

1. Dataset Completo

Descrição:

O conjunto de dados analisado contém informações sobre folhas, incluindo comprimento (comp), largura (larg), área foliar (foliar) e número de folhas (folhas). A visualização dos dados por meio de gráficos de dispersão permite identificar padrões e possíveis relações entre as variáveis.

2. Correlação e Coeficiente de Determinação

Análise:

- A correlação entre largura e área foliar foi de **0,94**, indicando uma forte relação linear positiva.
 - O coeficiente de determinação ($R^2 = 0,88$) mostra que **88% da variação** em área foliar pode ser explicada pela largura da folha.
 - Isso sugere que folhas mais largas tendem a ter áreas foliares maiores.
-

3. Modelo Linear Simples

Ajuste e Resultados:

- Foi ajustado um modelo linear simples: foliar ~ larg.
 - O intercepto estimado foi **-14,803** e o coeficiente de largura foi **11,453**.
 - Isso significa que, para cada aumento unitário em largura, espera-se um aumento de 11,453 na área foliar.
 - O p-valor associado ao coeficiente de largura foi significativo, indicando forte evidência estatística da relação.
 - O modelo apresentou bom ajuste (R^2 ajustado = **0,86**).
-

4. Validação do Modelo

Diagnóstico dos Resíduos:

- O teste de normalidade de Shapiro-Wilk resultou em **p = 0,3079**, não rejeitando a hipótese de normalidade dos resíduos.
- O histograma dos resíduos padronizados mostra distribuição aproximadamente normal, validando o uso do modelo linear.

5. Modelo Linear Múltiplo

Ajuste e Interpretação:

- Um modelo múltiplo foi ajustado: $\text{foliar} \sim \text{comp} + \text{larg} + \text{folhas}$.
- Apenas o coeficiente de largura (**7,85**) foi significativo (**$p = 0,0182$**), enquanto comprimento e número de folhas não apresentaram significância.
- O modelo múltiplo apresentou excelente ajuste (**$R^2 \text{ ajustado} = 0,951$**), explicando **95,1% da variação** em área foliar.

6. Modelo Alternativo

Explorando Outras Relações:

- Um modelo alternativo foi testado com o número de folhas como variável dependente: $\text{folhas} \sim \text{comp} + \text{larg} + \text{foliar}$.
- Nenhum coeficiente foi significativo e o ajuste do modelo foi fraco (**$R^2 \text{ ajustado} = 0,4319$**).
- Isso indica que, para este conjunto de dados, o número de folhas não é bem explicado pelas demais variáveis.

7. Conclusão

- A largura da folha é o principal fator explicativo para a área foliar neste conjunto de dados.
- O modelo linear simples já apresenta bom desempenho, mas o modelo múltiplo reforça a importância da largura.
- Os resíduos seguem distribuição normal, validando as inferências estatísticas.
- Modelos alternativos podem ser testados, mas é importante considerar a significância estatística e o ajuste do modelo.
- **Próximos passos:** coletar mais dados, testar outras variáveis ou modelos não-lineares, e investigar possíveis fatores externos.