

Avaliação 1

Matéria: Estatística Aplicada 1

Professor: Dr. Arno Schmitz

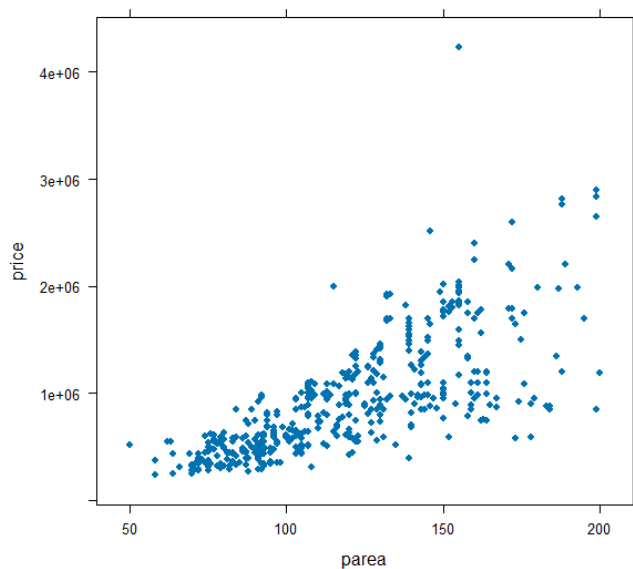
Aluno: Victor Carvalho de Andrade Lima

Data: 28/05/2023

Com a base de dados “imoveiscwbav” obter os seguintes resultados com o auxílio do “R”

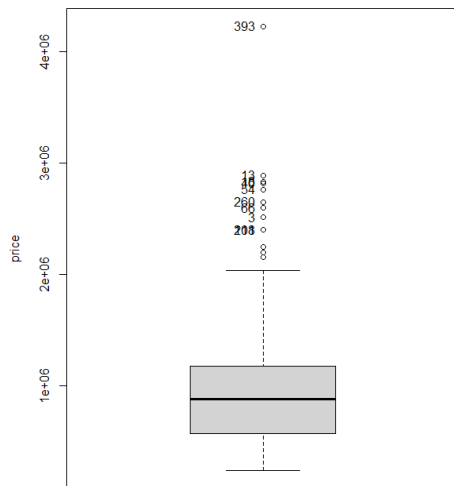
- a) Elaborar o histograma, *scatterplot* e o *boxplot* das variáveis “*price* e *parea*”.

Scarterplot

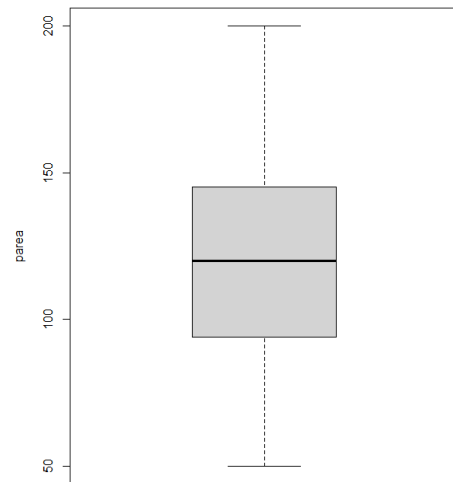


Interpretação do gráfico: O gráfico do *Scarterplot* mostra uma relação quantitativa entre 2 elementos, assim podemos afirmar inicialmente que quanto maior for a área de um imóvel, maior vai ser seu preço, quase que de uma forma linear. Além disso, podemos ver que quanto maior o preço, maior a chance de discrepância em relação a uma regra linear.

Boxplot “price”



Boxplot “area”



Interpretação dos gráficos: Nos gráficos *Bloxspot*, vemos uma distribuição por quartis, onde podemos fazer uma análise de como os valores se distribuem entre si. No gráfico de “price” (preço dos imóveis), vemos uma relação significativa de outliers para cima, assim podemos afirmar que alguns preços fogem consideravelmente da regra dos demais. Unindo com a análise do gráfico *Scatterplot*, vemos essa discrepância justamente nos maiores valores de área do imóvel, que tendem ter uma dispersão bem maior dos dados. Já no Boxplot de “area” (área de imóvel). Vemos uma relação uniforme e clara, com valores bem distribuídos entre os quartis do *Bloxspot*.

b) Elaborar a tabela de distribuição de frequências da variável “price” (preço dos imóveis);

	Class.limits	f	rf	Rf(%)	cf	Cf(%)
1	[237600,604380.9009)	161	0,297597	29,7597	161	29,7597
2	[604380.9009,971161.8018)	170	0,314233	31,42329	331	61,18299
3	[971161.8018,1337942.703)	108	0,19963	19,96303	439	81,14603
4	[1337942.703,1704723.604)	46	0,085028	8,502773	485	89,6488
5	[1704723.604,2071504.505)	42	0,077634	7,763401	527	97,4122
6	[2071504.505,2438285.405)	6	0,011091	1,109057	533	98,52126
7	[2438285.405,2805066.306)	4	0,007394	0,739372	537	99,26063
8	[2805066.306,3171847.207)	3	0,005545	0,554529	540	99,81516
9	[3171847.207,3538628.108)	0	0	0	540	99,81516
10	[3538628.108,3905409.009)	0	0	0	540	99,81516
11	[3905409.009,4272189.91)	1	0,001848	0,184843	541	100

Interpretação da tabela: A tabela de frequência o comportamento das variáveis, e podemos ver um forte acumulo de resultados na base dos dados, mostrando uma tendência de redução praticamente exponencial de frequência quanto maior o valor absoluto dos elementos.

c) Para a variável “price” calcular os seguintes indicadores: média; mediana; moda

Max: 4229891

Min: 240000

Média: 953186.7

Mediana: 880000

Moda: 880000

Interpretação dos dados: Utilizando referência como o valor máximo e mínimo junto com os indicadores de média, mediana e moda, vemos que os 3 valores estão bem mais próximos do mínimo, o que justifica sua maior frequência de dados como vista na tabela de distribuição de frequência