

Abe Porschet

1) Correct answers are a and d

2) The correct answers would be a and b because the others each have one indicator variable which doesn't make sense because the indicator variable has 3 levels.

5c) The average difference in lifespan between smokers and non-smokers is approximately 23.44 years. Smoking correlates with a decrease in lifespan but there is not necessarily evidence that smoking will lead to a 23.44 year decrease in lifespan only that there is an average of that decrease.

7b) The QQ plot and fitted vs residual plots both display favorable data that seems to show linearity and homoscedasticity which is positive. I do think that it is probably more reliable to use the test from 7a than the one from 6b because there could be multicollinearity in the data which could lead to the test for a linear relationship being compromised because the coefficients of the different variables would not be independent of each other leading to problems in the test, therefore decreasing reliability there is also the fact that testing for slope requires fewer assumptions than the test for 6b which means that the test from 7a will probably be more reliable more frequently given the fact that there are more opportunities to use it well than with the other test.