For the following examples, discuss whether each satisfy the four assumptions we use for linear regression. If not, which assumptions are violated, and why?

a) I wish to examine the relationship between wages and education for residents of California. To do so, I collect data from a random Econ 113 class on years of post-secondary education \( (educ) \), the most recent wage offer of each student \( (wage) \), and estimate the following equation:

\[
\text{wage} = \beta_0 + \beta_1 \text{educ} + \varepsilon
\]

b) I run the same regression as in (a) but instead I decide to use only a sample of seniors, who all have been in school for four years.

c) Finally, I run the following adjusted specification, where \( \text{score} \) is the score on the first midterm.

\[
\text{wage} = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{score} + \varepsilon
\]

Does this solve any of the previous problems in (a) or (b)

d.) What happens to the estimates in (c) if I measure education in months rather than years?