

Online HW1

Fechar Ourotcha

3/26/2020

Question 1

Come up with a 4 variable MLR using the CASchool set that does not include

income as the variable.

```
summary(lm(CASchools$score ~ CASchools$lunch + CASchools$computer + CASchools$expenditure + CASchools$teachers))
```

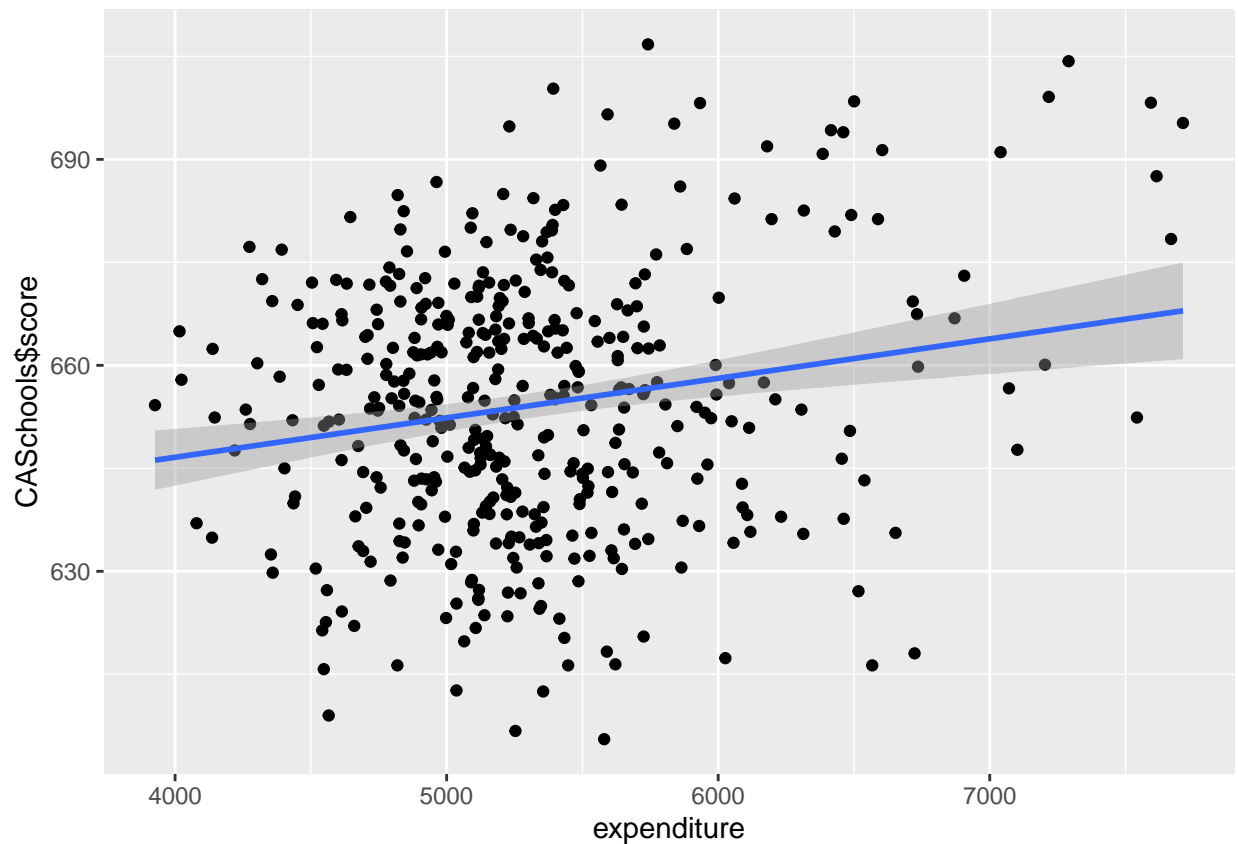
```
##
## Call:
## lm(formula = CASchools$score ~ CASchools$lunch + CASchools$computer +
##     CASchools$expenditure + CASchools$teachers)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -33.533  -5.686   0.032   5.753  33.069
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    6.596e+02  3.897e+00 169.269  < 2e-16 ***
## CASchools$lunch    -5.982e-01  1.668e-02 -35.861  < 2e-16 ***
## CASchools$computer    4.505e-03  2.917e-03   1.544   0.1233
## CASchools$expenditure  4.056e-03  7.032e-04   5.767 1.58e-08 ***
## CASchools$teachers   -1.257e-02  6.904e-03  -1.821   0.0694 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.064 on 415 degrees of freedom
## Multiple R-squared:  0.7759, Adjusted R-squared:  0.7737
## F-statistic: 359.1 on 4 and 415 DF,  p-value: < 2.2e-16
```

