:64.83	1st Qu.:0.4700	1st Qu.:2.21	1st Qu.: 1.690	1st Qu.:17.36	1st Qu.:0.4090	1st Qu.: 182.10
Median :69.19	Median :0.5200	Median :2.63	Median : 6.120	Median :38.45	Median :0.5390	Median : 299.38
Mean :68.41	Mean :0.5223	Mean :2.93	Mean : 9.486	Mean :40.32	Mean :0.5213	Mean : 355.66
3rd Qu.:72.65	3rd Qu.:0.5700	3rd Qu.:3.36	3rd Qu.:14.360	3rd Qu.:61.73	3rd Qu.:0.6385	3rd Qu.: 487.24
Max. :78.64	Max. :0.9200	Max. :8.68	Max. :76.780	Max. :98.27	Max. :0.8620	Max. :2043.74

```
Exibindo a MODA para cada um dos campos do dataframe:  
> cat("ESPVIDA:", statmod(df_01$ESPVIDA))  
ESPVIDA: 66.42  
> cat("GINI:", statmod(df_01$GINI))  
GINI: 0.53  
> cat("FECTOT:", statmod(df_01$FECTOT))  
FECTOT: 2.37  
> cat("T_FBSUPER:", statmod(df_01$T_FBSUPER))  
T_FBSUPER: 0  
> cat("PMPOB:", statmod(df_01$PMPOB))  
PMPOB: 6.84 25.17  
> cat("IDHM:", statmod(df_01$IDHM))  
IDHM: 0.592  
> cat("RDPC:", statmod(df_01$RDPC))  
RDPC: 142.25 150.04 244.39 292.96
```

Análise exploratória nas medidas de dispersão:

```
Exibindo a AMPLITUDE para cada um dos campos do dataframe:  
> cat("ESPVIDA:", max(df_01$ESPVIDA) - min(df_01$ESPVIDA))  
ESPVIDA: 27.67  
> cat("GINI:", max(df_01$GINI) - min(df_01$GINI))  
GINI: 0.65  
> cat("FECTOT:", max(df_01$FECTOT) - min(df_01$FECTOT))  
FECTOT: 7.47  
> cat("T_FBSUPER:", max(df_01$T_FBSUPER) - min(df_01$T_FBSUPER))  
T_FBSUPER: 76.78  
> cat("PMPOB:", max(df_01$PMPOB) - min(df_01$PMPOB))  
PMPOB: 98.27  
> cat("IDHM:", max(df_01$IDHM) - min(df_01$IDHM))  
IDHM: 0.742  
> cat("RDPC:", max(df_01$RDPC) - min(df_01$RDPC))  
RDPC: 2010.5
```

```

Exibindo a VARIANCIA para cada um dos campos do dataframe:
> cat("ESPVIDA:", var(df_01$ESPVIDA, na.rm = FALSE))
ESPVIDA: 29.65313
> cat("GINI:", var(df_01$GINI, na.rm = FALSE))
GINI: 0.005222114
> cat("FECTOT:", var(df_01$FECTOT, na.rm = FALSE))
FECTOT: 1.167431
> cat("T_FBSUPER:", var(df_01$T_FBSUPER, na.rm = FALSE))
T_FBSUPER: 101.6045
> cat("PMPOB:", var(df_01$PMPOB, na.rm = FALSE))
PMPOB: 653.4381
> cat("IDHM:", var(df_01$IDHM, na.rm = FALSE))
IDHM: 0.02176648
> cat("RDPC:", var(df_01$RDPC, na.rm = FALSE))
RDPC: 50246.22

```

```

Exibindo o DESVIO PADRAO (= variancia ao quadrado) para cada um dos campos do dataframe:
> cat("ESPVIDA:", sd(df_01$ESPVIDA, na.rm = FALSE))
ESPVIDA: 5.445469
> cat("GINI:", sd(df_01$GINI, na.rm = FALSE))
GINI: 0.07226419
> cat("FECTOT:", sd(df_01$FECTOT, na.rm = FALSE))
FECTOT: 1.080477
> cat("T_FBSUPER:", sd(df_01$T_FBSUPER, na.rm = FALSE))
T_FBSUPER: 10.0799
> cat("PMPOB:", sd(df_01$PMPOB, na.rm = FALSE))
PMPOB: 25.56244
> cat("IDHM:", sd(df_01$IDHM, na.rm = FALSE))
IDHM: 0.1475347
> cat("RDPC:", sd(df_01$RDPC, na.rm = FALSE))
RDPC: 224.1567

```

```

Exibindo o COEFICIENTE DE VARIACAO em % (= % desvio padrao em relacao a media) para cada um dos campos do dataframe:
> cat("ESPVIDA:", 100*sd(df_01$ESPVIDA)/mean(df_01$ESPVIDA))
ESPVIDA: 7.95986
> cat("GINI:", 100*sd(df_01$GINI)/mean(df_01$GINI))
GINI: 13.83579
> cat("FECTOT:", 100*sd(df_01$FECTOT)/mean(df_01$FECTOT))
FECTOT: 36.88035
> cat("T_FBSUPER:", 100*sd(df_01$T_FBSUPER)/mean(df_01$T_FBSUPER))
T_FBSUPER: 106.2602
> cat("PMPOB:", 100*sd(df_01$PMPOB)/mean(df_01$PMPOB))
PMPOB: 63.39289
> cat("IDHM:", 100*sd(df_01$IDHM)/mean(df_01$IDHM))
IDHM: 28.2993
> cat("RDPC:", 100*sd(df_01$RDPC)/mean(df_01$RDPC))
RDPC: 63.02563

```

Para verificar os valores **OUTLIERS** (fora do padrão), foi utilizado o “Teste Z” ou “Z Score” (escore padronizado). Como forma de classificar se determinado valor é outlier, foi considerada a seguinte regra:

Se o valor do cálculo do z-score for  $\leq -3.0$  OU  $\geq +3.0$ , é considerado um valor fora da curva. Essa regra é indicada no livro “Estatística Aplicada à administração e economia” (Editora CENGAGE Learning, sexta edição norte-americana. Autores: Dennis J. Sweeney, Thomas A. Williams, David R. Anderson).

