

# Homework 3

## Type all homework.

Due date: See syllabus. Turn in on Bruin Learn.

You may use any R package to answer these questions.

**Dental data.** The Dental data set is a classic data set for longitudinal data analysis. The response is the length in millimeters from the center of the pituitary gland to the pterygomaxillary fissure for 11 girls and 16 boys. The measurements were taken every two years at ages 8, 10, 12, and 14. There is a single covariate, gender.

All data sets are available on the book web site <https://robweiss.faculty.biostat.ucla.edu/book-data-sets> in text formats. Data sets are described briefly in the book appendix A.

1. **Graphical Analysis.** The purpose of this analysis is to use graphics (exploratory data analysis) to describe the important characteristics of the data.
  - (a) Create a profile plot of the data. Use separate line types (or colors) for the boys and girls.
  - (b) Briefly report your findings. What is the overall pattern? Are boys and girls different? In what ways?
  - (c) Calculate the correlations among the observations, and draw a scatterplot matrix. Use separate plotting characters for boys and girls.
  - (d) Report your additional findings.
  - (e) Draw an empirical summary plot.
  - (f) Draw a second empirical summary plot with separate summaries for boys and girls. What are your conclusions? In particular:
    - i. Is there a difference in level between boys and girls?
    - ii. Is there a difference in slope between boys and girls?
  - (g) Inspect the profile plot of empirical within-subject residuals  $Y_{ij} - \bar{Y}_i$ . What do you learn about the data?
  - (h) There are four important inferences to draw regarding this data set. What are they?
2. **Data Summaries.** Dental data again. We analyze subject summaries.
  - (a) Calculate averages for each child. Plot these 27 numbers, distinguishing boys from girls. Describe the distributions of the two samples (one sentence). Is there a difference between groups? Present (1) a formal inference (i.e., a test and  $p$ -value) and (2) a plot of the data illustrating the conclusion. Do not present a plot with just two points on it!
  - (b) Repeat the previous item for the slopes for each child.
  - (c) Are the variances the same for boys and girls? Your answer may differ depending on whether you based your answer on different parts of the data. Consider all of
    - i. The averages
    - ii. The slopes

- iii. A profile plot distinguishing boys and girls.
- iv. The profile plot of  $Y_{ij} - \bar{Y}_i$ .
- (d) Draw a scatterplot of the intercepts and slopes, distinguishing boys and girls. Draw conclusions. Anything new here?
- (e) Compare the analysis of **data summaries** to the **graphical** analysis. Which analysis shows more about the data? In a short, itemized list, list what we learn in each analysis and whether we can learn it from both analyses or just from one analysis.

**Important!** Don't answer questions with only a YES/NO answer. Discuss briefly. Properly label all plots:  $X$  and  $Y$  axis labels, reasonable tick marks and tick labels, title and caption.