# M13 Problem Set: Multiple Linear Regression for TAs

Pablo E. GutiÃ©rrez-Fonseca

2024-03-28 11:20:09

#1. Import libraries and load packages

```
library(tidyverse)
library(dplyr)
library(readxl)
```

#2. importing our data

```
Plant_height <- read.csv(file = "Plant_height.csv", header = TRUE)
head(Plant_height)
```

```
##   sort_number site         Genus_species        Family growthform height
## 1        1402  193    Acer_macryophyllum   Sapindaceae       Tree   28.0
## 2       25246  103    Quararibea_cordata     Malvaceae       Tree   26.6
## 3       11648   54    Eragrostis_dielsii       Poaceae       Herb    0.3
## 4        8168  144    Cistus_salvifolius     Cistaceae      Shrub    1.6
## 5       22422  178          Phlox_bifida Polemoniaceae       Herb    0.2
## 6       15925   59 Homalium_betulifolium    Salicaceae      Shrub    1.7
##        loght       Country              Site    lat     long entered.by alt
## 1  1.4471580           USA    Oregon - McDun 44.600 -123.334     Angela 179
## 2  1.4248816          Peru              Manu 12.183  -70.550     Angela 386
## 3 -0.5228787     Australia Central Australia 23.800  133.833   Michelle 553
## 4  0.2041200        Israel           Hanadiv 32.555   34.938     Angela 115
## 5 -0.6989700           USA     Indiana Dunes 41.617  -86.950   Michelle 200
## 6  0.2304489 New Caledonia              <NA> 21.500  165.500      Laura  95
##   temp diurn.temp isotherm temp.seas temp.max.warm temp.min.cold temp.ann.range
## 1 10.8       11.8      4.4       5.2          27.0           0.3           26.7
## 2 24.5       10.8      7.4       0.9          31.2          16.7           14.5
## 3 20.9       16.3      4.8       6.0          37.0           3.6           33.4
## 4 19.9        9.7      4.4       4.9          30.7           8.7           22.0
## 5  9.7       10.7      2.8       9.7          28.6          -9.5           38.1
## 6 22.6        7.4      5.4       2.2          29.0          15.5           13.5
##   temp.mean.wetqr temp.mean.dryqr temp.mean.warmqr temp.mean.coldqr rain
## 1             4.9            17.4             17.6              4.5 1208
## 2            25.1            23.2             25.3             23.1 3015
## 3            28.1            14.8             28.1             12.8  278
## 4            13.6            25.3             25.7             13.6  598
## 5            21.6            -3.3             21.6             -3.3  976
## 6            25.4            20.4             25.4             19.7 1387
##   rain.wetm rain.drym rain.seas rain.wetqr rain.dryqr rain.warmqr rain.coldqr
## 1       217        13        69        601         68          75         560
```

```
## 2       416        99        45       1177        340         928         359
## 3        37         9        42        109         35         109          42
## 4       159         0       115        408          0           2         408
## 5       104        44        23        299        165         299         165
## 6       216        59        46        600        186         600         212
##    LAI  NPP hemisphere
## 1 2.51  572          1
## 2 4.26 1405         -1
## 3 1.32  756         -1
## 4 1.01  359          1
## 5 3.26 1131          1
## 6 6.99 1552         -1
```

#3. Run a Simple Linear Regresion

```
model <- lm(loght ~ temp + rain, data = Plant_height)
summary(model)
```

```
##
## Call:
## lm(formula = loght ~ temp + rain, data = Plant_height)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.84556 -0.49216  0.00175  0.40639  1.62168
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.295e-01  1.031e-01  -3.197 0.001648 **
## temp         2.832e-02  6.441e-03   4.396 1.91e-05 ***
## rain         2.463e-04  6.208e-05   3.968 0.000106 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6578 on 175 degrees of freedom
## Multiple R-squared:  0.3085, Adjusted R-squared:  0.3006
## F-statistic: 39.03 on 2 and 175 DF,  p-value: 9.616e-15
```