ST 412/512
Homework 2 – Solutions

Q1:

1. $\mu\{Flowers|Intensity, Time + 1\} - \{Flowers|Intensity, Time\} =$
   $(\gamma_0 \ + \ \gamma_1 Intensity + \gamma_2(Time + \ 1)) - (\gamma_0 \ + \ \gamma_1 Intensity + \gamma_2(Time)) = \ \gamma_2$

   Effect of time = $\gamma_2$

2. In model 1 we have, the change in mean flowers moving from early to late timing is: $\mu\{Flowers|Intensity, Early = 1\} - \{Flowers|Intensity, Early = 0\}$
   $= (\beta_0 + \beta_1 Intensity + \beta_2 1) \ - \ (\beta_0 + \beta_1 Intensity + \beta_2 0) = \ \beta_2$

   And that should be equivalent to, in model 2, time moving from 24 to 0:
   $\mu\{Flowers|Intensity, Time = 24\} - \{Flowers|Intensity, Time = 0\} =$
   $(\gamma_0 \ + \ \gamma_1 Intensity + \gamma_2(24)) - (\gamma_0 \ + \ \gamma_1 Intensity + \gamma_2(0)) = \ 24\gamma_2$

   These two changes in mean should be equal, so have: $\beta_2 = 24\gamma_2$

3. Model with early variable:

```
Call:
lm(formula = Flowers ~ Intens + Time, data = case0901)

Residuals:
   Min     1Q Median     3Q    Max 
-9.652 -4.139 -1.558  5.632 12.165 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 71.305834   3.273772  21.781 6.77e-16 ***
Intens      -0.040471   0.005132  -7.886 1.04e-07 ***
TimeEarly   12.158333   2.629557   4.624 0.000146 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.441 on 21 degrees of freedom
Multiple R-squared: 0.7992, Adjusted R-squared:  0.78 
F-statistic: 41.78 on 2 and 21 DF,  p-value: 4.786e-08
```

   Model with time variable:

```
Call:
lm(formula = Flowers ~ Intens + time, data = case0901)

Residuals:
   Min     1Q Median     3Q    Max 
-9.652 -4.139 -1.558  5.632 12.165 

Coefficients:
```

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 71.305834   3.273772  21.781 6.77e-16 ***
Intens      -0.040471   0.005132  -7.886 1.04e-07 ***
time         0.506597   0.109565   4.624 0.000146 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.441 on 21 degrees of freedom
Multiple R-squared: 0.7992, Adjusted R-squared:  0.78
F-statistic: 41.78 on 2 and 21 DF,  p-value: 4.786e-08
```

Note: the coefficient estimates for intercept and intensity are equal for the two models. $\hat{\beta}_2 = 12.1583; \hat{\gamma}_2 = 0.5066 \rightarrow 24 * 0.5066 \approx 12.1583.$

R Input:

```
#### Question 1 ####

head(case0901)
case0901$time = ifelse(case0901$Time == "Early", 24, 0)

# Fit model using early variable
m.early = lm(Flowers ~ Intensity + Time, data = case0901)
summary(m.early)

# Fit model using time variable
m.time = lm(Flowers ~ Intensity + time, data = case0901)
summary(m.time)
```

Q2:

1.

a) $\mu\{Flowers|Intensity,\ early\} = \beta_0 + \beta_6$

b) $\mu\{Flowers|Intensity,\ early\} = \beta_0$

c) $\mu\{Flowers|Intensity,\ early\} = \beta_0 + \beta_2 + \beta_6$

d) $\mu\{Flowers|Intensity,\ early\} = \beta_0 + \beta_2$

2.

a) $\mu\{Flowers|Intensity,\ early\} = \beta_0 + \beta_6$

b) $\mu\{Flowers|Intensity,\ early\} = \beta_0$

c) $\mu\{Flowers|Intensity,\ early\} = \beta_0 + \beta_2 + \beta_6 + \beta_8$

d) $\mu\{Flowers|Intensity,\ early\} = \beta_0 + \beta_2$

3. Plot A = no interaction variable present in model;
   Plot B = interaction variable present in model

Q3:

For problem three, let:

I = college; II = high school/some college; III = high school only

$v.II = \begin{cases} 1\ if\ in\ group\ II \\ 0\ if\ in\ group\ I\ or\ III \end{cases}$ & $v.III = \begin{cases} 1\ if\ in\ group\ III \\ 0\ if\ in\ group\ I\ or\ II \end{cases}$

1. $\mu\{Score|Age, Educ\} = \beta_0 + \beta_1 Age + \beta_2(v.II) + \beta_3(v.III) + \beta_4[(Age) * (v.II)] + \beta_5[(Age) * (v.III)]$

   Diverging gap measured by: $\beta_5$

2. $\mu\{Score|Age, Educ\} = \beta_0 + \beta_1 Age + \beta_2(v.II) + \beta_3(v.III) + \beta_4[(Age) * (v.III)]$

   Diverging gap measured by: $\beta_4$