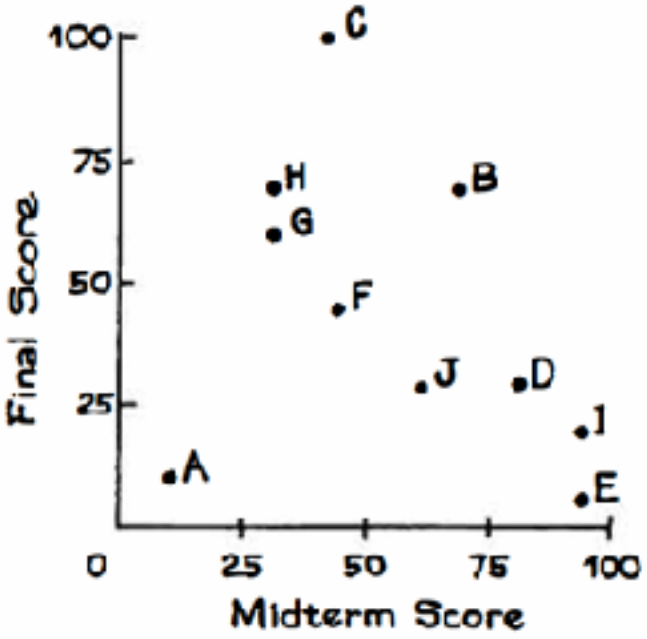
Stat 20 Homework 9

Please do this assignment in R Markdown and submit the knitted document.

1. Ch 8 A5:Students named A, B, C, D, E, F, G, H, I, and J took a midterm and a final in a certain course. A scatter diagram for the scores is shown below.
   a) Which students scored the same on the midterm as on the final?
   b) Which students scored higher on the final?
   c) Was the average score on the final around 25, 50, or 75?
   d) Was the SD of the scores on the final around 10, 25, or 50?
   e) For the students who scored over 50 on the midterm, was the average score on the final around 30, 50, or 70?
   f) True or false: on the whole, students who did well on the midterm also did well on the final.
   g) True or false: there is strong positive association between midterm scores and final scores.



2. Ch 8 Rev 3: Suppose men always married women who were exactly 8% shorter. What would the correlation between their heights be?
3. Ch 8 Rev 4: Is the correlation between the heights of husbands and wives in the U.S. around -0.9, -0.3, 0.3, or 0.9? Explain briefly.
4. Based on Ch 9 A10: Six data sets are shown on the next page. Calculate the correlation for (i) and (ii) in two ways. First, use the method in lecture code for 11/17: take the average of the products of the standard units. Second, check that you get the same answer if you use the cor() function. Find the correlations for the remaining data sets without doing any arithmetic, the other datasets are related to (i) and (ii) in ways that should make that easy.

(i)		(ii)		(iii)		(iv)		(v)		(vi)	
x	y	x	y	x	y	x	y	x	y	x	y
1	2	1	2	2	1	2	2	1	4	0	6
2	3	2	3	3	2	3	3	2	6	1	9
3	1	3	1	1	3	4	1	3	2	2	3
4	4	4	4	4	4	5	4	4	8	3	12
5	6	5	6	6	5	6	6	5	12	4	18
6	5	6	7	7	6	7	5	6	10	5	21
7	7	7	5	5	7	8	7	7	14	6	15

5. Based on Ch 9 B2: The National Health and Nutrition Examination Survey collects data on children. In this dataset, at each age from 6 to 11, the correlation between height and weight was just about 0.67. For all the children together, would the correlation between height and weight be just about 0.67, somewhat more than 0.67, or somewhat less than 0.67? Choose one option and explain. (Hint: consider drawing a scatter plot of what you think the dataset will look like.)
6. Ch 9 C4: For a certain data set, $r = 0.57$. Say whether each of the following statements is true or false, and explain briefly. If you need more information, say what you need and why.
   a) There are no outliers.
   b) There is a non-linear association.
7. Ch 9 E3: The correlation between height and weight among men age 18-74 in the U.S. is about 0.40. Say whether each conclusion below follows from the data; explain your answer.
   a) Taller men tend to be heavier.
   b) The correlation between weight and height for men age 18-74 is about 0.40.
   c) Heavier men tend to be taller.
   d) If someone eats more and puts on 10 pounds, he is likely to get somewhat taller.
8. Ch 9 Rev 10: In a study of 2005 Math SAT scores, the Educational Testing Service computed the average score for each of the 51 states, and the percentage of the high-school seniors in that state who took the test.14 (For these purposes, D.C. counts as a state.) The correlation between these two variables was equal to -0.84.
   a) True or false: test scores tend to be lower in the states where a higher percentage of the students take the test. If true, how do you explain this? If false, what accounts for the negative correlation?
   b) In Connecticut, the average score was only 517. But in Iowa, the average was 608. True or false, and explain: the data show that on average, the schools in Iowa are doing a better job at teaching math than the schools in Connecticut.

9. Ch 10 A2: For the men age 18 and over in HANES: average height = 69 inches, SD =3 inches, average weight =190 pounds, SD = 42 pounds, r = 0.41. Estimate the average weight of the men whose heights were each of the following, and comment on the answers to c) and d).

   a) 69 inches; b) 66 inches; c) 24 inches; d) 0 inches

10. Ch 10 C2: For the first-year students at a certain university. the correlation between SAT scores and first-year GPA was 0.60. The scatter diagram is football-shaped. Predict the percentile rank on the first-year GPA for a student whose percentile rank on the SAT was:

    a) 90%; b) 30%; c) 50%; d) unknown

11. Ch 10 Rev 3: Pearson and Lee obtained the following results in a study of about 1,000 families: average height of husband = 68 inches, SD = 2. 7 inches, average height of wife = 63 inches, SD = 2.5 inches, r = 0.25. Predict the height of a wife when the height of her husband is: a) 72 inches; b) 64 inches; c) 68 inches; d) unknown

12. For the data in the file family.csv (in the Data folder under Files) make a plot using ggplot() with height on the horizontal axis and weight on the vertical axis. Look at lecture code from 4/8 for examples. Include this in a code chunk as in the .rmd file to make it appear in the file you submit.