

Projeto Floricultura

```
1 install.packages("repr")
2 install.packages("effects")
3 install.packages("scatterplot3d")

Installing package into '/usr/local/lib/R/site-library'
(as 'lib' is unspecified)

Installing package into '/usr/local/lib/R/site-library'
(as 'lib' is unspecified)

Installing package into '/usr/local/lib/R/site-library'
(as 'lib' is unspecified)

1 library(repr)
2 options(repr.plot.width = 16, repr.plot.height = 8)
```

```
1 #Carregar a base iris
2 data(iris)
3
4 #Visualização geral
5 head(iris)
6 str(iris)
7 summary(iris)
8
9 #Tabela de frequência para a variável categórica
10 table(iris$Species)
11
12 #Gráficos exploratórios
13 boxplot(Petal.Length ~ Species, data = iris, main = "Petal Length por Espécie", col = "lightblue")
14 pairs(iris[1:4], main = "Matriz de dispersão")
```

A data.frame: 6 x 5

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
	<dbl>	<dbl>	<dbl>	<dbl>	<fct>
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa

'data.frame': 150 obs. of 5 variables:

```
$ Sepal.Length: num 5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
$ Sepal.Width: num 3.5 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
$ Petal.Length: num 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
$ Petal.Width: num 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
$ Species: Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1 ...
```

Sepal.Length Sepal.Width Petal.Length Petal.Width

Min. :4.300 Min. :2.000 Min. :1.000 Min. :0.100

1st Qu.:5.100 1st Qu.:2.800 1st Qu.:1.600 1st Qu.:0.300

Median :5.800 Median :3.000 Median :4.350 Median :1.300

Mean :5.843 Mean :3.057 Mean :4.758 Mean :1.199

3rd Qu.:6.400 3rd Qu.:3.300 3rd Qu.:5.100 3rd Qu.:1.800

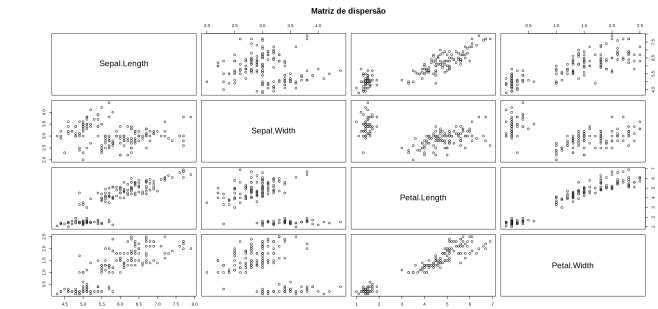
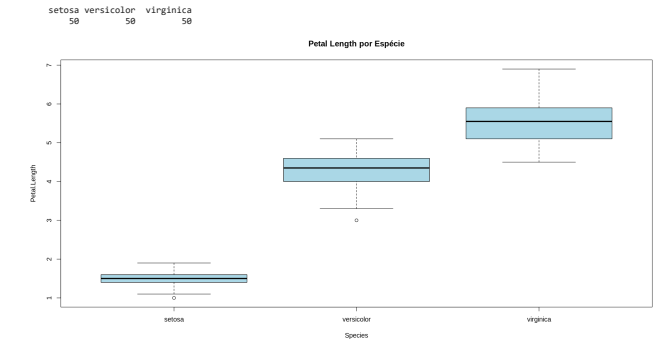
Max. :7.900 Max. :4.400 Max. :6.900 Max. :2.500

Species

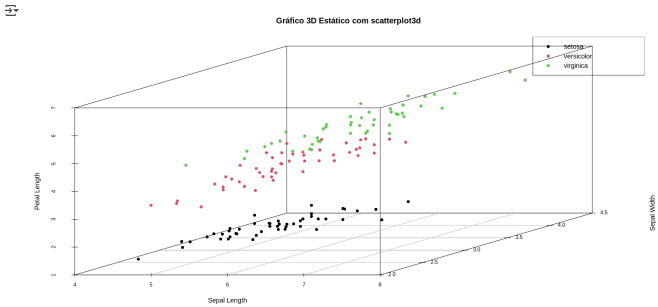
setosa :50

versicolor:50

virginica :50



```
1 library(scatterplot3d)
2
3 #Atribuir cores por espécie
4 colors <- as.numeric(iris$Species)
5
6 scatterplot3d(iris$Sepal.Length,
7               iris$Sepal.Width,
8               iris$Petal.Length,
9               color = colors,
10              pch = 16,
11              xlab = "Sepal Length",
12              ylab = "Sepal Width",
13              zlab = "Petal Length",
14              main = "Gráfico 3D Estático com scatterplot3d")
15 legend("topright", legend = levels(iris$Species), col = 1:3, pch= 16)
```



```

1 #-----
2 # Regressão Linear Simples
3 #-----
4
5 modelo_simples <- lm(Petal.Length ~ Sepal.Length, data = iris)
6 summary(modelo_simples)
7
8 # Gráfico com a reta de regressão
9 plot(iris$Sepal.Length,
10      iris$Petal.Length,
11      main = "Petal.Length vs Sepal.Length",
12      xlab = "Sepal.Length",
13      ylab = "Petal.Length",
14      pch = 16)
15 abline(modelo_simples, col = "blue", lwd = 2)
16 grid()

