

Trabalho 2 de TGT410029 – Leg. Territorial e PVG

Willian Zonato

Luiz Fernando Palin Droubi

20 de agosto de 2018

1 Importação dos dados

a. Coordenadas

As coordenadas foram extraídas de arquivo .kml¹ diretamente para o R version 3.5.1 (2018-07-02).

```
source("E:\\Documents\\appraiseR\\R\\kml.R")
df <- read.kml("Sto_Amaro_4.kml", "Meus lugares")

## OGR data source with driver: KML
## Source: "E:\\Documents\\UFSC\\Planta de Valores\\Trabalho\\Sto_Amaro_4.kml", layer: "Meus lugares"
## with 30 features
## It has 2 fields
```

ID	N	E
LT_01	-27,68	-48,76
LT_02	-27,69	-48,78
LT_03	-27,69	-48,73
LT_04	-27,68	-48,76
LT_05	-27,68	-48,77
LT_06	-27,68	-48,76

b. Dados do Excel

Os dados da pesquisa de mercado foram lidos diretamente no R version 3.5.1 (2018-07-02).

```
Dados <- read_excel("Dados.xlsx")
```

ID	Area	Valor	VU	Geral	topografia	pavimentacao	pavimentado	situacao
LT_01	28.750,00	3.800.000	132,17	sim	plano	asfalto	sim	esquina
LT_02	1.800,00	2.700.000	1.500,00	sim	plano	asfalto	sim	meio
LT_03	5.373,72	250.000	46,52	sim	acidentado	asfalto	sim	meio
LT_04	6.188,00	880.000	142,21	nao	plano	asfalto	sim	meio
LT_05	941,00	800.000	850,16	sim	plano	asfalto	sim	meio
LT_06	838,72	850.000	1.013,45	sim	plano	asfalto	sim	esquina

c. Aglutinação dos dados

Posteriormente, os dados da pesquisa foram mesclados com as coordenadas dos dados. O conjunto de dados assim obtido pode ser visto na tabela 1

```
data <- inner_join(df, Dados, by = "ID")
```

¹https://github.com/lfpdroubi/planta_valores/blob/master/Sto_Amaro_4.kml

2 Espacialização

a. Criação do conjunto de dados espaciais

Para a transformação dos dados em um conjunto de dados espaciais, basta informar ao R as colunas das coordenadas e o seu sistema de referência.

```
coordinates(data) <- ~E+N  
proj4string(data) <- CRS("+init=epsg:4326") # WGS 84
```

Tabela 1: Dados com coordenadas geográficas (WGS-84)

