

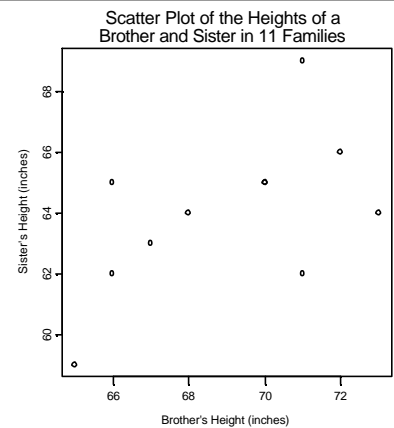
### The Sample Linear Correlation Coefficient

- $r_{XY}$  (or just  $r$  for short) is the **sample linear correlation coefficient**
- $r_{XY}$  measures the **strength** and **direction** of **linear association** between two quantitative variables  $X$  and  $Y$ .

### Formula for the Sample Linear Correlation Coefficient

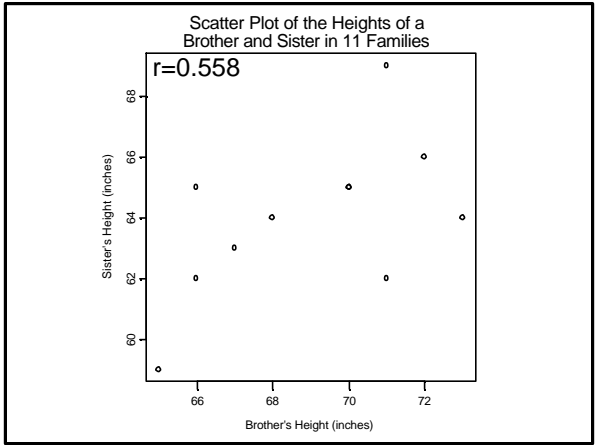
$$r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{s_X s_Y (n-1)}$$

For example, consider 11 families randomly selected from the population of families with one brother and one sister, both full grown. Let  $X_i$  denote the height (in inches) of the brother in the  $i$ -th family. Let  $Y_i$  denote the height (in inches) of the sister in the  $i$ -th family.

