

## 3ª Etapa

Eniedson Fabiano Pereira da Silva Junior - 119110512  
João Paulo Alves Dantas - 119210258

1)

A base de dados usada representa valores de alugueis de imoveis em algumas cidades do Brasil, dentre os dados se encontram a área do imovel e os valores dos alugueis dos mesmos, desse modo, a inteção do seguinte estudo é encontrar uma reação linear entre essas duas variaveis quantitativas de modo a identificar preços de novos apartamentos a partir das suas áreas, em metros quadrados.

Base de dados - <https://www.kaggle.com/datasets/shwaubh/updated-brasilian-housing-to-rent>

2)

**Dados**

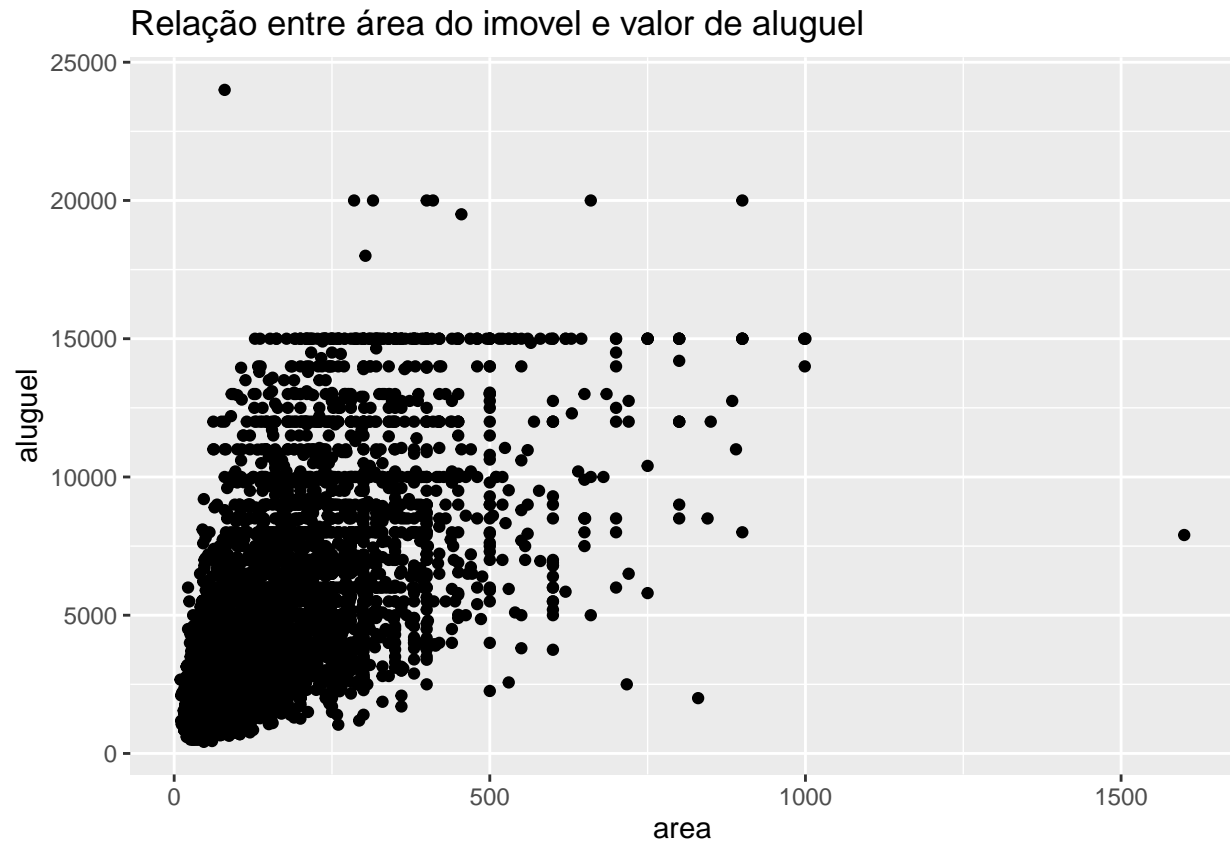
## Rows: 6,077

## Columns: 2

## \$ area <int> 240, 64, 443, 73, 19, 13, 55, 55, 82, 32, 60, 64, 20, 375, 92,~

## \$ aluguel <int> 8000, 820, 7000, 1250, 1200, 2200, 5000, 1300, 2000, 2900, 720~

## Analise de correlação



## Teste de hipotese sobre correlacao nula

```
##  
## Pearson's product-moment correlation  
##  
## data: dados$area and dados$aluguel  
## t = 70.09, df = 6075, p-value < 0.00000000000000022  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.6545215 0.6823323  
## sample estimates:  
## cor  
## 0.6686607
```