ID	Area	Valor	VU	Geral	topografia	pavimentacao	pavimentado	situacao	E	N
LT_01	28.750,00	3.800.000	132,17	sim	plano	asfalto	sim	esquina	-48,76	-27,68
LT_02	1.800,00	2.700.000	1.500,00	sim	plano	asfalto	sim	meio	-48,78	-27,69
LT_03	5.373,72	250.000	46,52	sim	acidentado	asfalto	sim	meio	-48,73	-27,69
LT_04	6.188,00	880.000	142,21	nao	plano	asfalto	sim	meio	-48,76	-27,68
LT_05	941,00	800.000	850,16	sim	plano	asfalto	sim	meio	-48,77	-27,68
LT_06	838,72	850.000	1.013,45	sim	plano	asfalto	sim	esquina	-48,76	-27,68
LT_07	777,00	690.000	888,03	sim	plano	asfalto	sim	meio	-48,75	-27,68
LT_08	630,00	600.000	952,38	sim	plano	blokret	sim	esquina	-48,75	-27,68
LT_09	1.740,00	380.000	218,39	nao	plano	blokret	sim	meio	-48,78	-27,69
LT_12	585,00	240.000	410,26	nao	plano	asfalto	sim	esquina	-48,78	-27,68
LT_13	668,29	230.000	344,16	nao	plano	blokret	sim	meio	-48,75	-27,68
LT_16	700,00	150.000	214,29	nao	acidentado	blokret	sim	meio	-48,78	-27,69
LT_17	360,00	130.000	361,11	nao	plano	blokret	sim	meio	-48,76	-27,68
LT_22	385,00	95.000	246,75	nao	acidentado	blokret	sim	meio	-48,75	-27,71
LT_26	472,00	380.000	805,08	sim	plano	asfalto	sim	meio	-48,76	-27,68
LT_27	380,00	285.000	750,00	nao	plano	blokret	sim	meio	-48,78	-27,69
LT_29	2.261,00	398.000	176,03	nao	plano	asfalto	sim	meio	-48,73	-27,69
LT_30	30.000,00	850.000	28,33	nao	plano	asfalto	sim	meio	-48,77	-27,68
LT_31	360,00	180.000	500,00	nao	plano	asfalto	sim	meio	-48,75	-27,68
LT_32	384,00	95.000	247,40	nao	acidentado	blokret	sim	meio	-48,77	-27,68
LT_33	2.000,00	2.000.000	1.000,00	nao	plano	blokret	sim	meio	-48,78	-27,68
LT_35	360,00	105.000	291,67	nao	plano	blokret	sim	esquina	-48,74	-27,68
LT_36	400,00	380.000	950,00	sim	plano	asfalto	sim	esquina	-48,80	-27,70
LT_37	1.507,17	420.000	278,67	nao	plano	sem	nao	meio	-48,76	-27,68
LT_38	808,92	370.000	457,40	nao	acidentado	asfalto	sim	esquina	-48,78	-27,69
AVAL_1	360,00	NA	NA	sim	plano	NA	NA	NA	-48,78	-27,69
AVAL_2	360,00	NA	NA	sim	acidentado	NA	NA	NA	-48,78	-27,69
AVAL_3	360,00	NA	NA	nao	acidentado	NA	NA	NA	-48,78	-27,69
AVAL_4	360,00	NA	NA	nao	plano	NA	NA	NA	-48,78	-27,69

b. Escrita do Shapefile no disco

Foi escrito um *shapefile* no disco à partir do conjunto de dados espaciais, através da função `writeOGR`, do pacote **rgdal**.

```
# Para escrever o shapefile no disco:
writeOGR(data,
  dsn = "E:\\Documents\\UFSC\\Planta de Valores\\Trabalho",
  layer = "dados",
  driver = "ESRI Shapefile",
  overwrite_layer = TRUE,
  delete_dsn = TRUE)
```

c. Conversão de unidades

O sistema de referência pode ser alterado através da função `spTransform`, do pacote **sp**. Por exemplo, para alterar para SIRGAS2000, basta informar o código EPSG deste sistema de referência (31997). Os dados com as coordenadas transformadas podem ser vistos na tabela 2.

```
# Conversão de coordenadas para SIRGAS2000
CRS.new <- CRS("+init=epsg:31997")
```