Calculando valores fora da curva (OUTLIERS) com Z-SCORE (valor da posição - média / desvio padrão) foram encontrados resultados. Seguem amostras do resultado para cada uma das variáveis:

ESPVIDA:

ESPVIDA	GINI	FECTOT	T_FBSUPER	PMPOB	IDHM	RDPC	Z_SCORE_ESPVIDA
52.04	0.39	5.61	0.00	87.10	0.213	80.40	-3.006465
52.04	0.40	5.91	0.00	95.47	0.208	60.72	-3.006465
51.18	0.40	5.24	0.00	79.81	0.241	90.32	-3.164395
51.69	0.59	5.84	0.00	87.35	0.277	95.79	-3.070739
51.42	0.58	4.72	1.63	86.98	0.251	96.45	-3.120322
51.87	0.49	7.34	0.00	89.12	0.168	54.80	-3.037684
51.87	0.55	7.34	1.68	90.36	0.246	76.97	-3.037684
51.86	0.44	5.60	2.09	81.03	0.256	99.53	-3.039520
51.26	0.47	6.12	0.00	91.78	0.217	90.92	-3.149704
51.86	0.42	4.75	0.96	80.23	0.268	107.62	-3.039520
51.61	0.49	5.56	0.00	85.25	0.227	87.53	-3.085430
51.47	0.44	5.50	0.00	85.55	0.133	84.28	-3.111140
50.97	0.53	6.67	0.62	90.08	0.204	63.20	-3.202959
51.32	0.53	4.69	0.00	86.03	0.230	82.00	-3.138685
51.32	0.56	4.82	1.30	76.63	0.290	137.65	-3.138685
51.32	0.60	5.94	0.00	84.35	0.256	110.58	-3.138685
51.48	0.48	5.85	0.00	90.52	0.191	71.78	-3.109303

GINI:

ESPVIDA	GINI	FECTOT	T_FBSUPER	PMPOB	IDHM	RDPC	Z_SCORE_GINI
63.27	0.78	4.53	1.29	72.76	0.277	239.14	3.566097
59.95	0.75	4.00	0.00	82.12	0.216	104.05	3.150953
63.41	0.76	7.81	0.00	80.53	0.246	95.57	3.289334
56.56	0.83	6.79	0.00	86.64	0.249	69.24	4.258002
56.56	0.81	6.12	1.23	80.88	0.278	109.61	3.981240
59.44	0.86	4.38	0.00	88.67	0.201	56.02	4.673146
56.38	0.28	4.47	0.00	90.14	0.314	89.16	-3.352959
55.74	0.27	4.72	0.00	91.16	0.294	76.01	-3.491340
62.81	0.77	5.16	0.19	77.64	0.376	202.48	3.427716
60.99	0.76	4.21	0.67	91.96	0.292	57.45	3.289334
53.60	0.83	6.12	0.00	96.77	0.131	63.89	4.258002
61.76	0.92	3.26	0.67	93.41	0.321	118.34	5.503432
59.64	0.81	5.01	0.00	96.31	0.230	77.42	3.981240
57.40	0.84	4.60	0.84	90.53	0.273	174.75	4.396383
60.57	0.74	5.66	0.00	77.86	0.376	193.84	3.012572
59.59	0.79	5.89	1.81	84.85	0.305	245.04	3.704478

FECTOT:

ESPVIDA	GINI	FECTOT	T_FBSUPER	PMPOB	IDHM	RDPC	Z_SCORE_FECTOT
64.10	0.57	7.06	3.21	62.55	0.398	194.63	3.822678
60.04	0.55	6.26	0.00	74.77	0.258	139.30	3.082264
61.64	0.39	6.31	0.00	93.79	0.170	63.50	3.128540
62.33	0.47	7.84	2.01	75.10	0.298	115.12	4.544581
60.74	0.42	7.84	0.00	88.53	0.186	75.84	4.544581
60.74	0.45	7.84	0.00	89.57	0.176	72.49	4.544581
62.33	0.48	7.84	0.00	84.30	0.245	88.55	4.544581
61.68	0.57	6.31	0.59	77.69	0.274	131.44	3.128540
63.41	0.63	7.30	0.46	68.13	0.306	165.64	4.044802
60.91	0.70	7.81	0.00	78.50	0.254	112.32	4.516815
60.40	0.53	6.90	0.00	64.56	0.307	184.60	3.674595
60.71	0.46	6.86	0.00	59.72	0.357	168.24	3.637574
63.41	0.76	7.81	0.00	80.53	0.246	95.57	4.516815
61.87	0.43	6.63	0.00	48.02	0.317	200.70	3.424706
62.17	0.52	6.18	0.00	70.39	0.293	122.12	3.008223
62.31	0.73	6.73	1.02	96.13	0.300	98.76	3.517257
60.40	0.49	7.64	0.80	81.75	0.272	109.23	4.359478
60.95	0.49	7.43	0.64	77.19	0.287	127.87	4.165119

T\_FBSUPER:

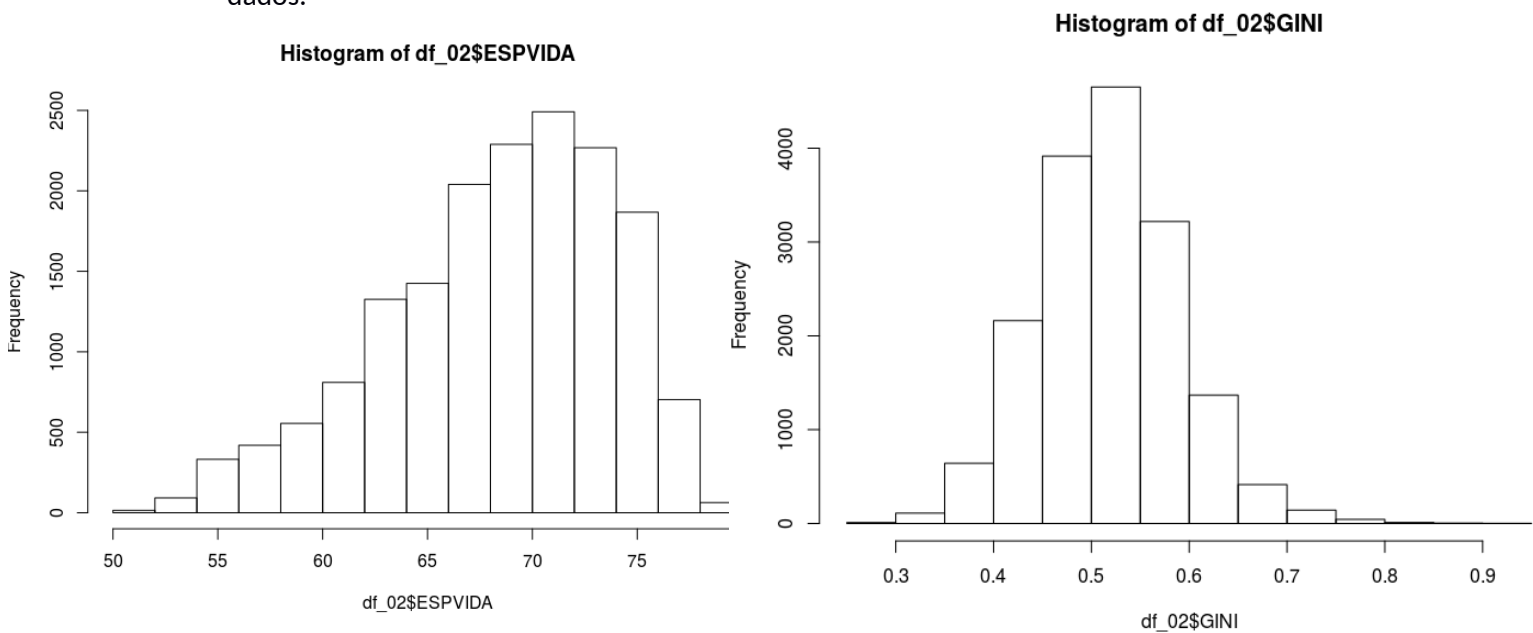
ESPVIDA	GINI	FECTOT	T_FBSUPER	PMPOB	IDHM	RDPC	Z_SCORE_T_FBSUPER
72.26	0.58	1.61	44.60	6.60	0.771	1596.51	3.483559
72.16	0.54	2.02	41.40	18.44	0.668	490.25	3.166096
77.18	0.50	1.56	46.71	1.38	0.820	1639.93	3.692886
74.35	0.56	1.68	47.44	5.31	0.766	1383.78	3.765308
73.79	0.54	2.18	41.61	18.82	0.683	582.20	3.186929
73.65	0.60	1.79	42.01	8.49	0.744	1399.50	3.226612
73.81	0.49	1.89	41.96	15.76	0.685	535.41	3.221652
73.95	0.58	2.18	45.32	13.61	0.752	786.55	3.554988
74.23	0.56	1.93	42.51	11.18	0.752	737.63	3.276216
75.60	0.54	1.97	45.00	7.14	0.759	778.90	3.523242
74.56	0.54	2.24	42.83	13.79	0.740	621.10	3.307962
74.61	0.58	2.04	53.43	6.91	0.788	1087.35	4.359559
73.00	0.56	1.87	40.04	17.93	0.698	563.88	3.031174
74.22	0.61	1.42	43.42	14.60	0.751	757.57	3.366494
74.52	0.55	1.62	48.47	10.91	0.766	850.44	3.867491
75.08	0.61	1.58	43.51	10.50	0.763	950.34	3.375423
74.31	0.70	1.85	43.14	15.61	0.748	1036.21	3.338716
73.73	0.58	1.72	39.98	16.34	0.720	630.03	3.025221
74.89	0.62	1.49	48.58	11.59	0.763	964.82	3.878404
75.36	0.46	1.43	48.26	2.20	0.788	1034.14	3.846658
74.50	0.68	1.35	43.71	13.20	0.772	1144.26	3.395264
74.36	0.62	1.44	56.94	11.11	0.770	1052.03	4.707777

RDPC:

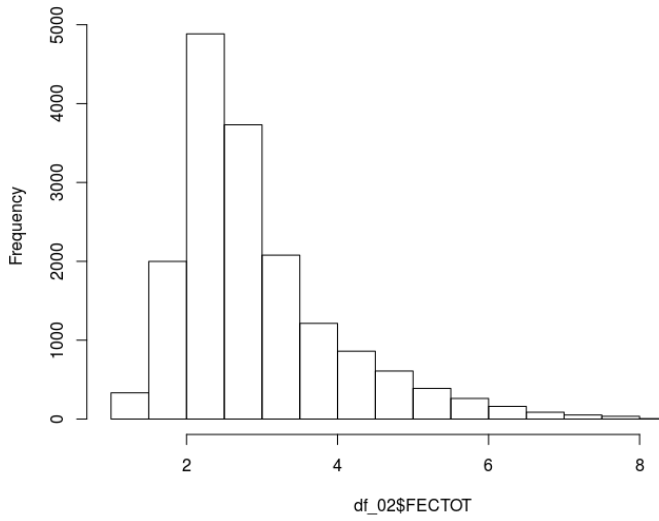
ESPVIDA	GINI	FECTOT	T_FBSUPER	PMPOB	IDHM	RDPC	Z_SCORE_RDPC
68.03	0.57	1.76	30.58	12.07	0.681	1090.93	3.280163
72.08	0.54	2.56	12.13	2.80	0.634	1185.28	3.701074
69.79	0.71	3.56	6.45	15.74	0.554	1167.62	3.622290
71.52	0.52	1.84	24.16	4.17	0.689	1075.13	3.209676
72.11	0.48	2.03	24.98	3.48	0.697	1107.53	3.354218
69.51	0.56	2.04	13.37	6.29	0.626	1050.76	3.100958
74.75	0.50	3.55	1.11	0.98	0.694	1104.89	3.342441
72.03	0.61	1.65	24.97	10.57	0.726	1101.96	3.329369
70.74	0.60	1.69	37.42	10.30	0.759	1315.63	4.282587
72.26	0.58	1.61	44.60	6.60	0.771	1596.51	5.535639
70.26	0.61	1.87	28.20	8.85	0.716	1187.08	3.709104
77.18	0.57	1.91	30.77	0.70	0.791	1759.76	6.263924
73.07	0.57	1.78	27.55	6.24	0.735	1223.77	3.872784
74.34	0.53	1.96	21.84	4.90	0.744	1089.92	3.275657
72.46	0.56	1.89	24.59	5.95	0.733	1070.28	3.188040
73.29	0.71	2.31	18.86	11.53	0.713	1565.51	5.397343
72.46	0.52	2.00	24.39	6.02	0.738	1029.13	3.004463
73.58	0.53	1.57	37.02	3.49	0.785	1441.86	4.845719
77.18	0.50	1.56	46.71	1.38	0.820	1639.93	5.729342