```

```

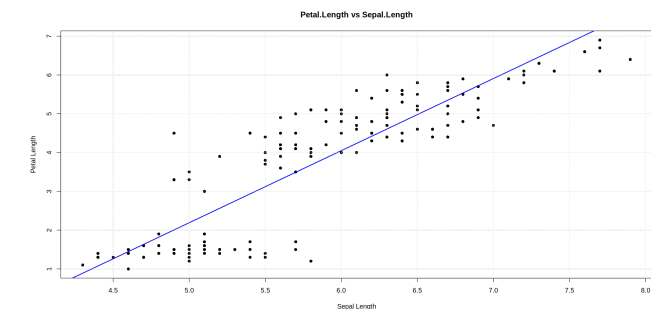
Call:
lm(formula = Petal.Length ~ Sepal.Length, data = iris)

Residuals:
    Min       1Q   Median       3Q      Max
-2.47747 -0.59072 -0.00668  0.60484  2.49512

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -7.10144    0.50666  -14.02  <2e-16 ***
Sepal.Length  1.85043    0.08580   21.65  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8678 on 148 degrees of freedom
Multiple R-squared:  0.76,    Adjusted R-squared:  0.7583
F-statistic: 468.6 on 1 and 148 DF,  p-value: < 2.2e-16

```



```

1 #-----
2 # Regressão Linear múltipla
3 #-----
4
5 # Ajustar modelo de regressão linear múltipla
6 modelo_multiplo <- lm(Petal.Length ~ Sepal.Length + Sepal.Width, data = iris)
7 summary(modelo_multiplo)
8
9 # Extrair coeficientes
10 coef <- round(coef(modelo_multiplo), 2)
11 eq_text <- paste0("Petal.Length = ", coef[1], "+",
12                  coef[2], " * Sepal.Length +",
13                  coef[3], " * Sepal.Width")
14
15 # Criar gráfico 3D com pontos
16 grafico <- scatterplot3d(iris$Sepal.Length,
17                          iris$Sepal.Width,
18                          iris$Petal.Length,
19                          pch = 16, color = as.numeric(iris$Species),
20                          xlab = "Sepal.Length", ylab = "Sepal.Width", zlab = "Petal.Length",
21                          main = "Regressão Linear Múltipla: Petal.Length ~ Sepal.Length + Sepal.Width")
22
23 # Adicionar plano de regressão
24 grafico$plane3d(modelo_multiplo)
25
26 # Adicionar a equação no gráfico (posição X, Y, Z ajustável)
27 text(x = 4.5, y = 4.2, labels = eq_text, cex = 0.8, pos = 4)

```

```

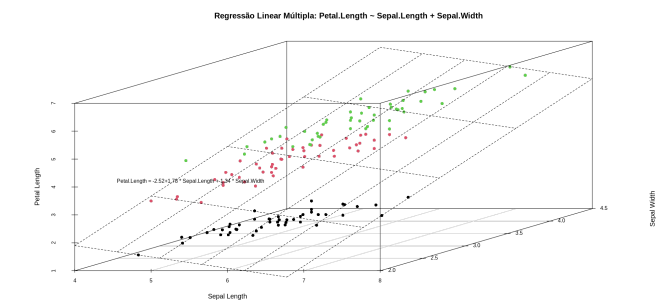
Call:
lm(formula = Petal.Length ~ Sepal.Length + Sepal.Width, data = iris)

Residuals:
    Min       1Q   Median       3Q      Max
-1.25582 -0.46922 -0.05741  0.45530  1.75999

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.52476    0.56344   -4.481 1.48e-05 ***
Sepal.Length  1.77559    0.06441  27.569 < 2e-16 ***
Sepal.Width  -1.33862    0.12236 -10.948 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6465 on 147 degrees of freedom
Multiple R-squared:  0.8677,    Adjusted R-squared:  0.8659
F-statistic:  482 on 2 and 147 DF,  p-value: < 2.2e-16

```



```

1 # Comparação com o modelo anterior
2 anova(modelo_simples, modelo_multiplo)

```

```

A anova: 2 x 6
  Res.Df    RSS    DF Sum of Sq    F    Pr(>F)
<dbl>    <dbl> <dbl>    <dbl>    <dbl>    <dbl>
1     148 111.45916    NA      NA      NA      NA
2     147  61.43675     1  50.02241 119.6809 9.429194e-21

1 #-----
2 # Inclusão de variável categórica e efeitos
3 #-----
4
5 is.factor(iris$Species) # Deve retornar TRUE
6 modelo_species <- lm(Petal.Length ~ Sepal.Length + Species, data = iris)
7 summary(modelo_species)
8
9 # Visualização dos efeitos
10 library(effects)
11 efeitos <- allEffects(modelo_species)
12 plot(efeitos)

