For this assignment use the Pulse.txt dataset used in the class examples this week. (The dataset is part of the Stat2Data library.) Use Resting pulse rate as the response variable (Y = Rest). The explanatory variables you will use are:
X₁ = Height (Hgt)
X₂ = Weight (Wgt)
X₃ = Smoke (1 for Smokers, 0 for Nonsmokers)

1. For simple linear regression there are three equivalent ways to test whether the slope is 0. In this problem, you will use them to determine whether height (Hgt) is a significant predictor of resting pulse rate (Rest) in the absence of any other predictors. For each of those three tests, state the hypotheses. Make sure you define any parameters you use in your hypotheses. Then use R to find the test statistic and p-value for each of these three tests. Show the R output, with the test statistic and p-value highlighted or underlined for each one.

2. Now include all 3 explanatory variables. Write the population model for this situation, including any necessary conditions. Use standard notation for all parameters, and use X₁, etc., instead of the names of the variables.

3. Fit the regression model using all 3 explanatory variables and show the results of the “summary” command. Use it to answer the following questions:
   a. Write the (sample) regression equation that would be used to predict resting pulse rate using these 3 explanatory variables. Fill in the numerical values for the coefficients; round each one to 3 decimal places. Use the names of the variables (Hgt, etc.) instead of X₁, etc.
   b. Interpret the value of the coefficient for “Smoke.”
   c. Using R or a calculator, find the predicted resting pulse rate for a smoker, 65 inches tall and 150 pounds.
   d. State and test the hypotheses to determine whether any one or more of the three predictors are useful in predicting resting pulse rate. Use appropriate parameter notation in your hypotheses (not words).

4. Show the results of the “anova” command for the model you fit in #3, and use it to answer the following questions.
   a. What are the numerical values of SSModel, SSE and SSTotal?
   b. If Height was the only predictor in the model, what would be the values of SSModel, SSE and SSTotal? (You do not need to rerun the model. You can find the appropriate values from the results of the anova command.)

5. For this question, you will continue the investigation from #1 of the role of height in this situation. In #1 you tested whether it was a significant predictor of resting pulse rate without any other explanatory variables.
   a. State and test the hypotheses to determine whether height is a significant predictor of resting pulse rate after accounting for weight and smoking status. Make sure you define any parameters used.
   b. Discuss whether your results differed for Question 1 and part (a) of this question. If you conclude that your results were similar, justify your conclusion statistically. If instead you conclude that they were different, explain why the results would differ in this situation.

Now you will use a model with only Weight and Smoke as predictors (not Height).

6. Write the population model. (Use variable names.) Identify what each coefficient in the model represents. As a hint to get you started, β₀ = the intercept of the regression line relating Y = Rest and X = Weight for non-smokers.

7. Fit the model in R, and show the results of the summary command. State and test the hypotheses to determine if the intercepts differ for the regression lines relating Y = Rest and X = Weight for smokers and non-smokers.

8. Now write the population model that would include different intercepts and different slopes for smokers and non-smokers in the regression lines relating Y = Rest and X = Weight. Interpret each of the coefficients.

9. Using the notation from your model in question 8, write the hypotheses that would be used to test whether the regression lines are different (intercepts and/or slopes) for smokers and non-smokers. (You do not have to carry out the test.)