As variáveis PMPOB e IDHM não apresentaram valores fora da curva.

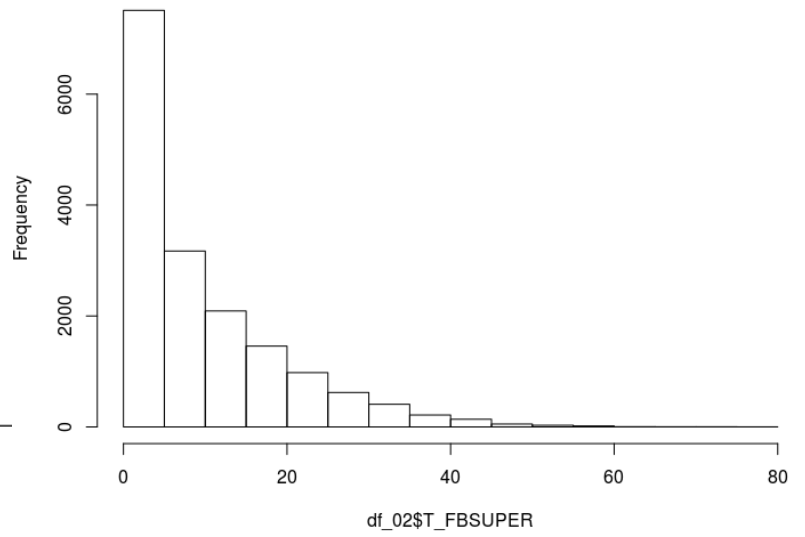
Plotando **gráficos de frequência** para cada uma das variáveis, para avaliar a distribuição dos dados:



