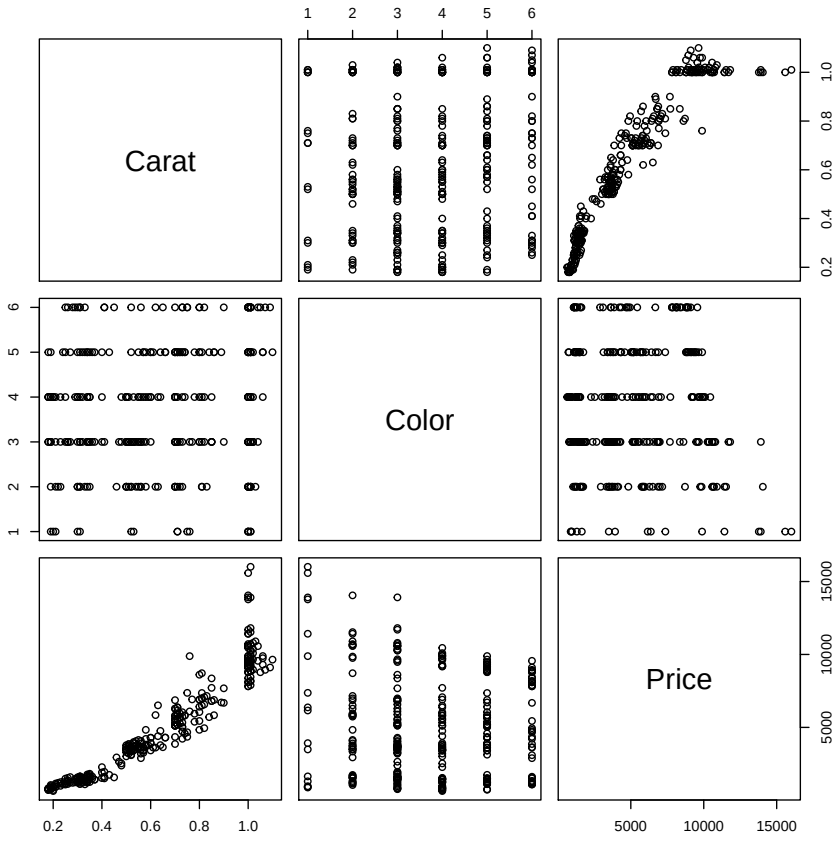
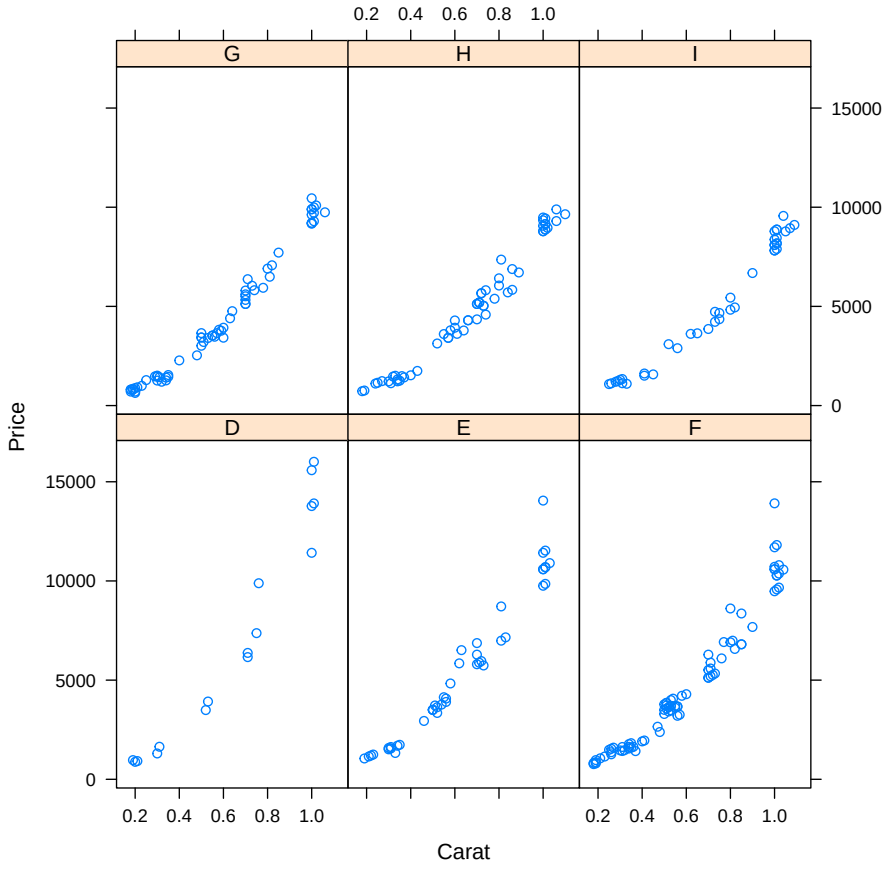
```
diamonds <- read.csv("~/Documents/Stat217Fall2014/Projects/Project4/diamonds.csv")
##You only need to run this first chunk of code once after importing the data##
diamonds <- diamonds[,c(2,3,6)]
diamonds$Color <- as.factor(diamonds$Color)
require(mosaic)
##You may need to install the lattice package##
