Math 116 ACTIVITY 6: Understanding Relationships between variables - scatterplots and correlation

Why

Much of the use of statistics involves relationships between variables. It is important to understand the terms and ideas involved in describing such relationships in the same way that it is important to understand the terms and ideas used in describing distributions of single variables

LEARNING OBJECTIVES

1. Learn the terminology and ideas for looking at association between two quantitative variables
2. Be able to find and interpret the correlation coefficient for a set of bivariate (two-variable) data.
3. Work as a team, using the team roles.

CITERIA

1. Success in working as a team and in fulfilling the team roles.
2. Success in involving all members of the team in the conversation.
3. Success in completing the exercises

RESOURCES

1. The Statistics Handbook chapter on Organizing an Describing data - especially pp.4–5
2. The notes on Association and Correlation handed out in class
3. Your calculators and the sheet with instructions for your calculator (different for different individuals)
4. 40 minutes

PLAN

1. Select roles, if you have not already done so, and decide how you will carry out steps 2 and 3 (5 minutes)
2. Work through the group exercises given here - be sure everyone understands all results & procedures(25 minutes)
3. Assess the team's work and roles performances and prepare the Reflector's and Recorder's reports including team grade (5 minutes).

EXERCISE

1. In a study of diet and exercise, data have been collected on the lean body mass (in kilograms- body mass excluding all fat) and metabolic rate (calories burned per 24 hours) of 12 women . The researchers hope to show that lean body mass [which is relatively easy to measure] is useful in predicting metabolic rate [which is harder to measure]. The data are given here:

Subject #	Mass	Rate	Subject#	Mass	Rate
1	40.3	1189	7	41.2	1204
2	33.1	913	8	50.6	1502
3	36.1	995	9	42.0	1256
4	54.6	1425	10	42.4	1124
5	48.5	1396	11	34.5	1052
6	42.0	1418	12	51.1	1347

   (a) Which variable (metabolic rate, lean body mass) should be considered the explanatory variable, and which should be considered the response, [or does this distinction not make sense in this situation]?

   (b) Plot the data. Be sure to label the axes, indicating which variable is on which axis (Mass ranges from about 30 Kg to about 55 Kg. Rate ranges from about 900 calories to about 1600 calories).

   (c) Describe the association between the variables:
   Is it positive or negative?
   What is the form (linear, curved, complicated)?
   Does the association appear to be strong (a clear pattern, points close to a line or curve) or weak (no clear pattern, or lots of scatter away from the pattern)?

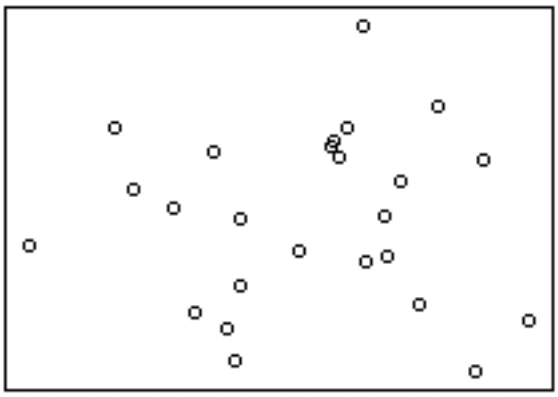
(d) Find the correlation coefficient r [use your calculator - do not work from the formula]. Does the value seem reasonable for the graph?

2. Below there are 5 scatterplots, labeled a— e. The correlation coefficients for the plots are $0, -.3, -.65, .8,$ and $.4$. Which diagram corresponds to which correlation coefficient? How can you tell?

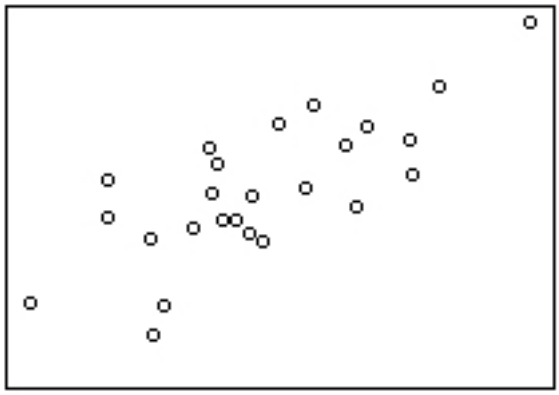
READING ASSIGNMENT (in preparation for next class)
Read the section on the regression line (pp 5–6)

SKILL EXERCISES:(hand in - individually - with assignment for this week)
Exercises for Chapter 1, 9-11

a.)

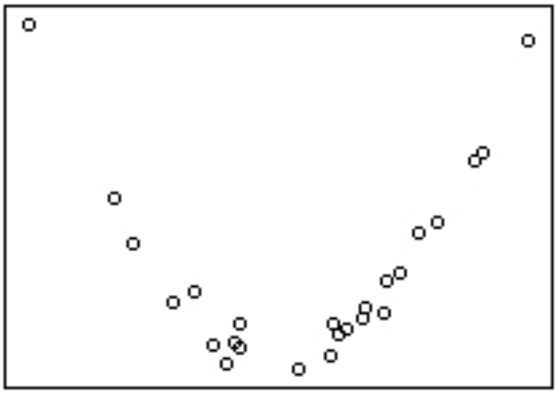


b.)

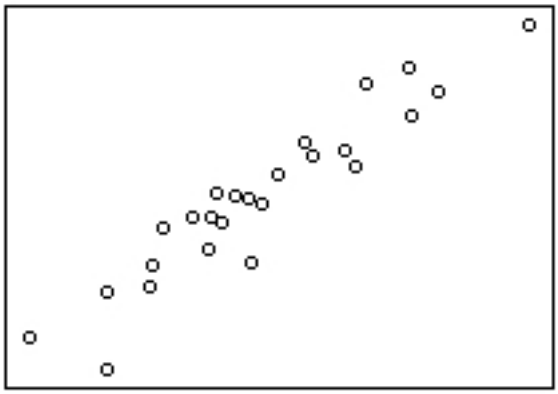


a

c.)



d.)



c

e.)