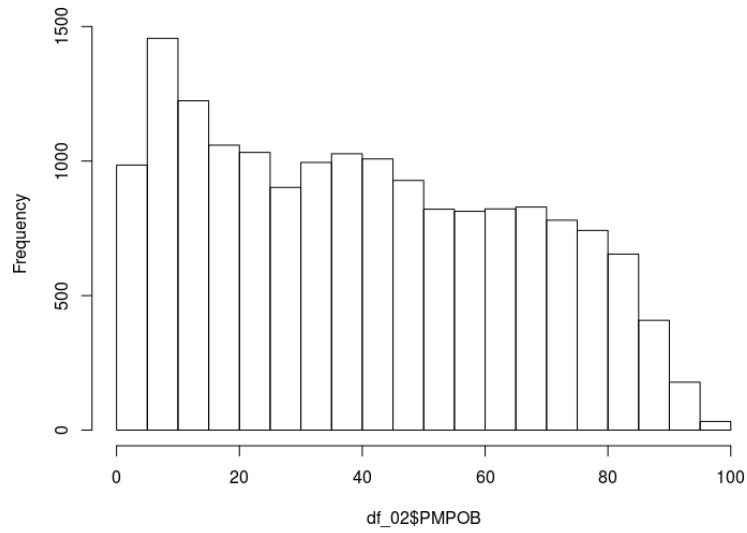
**Histogram of df\_02\$FECTOT**



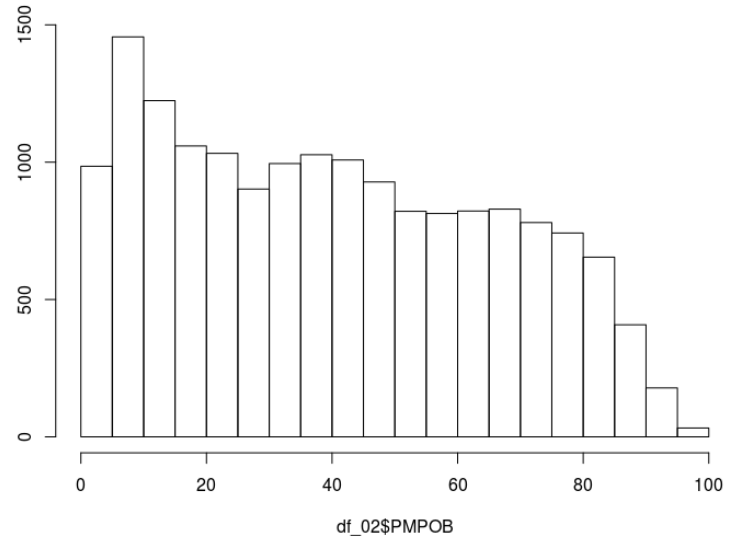
**Histogram of df\_02\$T\_FBSUPER**



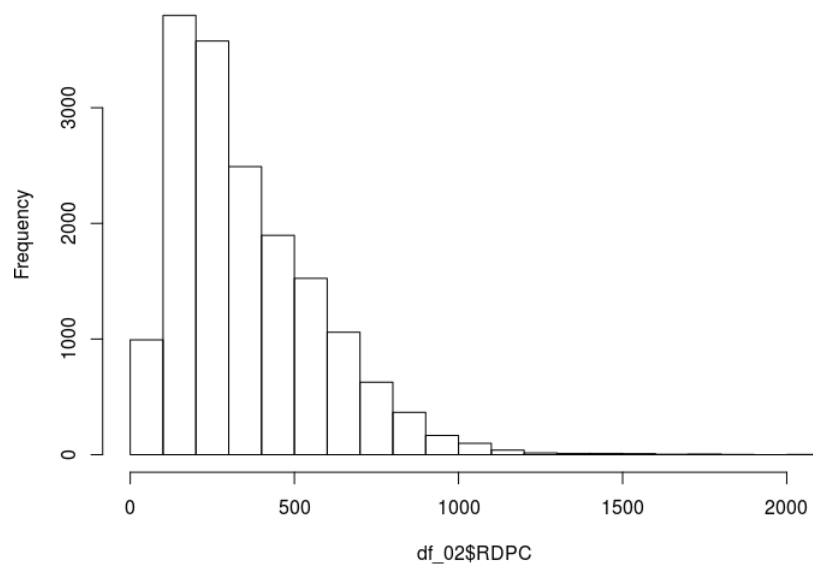
**Histogram of df\_02\$PMPOB**



**Histogram of df\_02\$PMPOB**



**Histogram of df\_02\$RDPC**

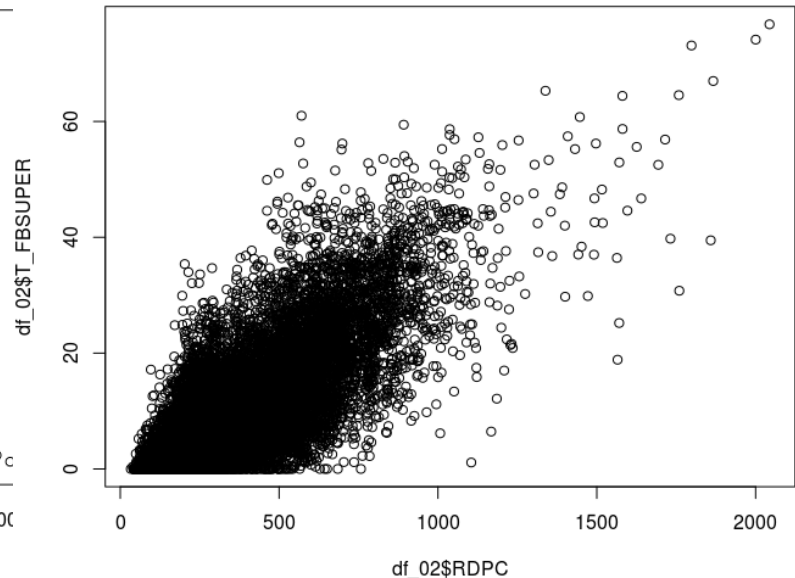
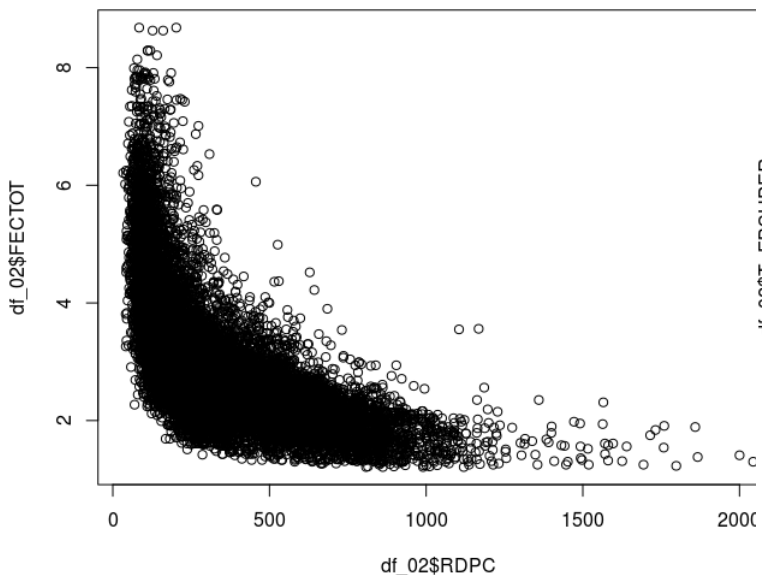
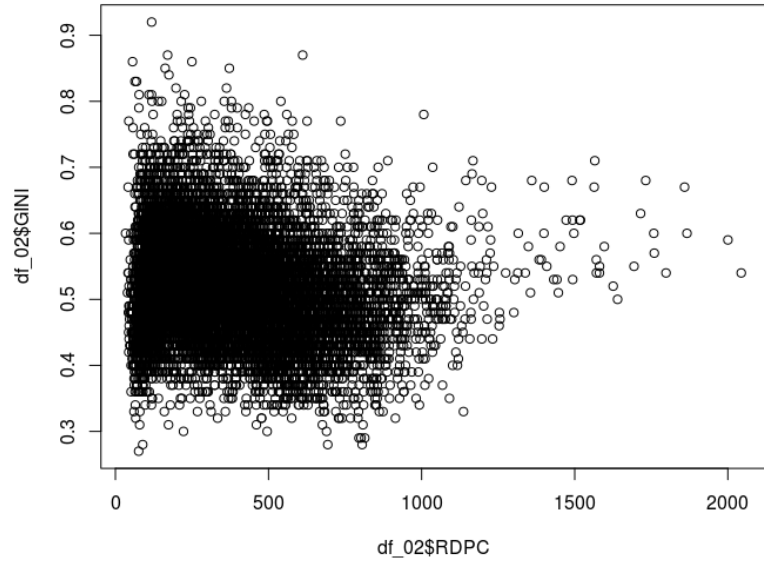
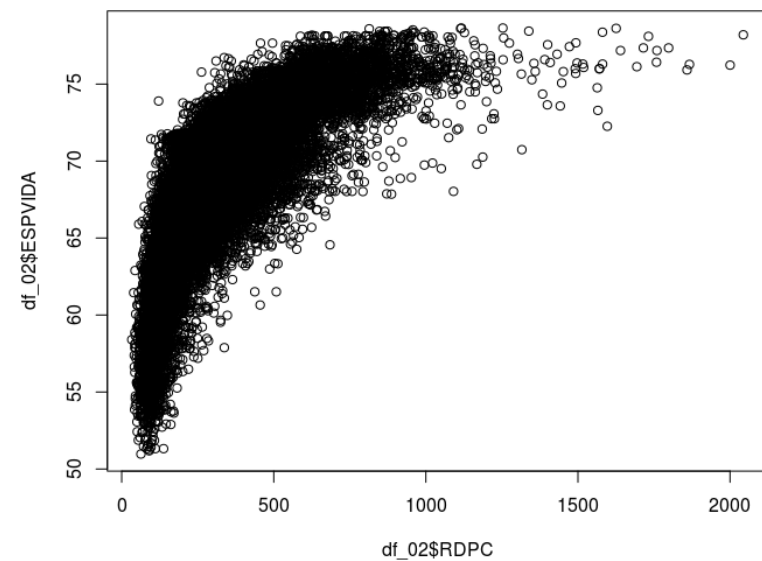