require(lattice)
options(show.signif.stars = F)
```

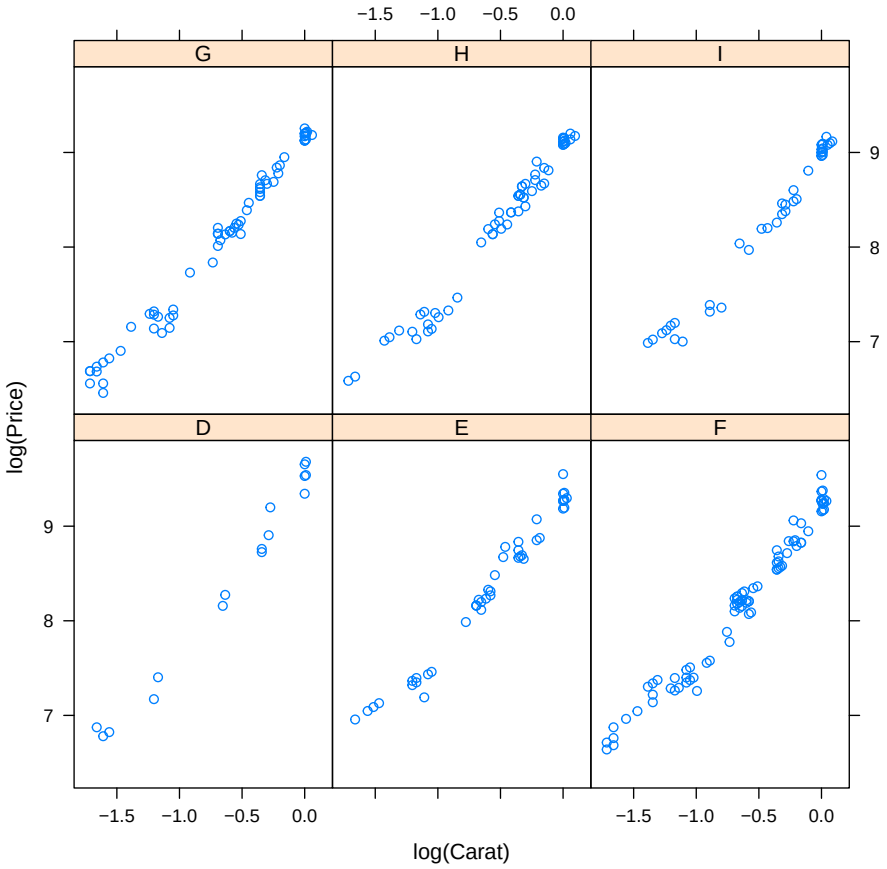
```
#Look at and describe scatterplots
pairs(diamonds)
```



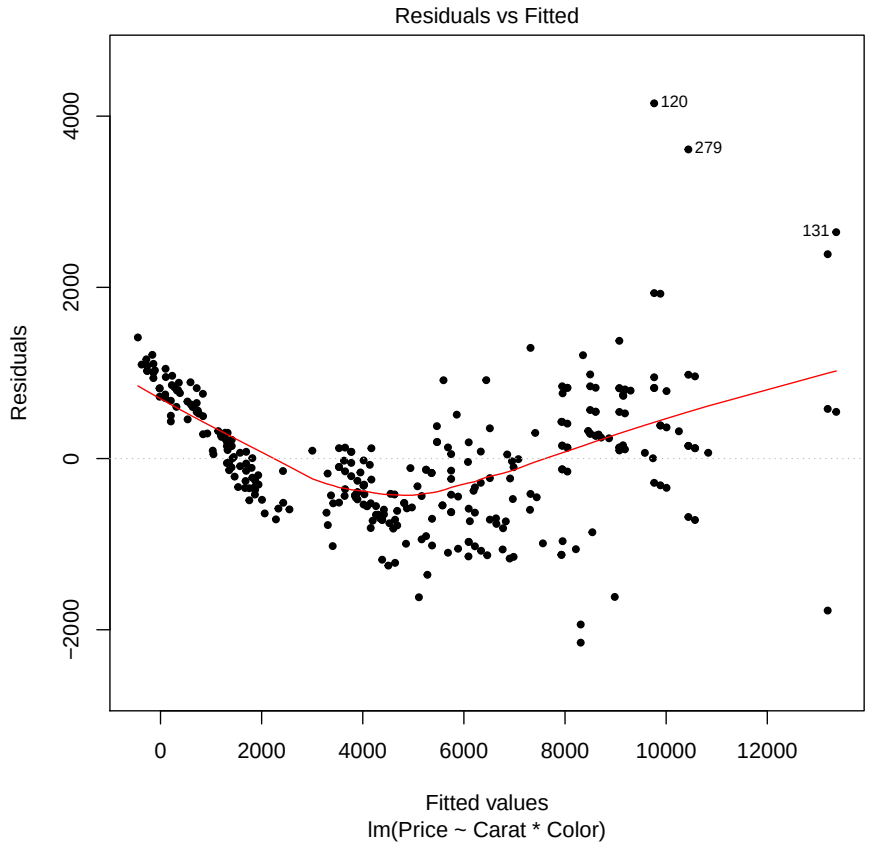

```
xyplot(Price~Carat|Color,data=diamonds)
```




```
xyplot(log(Price)~log(Carat)|Color,data=diamonds)
```




```
#Why should you consider a log transformation?
fit.first <- lm(Price~Carat*Color, data = diamonds)
plot(fit.first, which = 1, pch = 20)
```



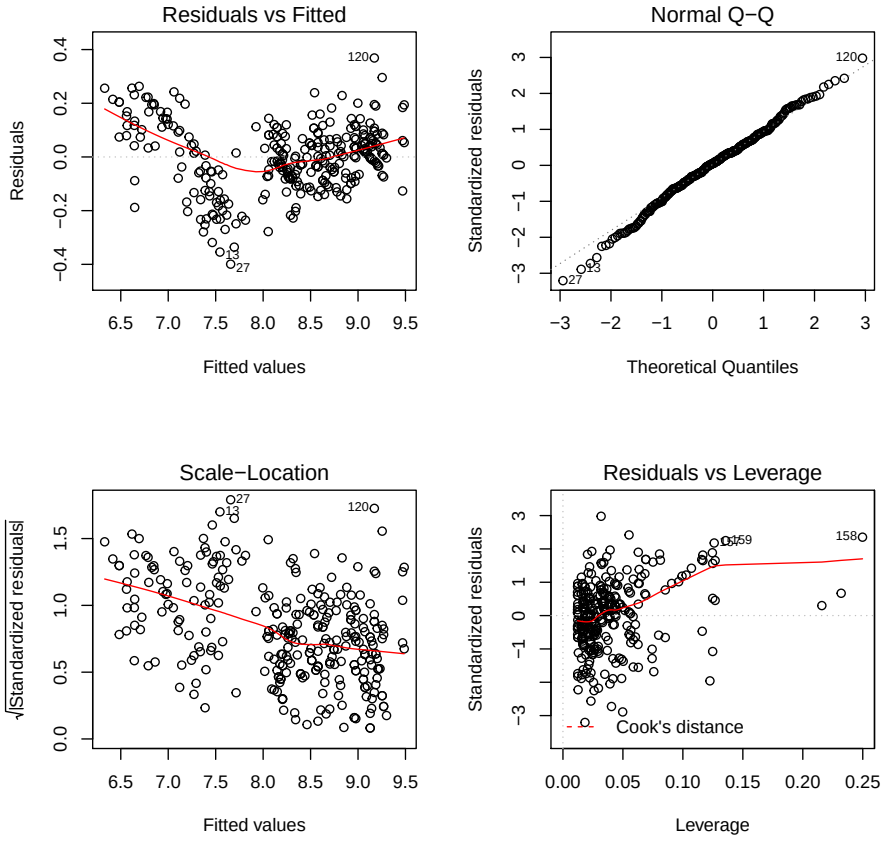

```
#Look at the interaction model
fit.Int <- lm(log(Price)~I(log(Carat))*Color, data = diamonds)
summary(fit.Int)

##
## Call:
## lm(formula = log(Price) ~ I(log(Carat)) * Color, data = diamonds)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3994 -0.0732  0.0065  0.0775  0.3687
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)             9.4697     0.0444  213.16  < 2e-16
## I(log(Carat))           1.7172     0.0517   33.23  < 2e-16
```

```
## ColorE                -0.2151     0.0533   -4.03  7.0e-05
## ColorF                -0.2978     0.0498   -5.98  6.3e-09
## ColorG                -0.3416     0.0508   -6.72  9.3e-11
## ColorH                -0.4132     0.0501   -8.24  5.6e-15
## ColorI                -0.5087     0.0516   -9.87  < 2e-16
## I(log(Carat)):ColorE  -0.1750     0.0648   -2.70   0.0073
## I(log(Carat)):ColorF  -0.1943     0.0586   -3.32   0.0010
## I(log(Carat)):ColorG  -0.1753     0.0588   -2.98   0.0031
## I(log(Carat)):ColorH  -0.1273     0.0619   -2.06   0.0406
## I(log(Carat)):ColorI  -0.1329     0.0659   -2.02   0.0446
##
## Residual standard error: 0.126 on 296 degrees of freedom
## Multiple R-squared:  0.977,Adjusted R-squared:  0.976
## F-statistic: 1.14e+03 on 11 and 296 DF,  p-value: <2e-16
```

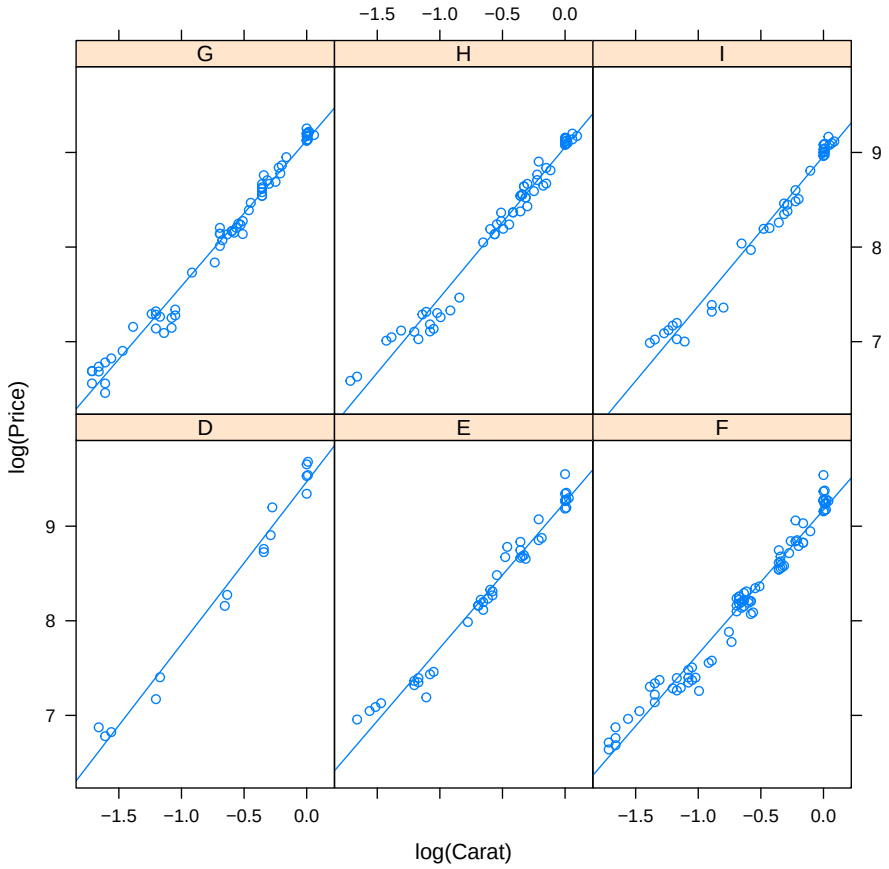
```
anova(fit.Int)
```

```
## Analysis of Variance Table
##
## Response: log(Price)
##                      Df Sum Sq Mean Sq  F value Pr(>F)
## I(log(Carat))         1  194.9   194.9 12328.86 <2e-16
## Color                 5    3.8     0.8    48.00 <2e-16
## I(log(Carat)):Color   5    0.2     0.0     2.55  0.028
## Residuals           296    4.7     0.0
```

```
par(mfrow=c(2,2))
plot(fit.Int)
```




```
xyplot(log(Price) ~ log(Carat) | Color, type = c("p","r"),data=diamonds)
```




```
fit.int.cellmeans <- lm(log(Price)~Color+I(log(Carat)):Color-1, data = diamonds)
confint(fit.int.cellmeans)

##                       2.5 % 97.5 %
## ColorD                9.382  9.557
## ColorE                9.197  9.313
## ColorF                9.128  9.216
## ColorG                9.079  9.177
## ColorH                9.011  9.102
## ColorI                8.910  9.013
## ColorD:I(log(Carat)) 1.615  1.819
## ColorE:I(log(Carat)) 1.465  1.619
## ColorF:I(log(Carat)) 1.468  1.577
## ColorG:I(log(Carat)) 1.486  1.597
## ColorH:I(log(Carat)) 1.523  1.657
## ColorI:I(log(Carat)) 1.504  1.665
```