Tabela 2: Dados com coordenadas em SIRGAS2000

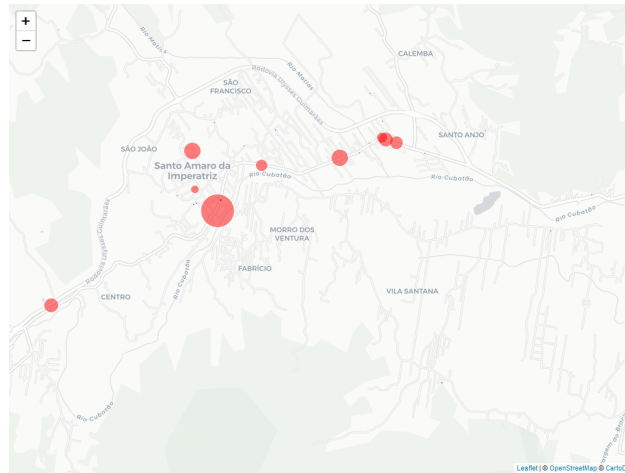
ID	Area	Valor	VU	Geral	topografia	pavimentaca	pavimentado	situacao	E	N
LT_01	28.750	3.800.000	132	sim	plano	asfalto	sim	esquina	721.057	6.935.883
LT_02	1.800	2.700.000	1.500	sim	plano	asfalto	sim	meio	719.145	6.935.146
LT_03	5.374	250.000	47	sim	acidentado	asfalto	sim	meio	724.179	6.935.262
LT_04	6.188	880.000	142	nao	plano	asfalto	sim	meio	721.200	6.936.007
LT_05	941	800.000	850	sim	plano	asfalto	sim	meio	719.747	6.935.745
LT_06	839	850.000	1.013	sim	plano	asfalto	sim	esquina	720.794	6.935.826
LT_07	777	690.000	888	sim	plano	asfalto	sim	meio	721.566	6.936.011
LT_08	630	600.000	952	sim	plano	blokret	sim	esquina	721.423	6.936.059
LT_09	1.740	380.000	218	nao	plano	blokret	sim	meio	718.600	6.935.355
LT_12	585	240.000	410	nao	plano	asfalto	sim	esquina	719.123	6.936.329
LT_13	668	230.000	344	nao	plano	blokret	sim	meio	721.496	6.936.081
LT_16	700	150.000	214	nao	acidentado	blokret	sim	meio	718.823	6.935.234
LT_17	360	130.000	361	nao	plano	blokret	sim	meio	721.383	6.936.302
LT_22	385	95.000	247	nao	acidentado	blokret	sim	meio	722.114	6.932.789
LT_26	472	380.000	805	sim	plano	asfalto	sim	meio	721.370	6.936.080
LT_27	380	285.000	750	nao	plano	blokret	sim	meio	718.846	6.935.437
LT_29	2.261	398.000	176	nao	plano	asfalto	sim	meio	723.612	6.934.793
LT_30	30.000	850.000	28	nao	plano	asfalto	sim	meio	719.513	6.936.004
LT_31	360	180.000	500	nao	plano	asfalto	sim	meio	721.938	6.935.865
LT_32	384	95.000	247	nao	acidentado	blokret	sim	meio	719.942	6.936.207
LT_33	2.000	2.000.000	1.000	nao	plano	blokret	sim	meio	718.821	6.935.955
LT_35	360	105.000	292	nao	plano	blokret	sim	esquina	722.757	6.936.153
LT_36	400	380.000	950	sim	plano	asfalto	sim	esquina	716.888	6.933.926
LT_37	1.507	420.000	279	nao	plano	sem	nao	meio	721.221	6.936.674
LT_38	809	370.000	457	nao	acidentado	asfalto	sim	esquina	718.864	6.935.252
AVAL_1	360	NA	NA	sim	plano	NA	NA	NA	719.186	6.935.292
AVAL_2	360	NA	NA	sim	acidentado	NA	NA	NA	719.186	6.935.291
AVAL_3	360	NA	NA	nao	acidentado	NA	NA	NA	719.185	6.935.291
AVAL_4	360	NA	NA	nao	plano	NA	NA	NA	719.185	6.935.291

3 Confecção de mapas temáticos

Foram elaborados mapas temáticos de algumas variáveis pesquisadas, também com o auxílio do R. Nas figuras abaixo, o tamanho dos pontos foi escalonado de acordo com a escala de valor unitário para cada dado.

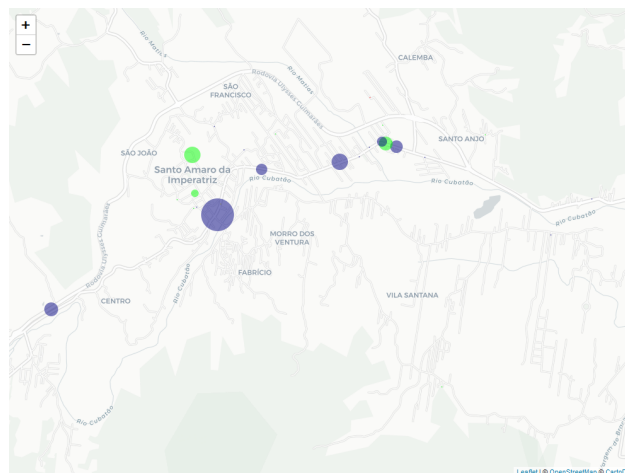
a. Topografia

Os terrenos foram classificados em plano (vermelho) e acidentado (azul), conforme pode ser visto na figura abaixo:



