

## Methods of Data Analysis II

### Homework 4

Name:

OSU-ID:

#### Instructions

- The homework is due on Wednesday, May 20th (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. The Chernobyl Fallout data set consists of cesium concentrations (in  $Bq/Kg$ ) in soil and mushrooms at 17 wooded locations in Umbria (central Italy), from August 1986 to November 1989. Researchers want to investigate the cesium transfer from contaminated soil to plants following the Chernobyl nuclear plant accident in 1986. Specifically, they want to describe the distribution of the mushroom cesium concentration as a function of the soil concentration.
  - (a) Construct a scatter-plot (including the regression line) for *Mushroom* vs *Soil*. Is there any indication of possible influential observations? If so, identify those observations.
  - (b) Fit a simple linear regression model of *Mushroom* vs *Soil* and obtain a summary of the model.
  - (c) Construct the residual plot for your model and comment on the results.
  - (d) Construct the Cook's distance and the studentized residuals plots. Identify any unusual observations.
  - (e) Regardless of your answer in part (d), remove observation 17 and re-fit the model. Do you see any relevant changes from the output obtained in part (b)?
  - (f) Re-do the analysis but considering the transformed variables  $\log(Mushroom)$  vs  $\log(Soil)$ .

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

```
> library('Sleuth3')  
> head(ex1111) # To see the data structure  
> ex1111 # To read the full data set
```