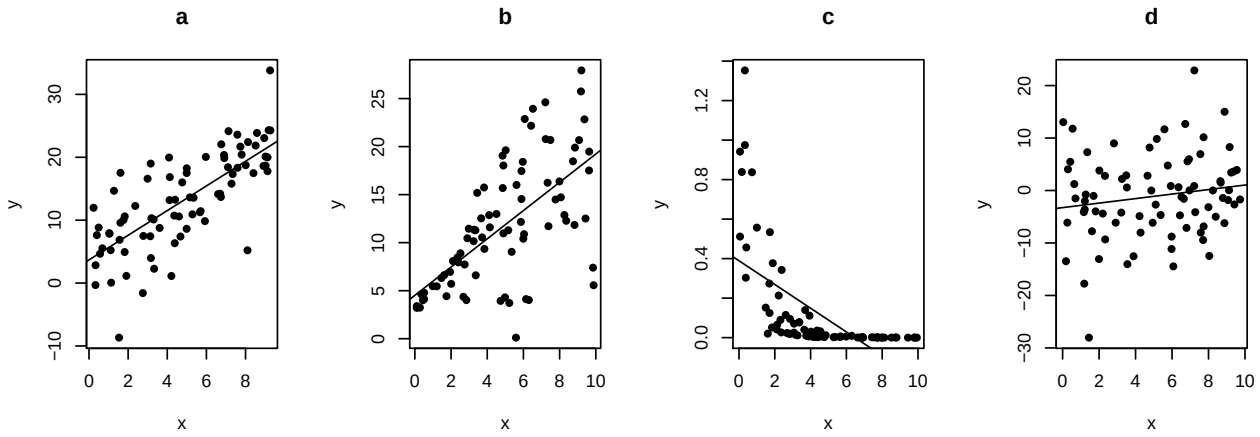
Quiz 8

Name:

Remember to state the logic behind your answers. If you use any variables, please state what the variables represent.

1. Which of the following plots show signs of heteroscedasticity?



2. From the below regression summary answer the following questions.

```
credit = read.csv(file = "http://www-bcf.usc.edu/~gareth/ISL/Credit.csv")
income_vs_education.lm = lm(Income ~ Education, data = credit)
summary(income_vs_education.lm)
```

```
##
## Call:
## lm(formula = Income ~ Education, data = credit)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -36.83 -23.88 -12.57  11.89 141.59
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  49.4192     7.8022   6.334 6.46e-10 ***
## Education    -0.3123     0.5651  -0.553    0.581
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 35.27 on 398 degrees of freedom
## Multiple R-squared:  0.0007668,  Adjusted R-squared:  -0.001744
## F-statistic: 0.3054 on 1 and 398 DF,  p-value: 0.5808
```

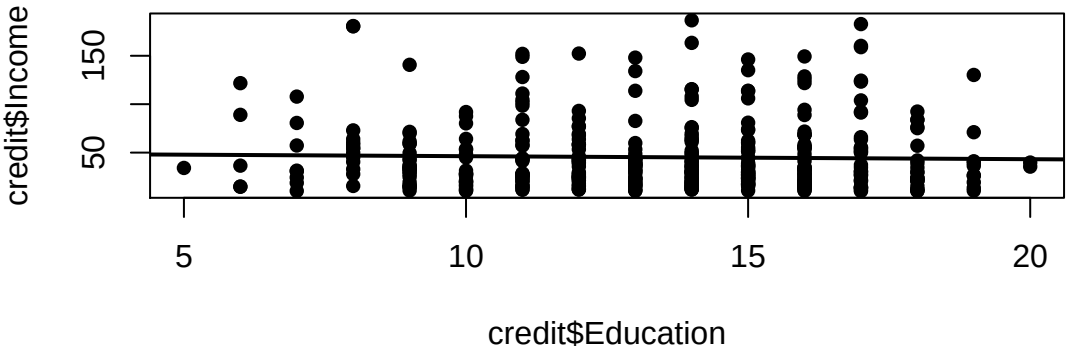
a. What is the R^2 of this regression and what is correlation between the fitted income and the observed income?

b. Is the coefficient for Education significantly different from 0 at the 95% confidence level?

c. Assuming the linear model is correct, an increase in 1 year of Education corresponds to what expected increase of Income?

d. Below are diagnostic plots for the above regression. Does the fit of this regression look good and why or why not?

```
plot(x = credit$Education, y = credit$Income, type = "p", pch = 16)
abline(income_vs_education.lm, lwd = 2)
```




```
par(mfrow = c(1, 4))
plot(income_vs_education.lm)
```