

Dynamic Documents for Reproducible Research

Use R and Quarto (via RStudio) to produce a reproducible research document for a simple linear model analysis of the finch beak data (Sleuth3 package: case0201), following the requirements given below. The assignment will be assessed on code (functionality, clarity etc – 50%) and content (text, tables, figures) – 50%). You should submit two files:

- 1) A rendered document (HTML).
- 2) The Quarto (Qmd) file that produces the rendered document.

As part of a study of natural selection, biologists have regularly measured individuals of the medium ground finch (*Geospiza fortis*) on the Galapagos Island of Daphne Major. Beak morphology is thought to be under selection in relation to food availability. During a severe drought in 1977 the only food source was a large, tough seed which the finches ordinarily ignore. The data collected included beak depths (depth of the beak at its base, mm) for a random sample of 89 finches caught before the drought (in 1976) and for all 89 finches captured the year after the drought (1978). Are the data consistent with natural selection of beak depth in response to drought?

The report should be functional and understandable by others. Aim to be complete (include all steps in the workflow) but concise (e.g. avoid purely cosmetic formatting and unnecessarily long and inefficient documents). Follow current best practice using appropriate models (e.g. *Journal of Animal Ecology* etc) where appropriate.

CONTENT

Text sections

- **Introduction:** A very short paragraph explaining the analysis aims in your own words.
- **Methods – statistical analysis:** A short paragraph explaining the data, statistical model and analysis.
- **Results:** A short paragraph reporting the results of the analysis.
- **Conclusion:** A concise conclusion with respect to the main aims and questions.

Table(s): Include a small table of some appropriate output generated by your analysis.

Figures (single or multi-panel with explanatory text captions)

- A descriptive plot of the data (including the raw data values).
- A figure displaying the key result of the linear model analysis.
- A multi-panel figure allowing graphical assessment of the normality and heteroscedasticity of the residuals.

CODE

- Do not reveal your identity in the document, including in the code (directory path etc).
- Include ('echo') your code in the report so that it can be assessed.
- Consider sparing use of appropriate options in the YAML header.
- Apply appropriate formatting options to your text.
- Use the `lm` function to perform an appropriate linear model analysis.
- Use the appropriate visualization capabilities in R (e.g. `ggplot2`) to produce publication quality figures.