Homework #7

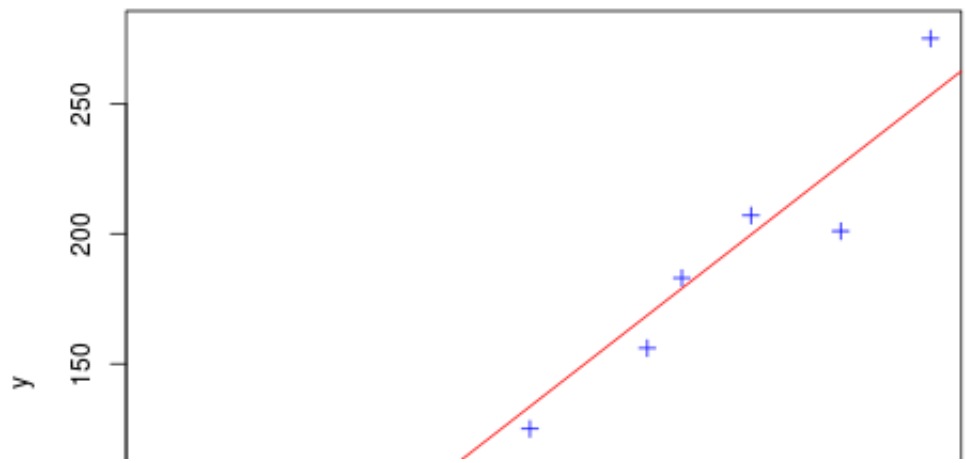
```
Populating the interactive namespace from numpy and matplotlib
The rpy2.ipython extension is already loaded. To reload it, use:
  %reload_ext rpy2.ipython
```

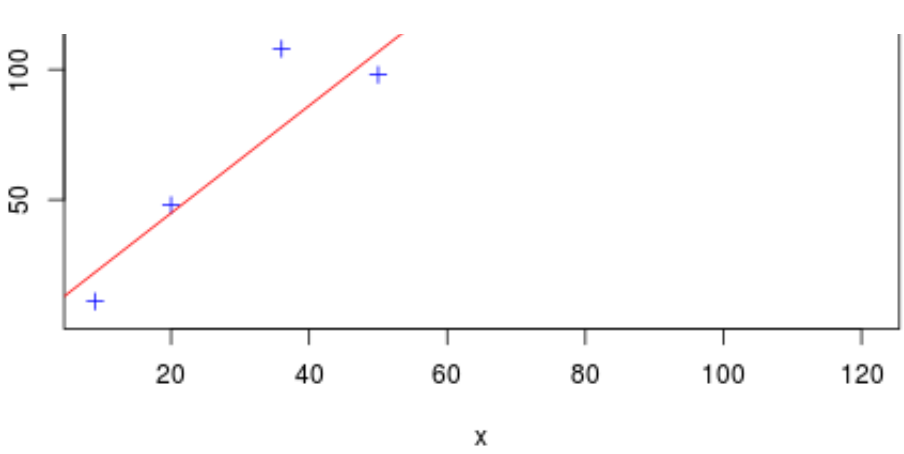
The raw code for this IPython notebook is by default hidden for easier reading.

To toggle on/off the raw code, click here.

11.6 ¶

```
Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-25.400 -10.562  -2.713   6.575  30.133

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.6003    11.9602   0.301    0.771
x             2.0630     0.1581  13.045 1.13e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 17.83 on 8 degrees of freedom
Multiple R-squared:  0.9551,    Adjusted R-squared:  0.9495
F-statistic: 170.2 on 1 and 8 DF,  p-value: 1.132e-06
```