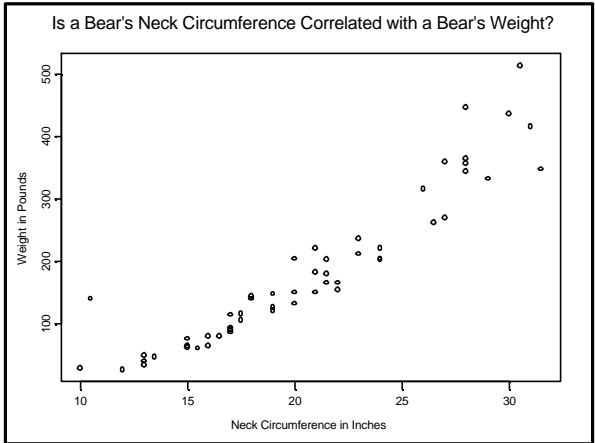


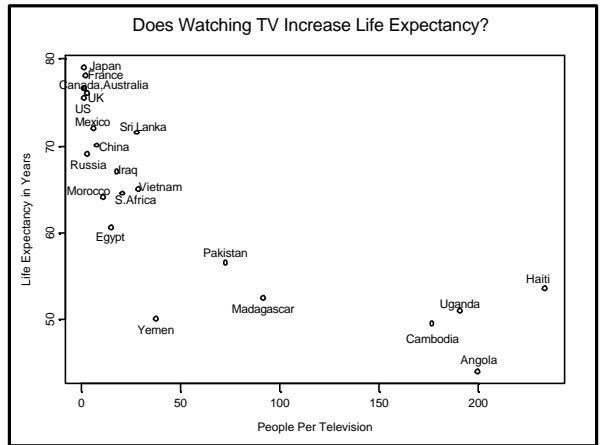
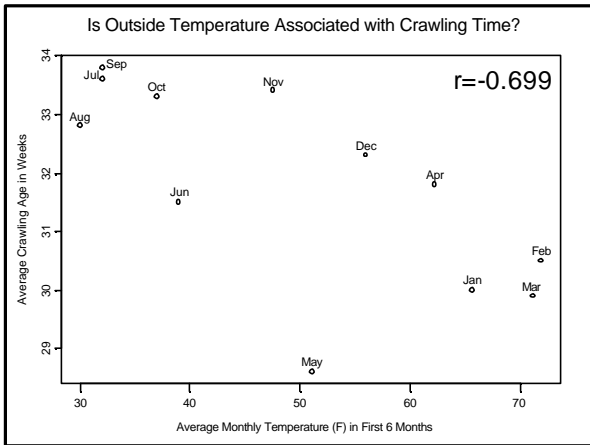
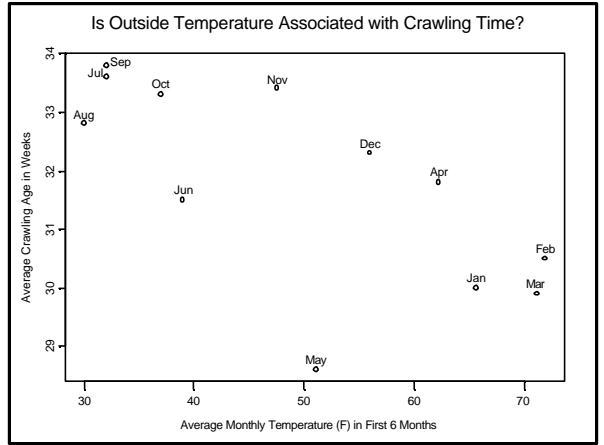
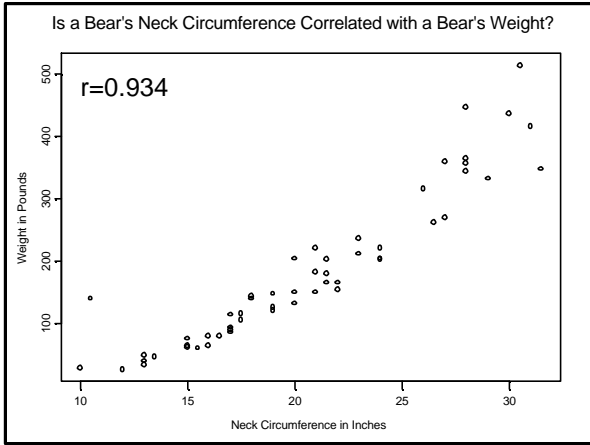
$i$	$X_i$	$Y_i$	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$		
1	71	69	2	5	10	4	25		
2	68	64	-1	0	0	1	0		
3	66	65	-3	1	-3	9	1		
4	67	63	-2	-1	2	4	1		
5	70	65	1	1	1	1	1		
6	71	62	2	-2	-4	4	4		
7	70	65	1	1	1	1	1		
8	73	64	4	0	0	16	0		
9	72	66	3	2	6	9	4		
10	65	59	-4	-5	20	16	25		
11	66	62	-3	-2	6	9	4		
			759	704	0	0	39	74	66

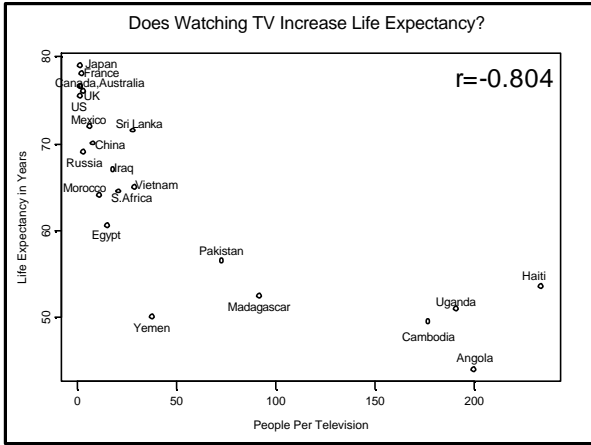
$\bar{X} = 759/11 = 69$      $\bar{Y} = 704/11 = 64$      $S_x = \sqrt{\frac{74}{11-1}}$      $S_y = \sqrt{\frac{66}{11-1}}$   
 $r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) / (n-1)}{s_x s_y} = \frac{3.9}{\sqrt{(7.4)(6.6)}} = 0.558$



- ### Some Properties of $r_{XY}$
- $r_{XY}$  estimates the population linear correlation coefficient  $\rho_{XY}$ .
  - $r_{XY}$  is dimensionless and is always between 1 and -1.
  - $r_{XY}=1$  if and only if all data points fall perfectly on a line with positive slope.
  - $r_{XY}=-1$  if and only if all data points fall perfectly on a line with negative slope.
  - $r_{XY}=0$  suggests that there is no *linear* association between X and Y.







**Guess the correlation for each pair of variables.**

A) Stalk Diameter of Corn Plant	Weight of Corn Plant
B) Person's Age	Person's Year of Birth
C) Daily Dow Jones Industrial Average	Daily Rainfall in Seattle
D) # of Ultrasounds During Pregnancy	Birth Weight of Baby
E) U.S. Monthly Ice Cream Cone Sales	Drowning per Month in U.S.
F) Age of Wife	Age of Husband