*C.1.*

*Consider the relationship between yearly wine consumption (liters of alcohol from drinking wine, per person) and yearly deaths from heart disease (deaths per 100,000 people) in 19 developed countries. Suppose that you read a newspaper article in which the reporter states the following:*

*Researchers find that the correlation between yearly wine consumption and yearly deaths from heart disease is –0.84. Thus, it is reasonable to conclude that increased consumption of alcohol from wine causes fewer deaths from heart disease in industrialized societies.*

Comment on the reporter’s interpretation of the correlation in this situation.

- Wine is not inclusive of all alcohol

- Wine is not a controlled variable

- Alcohol is not a controlled variable

- ‘Industrialized societies’ is not a controlled variable

*C.2.*

*“It is generally appropriate to delete all outliers in a data set that are apparent in a scatterplot.” Do you agree with this statement? Explain.*

No, outliers can be the matter of the opinion of the analyst. If enough ‘outliers’ are removed from the dataset nearly any conclusion can be reached.

*C.3.*

*How would you interpret the relationship between two numeric variables when the estimated least squares regression line for them is essentially horizontal (slope 0)?*

Data does not show a relationship; possibly no relationship.

*C.4.*

*Suppose that you generate a scatterplot of residuals versus fitted values of the dependent variable for an estimated regression equation. Furthermore, you find the correlation between the residuals and fitted values to be 0.829. Does this provide a good indication that the estimated regression equation is satisfactory? Explain why or why not.*

It does not. The errors should not align with the fitted values.

*C.5.*

*Suppose that you have generated three alternative multiple regression equations to explain the variation in a particular dependent variable. The regression output for each equation can be summarized as follows:*

A screenshot of a graph

Description automatically generated

Which of these equations would you select as “best”? Explain your choice.

Eq. 3, Adjusted R^2 is the lowest.