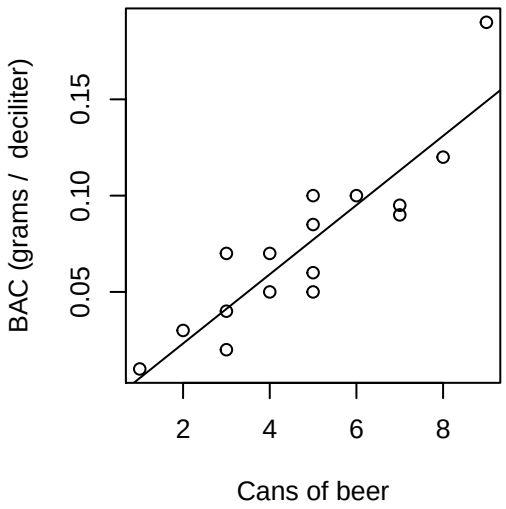
SLR Practice Problem

Many people believe that gender, weight, drinking habits, and many other factors are much more important in predicting blood alcohol content (BAC) than simply considering the number of drinks a person consumed. Here we examine data from sixteen student volunteers at Ohio State University who each drank a randomly assigned number of cans of beer. These students were evenly divided between men and women, and they differed in weight and drinking habits. Thirty minutes later, a police officer measured their blood alcohol content (BAC) in grams of alcohol per deciliter of blood. Below is the R output from fitting a simple linear regression model to this data. A scatter plot with the least squares line is also shown below.

```
library(openintro)
data("bac")
lm1 <- lm(BAC ~ Beers, data = bac)
summary(lm1)

##
## Call:
## lm(formula = BAC ~ Beers, data = bac)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.027118 -0.017350  0.001773  0.008623  0.041027
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.012701   0.012638  -1.005    0.332
## Beers        0.017964   0.002402   7.480 2.97e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02044 on 14 degrees of freedom
## Multiple R-squared:  0.7998,Adjusted R-squared:  0.7855
## F-statistic: 55.94 on 1 and 14 DF,  p-value: 2.969e-06

par(mar=c(4,4,1,1), cex=0.8) # adjust margins and font size
plot(bac$Beers, bac$BAC, xlab="Cans of beer", ylab="BAC (grams /  deciliter)")
abline(lm1)
```



(a) Describe the association between number of cans of beer and BAC.

(b) What are the explanatory and response variables for the linear regression model?

(c) Write the equation for the least squares line.

(d) Interpret the slope and the intercept in context.

(e) What is the predicted BAC for a person that drank 5 cans of beer?

(f) A student in this data set drank 9 beers and had a measured BAC of 0.19. Calculate the residual for this student.

(g) Interpret the coefficient of determination (R^2).

(h) Do the data provide strong evidence that drinking more cans of beer is associated with an increase in blood alcohol content? State the null and alternative hypotheses, report the test statistic and p-value (from the `summary()` command), and state your conclusion.

(i) Calculate a 95% confidence interval for β_1.

(j) Do the data provide evidence that the intercept is significantly different than 0? State the null and alternative hypotheses, report the test statistic and p-value (from the `summary()` command), and state your conclusion.

(k) Calculate a 95% confidence interval for β_0.

(l) Are the conditions for linear regression reasonably satisfied? In your assessment, comment on the plot of the residuals versus number of cans of beer (x), and the QQ plot of the residuals shown below.

```
par(mfrow=c(1,2), cex=0.7)
plot(bac$Beers, resid(lm1), xlab="Number of beers", ylab="Residuals")
abline(h=0)
qqnorm(resid(lm1))
qqline(resid(lm1))
```