

# exercicio1

September 16, 2021

## 0.1 Exercício 1 - Regressão linear múltipla

Gabriel De Pádua RA 200749

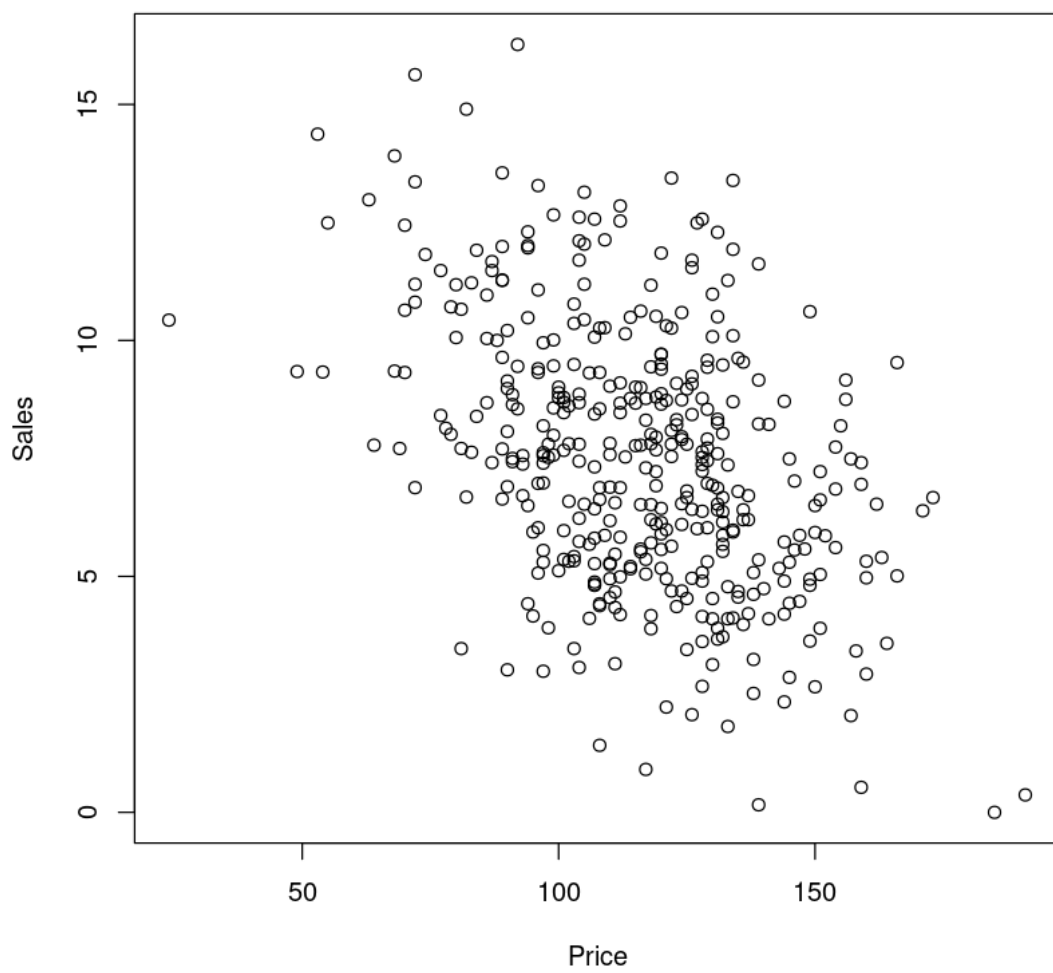
```
[ ]: ##install.package("ISLR")
```

```
[2]: library(ISLR)
```

