

# modelo de regressão e análise - conclusão precária....

Lembrete os dados são todos probabilísticos, há suposições e zero certezas.

QuartilHobus2008 que verás abaixo é uma divisão dos eleitores simulados em quatro grupos de tamanho igual, divididos pela probabilidade simulada de terem votado em Hobus em 2008.

```
by(df$Hobus2008, df$QuartisHobus2008, mean)
```

```
## df$QuartisHobus2008: 1
```

```
## [1] 71.26164
```

```
## -----
```

```
## df$QuartisHobus2008: 2
```

```
## [1] 78.60998
```

```
## -----
```

```
## df$QuartisHobus2008: 3
```

```
## [1] 83.06652
```

```
## -----
```

```
## df$QuartisHobus2008: 4
```

```
## [1] 89.16883
```

Acima você vê como a minha criação de quartis deu certo, pois a media de Hobus2008 é crescente entre os quartis. Quanto maior, maior a média.

## VAMOS AO MODELO

Gariba2012 explicado por Quartis + enchente + Quartis\*enchente (termo interativo, leia abaixo a interpretação que vai aclarar)

```
simulado55 <- lm(Gariba2012 ~ QuartisHobus2008 +
```

```

enchente + QuartisHobus2008*enchente,
data=df)
library(huxtable)
huxreg(simulado55)

```

	(1)
(Intercept)	59.907 ***
	(0.175)
QuartisHobus20082	-2.074 ***
	(0.285)
QuartisHobus20083	-5.144 ***
	(0.303)
QuartisHobus20084	-6.687 ***
	(0.299)
enchenteSim	-3.145 ***
	(0.194)
QuartisHobus20082:enchenteSim	-0.170
	(0.308)
QuartisHobus20083:enchenteSim	1.004 **
	(0.324)
QuartisHobus20084:enchenteSim	0.800 *
	(0.321)
N	37236
R2	0.106

Supondo que as suposições feitas com `rnorm` façam sentido (o que não sabemos se ocorre) Estima-se que a chance de um eleitor ao acaso, que seja do quartil inferior de Hobus2008 e que não tenha sido atingido por enchente, seja de 59,89% (veja intercept) a partir daí você tem comparações com esse carinho.

Se ao invés do 1, ele fosse do quartil 2, sua chance de 2,054% a menos do que 59,80%

Se fosse do 3, ao invés do 1, a redução na chance seria de 5,13%. Se fosse no maior quartil, a redução seria de 6,61%. Isso quer dizer que alguém não atingido pela enchente e com votação alta em Hobus 2008 teria chance de votar em Gariba 2012 de 53.13%

Se você mantivesse o quartil1 e alterasse o fato de o indivíduo ser atingido pela enchente, a probabilidade aumentaria em cerca de 3,2%. O que vai contra o senso comum de que Gariba foi melhor votado em lugar afetados pela enchente (dois fatos no indicam que pode haver problemas, o primeiro é que nosso dado de atingimento de enchente se deu por bairros e não por ruas e o segundo é que são dados simulados, probabilísticos e não reais).

as últimas três colunas, são as mais interessantes, elas vão subindo o quartil dentro dos atingidos pelas cheias. Aqui nota-se um fato que mereceria melhor investigação, Gariba tem um aumento em relação ao não enchente e quartil baixo nos grupos de atingidos com maior quartil.

Isso é muito louco, pq é de se supor que o maior quartil aumentasse a proporção de voto em Teixeira (O RIVAL DE GARIBA), apoiado por HOBUS.

as comparações só são significativas nas duas últimas colunas e não no quartil 2 que não tem significância

