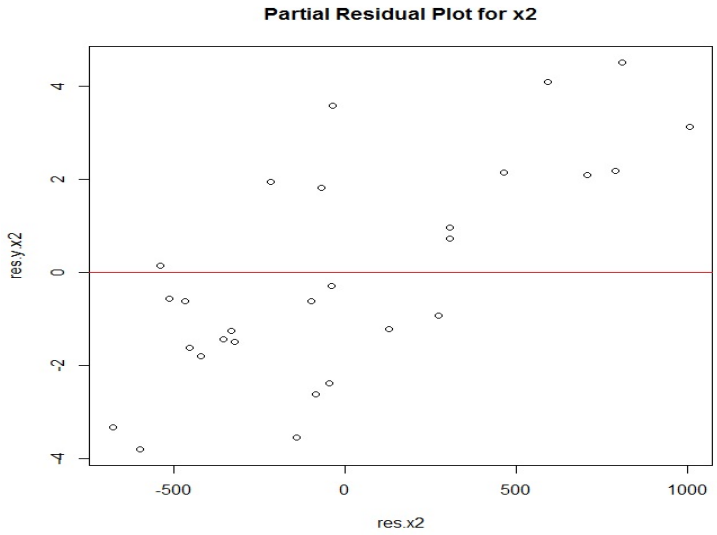
Stat 6021: Guided Question Set 8 Solutions

1. (a) The partial residual plot for x_2 informs us that a linear term for x_2 will be appropriate when x_7 and x_8 are already in the model, and that the estimated coefficient for x_2 would be positive in the MLR model with x_2, x_7, x_8 as predictors.



   (b) The estimated slope is 0.003598 and the estimated intercept is 0.

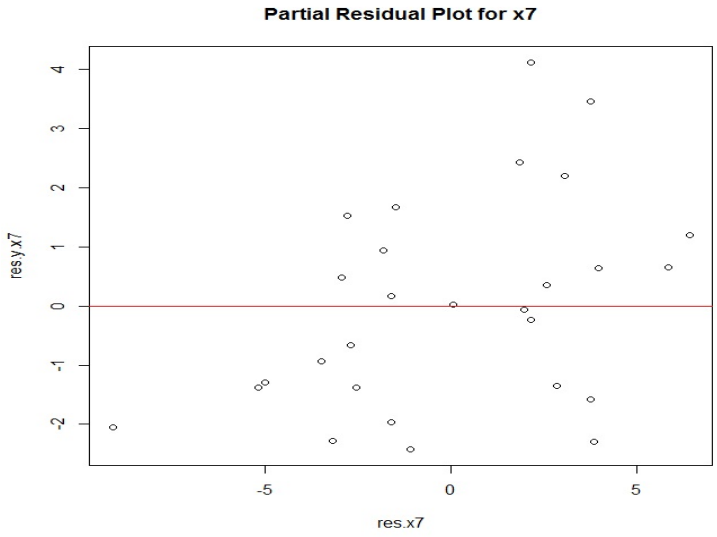
```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -6.714e-16  3.098e-01   0.000        1
res.x2       3.598e-03  6.677e-04   5.388 1.21e-05 ***
```

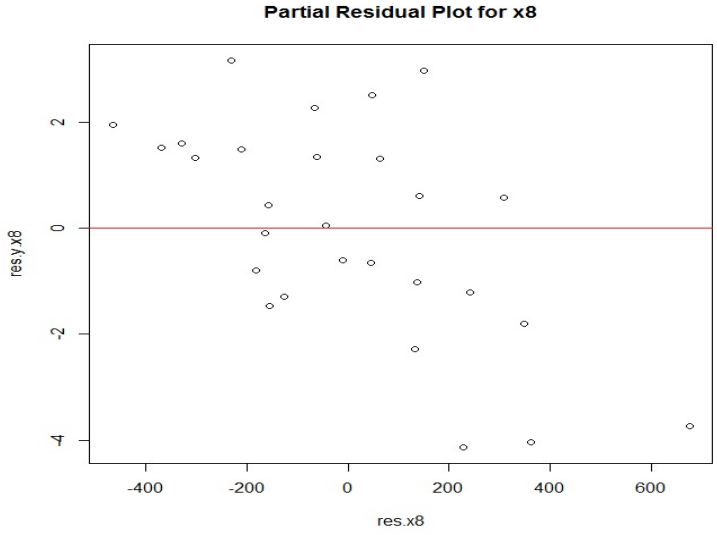
   (c) The estimated slope for x_2 is 0.003598, which is the same as the estimated slope for the partial residual plot from the previous part.