plot all of the independent variables separately against scores

plot the fitted value of your model against the dependent variable

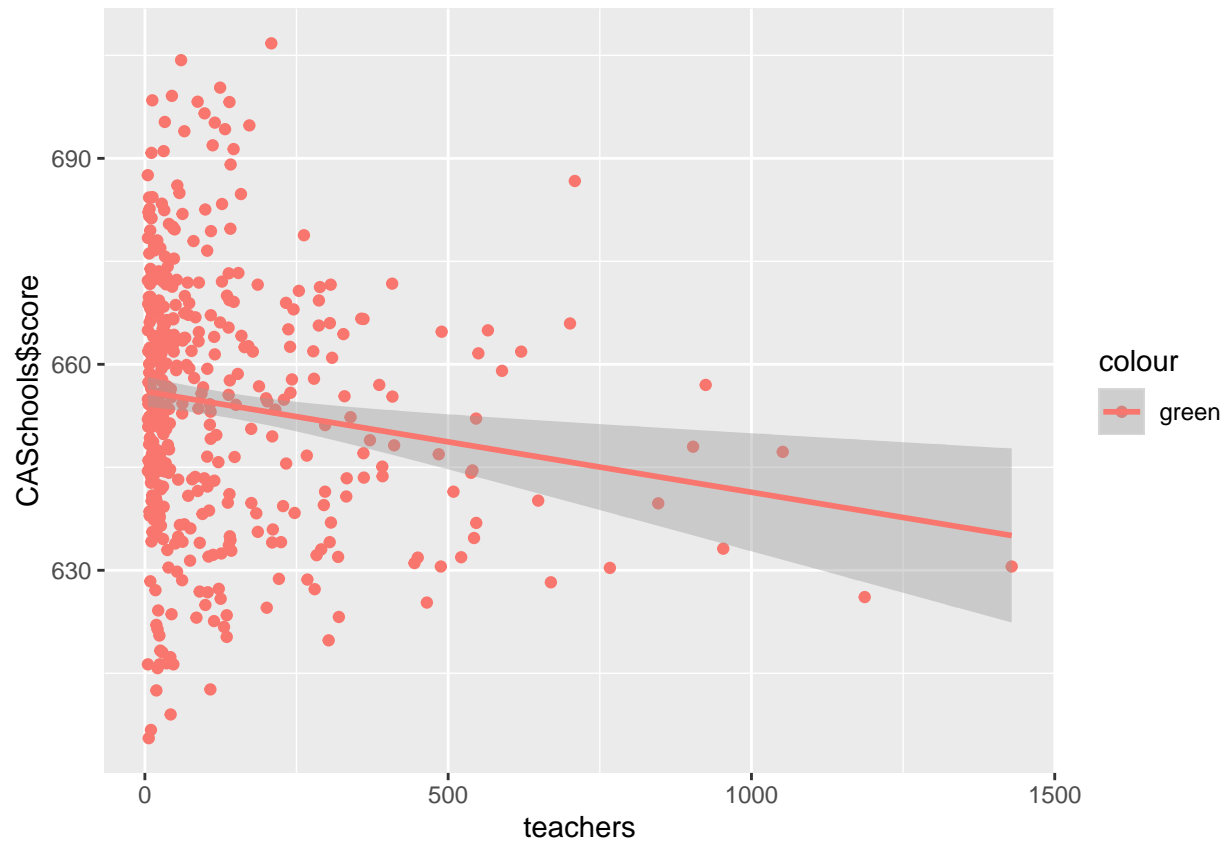
interpret your model in English.

```
ggplot(data = CASchools, aes(x = expenditure, y = CASchools$score)) +  
  geom_point() + geom_smooth(method = lm)
```

