3 pg 426

a.) if we take the derivative with respect to x , we get the equation $\mu_{y|x} = -3.2+1.4x$

The slope where x= 0,2,3 is rate = -3.2, -0.4, 1 respectively.

b.) if we plug in $\mu_x = 2$ into the provided equation we get :

$\mu_{y|x} = ß_0 + ß_1(x - 2) + ß_2(x-2)^2$ then multiplying out the coefficients

$\mu_{y|x} = ß_0 + xß_1 - 2ß_1 + ß_2(x^2 - 4x + 4)$

$\mu_{y|x} = ß_0 + xß_1 - 2ß_1 + ß_2x^2 - 4xß_2 + 4ß_2$ rearrange and group

$\mu_{y|x} = (ß_0 + - 2ß_1 + 4ß_2) + (xß_1 - 4xß_2) + ß_2x^2$

$\mu_{y|x} = (ß_0 + - 2ß_1 + 4ß_2) + (ß_1 - 4ß_2)x + ß_2x^2$

Since we have the $ß_2x^2$ at the end this directly correlates with the originals equations we can

see that $ßb_2 = 0.7$

With the $(ß_0 + - 2ß_1 + 4ß_2) = -8.5$

$ß_1 - 4ß_2 = -3.2$

We can solve this system of equations with the term we already have and find that

$ß_1 = \ = -0.4$

$ß_2 = 0.7$

$ß_0 = -12.1$

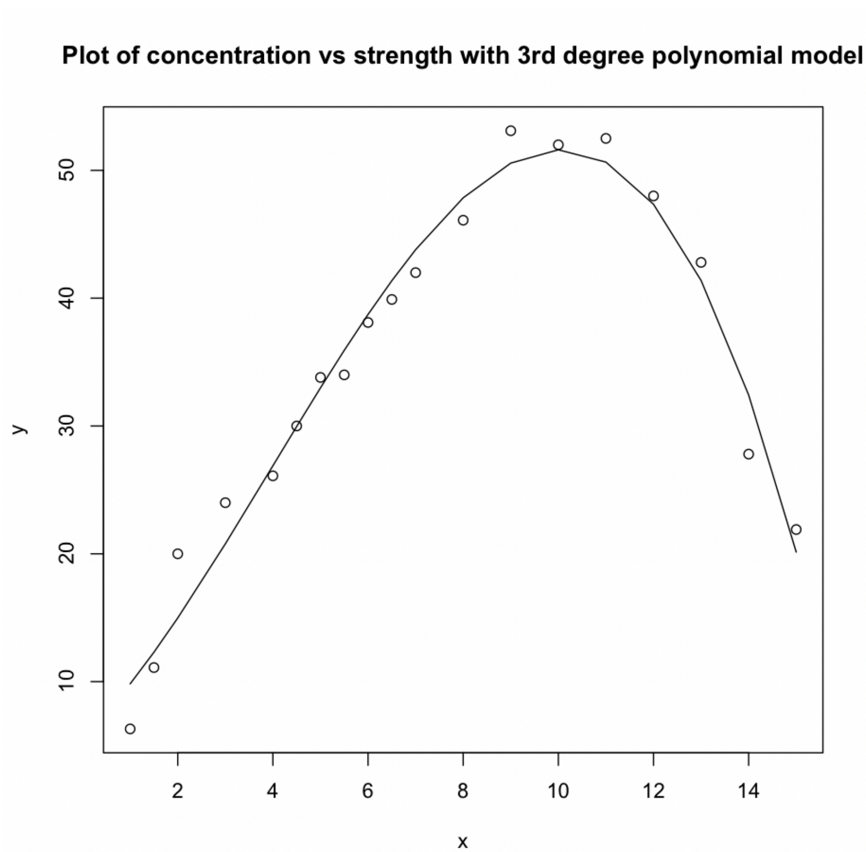
4 pg 444

a.)