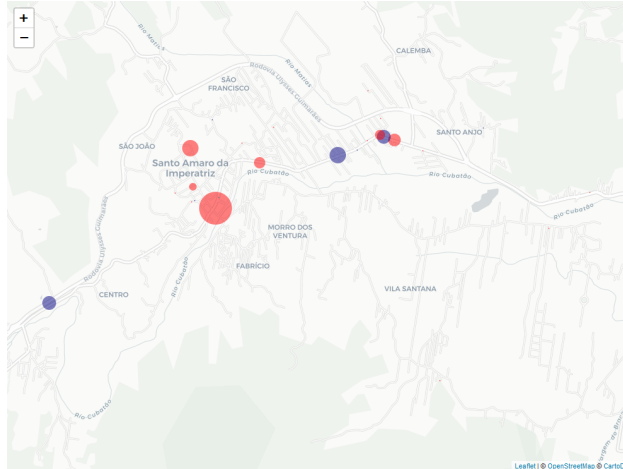
b. Pavimentação

Na figura abaixo, os dados podem ser vistos em função da pavimentação da frente do lote, se asfalto (azul), *bloquet* (verde) e sem pavimentação (vermelho).



c. Situação

Na figura abaixo, os dados de meio de quadra são vistos em azul e os dados de esquina em vermelho.



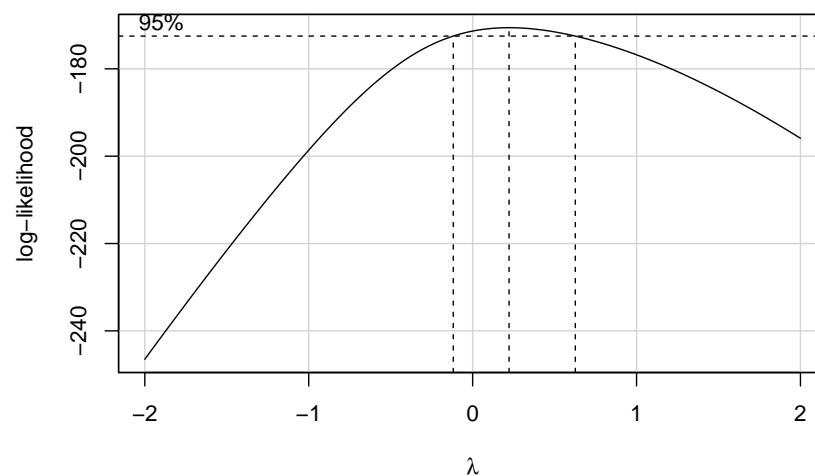
4 Ajuste do modelo OLS

Foi ajustado inicialmente um modelo linear com todas as variáveis pesquisadas:

```
fit <- lm(VU ~ Area + Geral + topografia + pavimentado + situacao, data = data)
```

5 Diagrama de Box-Cox

De posse do modelo linear, foi feito o diagrama de Box-Cox, para pesquisar melhores transformações para a variável dependente.



6 Modelo final

Após efetuadas as transformações necessárias, removidos os *outliers* e as variáveis insignificantes, chegou-se ao modelo descrito na tabela 3.

Tabela 3: Coeficientes do modelo final

	Dependent variable:
	log(VU)
log(Area)	-0.508 (-0.637, -0.380) t = -7.733 p = 0.00000***
Geralsim	1.101 (0.753, 1.448) t = 6.201 p = 0.00001***
topografiaplano	0.303 (-0.146, 0.753) t = 1.324 p = 0.202
Constant	8.825 (7.932, 9.717) t = 19.380 p = 0.000***
Observations	23
R ²	0.843
Adjusted R ²	0.818
Residual Std. Error	0.382 (df = 19)
F Statistic	33.962*** (df = 3; 19)
Note:	*p<0.1; **p<0.05; ***p<0.01