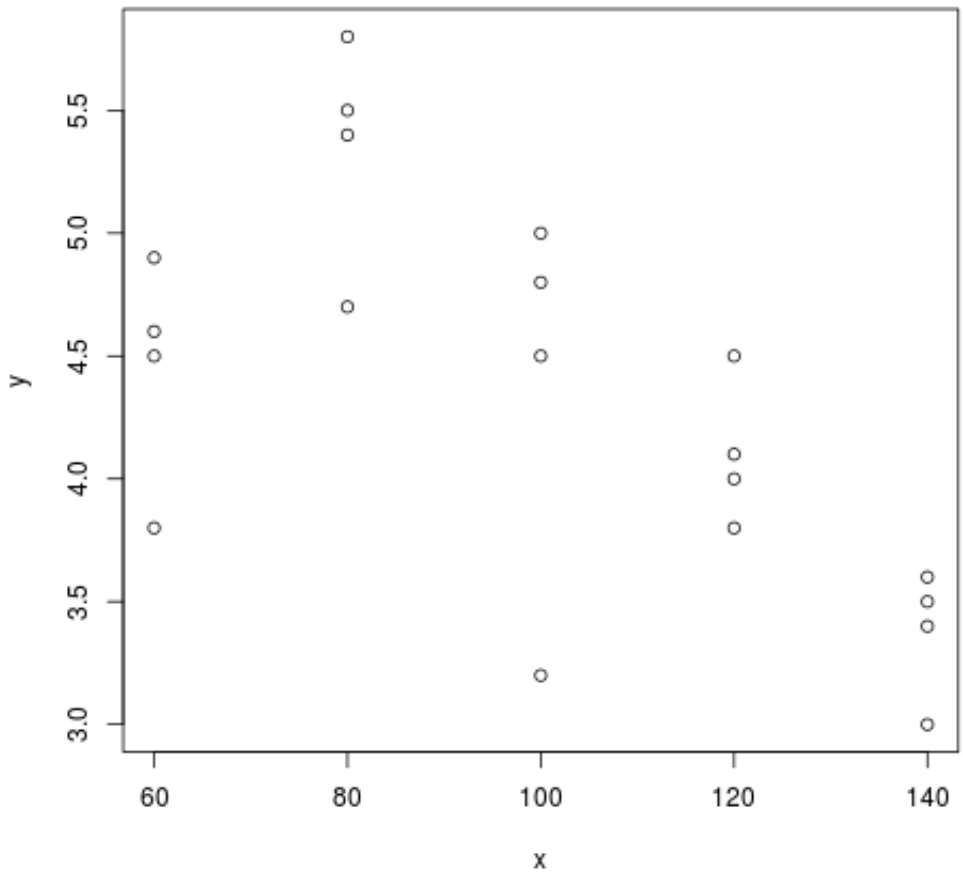


b.) Yes.
c.) 121.191

11.10

This is the standard deviation. Approximately 65% of observations wi
ll fall within this range from the mean

11.12(a)



b.) There are some outliers, but **no** they do not have high influence

11.13

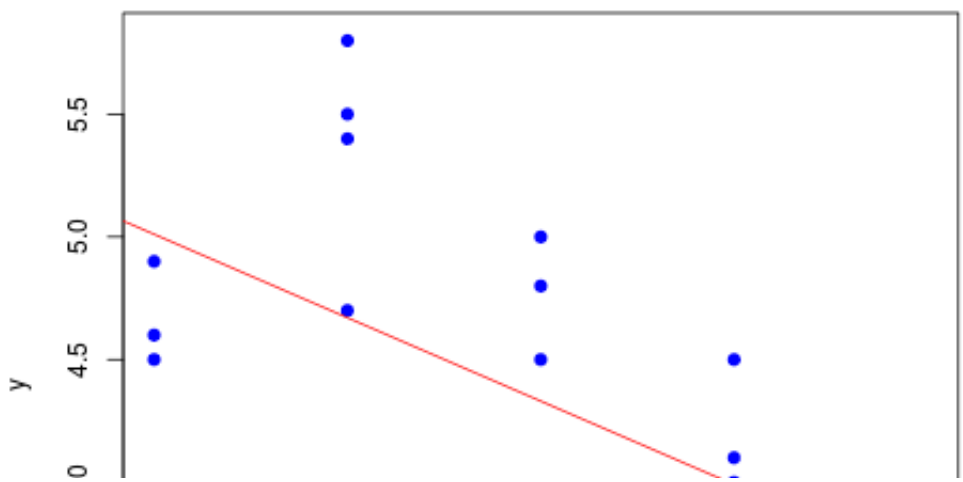
```
Call:
lm(formula = y ~ x)

Residuals:
   Min     1Q Median     3Q    Max
 -1.21  -0.29  -0.02   0.48   1.13

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.030000   0.519543   11.61 8.62e-10 ***
x           -0.017000   0.004999   -3.40  0.00319 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6324 on 18 degrees of freedom
Multiple R-squared:  0.3911,    Adjusted R-squared:  0.3573
F-statistic: 11.56 on 1 and 18 DF,  p-value: 0.003188
```

