MATH 3080 Lab Project 10

Aini Liang

March 15, 2018

Remember: I expect to see commentary either in the text, in the code with comments created using `#`*, or (preferably) both!* ***Failing to do so may result in lost points!***

Problem 1 (3.29 modified)

The data set `diamond` (**UsingR**) contains data about the price of 48 diamond rings. The variable price records the price in Singapore dollars and the variable carat records the size of the diamond.

- Make a scatterplot of carat versus price, use pch=5 to plot with diamonds. Do you see any linear relationship between the two variables? Also compute the Pearson's correlation coefficient.

```
# Your code here
library(UsingR)
```

```
## Loading required package: MASS
```

```
## Loading required package: HistData
```

```
## Loading required package: Hmisc
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:base':
##
##     format.pval, round.POSIXt, trunc.POSIXt, units
```

```
##
## Attaching package: 'UsingR'
```

```
## The following object is masked from 'package:survival':
##
##     cancer
```

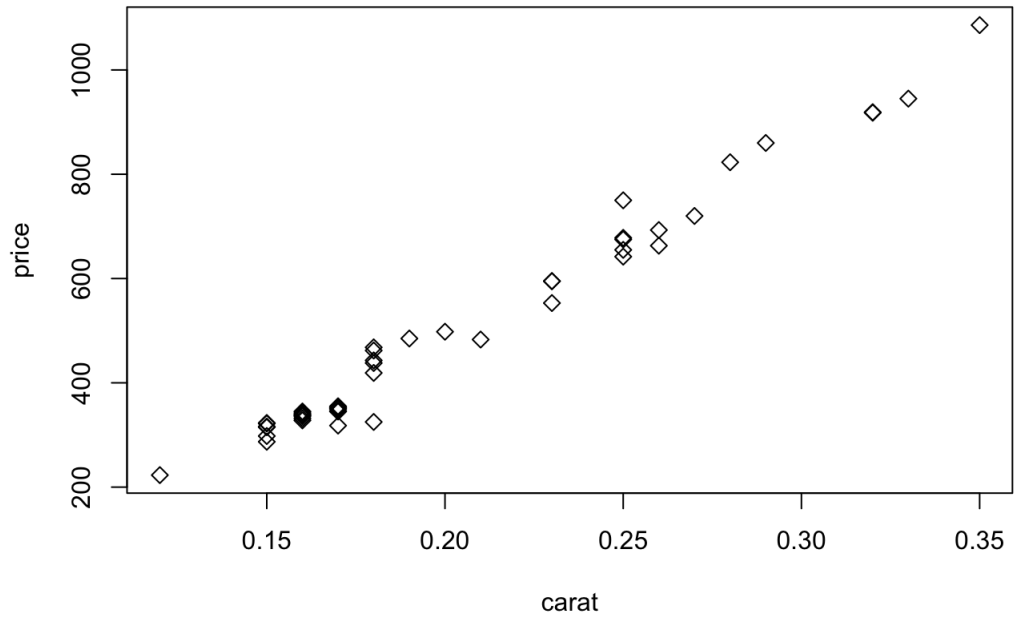
```
attach(diamond)
dim(diamond)
```

```
## [1] 48  2
```

```
cor(carat, price)
```

```
## [1] 0.9890707
```

```
plot(carat, price, pch = 5)
```



- Fit the data with a simple linear regression model. Add the regression line to the scatterplot. Does the regression line fit the data well? Check model assumptions and comment.

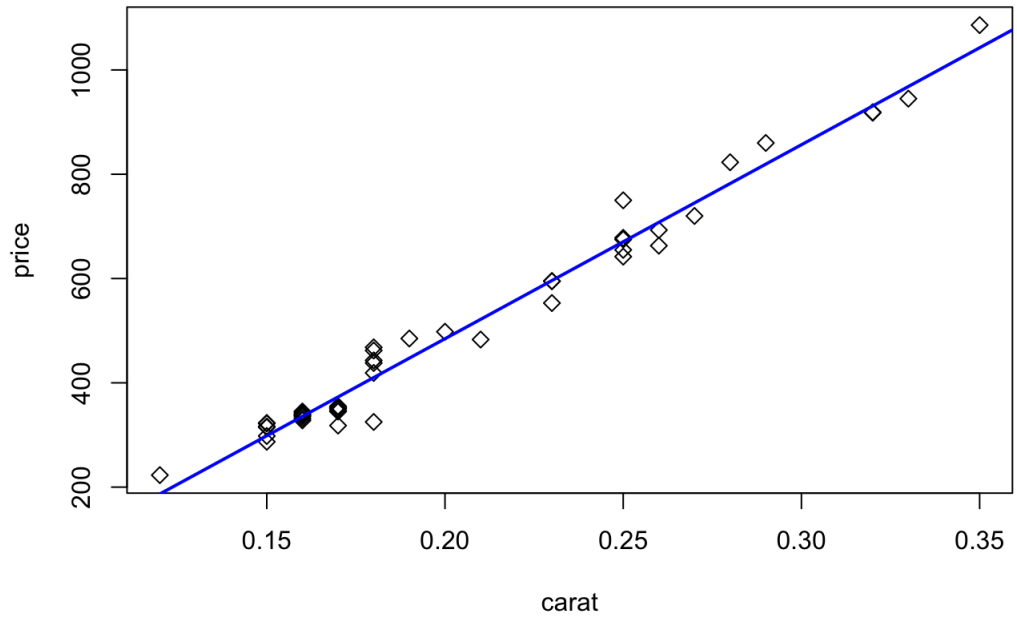
```
# Your code here

attach(diamond)
```

```
## The following objects are masked from diamond (pos = 3):
##
##     carat, price
```

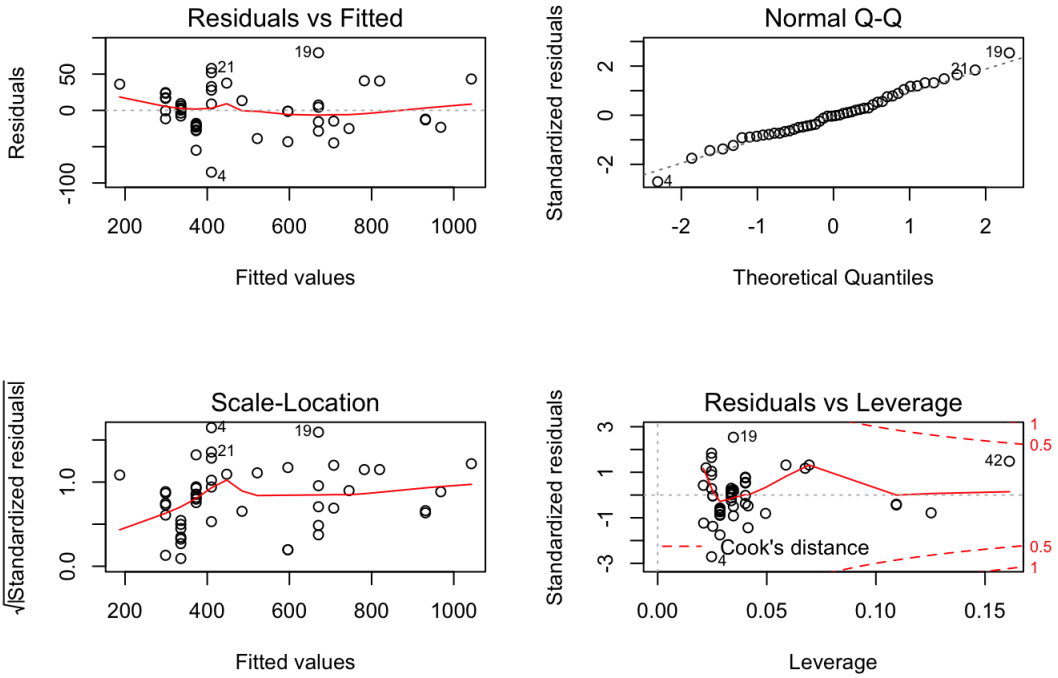
```
plot(carat, price, pch = 5)
fit <- lm(price ~ carat)
summary(fit)
```

```
##
## Call:
## lm(formula = price ~ carat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -85.159 -21.448  -0.869  18.972  79.370
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -259.63      17.32  -14.99   <2e-16 ***
## carat        3721.02      81.79   45.50   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 31.84 on 46 degrees of freedom
## Multiple R-squared:  0.9783, Adjusted R-squared:  0.9778
## F-statistic:  2070 on 1 and 46 DF,  p-value: < 2.2e-16
```

```
abline(fit, col = "blue", lwd = 2)
```




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



- Predict the amount a one-third carat diamond ring would cost.

```
# Your code here
predict(fit, data.frame(carat = 1/3))
```

```
##        1
## 980.7157
```