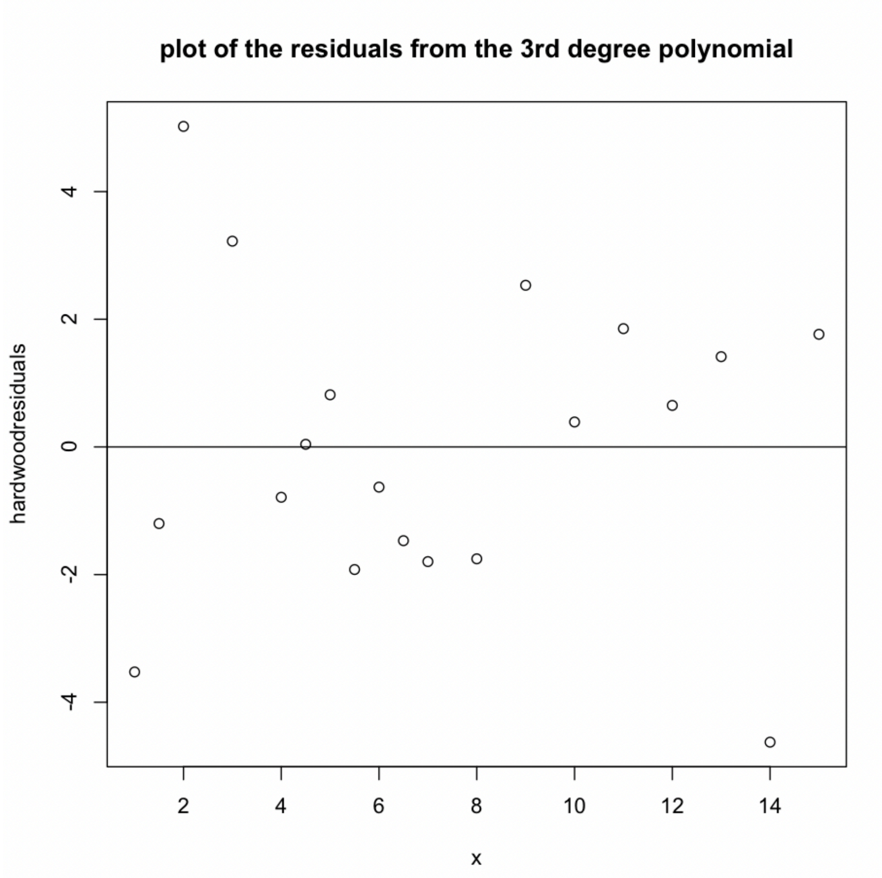
```
> summary(fit)

Call:
lm(formula = y ~ poly(x, 3))

Residuals:
    Min      1Q  Median      3Q     Max 
-4.6250 -1.6109  0.0413  1.5892  5.0216 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  34.1842     0.5931  57.641  < 2e-16 ***
poly(x, 3)1  32.3021     2.5850  12.496 2.48e-09 ***
poly(x, 3)2 -45.3963     2.5850 -17.561 2.06e-11 ***
poly(x, 3)3 -14.5740     2.5850  -5.638 4.72e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.585 on 15 degrees of freedom
Multiple R-squared:  0.9707,	Adjusted R-squared:  0.9648 
F-statistic: 165.4 on 3 and 15 DF,  p-value: 1.025e-11
```

As you can see from the output of the summary command , the adj R^2 value for the 3rd degree model is .964 which is very close to 1 and the p value is much less than .01 , which means the model should be useful.

b.)



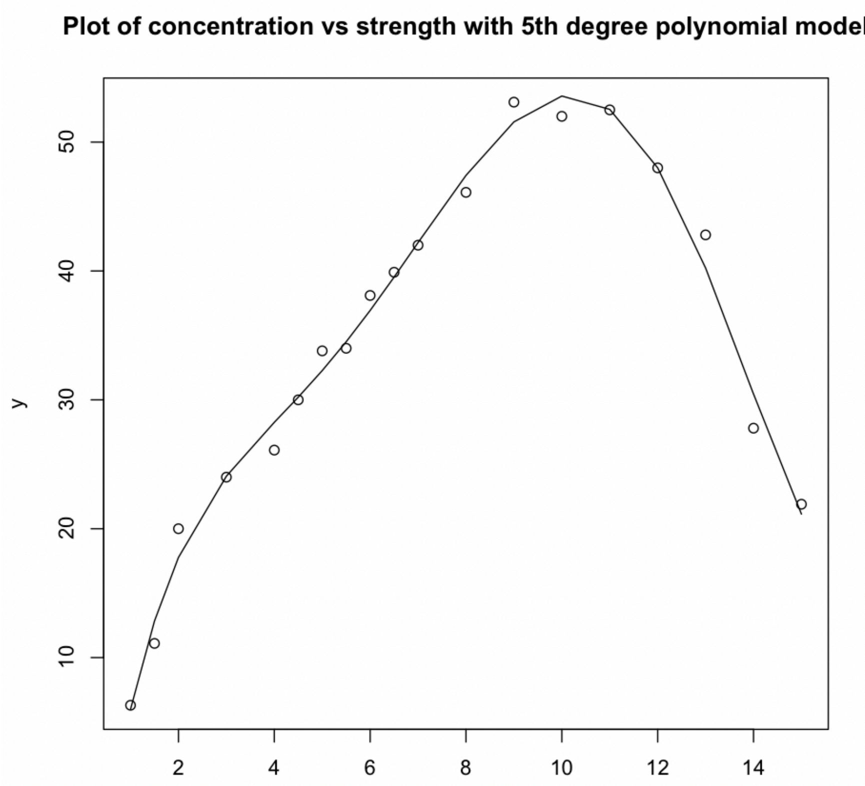
c.)