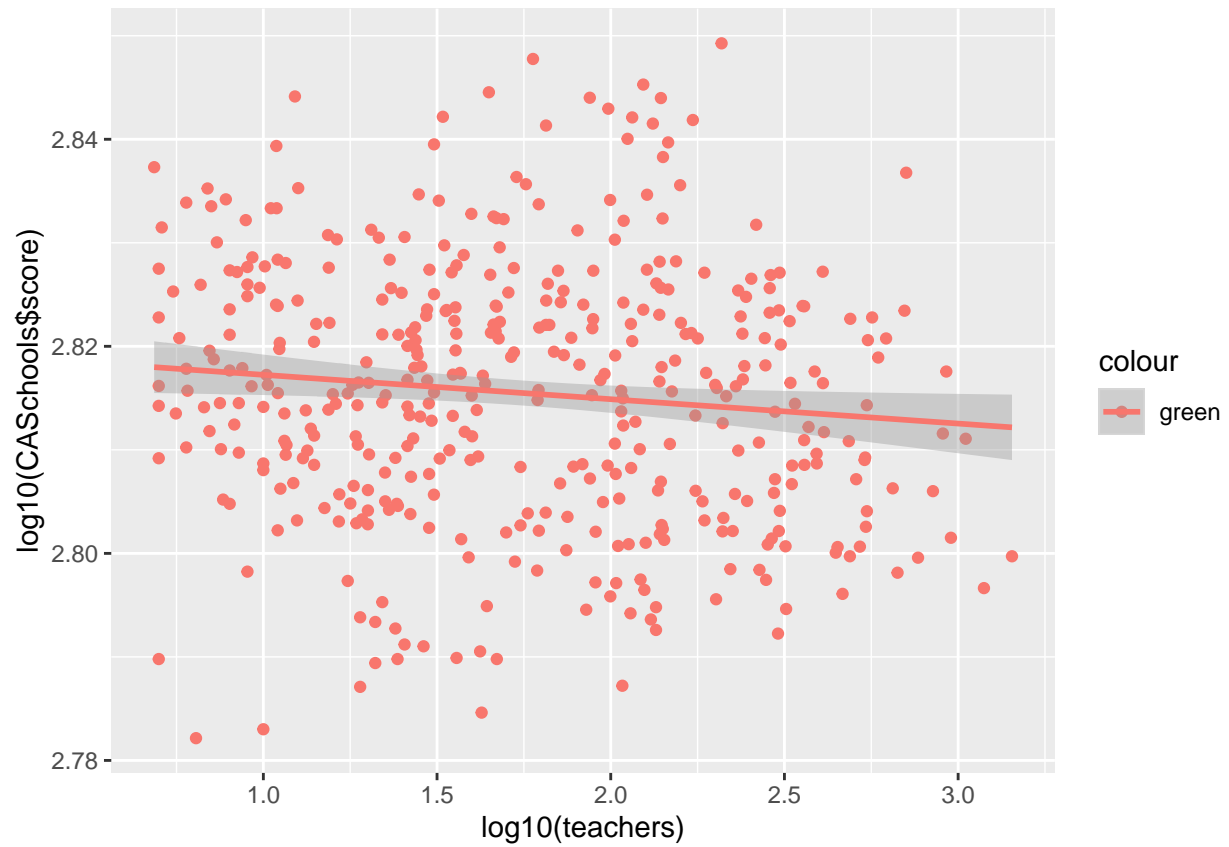


From this model, we can see that there is a small positive correlation between the expenditure spend per student and there test scores. However, the relationship is weak as you can tell from how spread out the data point are from each other. Because of this, we can not say this is a good model for any sort of insight.

```
ggplot(data = CASchools, aes(x = teachers, y = CASchools$score, color = "green")) +  
  geom_point() + geom_smooth(method = lm)
```

