

INSTITUTO SUPERIOR DE AGRONOMIA
ESTATÍSTICA E DELINEAMENTO EXPERIMENTAL – 2025-26

29 outubro 2025

Primeiro Teste

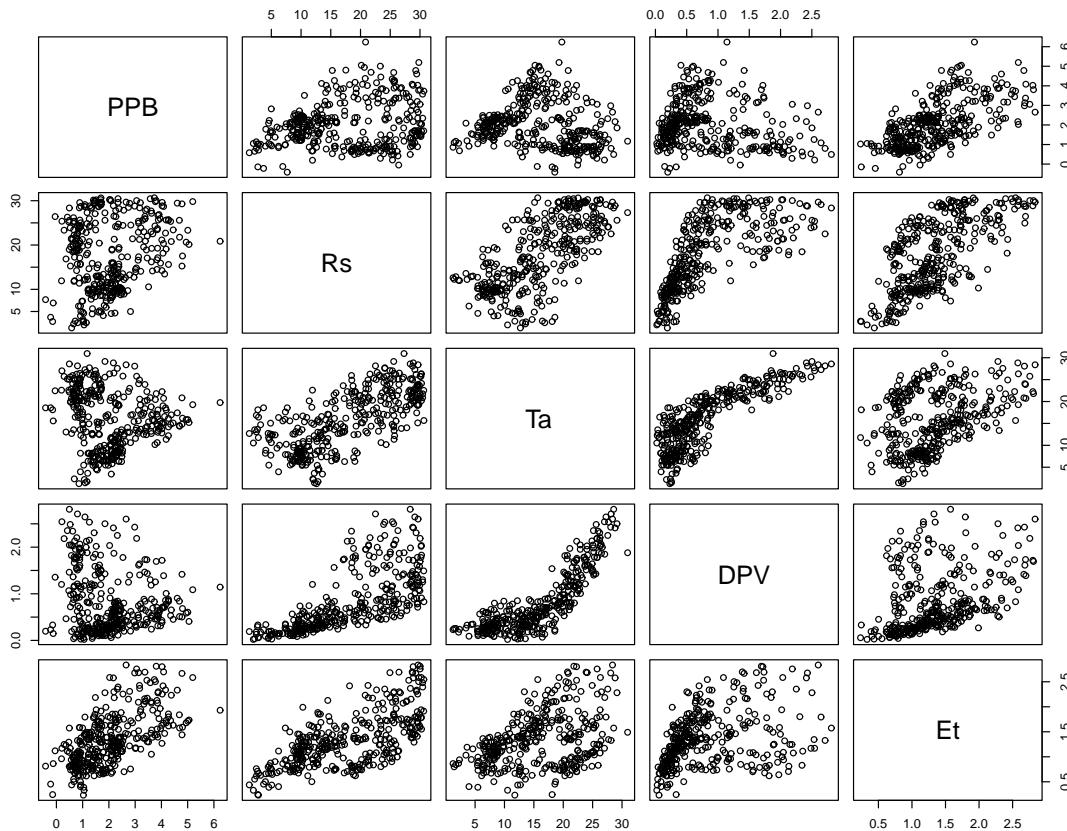
Duração: 2h (+15 min.)

Num estudo sobre os ciclos de água e de carbono numa floresta foram medidas, em 354 dias escolhidos ao acaso ($n = 354$), várias variáveis, nomeadamente, a transpiração das árvores (variável Et, em mm H₂O/dia), a produção primária bruta (PPB, em g C/m²/dia), a radiação solar (Rs, em MJ/m²/dia), a temperatura média do ar (Ta, em °C) e o défice médio de pressão de vapor do ar (DPV, em kPa). Eis os seguintes indicadores estatísticos, bem como a matriz de correlações e nuvens de pontos para todos os pares de variáveis.

Variável	Média	Variância
PPB	2.0412	1.3918
Rs	17.705	63.8911
Ta	15.80	44.6365
DPV	0.81528	0.4261
Et	1.3424	0.2889

```
> round(cor(floresta),4)
      PPB      Rs      Ta      DPV      Et
PPB  1.0000  0.2265 -0.1483 -0.1611  0.6359
Rs   0.2265  1.0000  0.7009  0.7082  0.6669
Ta  -0.1483  0.7009  1.0000  0.8390  0.3868
DPV -0.1611  0.7082  0.8390  1.0000  0.3750
Et   0.6359  0.6669  0.3868  0.3750  1.0000
```

```
> plot(floresta)
```



I [15 valores]