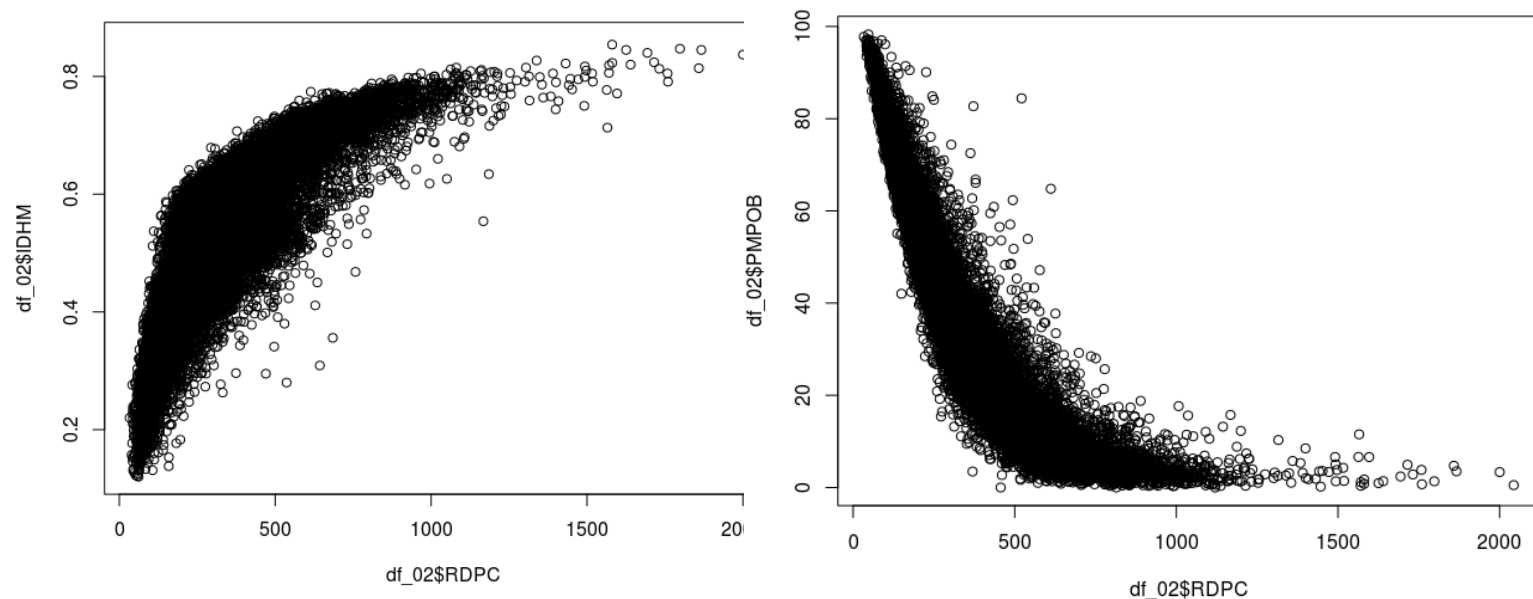
Averiguando a **correlação das variáveis** em relação à variável alvo de estudo – RDPC (renda percapita).

```
Correlacao entre a variavel ALVO <RDPC> com <ESPVIDA> 0.7627402  
Correlacao entre a variavel ALVO <RDPC> com <GINI> -0.179176  
Correlacao entre a variavel ALVO <RDPC> com <T_FBSUPER> 0.7778708  
Correlacao entre a variavel ALVO <RDPC> com <FECTOT> -0.6331079  
Correlacao entre a variavel ALVO <RDPC> com <PMP0B> -0.8737927  
Correlacao entre a variavel ALVO <RDPC> com <IDHM> 0.8268899
```

Plotando o **gráfico de pontos** (scatterplot) para averiguar graficamente a correlação de cada uma das variáveis com o alvo:







Após toda a análise exploratória sobre as variáveis em estudo, é hora de gerar a equação do modelo de regressão linear considerando as variáveis INDEPENDENTES → ESPVIDA, GINI, FECTOT, T\_FBSUPER, PMPOB, IDHM e a variável DEPENDENTE → RDPC.

```
reg_linear_multipla=lm(formula = df_01$RDPC ~ df_01$ESPVIDA + df_01$GINI +
df_01$FECTOT + df_01$T_FBSUPER + df_01$PMPOB + df_01$IDHM)
```

## ➔ Resultados das Análises

Após o modelo gerado, é hora de avaliar o modelo.

--->>> COEFICIENTE DE EXPLICAÇÃO do modelo <<<---

(Multiple R-Squared) r2= 0.8528305.

Indica que 85.28305 % da variação da variável DEPENDENTE é explicada pela variação das variáveis DEPENDENTES selecionadas no modelo gerado.