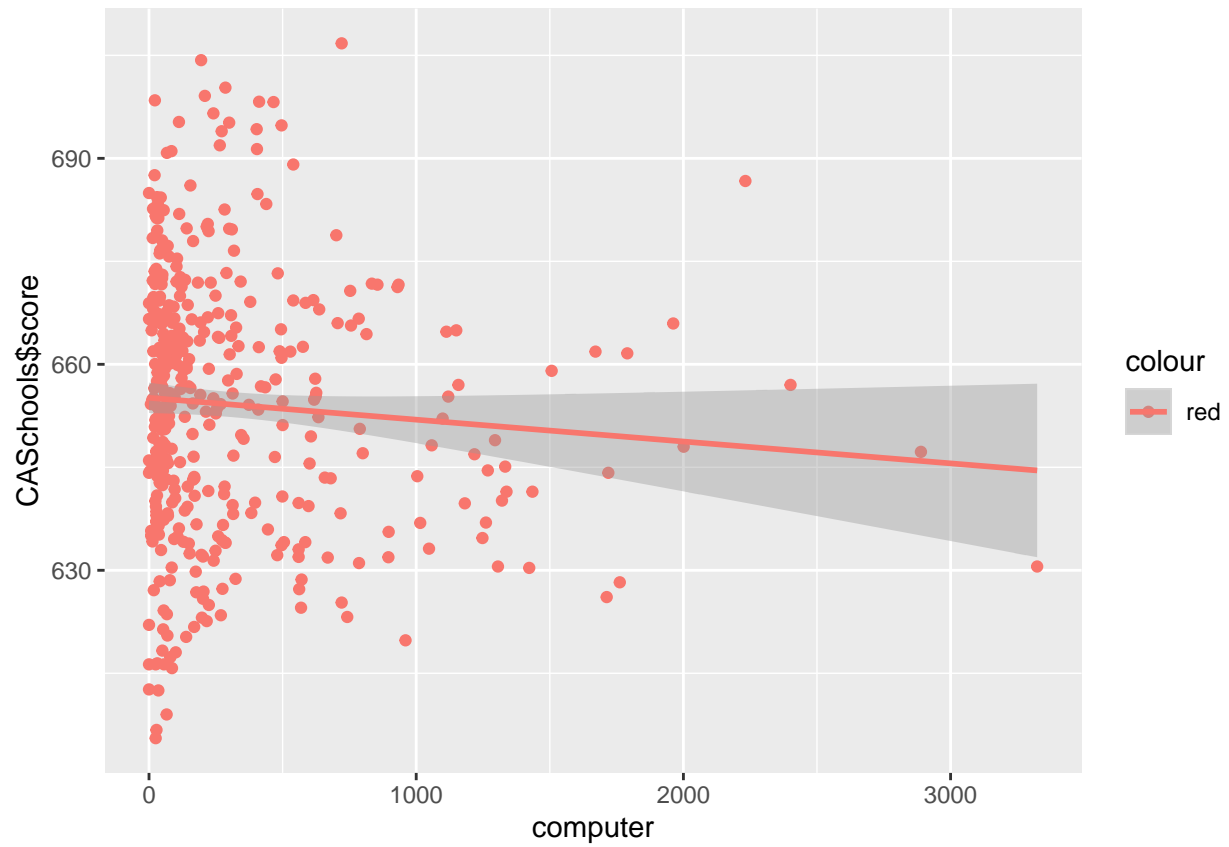


```
ggplot(data = CASchools, aes(x = log10(teachers), y = log10(CASchools$score), color = "green")) +
  geom_point() + geom_smooth(method = lm)
```

