

Classwork 06

Instruções:

1. Este Classwork (CW) poderá ser resolvido em duplas;
2. O CW precisará conter o(s) **nome(s) completo(s)** e a(s) **matrícula(s)** do(s) discente(s);
3. **Serão considerados apenas** os Classworks resolvidos por discentes presentes;
4. O CW deverá ser enviado via **Google Classroom até as 07h51 do dia 22/05/2026** em **um único arquivo** e em **PDF** (*Portable Document Format*);
5. **Não serão considerados** Classworks com péssima qualidade visual ou em desacordo com as instruções anteriores.

Exercício. Considere o banco de dados de Pétilion et al. (2025), construído para testar se o número de etimologias dedicadas por cada primeiro autor a cientistas (“**scientists**”) foi afetado pelo gênero do primeiro autor (“**gender**”), pelo ano da primeira descrição da espécie (“**year_min**”) e pelo total de anos de atividade (“**year_range**”). Assim como os autores, considere a variável **year_min** como quantitativa e, em seguida, utilize a escala normal de unidade, isto é, padronize a variável subtraindo sua média e dividindo o resultado pelo desvio padrão. A partir desses dados, ajuste um modelo de regressão linear múltipla da forma:

$$\text{scientists} \sim \beta_1 + \beta_2 \text{gender} + \beta_3 \text{year_min} + \beta_4 \text{year_range}.$$

Tarefas.

1. Realize uma análise exploratória dos dados. Apresente gráficos e medidas descritivas que permitam compreender o comportamento das variáveis. Discuta possíveis padrões, relações entre as variáveis e a presença de valores atípicos.
2. Ajuste o modelo completo e avalie a significância estatística dos coeficientes (considere, por exemplo, um nível de significância de 5%). Caso existam variáveis não significativas, proponha um modelo reduzido, justificando a exclusão dessas variáveis.

3. Verifique a adequação do modelo ajustado por meio da análise de resíduos e de diagnósticos apropriados. Discuta a presença de possíveis problemas, como observações influentes ou padrões sistemáticos nos resíduos. Quando apropriado, proponha uma transformação da variável resposta que possa melhorar o ajuste do modelo.
4. Interprete os resultados do modelo final, incluindo os coeficientes estimados. Discuta o efeito das variáveis explicativas sobre o número esperado de etimologias dedicadas a cientistas, mantendo as demais variáveis constantes.

Referências

Pétillon, J., C. Noël, A. Brescovit, Y. M. Marusik, C. A. Rheims, X. Xu, D. Fontaneto, and S. Mammola (2025). The use of eponyms can also promote gender equity in modern taxonomy. *BioScience* 75(9), 691–694.