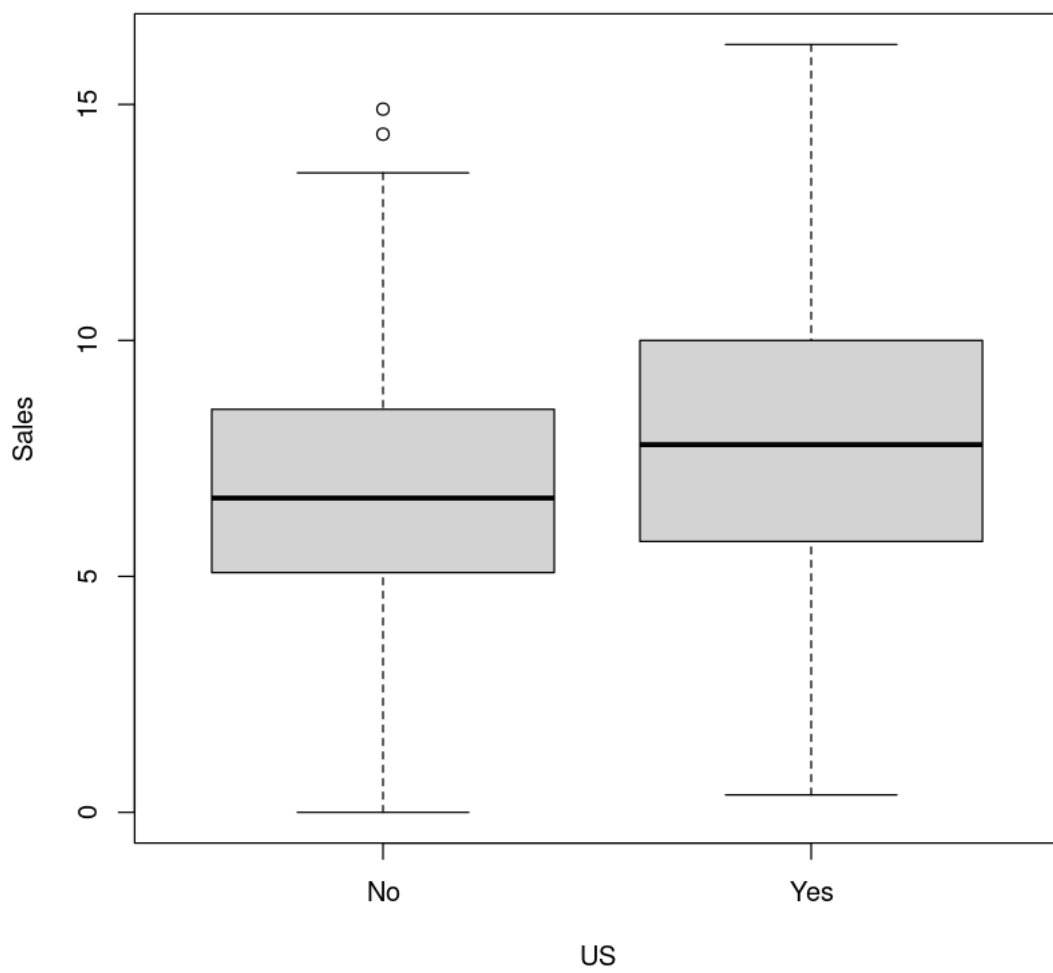
```
[5]: names(Carseats)
```

1. 'Sales' 2. 'CompPrice' 3. 'Income' 4. 'Advertising' 5. 'Population' 6. 'Price' 7. 'ShelveLoc' 8. 'Age'  
9. 'Education' 10. 'Urban' 11. 'US'