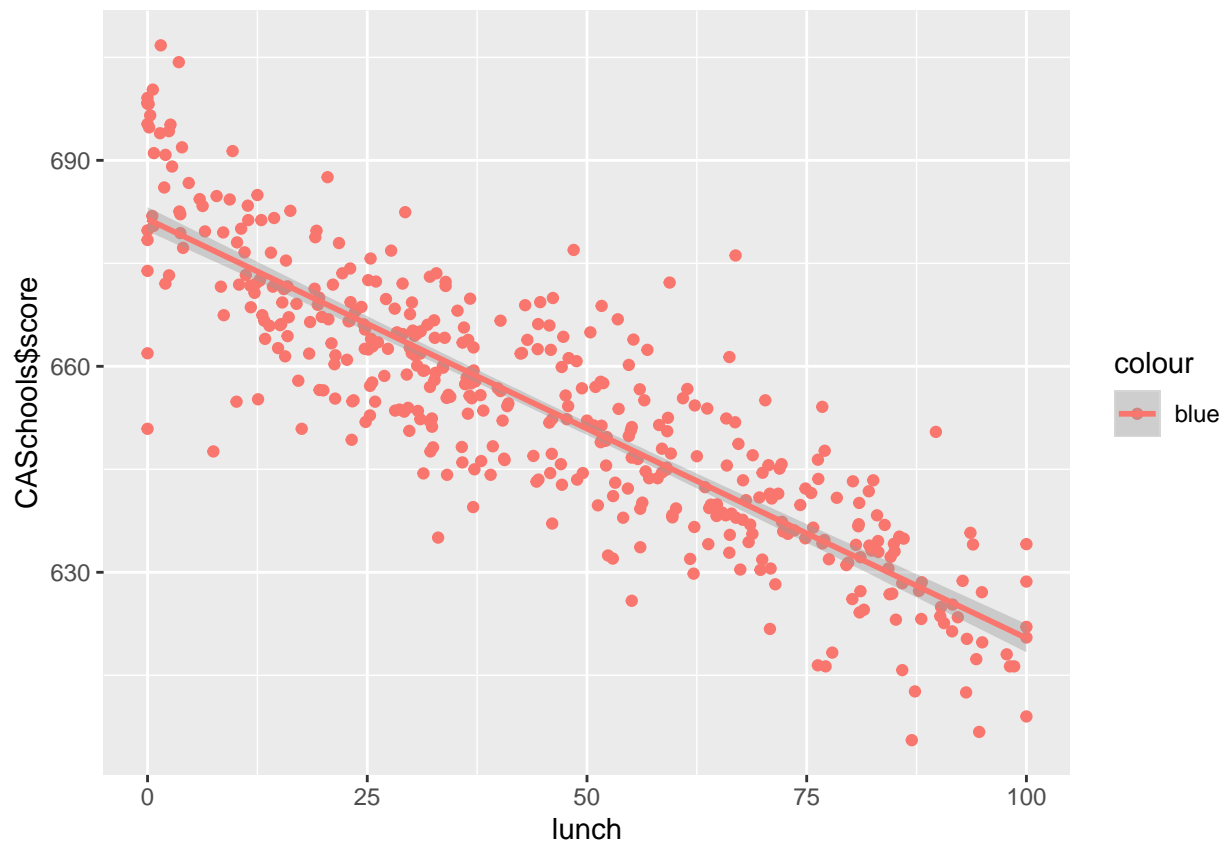


#

```
ggplot(data = CASchools, aes(x = computer, y= CASchools$score, color = "red"))+
  geom_point() + geom_smooth(method = lm)
```



```
ggplot(data = CASchools, aes(x = lunch, y= CASchools$score, color = "blue"))+  
  geom_point() + geom_smooth(method = lm)
```



Question 2

Thoughts/feedback on this new format. How can I make it better?

I like the online learn platform. The only problem for me is that Its hard for me to code and watch the session at the same time. Right now, I am trying to see if I can create two screens. Other than this I don;t have any problems with the online learning environment.