Regression Review

NHES Data:
Average Weight: 122 lbs, SD: 30 lbs
Average Systolic BP 127 mmHg, SD 14 mmHg
r = 0.5, Data is homeoscedastic

1. Estimate the percentage of kids who weigh 128 lbs and have a BP above the population's average.
2. Estimate the percentage of kids who weigh 116 lbs and have a BP above the population's average.
3. Guess the weight of a kid with BP 148 mmHg.
4. Estimate the percentage of kids with BP over 106 mmHg
5. Estimate the percentage of kids with BP and weight above average (you can use Excel).

NHES III

1. Find the correlation between systolic BP at beginning and diastolic BP at beginning. Is there a linear relationship between the two?
2. Make a scatter diagram of systolic BP at the beginning vs. systolic BP at the end of an exam. What is the correlation coefficient for these variables?
3. For kids with a systolic BP of 100 at the beginning what is their average systolic BP at the end? What would you have guessed? Why is there a difference?
4. Repeat the last question for kids with a systolic BP of 184 at the beginning.
5. Measure the effect that relaxation over the course of the exam has on BP. Does this account for the results in the previous 2 questions?
6. Compute the RMS error for predicting end systolic BP simply from the effect of being more relaxed on the second reading. Also compute the RMS error for predicting end systolic BP by the regression line. Which is better?

BMI

1. Is there any evidence in this data set that patients became more relaxed over the course of the exam?
2. Create a column that has BMI.
3. Create a scatter plot with 3 series:
   - people who have a "normal" BMI
   - people who have an "overweight" BMI
   - people who have an "obese" BMI
4. Is the BMI index gender neutral?
5. Does BMI favor people who are tall or short?
6. Would it be more fair to classify people who are 1 SD above the average for their height to be overweigh? What percentage of people would fall into this category?