Então vamos à tabela abaixo que talvez explique melhor

```
ChanceGariba2012 <- c(59.89, 59.89+0.980, 59.89+0.827)
```

```
ApoioQuartilHobus2008 <- c(1,3,4)
```

```
Enchente <- c("Não", "Sim", "Sim")
```

```
cor <- data.frame(ChanceGariba2012, ApoioQuartilHobus2008, Enchente)
```

```
b5 <- cor %>%
```

```
dplyr::select(ApoioQuartilHobus2008, Enchente, ChanceGariba2012) %>%
```

```
arrange((ChanceGariba2012))
```

```
b5 %>%
```

```
kbl(caption = "Resumo dos dados do modelo de regressão Intercept x casos com enchente e  
alto apoio à Hobus 2008, dados simulados") %>%
```

```
kable_classic(full_width = F, html_font = "Garamond")
```

Resumo dos dados do modelo de regressão Intercept x casos com enchente e alto apoio à Hobus 2008, dados simulados

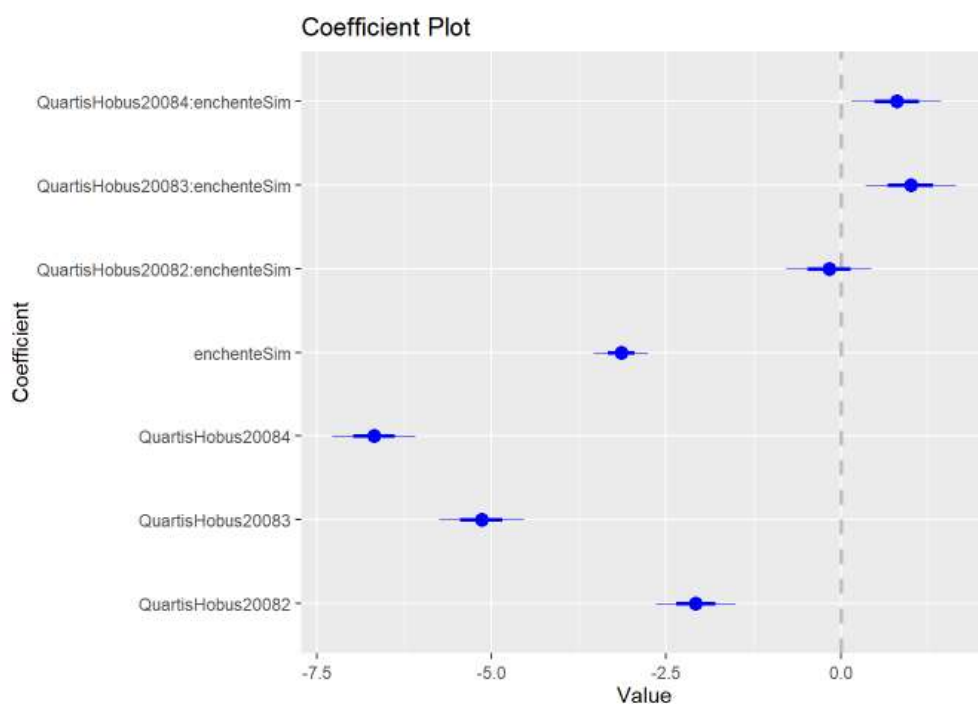
ApoioQuartilHobus2008	Enchente	ChanceGariba2012
1	Não	59.890
4	Sim	60.717
3	Sim	60.870

Essa tabela é extraída do modelo acima, é nota-se que pode ser aí que a influência de Hobus tenha sido insuficiente para ajudar Teixeira a repassar votos em regiões. As diferenças são pequenas e pode ser que foram significantes por que aumentei os números de casos imputando valores com rnorm, mas o tópico merece investigação (se for possível). O que eu descobri fica muito mais como hipótese do que como achado.

Abaixo os valores dentro do intervalo (nota-se que todos são comparados ao intercept)

```
library(coefplot)
```

```
coefplot(simulado55, intercept=FALSE, interactive=FALSE)
```



*A enchente não melhorou a votação, menos no caso dos bairros em que Hobus fez boa votação. É o que aparenta-se para que isso seja verdade seria necessário que os dados simulados fossem próximos da realidade e que os dados de enchente (por bairro) fossem bons o suficiente. O que ainda não sei quando escrevo essa análise.*