6.1 Diagnóstico básico

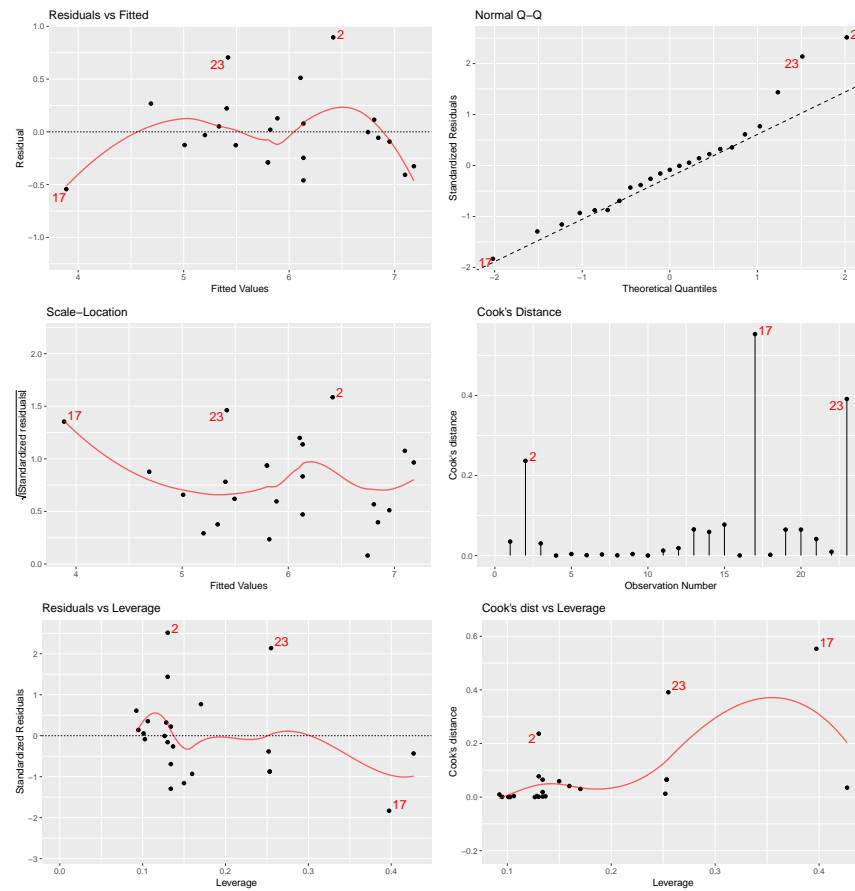


Figura 1: Gráficos básicos do modelo

6.2 Testes do modelo

6.2.1 Homoscedasticidade

```
##
## studentized Breusch-Pagan test
##
## data: fit
## BP = 1.0575801, df = 3, p-value = 0.787323
```

6.2.2 Normalidade

a. Teste de Pearson (χ^2)

```
##
## Pearson chi-square normality test
##
## data: resid(fit)
## P = 3.7826087, p-value = 0.58112
```

b. Teste de Lilliefors (Kolgomorov-Smirnov):

```
##  
##  Lilliefors (Kolmogorov-Smirnov) normality test  
##  
## data:  resid(fit)  
## D = 0.14183288, p-value = 0.2683846
```

c. Teste de Shapiro-Wilk:

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  resid(fit)  
## W = 0.94114636, p-value = 0.1901815
```

d. Teste de Anderson-Darling:

```
##  
##  Anderson-Darling normality test  
##  
## data:  resid(fit)  
## A = 0.45939696, p-value = 0.2384414
```

e. Teste de Jarque-Bera:

```
##  
##  Jarque-Bera test for normality  
##  
## data:  resid(fit)  
## JB = 2.9044081, p-value = 0.089
```

f. Histograma

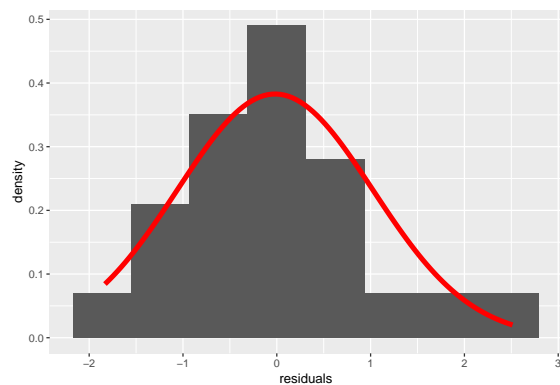


Figura 2: Histograma dos resíduos padronizados

g. Teste K-S (Kolgomorov-Smirnov) [KS]

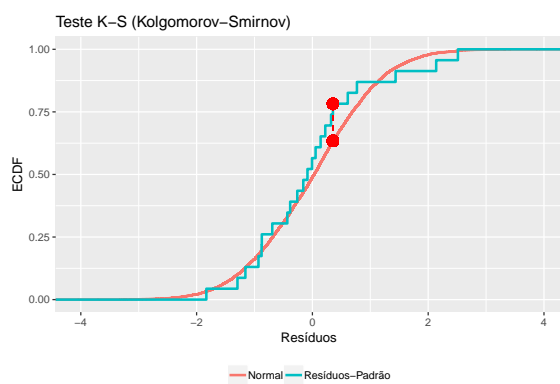


Figura 3: Curva da função de distribuição acumulada (FDA) empírica

6.2.3 Gráficos do modelo

a. Na mediana das variáveis

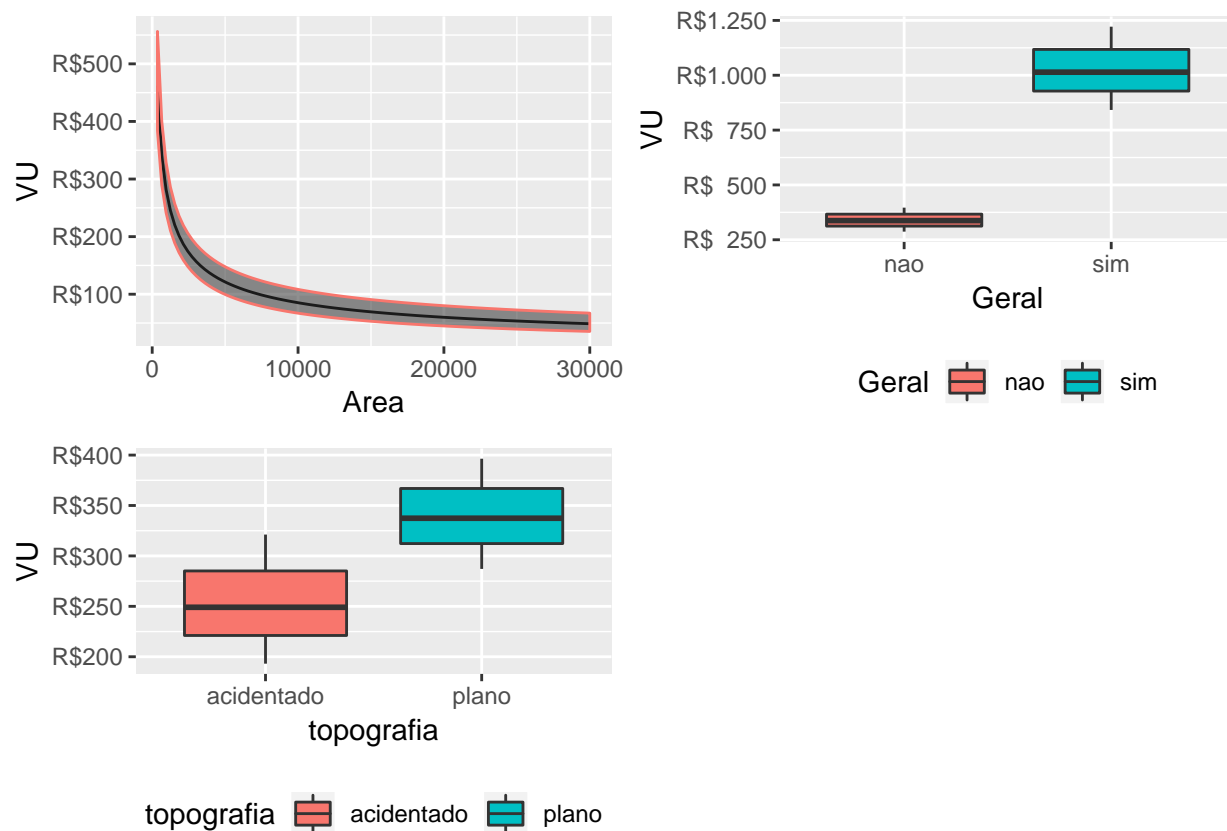


