extra credit

Tom Wilson

October 10, 2018

library(data.table)

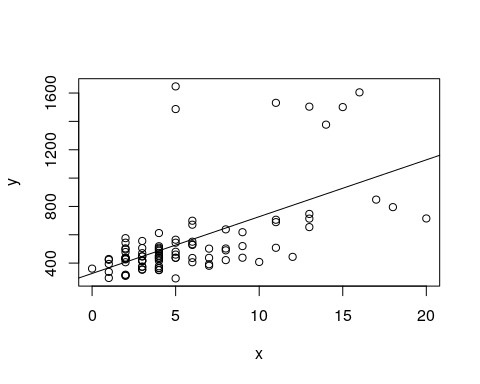
Y= Sales and X=Advert

1. Write the simple linear regression equation.
2. Interpret the slope and intercept terms.

bookstore <- fread('https://raw.githubusercontent.com/wilsonify/AppliedRegression/master/data/bookstore.txt')

x <- bookstore$Advert  
y <- bookstore$Sales  
fit <- lm(y~x)

plot(x,y)  
abline(fit)



summary(fit)

##   
## Call:  
## lm(formula = y ~ x)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -413.89 -109.29 -38.29 27.01 1119.61   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 326.89 43.12 7.582 2.79e-11 \*\*\*  
## x 40.10 6.05 6.628 2.36e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 247.1 on 91 degrees of freedom  
## Multiple R-squared: 0.3256, Adjusted R-squared: 0.3182   
## F-statistic: 43.93 on 1 and 91 DF, p-value: 2.355e-09

A unit change in advert is associated with a change in sales of 40.1. Under the null hypothesis that there is no relationship between advert and sales, we would observe a best fit slope at least this extreme 0.000000236% of the time.

The expected sales with zero advert is 326.89. Since there is data available with zero advert this is a reasonable estimate.