Com o objectivo de modelar a transpiração das árvores (variável Et), foi ajustado um modelo de regressão linear múltipla de Et sobre as restantes variáveis observadas. Eis os resultados obtidos com o ajustamento deste modelo:

```

> floresta.lm1<-lm(Et~PPB+Rs+Ta+DPV, data=floresta)
> summary(floresta.lm1)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.119291   0.054988   2.169   0.0307  
PPB          0.266494   0.015573  17.113 < 2e-16  
Rs           0.025346   0.003344   7.580 3.14e-13  
Ta           0.010751   0.004529   2.374   0.0181  
DPV          0.074280         A       B      0.1190  
---
Residual standard error: 0.2917 on 349 degrees of freedom
Multiple R-squared:  0.7088, Adjusted R-squared:  0.7055 
F-statistic: 212.4 on 4 and C DF,  p-value: < 2.2e-16

> round(vcov(floresta.lm1), 6)
            (Intercept)      PPB        Rs        Ta        DPV      
(Intercept) 0.003024 -0.000385 -1.3e-05 -0.000158  0.000897  
PPB         -0.000385  0.000243 -2.7e-05  0.000013  0.000188  
Rs          -0.000013 -0.000027  1.1e-05 -0.000005 -0.000062  
Ta          -0.000158  0.000013 -5.0e-06  0.000021 -0.000129  
DPV          0.000897  0.000188 -6.2e-05 -0.000129  0.002259  

> confint(floresta.lm1)
              2.5 %    97.5 %
(Intercept) 0.011141193 0.22743999
PPB         0.235865565 0.29712309
Rs          0.018769799 0.03192254
Ta          0.001843727 0.01965907
DPV         -0.019201090 0.16776097

```

Tendo em conta a informação disponibilizada, responda às seguintes questões.

1. Descreva, em pormenor, o modelo linear ajustado, não se esquecendo de mencionar os seus pressupostos. Quantos parâmetros tem este modelo?
2. No *output* do comando *summary* do R identifique e calcule, justificando, os valores em falta A, B e C.
3. Calcule a Soma de Quadrados Total (SQT), a Soma de Quadrados da Regressão (SQR) e a Soma de Quadrados dos Resíduos (SQRE) do modelo ajustado.
4. Sabendo que numa regressão linear a média dos valores observados da variável resposta é igual à média dos valores ajustados, prove que a soma dos resíduos é zero.
5. Usando notação matricial, prove que os estimadores dos parâmetros do modelo linear são estimadores centrados, ou seja, o seu valor esperado é igual ao parâmetro.
6. Interprete no contexto do problema o significado da estimativa do coeficiente associado ao preditor **Rs**. Sem fazer quaisquer cálculos, diga, com 95% de confiança, quais são os valores possíveis do coeficiente na população?
7. Será admissível afirmar que quando a temperatura média do ar (**Ta**) aumenta 1 °C, mantendo os restantes preditores constantes, a transpiração das árvores aumenta, em média, mais de 0.01 mm H₂O/dia? Responda através do teste de hipóteses adequado explicando, em pormenor, todos os seus passos.
8. Será admissível considerar que os coeficientes associados às variáveis preditoras temperatura média do ar (**Ta**) e radiação solar (**Rs**) são iguais? Responda através do teste de hipóteses adequado explicando, em pormenor, todos os seus passos.

9. Para um dia com produção primária bruta (PPB) de 3 g C/m²/dia, radiação solar (Rs) de 19 MJ/m²/dia, temperatura média do ar (Ta) de 16 °C e défice médio de pressão de vapor do ar (DPV) de 0.9 kPa, calcule:

- (a) O valor estimado da transpiração das árvores.
- (b) Um intervalo de confiança (95%) para o valor esperado da transpiração das árvores.

NOTA: o analista fez os seguintes cálculos no R:

```
florestavcov<-vcov(floresta.lm1)
a<-c(1,3,19,16,0.9)
sqrt(t(a)%%florestavcov%*%a)
[1,] 0.02115743
```

10. Explique por que o valor do R^2 modificado está próximo do valor do coeficiente de determinação.
11. É sugerido eliminar o preditor DPV. Sem fazer quaisquer cálculos, diga qual o intervalo de valores possível para o coeficiente de determinação desse submodelo.
12. Diga, justificando, qual o melhor modelo de regressão linear simples para prever a transpiração das árvores (variável Et).

II [5 valores]

Um especialista sugeriu que seria muito importante modelar o défice médio de pressão de vapor do ar (DPV, em kPa) a partir da temperatura média do ar (Ta, em °C). No entanto, dado o padrão da nuvem de pontos que relaciona estas duas variáveis, optou por fazer uma transformação linearizante, tendo ajustado o seguinte modelo:

```
> floresta.lm2<-lm(log(DPV)~Ta, data=floresta)
> summary(floresta.lm2)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.246028   0.077463 -29.00 <2e-16
Ta          0.106755   0.004518  23.63 <2e-16
---
Residual standard error: 0.5671 on 352 degrees of freedom
Multiple R-squared:  0.6134, Adjusted R-squared:  0.6123
F-statistic: ? on 1 and 352 DF, p-value: < 2.2e-16
```

1. Escreva a equação da recta ajustada e interprete o valor do coeficiente de determinação.
2. Para um $\alpha = 0.05$, o modelo adotado pelo analista é inútil? Responda através de um teste formal.
3. Deduza a relação não linear entre o défice médio de pressão de vapor do ar (DPV) e a temperatura média do ar (Ta). Escreva a equação da curva ajustada.
4. Sabendo que numa regressão linear simples a Soma dos Quadrados da Regressão (SQR) é dada por $SQR = \hat{\beta}_1^2(n-1)s_x^2$, prove que no contexto de uma regressão linear simples, a estatística do teste F de ajustamento global é o quadrado da estatística T do teste ao coeficiente β_1 ser zero.