ISyE 6739 – Group Activity 18
solutions

1.
```
> X=t(rbind(rep(1,9) , c(55, 46, 30, 35, 59, 61, 74, 38, 27),
+            c(2.1, 2.8, 3.3, 4.5, 2.0, 5.1, 5.5, 3.2, 3.1),
+            c(55, 46, 30, 35, 59, 61, 74, 38, 27)*
+              c(2.1, 2.8, 3.3, 4.5, 2.0, 5.1, 5.5, 3.2, 3.1) ))
> y=c(68, 77, 96, 80, 43, 44, 26, 88, 75)
> n=9;k=3
> # a------------------------------
> beta=solve(t(X)%*%X)%*%t(X)%*%y
> print(c(beta=beta))
       beta1        beta2        beta3        beta4
120.92810634  -1.02534094   1.29875381  -0.06265314
```

$$y = 120.92810634 - 1.02534094X_1 + 1.29875381X_2 - 0.06265314X_3.$$

   (a)
```
> SSE=t(y)%*%y-t(beta)%*%t(X)%*%y
> sigma2=SSE/(n-k)
> ci=NULL
> t=qt(0.05/2, n-k, lower.tail = F)
> C=solve(t(X)%*%X)
> for (i in 0:k){
+   ci=c(lower=beta[i+1]- t*sqrt(sigma2*C[i+1,i+1]),
+        upper=beta[i+1]+ t*sqrt(sigma2*C[i+1,i+1]))
+   print(c(i=i, CI=ci))
+ }
        i  CI.lower  CI.upper
  0.00000 -46.28013 288.13634
        i  CI.lower  CI.upper
 1.000000 -3.988563  1.937881
        i  CI.lower  CI.upper
  2.00000 -43.55879  46.15629
        i   CI.lower   CI.upper
 3.0000000 -0.8110655  0.6857592
```

95% CL's:

$$\begin{aligned} \beta_0 &: [-46.28013, 288.13634], \\ \beta_1 &: [-3.988563, 1.937881], \\ \beta_2 &: [-43.55879, 46.15629], \\ \beta_3 &: [-0.8110655, 0.6857592] \end{aligned} \tag{1}$$

The CI for age and for interaction are close to 0, so they do not impact satisfaction as much as anxiety.

   (b)
```
> SSR=t(beta)%*%t(X)%*%y-sum(y)^2/n
> F0=(SSR/k)/(SSE/(n-k))
> print(c(F0=F0))
      F0
10.11426
```

$$F_0 = 10.114 > 5.14 = F_{0.05,k-1,n-k}$$

Therefore, the regression is significant.

(c)
```
> SST=SSE+SSR
> R2=1-SSE/SST
> R2Adj=1-(SSE/(n-k))/(SST/(n-1))
> print(c(R2=R2, R2Adj=R2Adj))
       R2     R2Adj
0.8349053 0.7798737
```

$$R^2 = 0.8349, \quad R^2_{adj} = 0.7799.$$

(d)
```
> Xe=c(1,55,2.1,55*2.1)
> print(c(lower=t(Xe)%*%beta-t*sqrt(sigma2*t(Xe)%*%C%*%Xe) ,
+         upper=t(Xe)%*%beta+t*sqrt(sigma2*t(Xe)%*%C%*%Xe)))
   lower    upper
43.59305 76.45755
```

The CI for the mean response when $X = (1, 55, 2.1)$:

$$[43.59, 76.458]$$

```
> Xf=c(1, 52, 4, 52*4)
> print(c(lower=t(Xe)%*%beta-t*sqrt(sigma2*(1+t(Xe)%*%C%*%Xe)) ,
+         upper=t(Xe)%*%beta+t*sqrt(sigma2*(1+t(Xe)%*%C%*%Xe))))
   lower    upper
28.48837 91.56223
```

The prediction interval for a new observation with $X = (1, 52, 4.0)$:

$$[28.488, 91.562]$$