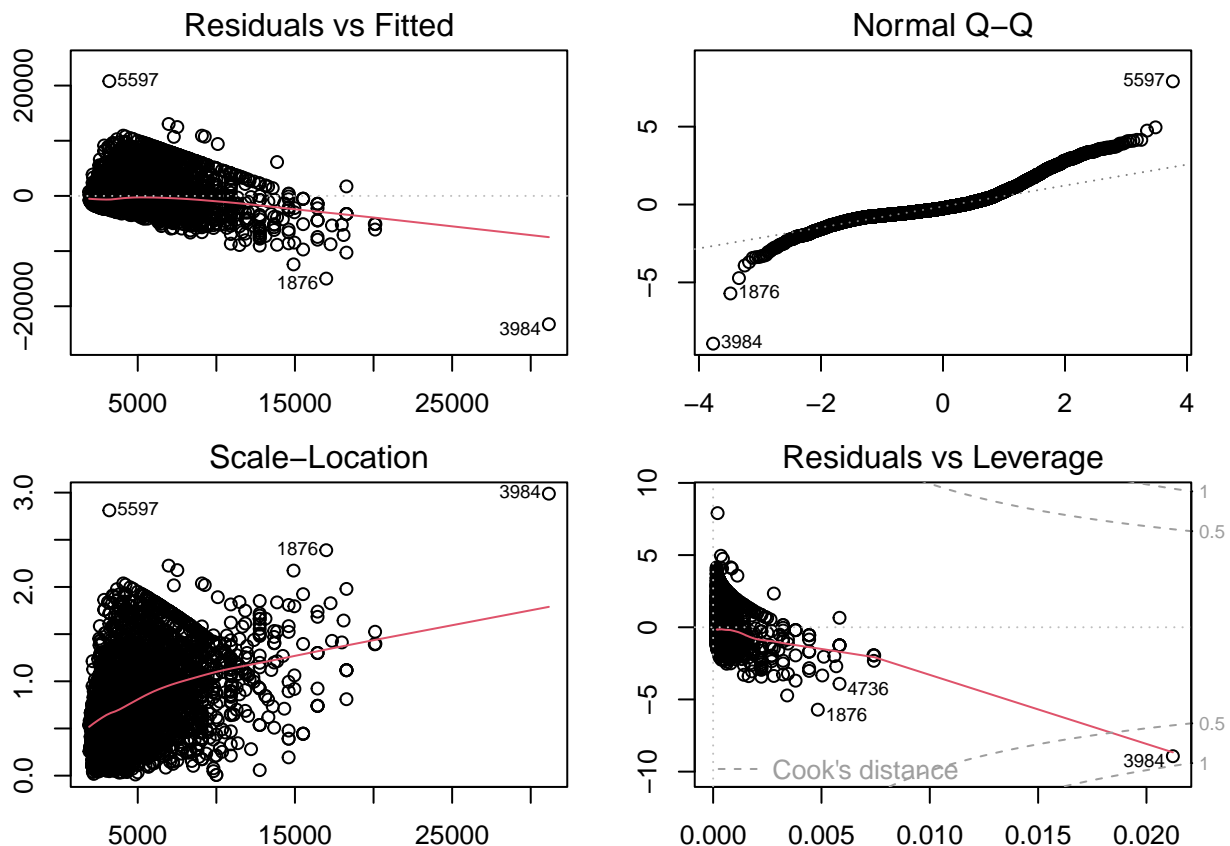
## Regressão Linear

### Coefficientes estimados:

```
##  
## Call:  
## lm(formula = aluguel ~ area, data = dados)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -23252.7  -1526.3   -633.5    861.3  20805.8
```

```
##
## Coefficients:
##           Estimate Std. Error t value      Pr(>|t|)
## (Intercept) 1722.6631    50.8605   33.87 <0.0000000000000002 ***
## area        18.3938     0.2624   70.09 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2632 on 6075 degrees of freedom
## Multiple R-squared:  0.4471, Adjusted R-squared:  0.447
## F-statistic: 4913 on 1 and 6075 DF, p-value: < 0.00000000000000022
```

