Project 1: Moisture Content of Tree Branches

Yingxue YU, Guanqun LIU, Yulun Jiang

2022-03-31

I. Introduction

This dataset is originally collected and used for J. Joseph McDermott's botanical research on the effect of cutting methods have on the moisture content of tree branch samples in 1941. In experiments, researchers use twig segments to confirm the moisture content of woody structure. If the sample is simultaneously cut at both ends, then the release of tension in both direction will not result in instantaneous water removal; whereas cutting at one single end will lead to rapid water loss in the vicinity of cut.

Grounded on this fact, we conduct a statistical analysis on validating how different cutting methods will affect the extent of water removal given different cut samples. We intend to find an estimation model of how each factor given in the dataset and their interaction terms with cutting methods are related to the mass of moisture content. We will carry out an exploratory data analysis in section II to examine and preprocess the data. Then we will process our data modeling and assessment results in section III, IV, and V. Finally we will conclude our analysis and which part in the original analysis should be criticized.

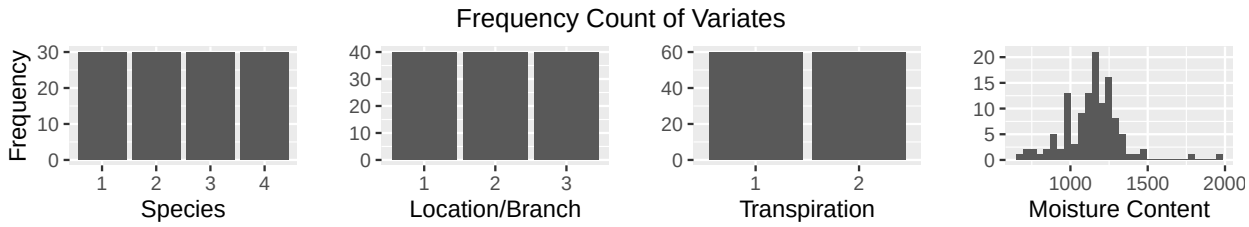
II. EDA

2.1 Data Validation

The dataset includes 120 branch samples, includes 4 variable columns: **species, branch/species, location/branch**, and **transpiration type**. The moisture content in the last column is expressed in $10 \times \%$ of its dry sample weight. Under each species, for each possible combination of cutting location and transpiration type, we have 5 sample measurements of moisture contents, which also reveals that the dataset is balanced. We will later exclude the **branch/species** column because it works as an indicator of the former combination. A variance summary table is provided below. We will keep the numerical values of the classes for further analysis.

Var Name	Type	Category	Classes	Detail
Species	num	categorical	4	1=
Location/Branch	num	categorical	3	
Transpiration	num	categorical	2	
Moisture Content	num(target)	continuous	N/A	

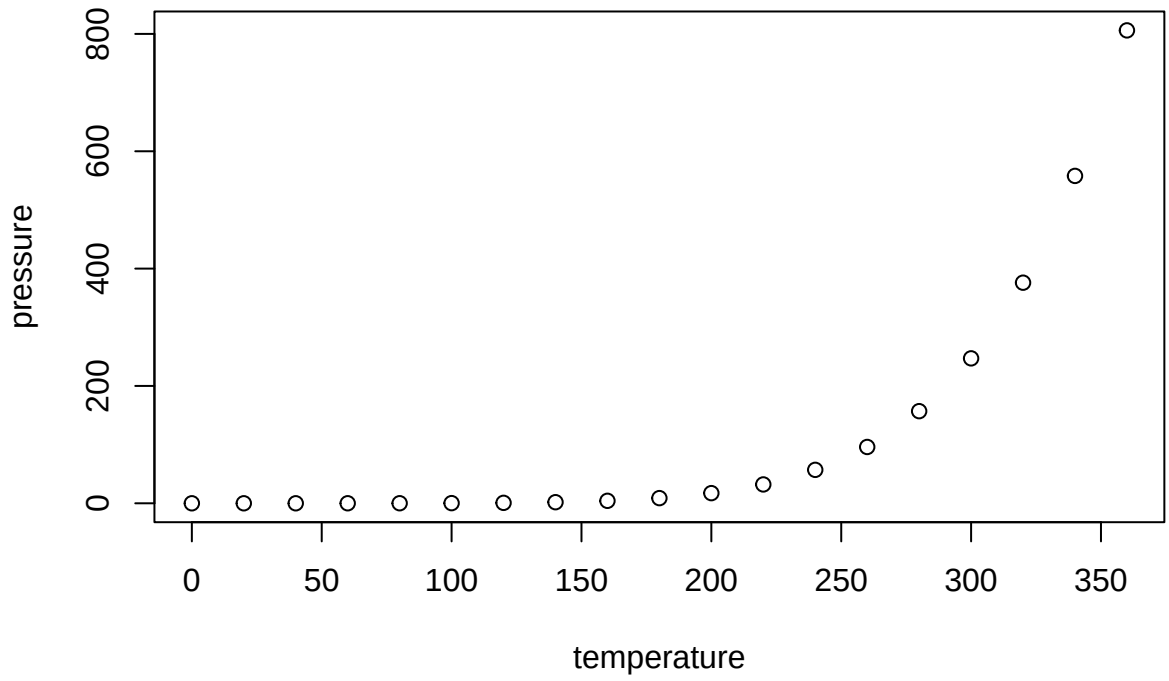
2.2 Univariate Analysis



2.3 Bivariate Analysis

A two way table for the dataset:

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

III. Model Fitting

IV. Assessment

V. Plots

VI. Conclusion