```

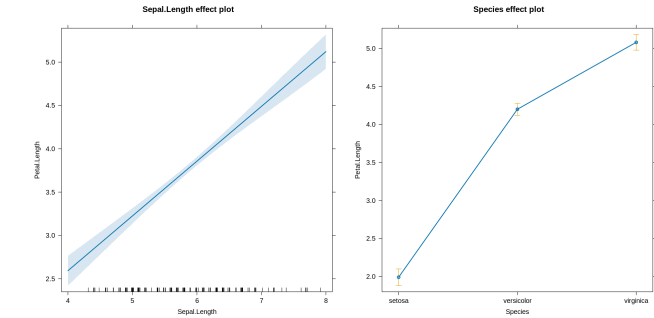
```
Call:
lm(formula = Petal.Length ~ Sepal.Length + Species, data = iris)

Residuals:
    Min       1Q   Median       3Q      Max
-0.76390 -0.17875  0.00716  0.17461  0.79954

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.78234    0.23013   -7.397 1.01e-11 ***
Sepal.Length   0.63211    0.04527  13.962 < 2e-16 ***
Speciesversicolor  2.21914    0.07047  31.362 < 2e-16 ***
Speciesvirginica   3.09000    0.09123  33.870 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2826 on 146 degrees of freedom
Multiple R-squared:  0.9749,    Adjusted R-squared:  0.9744
F-statistic: 1890 on 3 and 146 DF,  p-value: < 2.2e-16
Loading required package: carData

lattice theme set by effectsTheme()
See ?effectsTheme for details.
```



Questões

- Qual variável teve o maior impacto sobre o comprimento da pétala: o comprimento da sépala ou a largura da sépala?
 - Analisando o "modelo_múltiplo" (Regressão Linear Múltipla - célula uoKaKfmXmDuR), podemos ver os coeficientes de cada variável:
 - Sepal.Length: 1.78
 - Sepal.Width: -1.34
 - O coeficiente indica a magnitude do impacto da variável no comprimento da pétala, mantendo outras variáveis constantes. Neste caso, o comprimento da sépala (Sepal.Length) tem um coeficiente com valor absoluto maior (1.78) do que a largura da sépala (1.34), indicando que o comprimento da sépala teve o maior impacto no comprimento da pétala neste modelo.
- O modelo parece se ajustar bem visualmente aos dados?
 - Modelo Simples : O gráfico mostra uma relação positiva entre o comprimento da sépala e o comprimento da pétala. A reta de regressão azul parece capturar a tendência geral, mas há uma dispersão considerável dos pontos em torno da linha, especialmente nas extremidades. Isso sugere que o modelo simples não explica completamente a variação no comprimento da pétala.
 - Modelo Múltiplo : O gráfico 3D com o plano de regressão visualiza o ajuste. Embora o plano tente capturar a relação tridimensional, a dispersão dos pontos coloridos (representando as espécies) ao redor do plano indica que, embora a inclusão da largura da sépala melhore o ajuste (como visto no R-quadrado ajustado e na anova), ainda há variação não explicada, provavelmente devido às diferenças entre as espécies.
 - Modelo com Espécie: O gráfico de efeitos (plot(efeitos)) mostra claramente que a espécie tem um impacto muito significativo no comprimento da pétala. As diferentes linhas para cada espécie indicam que a relação entre o comprimento da sépala e o comprimento da pétala varia consideravelmente entre as espécies. Visualmente, este modelo (considerando os efeitos das espécies) parece se ajustar melhor aos dados, mostrando como a espécie é um fator crucial.
 - Em resumo, o modelo simples e o modelo múltiplo com Sepal.Width e Sepal.Length capturam algumas tendências, mas a inspeção visual e os modelos posteriores (especialmente o que inclui "Species") mostram que as diferenças entre as espécies são muito importantes e que os primeiros modelos, embora úteis, não se ajustam perfeitamente aos dados sem considerar a espécie.
- Como você interpretaria esse modelo para um colega da floricultura que não conhece estatística?
 - Analisamos nossos dados de flores para entender o que influencia o tamanho da pétala. Descobrimos que o comprimento da sépala (aquela parte verde logo abaixo da pétala) é o que mais influencia o comprimento da pétala. Quanto maior a sépala, geralmente maior a pétala. A largura da sépala também tem um impacto, mas não tanto quanto o comprimento. É como se o comprimento da sépala fosse o "motor" principal para o tamanho da pétala. Mas o mais importante que vimos é que o tipo da flor (se é Setosa, Versicolor ou Virginica) faz uma diferença enorme no tamanho da pétala. Cada tipo de flor tem um padrão diferente. Então, para saber o tamanho da pétala, precisamos saber tanto o comprimento da sépala quanto o tipo da flor. O comprimento da sépala nos ajuda a prever o tamanho da pétala dentro de cada tipo de flor, mas o tipo de flor em si já nos dá uma grande pista do tamanho que a pétala pode ter.