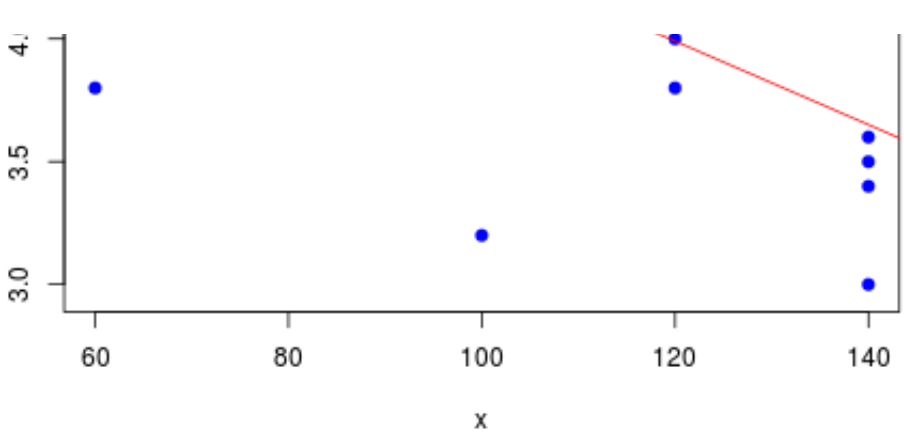
a.) The intercept is 6.0, and the slope is 0.0

b.) It indicates that there is almost zero correlation between the two, but if you really wanted to stretch the data, you might be able to say that there bit lifetime is negatively correlated with the speed of the drill

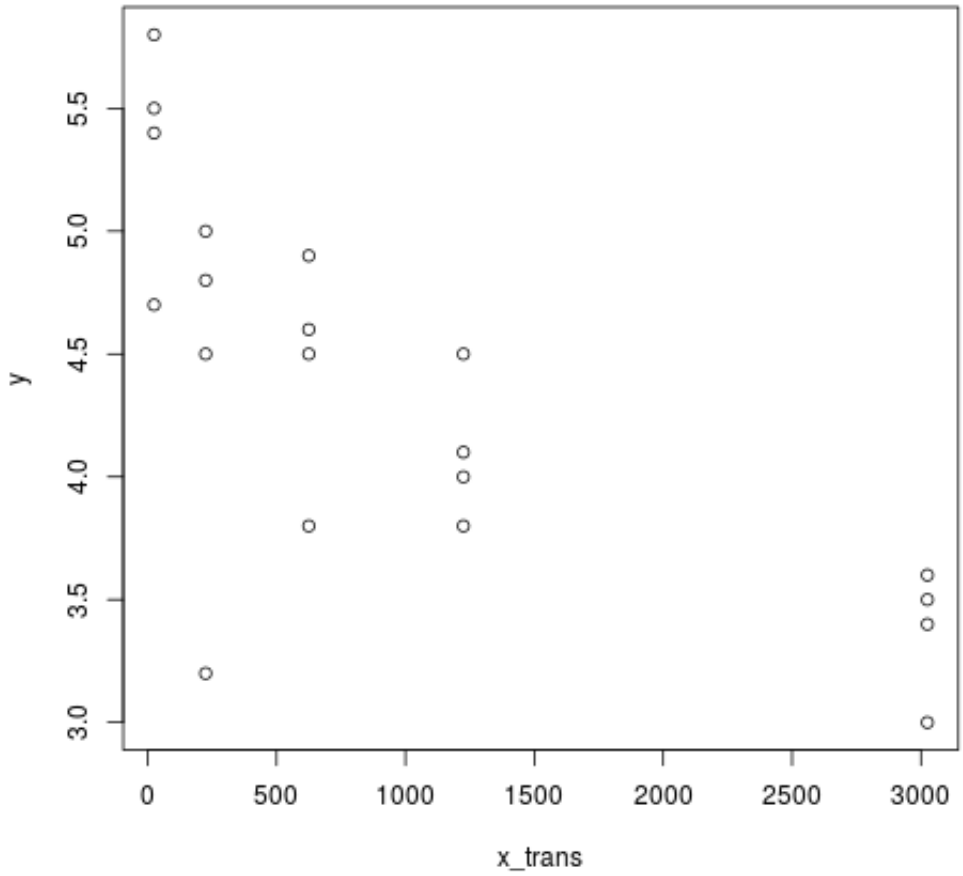
c.) hmmm... I have no idea. R does not have this output, and looking at the minitab data seems to suggest that the standard deviation is negative, which means nothing to me.

11.14(a)(b)






```
a.) [4.98, 4.64, 4.3, 3.96, -18.48]
b.) 80 and 100 are larger. 60 and 140 are smaller. It indicates a li
near-fit is not suitable
```



c.) Above, I squared and shifted the x variable. This leaves us with some slightly more linear looking data, but it is still not great.

11.60(a)(b)(d)

a.) I'm not sure what's bein asked of us

b.) Looking at the data, the fit of the two lines is very similar, with the top analysis having a slightly better fit

d.) The y-intercept is slightly different

11.63

b.) The slope coeff is apparently 2.46, this is lower than expected and must be because of that one outlier with high influence. No the intercept is basically meaningless. The 95% confidence level is very wide, and even then with that one outlier messing everything up, I would not take it too seriously.