Relacao linear entre a VD(Area) e a VI(Valor do Aluguel):



```
## Normalidade dos residuos:
ad.test(mod$residuals)
```

```
##
## Anderson-Darling normality test
##
## data: mod$residuals
## A = 224.07, p-value < 0.00000000000000022
```

```
## Outliers nos residuos:
summary(rstandard(mod))
```

```
##      Min.   1st Qu.   Median     Mean   3rd Qu.    Max.
## -8.931090 -0.580045 -0.240763 -0.000054  0.327305  7.906845
```

```
## Independencia dos residuos (Durbin-Watson):
# (from: car)
durbinWatsonTest(mod)

## lag Autocorrelation D-W Statistic p-value
## 1 -0.01029829 2.020494 0.414
## Alternative hypothesis: rho != 0

bptest(mod)

##
## studentized Breusch-Pagan test
##
## data: mod
## BP = 787.1, df = 1, p-value < 0.00000000000000022

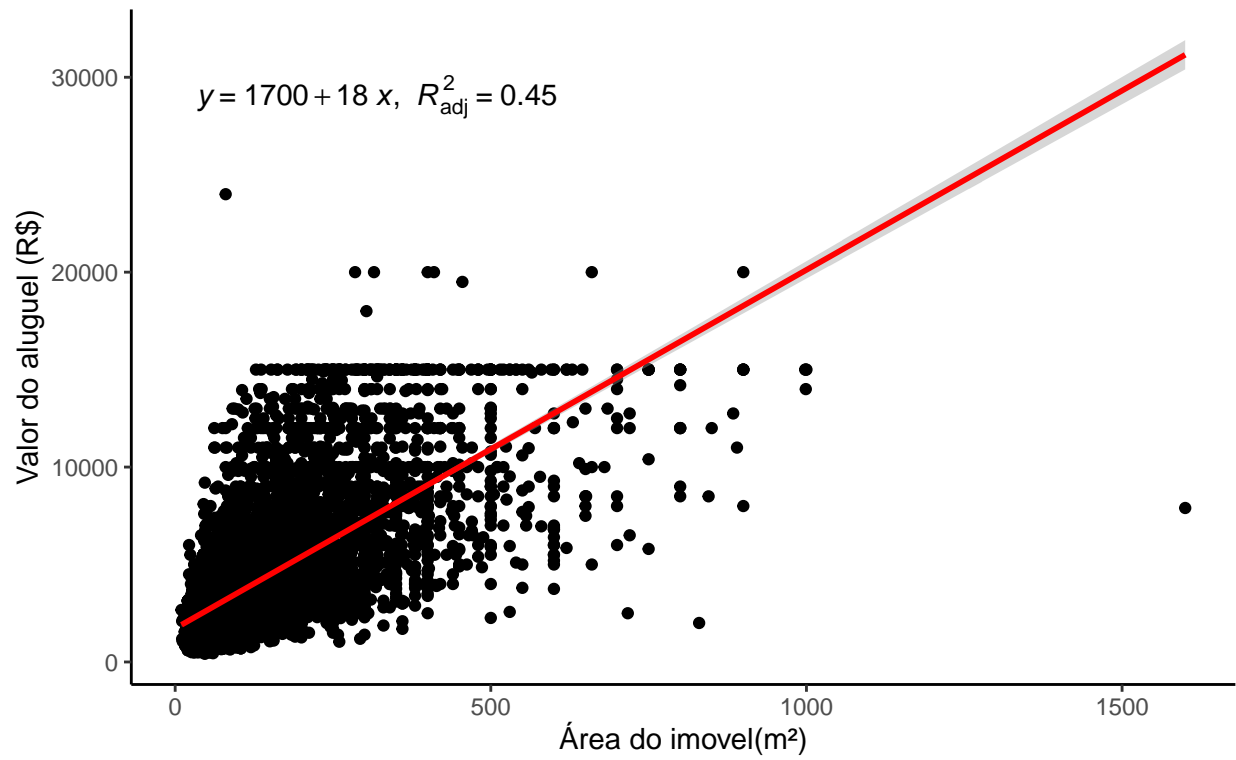
summary(mod)

##
## Call:
## lm(formula = aluguel ~ area, data = dados)
##
## Residuals:
## Min 1Q Median 3Q Max
## -23252.7 -1526.3 -633.5 861.3 20805.8
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1722.6631 50.8605 33.87 <0.0000000000000002 ***
## area 18.3938 0.2624 70.09 <0.0000000000000002 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2632 on 6075 degrees of freedom
## Multiple R-squared: 0.4471, Adjusted R-squared: 0.447
## F-statistic: 4913 on 1 and 6075 DF, p-value: < 0.00000000000000022
```

## Gráfico representativo da reta de regressão linear

### Ajuste de um Modelo de Regressão Linear Simples

Aluguel para x metros quadrados



### Valores de testes

```
## area
## 1 50
## 2 100
## 3 151
## 4 80
## 5 200
## 6 500
```

### Predições

```
##      1      2      3      4      5      6
## 2642.353 3562.043 4500.127 3194.167 5401.423 10919.563
```