--->>> COEFICIENTE DE EXPLICACAO das variáveis do modelo: <<<---

As variáveis que tiverem BAIXO NIVEL DE SIGNIFICÂNCIA (valor -p > 0,05) perante o modelo serão removidas. Após, o modelo de regreção sera recalculado. Seguem valores (Pr(>|t|)):

(Intercept)	df_01\$ESPVIDA	df_01\$GINI	df_01\$FECTOT	df_01\$T_FBSUPER	df_01\$PMPOB	df_01\$IDHM
4.401751e-152	9.519685e-34	0.000000e+00	2.160540e-118	0.000000e+00	0.000000e+00	6.521642e-12

Observa-se que as variáveis ESPVIDA, FECTOT e IDHM tiveram baixa significância para o modelo, logo, no recálculo do modelo, essas variáveis serão desconsideradas.

Novo modelo gerado:

```
reg_linear_multipla_NEW=lm(formula = df_01$RDPC ~ df_01$GINI + df_01$T_FBSUPER +
df_01$PMPOB)
```

--->>> COEFICIENTE DE EXPLICAÇÃO do modelo linear <<<---

(Multiple R-Squared)  $r^2 = 0.8390097$ .

Indica que 83.90097 % da variação da variável DEPENDENTE é explicada pela variação das variáveis DEPENDENTES selecionadas no modelo gerado.

Em uma nova avaliação dos COEFICIENTES DE EXPLICAÇÃO das variáveis do modelo, todas atendem o nível desejado de significância (valor  $-p \leq 0,05$ ):

(Intercept)	df_01\$GINI	df_01\$T_FBSUPER	df_01\$PMPOB
0	0	0	0

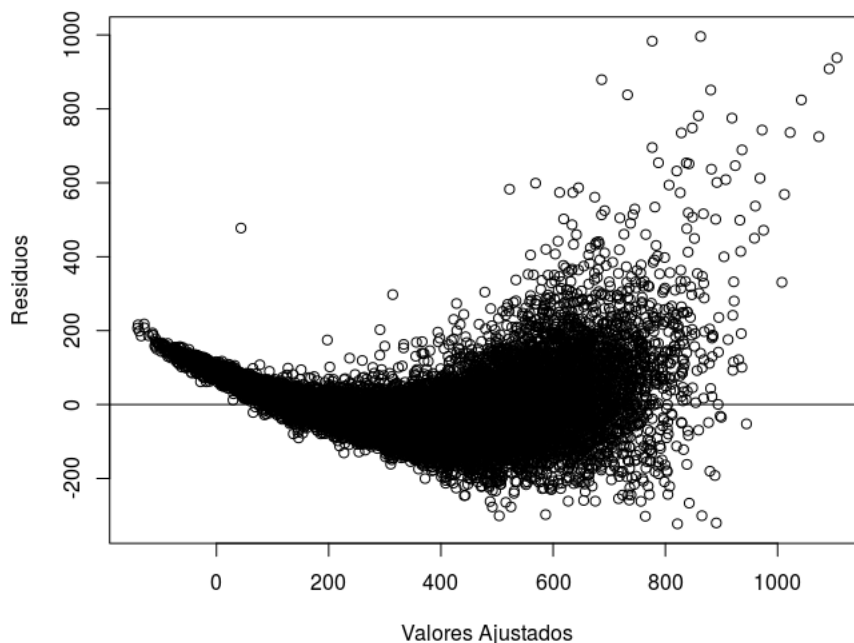
Os **resíduos** dos valores em relação à reta da equação gerada ficou com uma média de -7.01.

Residuals:				
Min	1Q	Median	3Q	Max
-322.71	-51.95	-7.01	40.27	995.91

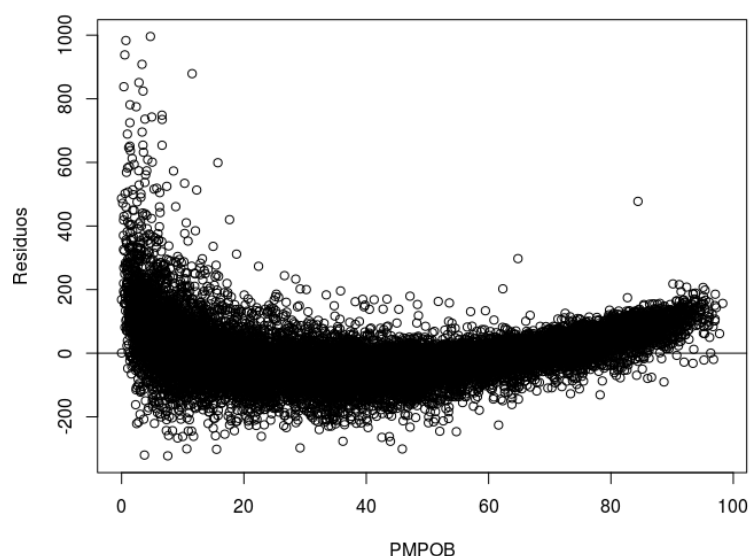
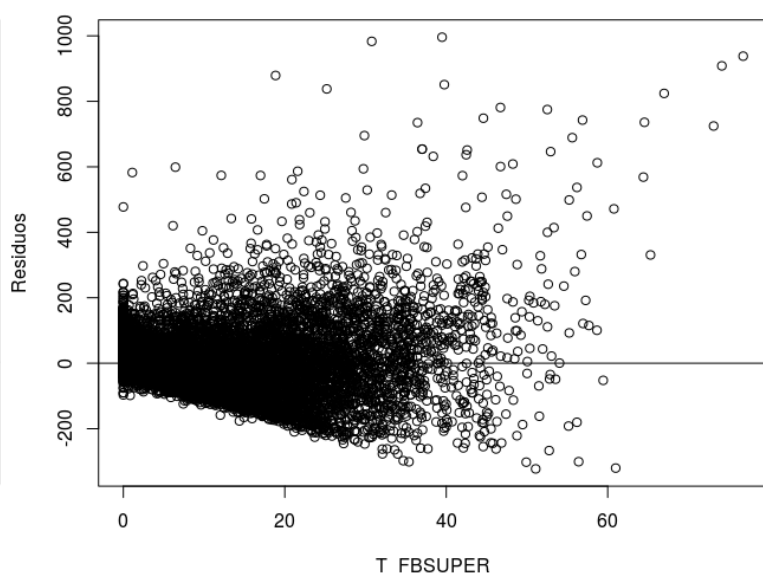
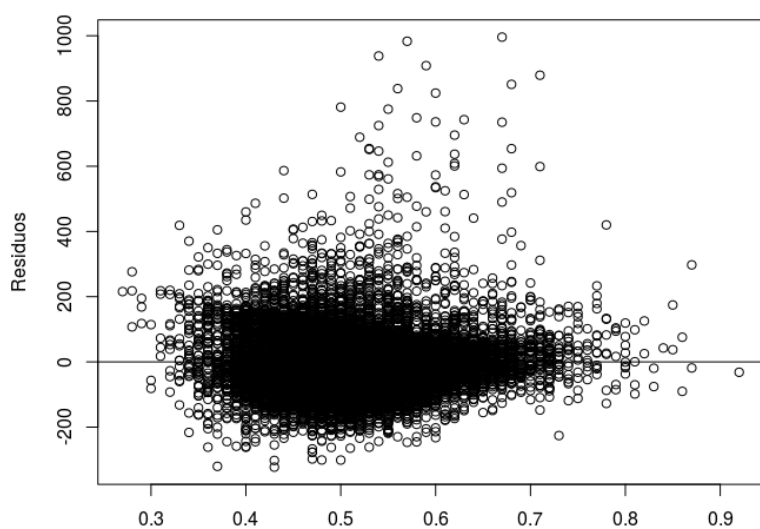
O **intervalo de confiança** para as variáveis em relação ao modelo gerado:

