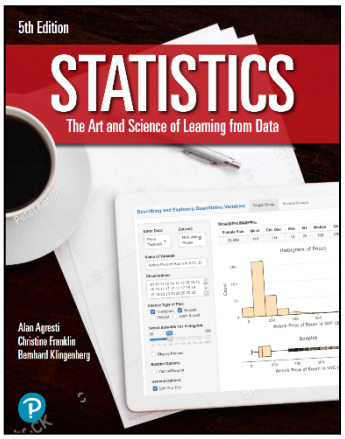

R Code for Examples in the book
"Statistics: The Art and Science of Learning from Data"
by Agresti, Franklin and Klingenberg, 5th edition

Chapter 14

Example 7: Telephone Holding Times – Regression Analysis

Reading in data

```
y <- c(5, 1, 11, 2, 8, 0, 1, 4, 6, 3, 13, 9, 8, 15, 7)
group <- rep(c('A', 'M', 'C'), times = c(5, 5, 5))
```

Fitting in regression model with advertisements as the baseline

```
lin.reg <- lm(y ~ group)
summary(lin.reg)

##
## Call:
## lm(formula = y ~ group)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##   -4.4   -2.6   -0.4    2.6    5.6
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.400      1.523   3.545  0.00403 **
## groupC         5.000      2.154   2.321  0.03868 *
## groupM        -2.600      2.154  -1.207  0.25068
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.406 on 12 degrees of freedom
## Multiple R-squared:  0.5173, Adjusted R-squared:  0.4369
## F-statistic: 6.431 on 2 and 12 DF,  p-value: 0.01264
```

To obtain ANOVA table

```
anova(lin.reg)

## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value  Pr(>F)
## group      2  149.2    74.6   6.431 0.01264 *
```

```
## Residuals 12  139.2    11.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note that R sets up the reference category alphabetically, so the baseline here is advertisements. You can change the reference category by using the `relevel()` function

```
group <- relevel(as.factor(group), ref = 'C')
```

Fitting in regression model with classical as the baseline

```
lin.reg <- lm(y ~ group)
summary(lin.reg)

##
## Call:
## lm(formula = y ~ group)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##   -4.4   -2.6   -0.4    2.6    5.6
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   10.400      1.523   6.828 1.83e-05 ***
## groupA        -5.000      2.154  -2.321  0.03868 *
## groupM        -7.600      2.154  -3.528  0.00416 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.406 on 12 degrees of freedom
## Multiple R-squared:  0.5173, Adjusted R-squared:  0.4369
## F-statistic: 6.431 on 2 and 12 DF,  p-value: 0.01264
```

You can verify that the ANOVA table output should be the same regardless of which reference category is used

```
anova(lin.reg)

## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value  Pr(>F)
## group      2  149.2    74.6   6.431 0.01264 *
## Residuals 12  139.2    11.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```