Linear Regression

FMU Biology Department

## Analysis

This documents includes code for creating a data object in R, creating an informative plot, and running a linear regression. A description of the data can be found in the “Correlation” tutorial.

#Create data frame  
dat3 = data.frame(sodium = c(95,104,88,97,102,89,104,105,111,88,  
 98,102,96,105,107,105,96,90,110,98),  
 elec = c(190,200,192,202,196,195,200,202,206,180,  
 205,199,193,205,208,210,204,197,206,200))  
#Print data frame  
dat3

## sodium elec  
## 1 95 190  
## 2 104 200  
## 3 88 192  
## 4 97 202  
## 5 102 196  
## 6 89 195  
## 7 104 200  
## 8 105 202  
## 9 111 206  
## 10 88 180  
## 11 98 205  
## 12 102 199  
## 13 96 193  
## 14 105 205  
## 15 107 208  
## 16 105 210  
## 17 96 204  
## 18 90 197  
## 19 110 206  
## 20 98 200

#Summary statistics  
#Average and standard deviation of sodium  
mean(dat3$sodium)

## [1] 99.5

sd(dat3$sodium)

## [1] 7.119284

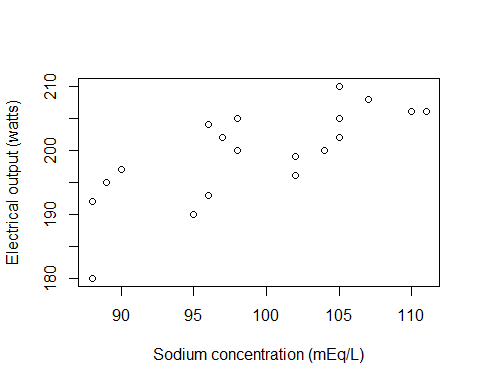
#Average and standard deviation of elec  
mean(dat3$elec)

## [1] 199.5

sd(dat3$elec)

## [1] 7.141428

#Create scatterplot  
plot(dat3$elec~dat3$sodium,  
 ylab = "Electrical output (watts)", xlab = "Sodium concentration (mEq/L)")



#Perform a linear regression analysis  
mod1=lm(dat3$elec ~ dat3$sodium)  
  
summary(mod1)

##   
## Call:  
## lm(formula = dat3$elec ~ dat3$sodium)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.9138 -2.8598 -0.1267 3.5963 7.1132   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 125.2108 15.7492 7.950 2.68e-07 \*\*\*  
## dat3$sodium 0.7466 0.1579 4.728 0.000168 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.9 on 18 degrees of freedom  
## Multiple R-squared: 0.554, Adjusted R-squared: 0.5292   
## F-statistic: 22.36 on 1 and 18 DF, p-value: 0.0001677

### Full code block

#Create data frame  
dat3 = data.frame(sodium = c(95,104,88,97,102,89,104,105,111,88,  
 98,102,96,105,107,105,96,90,110,98),  
 elec = c(190,200,192,202,196,195,200,202,206,180,  
 205,199,193,205,208,210,204,197,206,200))  
#Print data frame  
dat3  
  
#Summary statistics  
#Average and standard deviation of sodium  
mean(dat3$sodium)  
sd(dat3$sodium)  
  
#Average and standard deviation of elec  
mean(dat3$elec)  
sd(dat3$elec)  
  
#Create scatterplot  
plot(dat3$elec~dat3$sodium,  
 ylab = "Electrical output (watts)", xlab = "Sodium concentration (mEq/L)")  
  
#Perform a linear regression analysis  
mod1=lm(dat3$elec ~ dat3$sodium)  
  
summary(mod1)