Repairs

Cody Frisby

December 15, 2015

Let's get started with R. This first data set is small so we have the data loaded into two variables: minutes and units. We then put them together with the cbind function and store the result in a new variable called repair.

```
units <- c(1,2,3,4,4,5,6,6,7,8,9,9,10,10,3)
minutes <- c(23,29,49,64,74,87,96,97,109,119,149,145,154,166, NA)
repair <- as.data.frame(cbind(units, minutes))
repair <- repair[complete.cases(repair), ]
# note, the above line removes missing values.
describe(repair)
```

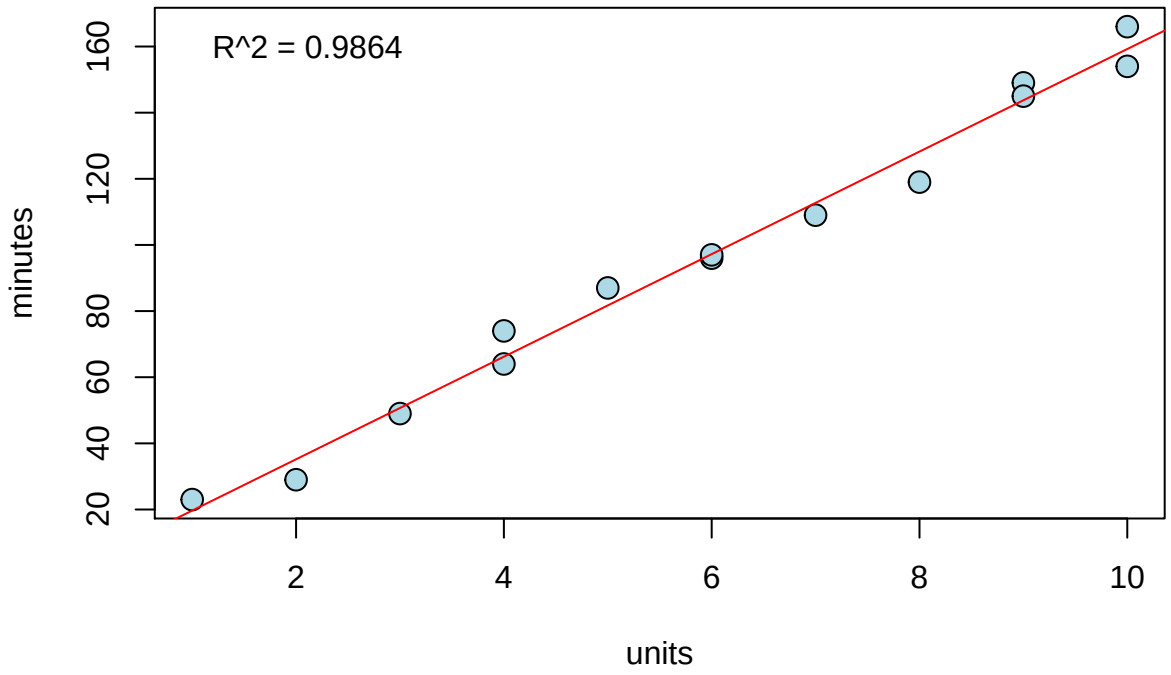
```
##         vars  n  mean    sd median trimmed   mad min max range  skew
## units      1 14  6.00  2.96    6.0    6.08  3.71   1  10     9 -0.12
## minutes    2 14 97.21 46.22   96.5   97.67 59.30  23 166   143 -0.09
##         kurtosis    se
## units      -1.43  0.79
## minutes    -1.37 12.35
```

Above is a summary of the data table "repair". Below is a table of the correlation coefficients.

```
cor(repair, method = "pearson", use = "complete.obs")
```

```
##             units   minutes
## units   1.0000000 0.9936987
## minutes 0.9936987 1.0000000
```

Now let's plot the data look at the variable minutes being modeled as a function of units.



Now to display a summary of the model. Above we fit a model with minutes being function of units. We stored this model in a variable called m.

```
anova(m) #ANOVA table
```

```
## Analysis of Variance Table
##
## Response: minutes
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## units      1 27419.5 27419.5   943.2 8.916e-13 ***
## Residuals 12   348.8    29.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(m) #summary of linear model
```

```
##
## Call:
## lm(formula = minutes ~ units, data = repair)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.2318 -3.3415 -0.7143  4.7769  7.8033
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.162      3.355    1.24    0.239
## units         15.509      0.505   30.71 8.92e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 5.392 on 12 degrees of freedom
## Multiple R-squared:  0.9874, Adjusted R-squared:  0.9864
## F-statistic: 943.2 on 1 and 12 DF,  p-value: 8.916e-13
```

```
t <- cbind(repair, fitted.values(m), residuals(m))
colnames(t) <- c("units", "observed", "predicted", "residuals")
print(t)
```

```
##    units observed predicted  residuals
## 1      1       23  19.67043  3.3295739
## 2      2       29  35.17920 -6.1791980
## 3      3       49  50.68797 -1.6879699
## 4      4       64  66.19674 -2.1967419
## 5      4       74  66.19674  7.8032581
## 6      5       87  81.70551  5.2944862
## 7      6       96  97.21429 -1.2142857
## 8      6       97  97.21429 -0.2142857
## 9      7      109 112.72306 -3.7230576
## 10     8      119 128.23183 -9.2318296
## 11     9      149 143.74060  5.2593985
## 12     9      145 143.74060  1.2593985
## 13    10      154 159.24937 -5.2493734
## 14    10      166 159.24937  6.7506266
```

Below is a plot of the residuals vs fitted, Normal Quantile, Scale-Location, Residuals vs Leverage.

```
#par(mfrow=c(2,2)) #setup your window to fit the next plots
#plot(m)
```