Figura 4: Gráficos dos regressores em função da variável dependente (em cada gráfico, os outros regressores estão em seus valores médios).

b. No ponto de avaliação

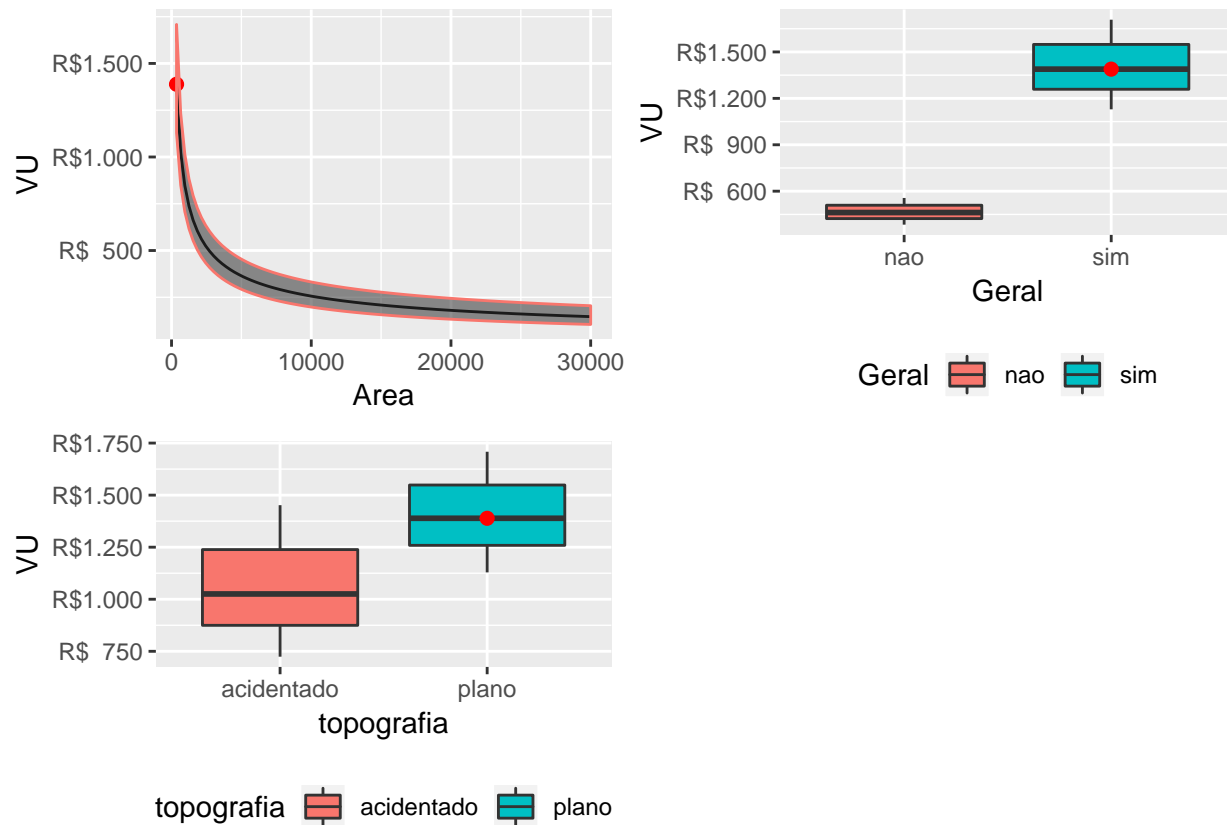
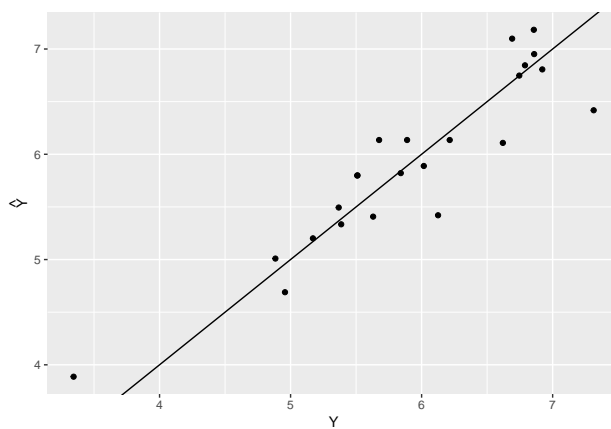


Figura 5: Gráficos dos regressores em função da variável dependente (em cada gráfico, os outros regressores estão com os valores reais do avaliando).

IV. Poder de Predição

*Figura 6: Poder de Predição.*