Lecture 13 Practicum

BIOF 339

November 28, 2017

You will create a report based on the diamonds dataset available in the ggplot2 package.

1. Open a R Markdown file.
2. Write a short introduction, displaying the first 5 rows of the diamonds data set.
3. Write a section titled “Descriptive data”
   * Provide information about the average price of the diamonds
   * Create a table showing the average price of diamonds by different clarities
   * Include a boxplot showing the prices by clarity categories. Add a caption (by using the option fig.caption in the R chunk options)
   * Provide a link to the diamonds dataset
4. Write a section titled “Modeling”
   * Fit a linear model to describe the relationship between carat and price
     + Print a table of the results. Add a caption (this is an option in the kable function)
   * Plot price vs carat and display the linear regression line through the data. In ggplot use geom\_smooth(method='lm'). Provide an appropriate caption
5. Write a section titled “Resources”
   * Link to the Duke University page which is a R Markdown generated description of the diamonds dataset and has further exercises to investigate this dataset. The URL is <https://www2.stat.duke.edu/courses/Fall14/sta112.01/app/app2.html>
6. Produce both a HTML and a Word document. Ensure that the actual R code is not visible in the document.
   * If messages and warnings from R appear, you can suppress them in the R chunk options using message = F and warning = F respectively