

## Methods of Data Analysis II

### Homework 2

**Name:**

**OSU-ID:**

#### Instructions

- The homework is due on Friday, Apr. 17th (in class). Late homeworks will not be accepted under any circumstances.
- You may work individually or in groups of 2 people. If you work in groups, you must include the names and OSU-ID numbers of all the students in the group and submit only one assignment per group.
- You must provide complete answers in order to receive full credit. The homework is worth 25 points.
- You must clearly indicate the problem that you are working.
- Homeworks must be stapled when submitted (including this sheet as a cover page). Do not use folders, paper clips or any other objects to keep the pages together.

1. Consider the meadowfoam example discussed in class to answer the following questions:
  - (a) Construct “side-by-side” scatter-plots (including the regression line) for the number of flowers vs. light intensity, with and without distinguishing for the the time effect. Do the plots suggest a linear relationship between these variables? Do the plots suggests an effect due to timing?
  - (b) Regardless your answer from part (a) define the indicator variable *early* taking the value 1, if the timing was before PFI. Use this indicator variable to fit a linear regression model to explain the mean number of flowers in terms of light intensity and timing. Report the output of the model.
  - (c) Write down the model equation and the fitted equation for the model in part (b). What is the difference between the model equation and the fitted equation?
  - (d) Using the output of the model, compute (by hand) the value of the t-statistic and the P-value for the hypothesis test on the coefficient multiplying the indicator variable *early*. Interpret for results.
  - (e) Construct a 95% confidence interval for the coefficient multiplying the indicator variable *before*. What is the interpretation of this interval? Does the interval agree with the results of the hypothesis test in part (c)?
  - (f) For a fix level of timing, use your fitted equation to determine the estimated difference between the mean number of flowers at 600 and 300  $\mu\text{mol}/\text{m}^2/\text{sec}$  of light intensity.

**Note:** The data set can be read in R using the following commands

```
> library('Sleuth3')  
> head(case0901) # To see the data structure  
> case0901 # To read the full data set
```