1. Use the data set "Sealing Strength of Wax-Polyethylene Blends" by Brown, Turner and Smith (*Tappi*,1958).

(a)
```
seal <- read.table("hw06.txt", header=T)
full <- lm(y~x1+x2+x3+I(x1^2)+I(x2^2)+I(x3^2)+x1*x2+x1*x3+x2*x3,data=seal)
library(MASS)
qqnorm(stdres(full),main="Normal Probability Plot",ylab="standardized residuals")
qqline(stdres(full))
```

(b)
```
reduced <- lm(y~x1+x2+x3,data=seal)
anova(reduced,full)

Analysis of Variance Table
Model 1: y ~ x1 + x2 + x3
Model 2: y ~ x1 + x2 + x3 + I(x1^2) + I(x2^2) + I(x3^2) + x1:x2 + x1:x3 +  x2:x3
  Res.Df    RSS Df Sum of Sq      F   Pr(>F)
1     16 51.209
2     10 11.859  6    39.349 5.5301 0.009126 **
```

F-ratio = 5.53 with p-value = 0.009. There is evidence that the response is not "planar" as a function of the x's.

(c)
```
new.x <- data.frame(-1.01,.26,.68)
names(new.x) <- c("x1","x2","x3")
predict(full,new.x,interval="confidence", level=0.90)
          fit      lwr      upr
[1,] 11.08156 10.05503 12.10809
predict(full,new.x,interval="prediction", level=0.90)
          fit      lwr      upr
[1,] 11.08156 8.856805 13.30632
```

Or you could have done:

```
x0 <- c(1,new.x$x1,new.x$x2,new.x$x3,new.x$x1^2,new.x$x2^2,new.x$x3^2,
        new.x$x1*new.x$x2,new.x$x1*new.x$x3,new.x$x2*new.x$x3)
y0 <- sum(full$coef*x0)
X <- cbind(rep(1,dim(seal)[1]),seal$x1,seal$x2,seal$x3,seal$x1^2,seal$x2^2,seal$x3^2,
           seal$x1*seal$x2,seal$x1*seal$x3,seal$x2*seal$x3)
MSE <- anova(full)$Mean[10]
cXXc <- x0%*%ginv(t(X)%*%X)%*%x0
# lower and upper confidence limits
ll <- y0 - qt(0.95,10)*sqrt(MSE*cXXc)
ul <- y0 + qt(0.95,10)*sqrt(MSE*cXXc)
# lower and upper prediction limits
ll <- y0 - qt(0.95,10)*sqrt(MSE*(1+cXXc))
ul <- y0 + qt(0.95,10)*sqrt(MSE*(1+cXXc))
```

2. Testing "Lack of Fit".

```
project <- function(X) {X%*%ginv(t(X)%*%X)%*%t(X)}
Xstar <- matrix(0,20,15)
Xstar[1:8,1:8] <- diag(rep(1,8))
Xstar[9:14,9] <- 1
Xstar[15:20,10:15] <- diag(rep(1,6))
F.ratio <- ((t(seal$y)%*%(project(Xstar)-project(X))%*%seal$y)/(15-10))/
           ((t(seal$y)%*%(diag(rep(1,20))-project(Xstar))%*%seal$y)/(20-15))
p.value <- 1-pf(F.ratio,5,5)
```

F-ratio = 1.39 with p-value = 0.363. There is no evidence of lack-of fit.

3. Use dogs data set.

```
lm.out <- lm(y~drug*disease,data=d)
anova(lm.out)             # Type I SS

Analysis of Variance Table
Response: y
             Df Sum Sq Mean Sq F value    Pr(>F)
drug          3 2992.8   997.6  9.0513 8.047e-05 ***
disease       2  365.7   182.9  1.6591    0.2015
drug:disease  6  737.9   123.0  1.1158    0.3680
Residuals    46 5070.0   110.2
```

4. Use table of fake 2-way factorial data.

(a)
```
Y <- c(12,14,10,12,9,11,12,6,7,10,11,7)
A <- c(rep(1,4),rep(2,5),rep(3,3))
B <- c(1,2,3,3,1,2,2,3,3,1,2,3)
d <- data.frame(Y,A,B)
d$A <- as.factor(d$A)
d$B <- as.factor(d$B)
lm.out <- lm(Y~A*B,data=d)
means <- tapply(d$Y,list(d$A,d$B),mean)
> means
   1    2    3
1 12 14.0 11.0
2  9 11.5  6.5
3 10 11.0  7.0

x.axis <- unique(d$A)
matplot(c(1,3),c(0,20),type="n",xlab="A",ylab="Mean Response")
matlines (x.axis,means,type="b",lty=c(1,3,7))
lm.out <- lm(Y~A*B,data=d)

> anova(lm.out)              # Type I SS
Analysis of Variance Table
Response: Y
          Df Sum Sq Mean Sq F value  Pr(>F)
A          2 22.250  11.125 11.1250 0.04095 *
B          2 37.039  18.520 18.5195 0.02051 *
A:B        4  2.628   0.657  0.6569 0.66182
Residuals  3  3.000   1.000
```