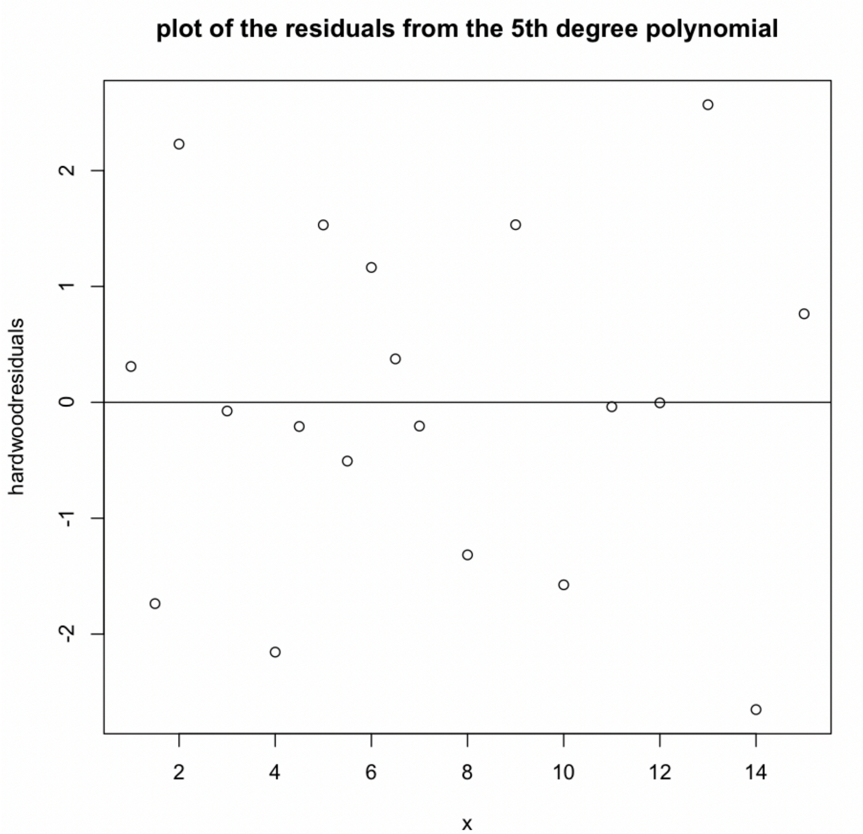


The residual plot of the 3rd degree polynomial seems to be random albeit there seems to be a slight positive trend/pattern suggesting that the model can be improved.

```
> fit5 = lm(y~poly(x,5))
> summary(fit5)

Call:
lm(formula = y ~ poly(x, 5))

Residuals:
     Min       1Q   Median       3Q      Max 
-2.65167 -0.91159 -0.03811  0.96396  2.56865 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  34.1842     0.3906  87.512  < 2e-16 ***
poly(x, 5)1  32.3021     1.7027  18.971 7.39e-11 ***
poly(x, 5)2 -45.3963     1.7027 -26.662 9.84e-13 ***
poly(x, 5)3 -14.5740     1.7027  -8.559 1.06e-06 ***
poly(x, 5)4  -3.1647     1.7027  -1.859 0.085859 .  
poly(x, 5)5   7.2479     1.7027   4.257 0.000935 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.703 on 13 degrees of freedom
Multiple R-squared:  0.989,     Adjusted R-squared:  0.9847 
F-statistic: 233.1 on 5 and 13 DF,  p-value: 3.022e-12
```

Here we can see that the adj R^2 value is 0.984 for the 5th degree model which is closer to one than that of the 3rd degree model however we can also see that the x^4 term is not achieving the .01 level of significance.





We can also notice that the residual plot is much more random than that of the 3rd degree model.

d.)

if we ommit the x^4 term than we can expect the adj R^2 value to increase as well as all terms to maintain significance. As compared to the plot in part b , there is a notable difference as the model fit tighter to the data without being as overfitted as the plot in part c was. Unfortunately I was not able to super impose the model to the data plot using the coefficient I obtained from the summary command. I have included my r code, if you could take a look at what I am doing wrong I would appreciate the feedback.