

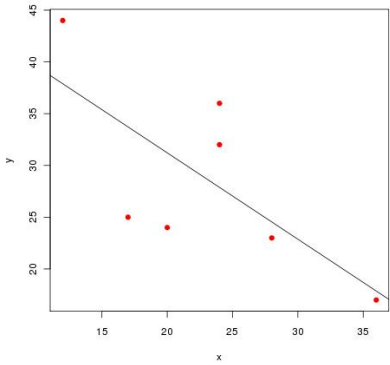
## SOC510 Homework #4

### Chapter 4.

A. Answer 1–6, using the following data.

work satisfaction( $x$ )	12	24	17	28	24	36	20
propensity to leave a job( $y$ )	44	36	25	23	32	17	24

- draw a scatterplot (ChangHwan: I added a regression line for your reference)



- $\bar{x} = 23, \bar{y} = 28.7$

- $s_x = \sqrt{\frac{SS(x)}{n-1}} = \sqrt{\frac{362}{6}} = 7.77,$

$$s_y = \sqrt{\frac{SS(y)}{n-1}} = \sqrt{\frac{503.4}{6}} = 9.16$$

- $SS(x) = \sum x^2 - \frac{(\sum x)^2}{n} = 4065 - \frac{161^2}{7} = 362$   
 $SS(y) = \sum y^2 - \frac{(\sum y)^2}{n} = 6275 - \frac{201^2}{7} = 503.4286$   
 $SS(xy) = \sum xy - \frac{\sum x \sum y}{n} = 4321 - \frac{161(201)}{7} = -302$

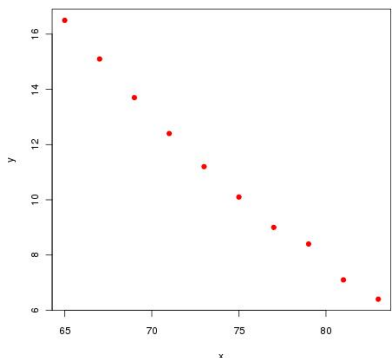
- correlation coefficient  $r = \frac{SS(xy)}{\sqrt{SS(x)SS(y)}} = \frac{-302}{\sqrt{362(503.43)}} = -.7074$

- Interpret  $r$ : Work satisfaction ( $x$ ) and propensity to leave a job ( $y$ ) are strongly negatively associated. That is, as the level of work satisfaction decreases, workers are more likely to leave the jobs.

B. People not only live longer today but also live longer independently, The May/June 1989 issue of *Public Health Reports* published an article titled “A Multistate Analysis of Active Life Expectancy.” Two of the variables studied were a person’s current age and the expected number of years remaining.

Age( $x$ )	65	67	69	71	73	75	77	79	81	83
Years Remaining ( $y$ )	16.5	15.1	13.7	12.4	11.2	10.1	9.0	8.4	7.1	6.4

1. draw a scatterplot



$$2. \quad s_x = \sqrt{\frac{SS(x)}{n-1}} = \sqrt{\frac{330}{9}} = 6.06,$$

$$s_y = \sqrt{\frac{SS(y)}{n-1}} = \sqrt{\frac{104.29}{9}} = 3.40$$

$$3. \quad SS(x) = \sum x^2 - \frac{(\sum x)^2}{n} = 55090 - \frac{740^2}{10} = 330$$

$$SS(y) = \sum y^2 - \frac{(\sum y)^2}{n} = 1312.09 - \frac{109.9^2}{10} = 104.289$$

$$SS(xy) = \sum xy - \frac{\sum x \sum y}{n} = 7947.9 - 740(109.9) = -184.7$$

$$4. \quad \text{correlation coefficient } r = \frac{SS(xy)}{\sqrt{SS(x)SS(y)}} = \frac{-184.7}{\sqrt{330(104.289)}} = -.996$$

C. From the textbook,

4.19; (a) This would be a perfect linear relationship: If the husbands age is y and the wives age is x, the linear relationship  $y = x + 2$  would hold, and hence  $r = 1$  (because the slope is positive).

$$4.21; r = \frac{392.878}{\sqrt{(161.988)(1238.968)}} = .877$$

4.33 (b only); (b) Although changing the scales (units) makes the scatterplot look very different, it has no effect on the correlation.

4.35; (a) Because gender has a nominal scale, we cannot compute the correlation between sex and anything. (There is a strong association between gender and income. Some writers and speakers use “correlation” as a synonym for association. It is much better to retain the more specific meaning.) (b) A correlation  $r = 1.09$  is impossible because  $-1 \leq r \leq 1$  always. (c) Correlation has no units, so  $r = 0.23$  bushel is incorrect.

4.38;  $\text{Sun} = -976.1409 + .6136 (\text{Year}) + e$ ; thus over time, sunlight has gotten brighter