(b)
```
> Xeffects
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]
 [1,]    1    1    0    1    0    1    0    0    0
 [2,]    1    1    0    0    1    0    1    0    0
 [3,]    1    1    0   -1   -1   -1   -1    0    0
 [4,]    1    1    0   -1   -1   -1   -1    0    0
 [5,]    1    0    1    1    0    0    0    1    0
 [6,]    1    0    1    0    1    0    0    0    1
 [7,]    1    0    1    0    1    0    0    0    1
 [8,]    1    0    1   -1   -1    0    0   -1   -1
 [9,]    1    0    1   -1   -1    0    0   -1   -1
[10,]    1   -1   -1    1    0   -1    0   -1    0
[11,]    1   -1   -1    0    1    0   -1    0   -1
[12,]    1   -1   -1   -1   -1    1    1    1    1
```

```
ssA <- t(d$Y)%*%(project(Xeffects[,1:3])-project(Xeffects[,1]))%*%d$Y      # R(alpha|mu)
ssB <- t(d$Y)%*%(project(Xeffects[,1:5])-project(Xeffects[,1:3]))%*%d$Y    # R(beta|mu,alpha)
ssAB <- t(d$Y)%*%(project(Xeffects)-project(Xeffects[,1:5]))%*%d$Y         # R(alphabeta|mu,alpha,beta)
```

Type I SS: $R(\alpha|\mu) = 22.250$, $R(\beta|\mu,\alpha) = 37.039$, $R(\alpha\beta|\mu,\alpha,\beta) = 2.628$.

```
> Xcells
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]
 [1,]    1    0    0    0    0    0    0    0    0
 [2,]    0    1    0    0    0    0    0    0    0
 [3,]    0    0    1    0    0    0    0    0    0
 [4,]    0    0    1    0    0    0    0    0    0
 [5,]    0    0    0    1    0    0    0    0    0
 [6,]    0    0    0    0    1    0    0    0    0
 [7,]    0    0    0    0    1    0    0    0    0
 [8,]    0    0    0    0    0    1    0    0    0
 [9,]    0    0    0    0    0    1    0    0    0
[10,]    0    0    0    0    0    0    1    0    0
[11,]    0    0    0    0    0    0    0    1    0
[12,]    0    0    0    0    0    0    0    0    1
```

Type III SS for testing main effect A: $H_0 : \bar{\mu}_{1.} = \bar{\mu}_{2.} = \bar{\mu}_{3.}$. That is, $H_0 : \bar{\mu}_{1.} - \bar{\mu}_{3.} = 0$ and $\bar{\mu}_{2.} - \bar{\mu}_{3.} = 0$.

```
C <- matrix(c(1/3,1/3,1/3,0,0,0,-1/3,-1/3,-1/3,0,0,0,1/3,1/3,1/3,-1/3,-1/3,-1/3),2,9,byrow=T)
```

Type III SS for testing main effect B: $H_0 : \bar{\mu}_{.1} = \bar{\mu}_{.2} = \bar{\mu}_{.3}$. That is, $H_0 : \bar{\mu}_{.1} - \bar{\mu}_{.3} = 0$ and $\bar{\mu}_{.2} - \bar{\mu}_{.3} = 0$.

```
C <- matrix(c(rep(c(1/3,0,-1/3),3),rep(c(0,1/3,-1/3),3)),2,9,byrow=T)
```

Then compute:

```
CXXC <- C%*%ginv(t(Xcells)%*%Xcells)%*%t(C)
Cb <- C%*%ginv(t(Xcells)%*%Xcells)%*%t(Xcells)%*%d$Y
SSH0 <- t(Cb)%*%solve(CXXC)%*%Cb
```

Type III SS A: $SS_{H0} = 25.108$. Type III SS B: $SS_{H0} = 32.331$.

(c)
```
Xincomp.full <- Xcells[-c(6,7),-5]
Xincomp.red <- Xeffects[-c(6,7),1:5]
Y.incomp <- d$Y[-c(6,7)]
full <- lm(Y.incomp~Xincomp.full-1)
reduced <- lm(Y.incomp~Xincomp.red-1)
anova(reduced,full)

Analysis of Variance Table
Model 1: Y.incomp ~ Xincomp.red - 1
Model 2: Y.incomp ~ Xincomp.full - 1
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1      5 3.8298
2      2 2.5000  3    1.3298 0.3546 0.7954
```

F-ratio = 0.35 with p-value= 0.7954. There is not enough evidence that there is interaction between factors A and B. (Note that in 4(a) the interaction term is not significant either.)

(d)
```
X <- Xeffects[-c(6,7),1:5]
C <- c(1,0,1,0,1)
mu22 <- C%*%ginv(t(X)%*%X)%*%t(X)%*%Y.incomp
> mu22
         [,1]
[1,] 10.08511
```