	2.5 %	97.5 %
(Intercept)	279.624530	301.317833
df_01\$GINI	438.949991	479.247520
df_01\$T_FBSUPER	7.241754	7.618435
df_01\$PMPOB	-6.154386	-6.001243

Plotando os **valores do modelo de regressão linear “treinado”** X **resíduos da equação do modelo**:



Plotando os **valores originais** X **resíduos gerados por cada uma das variáveis no modelo de regressão linear**:



Avaliação da homocedasticidade dos erros: teste de igualdade entre variâncias.

F test to compare two variances

```
data: residuals(reg_linear_multipla_NEW)[df_01$GINI > median_GINI] and residuals(reg_linear_multipla_NEW)[df_01$GINI < median_GINI]
F = 0.83284, num df = 8000, denom df = 7741, p-value = 5.02e-16
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.7968255 0.8704624
sample estimates:
ratio of variances
 0.8328373
```

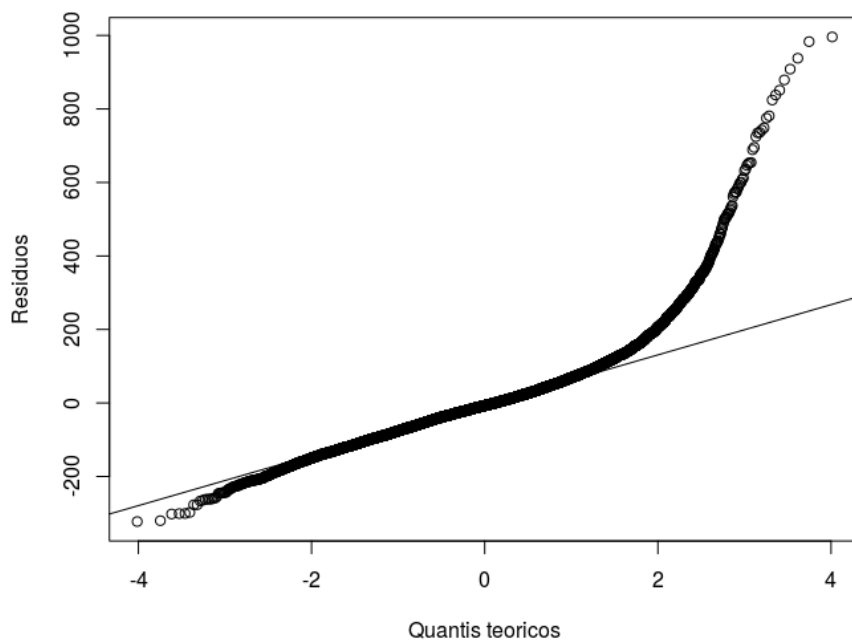
F test to compare two variances

```
data: residuals(reg_linear_multipla_NEW)[df_01$T_FBSUPER > median_T_FBSUPER] and residuals(reg_linear_multipla_NEW)[df_01$T_FBSUPER < median_T_FBSUPER]
F = 5.4039, num df = 8343, denom df = 8341, p-value < 2.2e-16
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 5.176843 5.640868
sample estimates:
ratio of variances
 5.403878
```

F test to compare two variances

```
data: residuals(reg_linear_multipla_NEW)[df_01$PMPOB > median_PMPOB] and residuals(reg_linear_multipla_NEW)[df_01$PMPOB < median_PMPOB]
F = 0.31593, num df = 8343, denom df = 8345, p-value < 2.2e-16
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.3026622 0.3297879
sample estimates:
ratio of variances
 0.315934
```

Avaliação da suposição de normalidade dos erros:



## ➔ Conclusões

A equação final gerada do modelo de regressão linear múltipla é a contida abaixo:

```
reg_linear_multipla_NEW=lm(formula = df_01$RDPC ~ df_01$GINI + df_01$T_FBSUPER + df_01$PMPOB)
```

Considerando o contexto das variáveis selecionadas e exploradas, os dados da população obtidos junto ao órgão “Atlas do Desenvolvimento Humano”, ao qual correspondem aos anos de 1991, 2000 e 2010, pode-se afirmar com 83,90% de certeza que a variação da renda percapta é explicada pelas variáveis independentes envolvidas no estudo - GINI (Índice de Gini), T\_FBSUPER (Taxa de frequência bruta ao ensino superior) e PMPOB (Proporção de pobres).

Como ponto de melhoria ao modelo desenvolvido, a sugestão é explorar outras variáveis que possam de alguma forma explicar com maior acurácia, as variações da renda percapta nos municípios brasileiros.