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.808372   7.900859  -0.229 0.820899
x2           0.003598   0.000695   5.177 2.66e-05 ***
x7           0.193960   0.088233   2.198 0.037815 *
x8          -0.004816   0.001277  -3.771 0.000938 ***
```

(d) The estimated slopes would be 0.1940 and -0.004816 respectively.

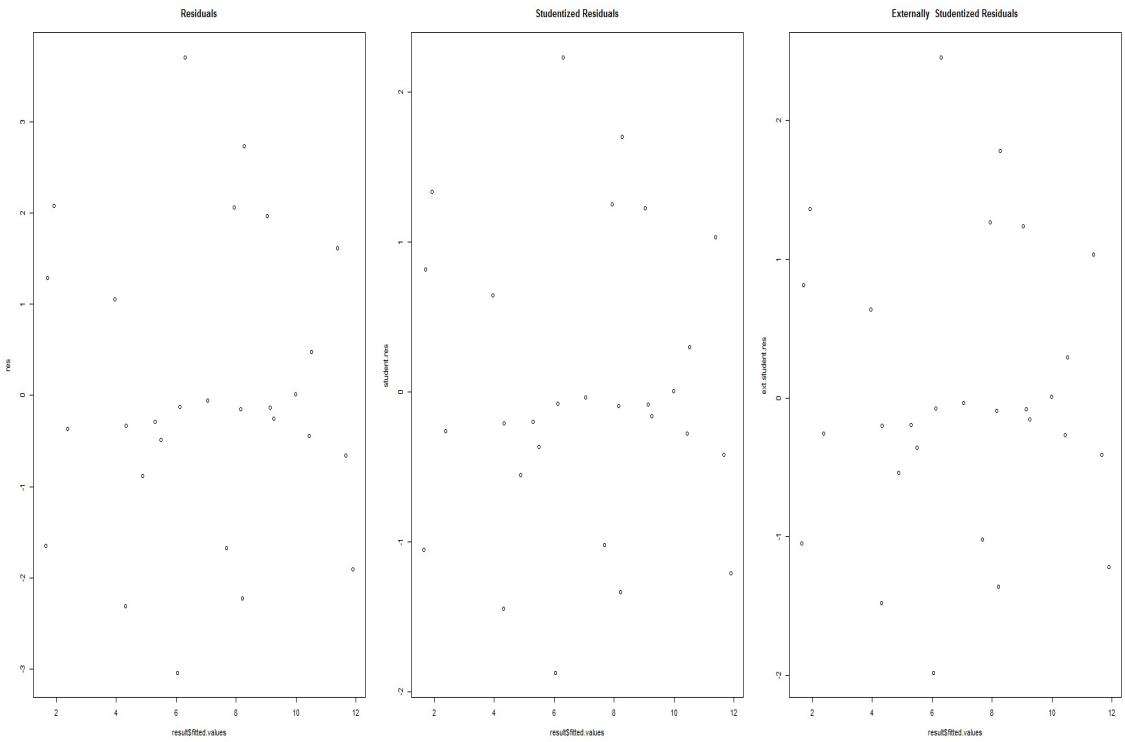
(e) The predictors x_7 and x_8 should be added as linear terms.





2. No outliers in terms of the response variable.

```
> n<-length(y)
> p<-4
> qt(1-0.05/(2*n), n-p-1)
[1] 3.53144
>
> ext.student.res[abs(ext.student.res)>qt(1-0.05/(2*n), n-p-1)]
named numeric(0)
```



3. Two teams that have high leverage.

```
> lev<-lm.influence(result)$hat ##leverages
>
> 2*p/n
[1] 0.2857143
>
> lev[lev>2*p/n]
       18        27
0.3928394 0.3192801
```

4. Observation 21 is influential in terms of $DFBETA_3$ and observation 10 is influential in terms of $DFBETA_4$.

```
> DFFITS<-dffits(result)
>
> DFFITS[abs(DFFITS)>2*sqrt(p/n)]
named numeric(0)
> DFBETAS<-dfbetas(result)
> DFBETA1<-DFBETAS[,1]
> DFBETA2<-DFBETAS[,2]
> DFBETA3<-DFBETAS[,3]
> DFBETA4<-DFBETAS[,4]
>
```

```
> DFBETA1[abs(DFBETA1)>2/sqrt(n)]
named numeric(0)
> DFBETA2[abs(DFBETA2)>2/sqrt(n)]
named numeric(0)
> DFBETA3[abs(DFBETA3)>2/sqrt(n)]
        21
-0.4148213
> DFBETA4[abs(DFBETA4)>2/sqrt(n)]
      10
-0.42602
> COOKS<-cooks.distance(result)
> COOKS[COOKS>qf(0.5,p,n-p)]
named numeric(0)
```