In-class Activity Chapter4_2

1. Suppose that you are interested in determining whether a relationship exists between the fluoride content in a public water supply and the dental caries experience of children using this water. The file water.Rdata contains the data from a study examining 7,257 children in 21 cities from the Flanders region in Belgium.
The fluoride content of the public water supply in each city, measured in parts per million (ppm), is saved under the variable name fluoride; the number of dental caries per 100 children examined is saved under the name caries. The total dental caries number is obtained by summing the numbers of filled teeth, teeth with untreated dental caries, teeth requiring extraction, and missing teeth.

a) Construct a two-way scatterplot for these data, with fluoride as the *x*-variable and caries as the *y*-variable.

b) Do fluoride and caries appear to be positively or negatively associated? Explain your answer.

2. This problem uses data from the Prevention of REnal and Vascular END-stage Disease (PREVEND) study, which took place between 2003 and 2006 in the Netherlands.
Body mass index (BMI) is a measure of body fat that is based on both height and weight. TheWorld Health Organization and National Institutes for Health define a BMI of over 25.0 as overweight; this guideline is typically applied to adults in all age groups. However, a recent study has reported that individuals of ages 65 or older with the greatest mortality risk were those with BMI lower than 23.0, while those with BMI between 24.0 and 30.9 were at lower risk of mortality. These findings suggest that the ideal weight-for-height in older adults may not be the same as in younger adults. Explore the relationship between BMI (BMI) and age (age), using the same sample of 500 individuals from the prevend data.

a) Create a plot that shows the association between BMI and age. Based on the plot, comment briefly on the nature of the association.

b) Fit a linear regression model to relate BMI and age.
c) Write the equation of the linear model.
d) Interpret the slope and intercept values in the context of the data. Comment on whether the intercept value has any interpretive meaning in this setting.
e) Is it valid to use the linear model to estimate BMI for an individual who is 30 years old? Explain your answer.
f) According to the linear model, estimate the average BMI for an individual who is 60 years old.
g) Based on the linear model, how much does BMI differ, on average, between an individual who is 70 years old versus an individual who is 50 years old?
h) Conduct a formal hypothesis test of no association between BMI and age, at the *alpha*= 05 significance level. Summarize your conclusions.

i) Report the R^2 of the linear model relating BMI and age. Based on the R^2 value, briefly comment on whether you think the estimated average BMI values calculated in part b) are accurate.

J) Create residual plots to assess the model assumptions of linearity, constant variability, and normally distributed residuals. In your assessment of whether an assumption is reasonable, be sure to clearly reference and interpret relevant features of the appropriate plot.

3. This problem uses data from the National Health and Nutrition Examination Survey (NHANES), a survey conducted annually by the US Centers for Disease Control (CDC). The data can be treated as if it were a simple random sample from the American population. The dataset nhanes.samp.adult.500 contains data for 500 participants ages 21 years or older that were randomly sampled from the complete NHANES dataset that contains 10,000 observations.
Regular physical activity is important for maintaining a healthy weight, boosting mood, and reducing risk for diabetes, heart attack, and stroke. In this problem, you will be exploring the relationship between weight (Weight) and physical activity (PhysActive) using the data in nhanes.samp.adult.500. Weight is measured in kilograms. The variable PhysActive is coded Yes if the participant does moderate or vigorous-intensity sports, fitness, or recreational activities, and No if otherwise.

a) Explore the data.
   i. Identify how many individuals are physically active.
   ii. Create a plot that shows the association between weight and physical activity. Describe what you see.

In-class Activity Chapter4_2