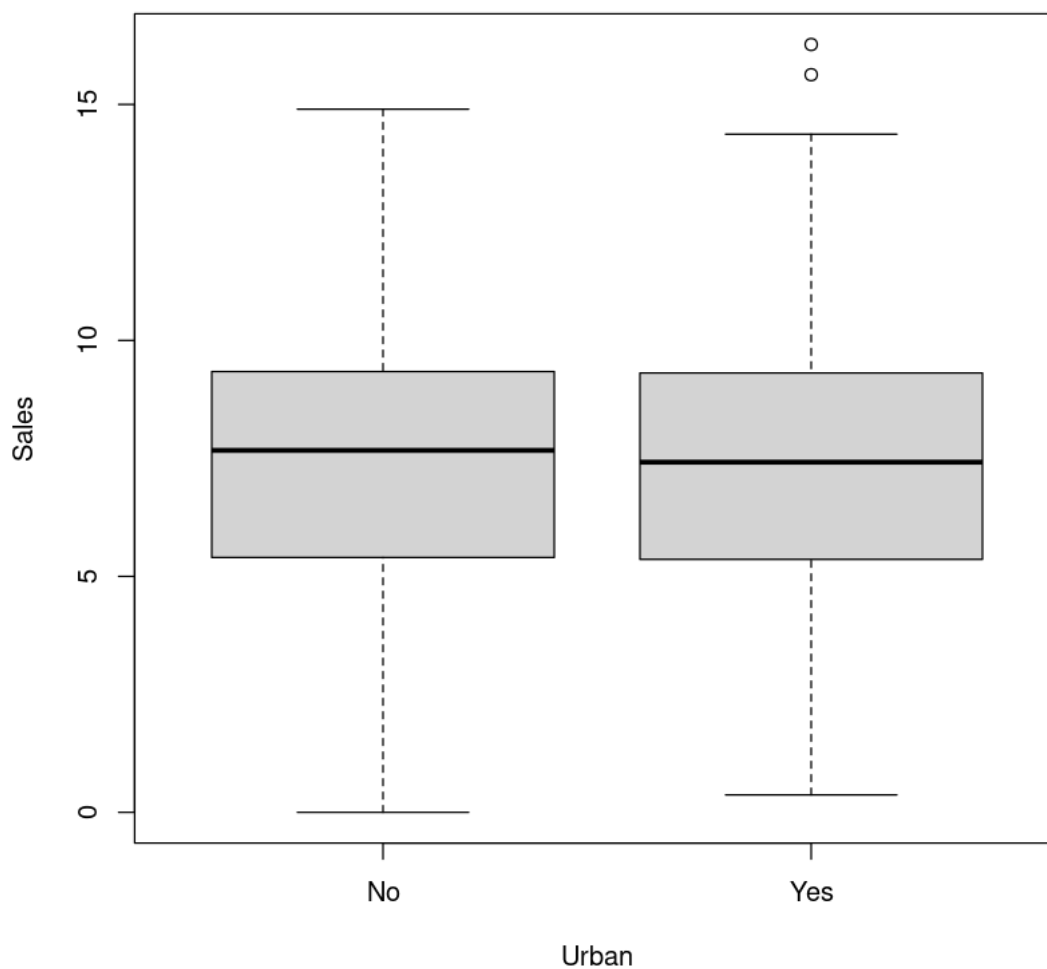
```
[8]: plot(Sales ~ Price, Carseats)
```



```
[10]: plot(Sales ~ US, Carseats)
```



```
[12]: plot(Sales ~ Urban, Carseats)
```



### 0.1.1 Modelo C

```
[14]: modelo <- lm(Sales ~ Price + Urban + US, Carseats)
      summary(modelo)
```

Call:

```
lm(formula = Sales ~ Price + Urban + US, data = Carseats)
```

Residuals:

Min	1Q	Median	3Q	Max
-6.9206	-1.6220	-0.0564	1.5786	7.0581

Coefficients:

```

              Estimate Std. Error t value Pr(>|t|)
(Intercept) 13.043469   0.651012  20.036 < 2e-16 ***
Price       -0.054459   0.005242 -10.389 < 2e-16 ***
UrbanYes    -0.021916   0.271650  -0.081  0.936
USYes       1.200573    0.259042   4.635 4.86e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.472 on 396 degrees of freedom
Multiple R-squared:  0.2393, Adjusted R-squared:  0.2335
F-statistic: 41.52 on 3 and 396 DF,  p-value: < 2.2e-16

```

- 1) No modelo aplicado como p-value é menor que 5% então a regressão existe e é significativa
- 2) As variáveis que explicam as vendas (Sales) são Price (sendo preço da empresa em cada local) e US (sendo uma variável qualitativa indicando se a loja está localizada nos Estados Unidos ou não).
- 3) Como R quadrado ajustado é de 0.2335 a relação entre as 3 variáveis é baixa

### 0.1.2 Equação matemática do modelo

$Sales = 13.043469 - 0.054459 * Price - 0.021916 * Urban (Yes) + 1.200573 * US (yes)$

Para as variáveis Price e US (yes) pode rejeitar a hipótese de  $\beta = 0$

### 0.1.3 Ajustando modelo - Modelo D

```
[23]: modelo <- lm(Sales ~ Price + US, Carseats)
      summary(modelo)
```

Call:

```
lm(formula = Sales ~ Price + US, data = Carseats)
```

Residuals:

```

      Min       1Q   Median       3Q      Max
-6.9269 -1.6286 -0.0574  1.5766  7.0515

```

Coefficients:

```

              Estimate Std. Error t value Pr(>|t|)
(Intercept) 13.03079    0.63098  20.652 < 2e-16 ***
Price       -0.05448    0.00523 -10.416 < 2e-16 ***
USYes       1.19964    0.25846   4.641 4.71e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 2.469 on 397 degrees of freedom
Multiple R-squared:  0.2393, Adjusted R-squared:  0.2354

```

F-statistic: 62.43 on 2 and 397 DF, p-value:  $< 2.2e-16$

- 1) O valor do R quadrado ajustado sofre um ligeiro aumento no modelo D, sendo assim o modelo D se mostra melhor do que o modelo C usando a variável Urban.

[ ]: