Homework 7 (Correlation)

Due: Monday, Dec 6th (10 pts)

1. The data set in Table 10.9 is in reference to Cartoon 10.1. As labeled, the independent variable represents the previously given number of complaints per month for Santa's reindeer. The dependent variable, however, represents the number of reindeer turned into venison.

a. Construct a scattergram with this data set. What type of relationship appears to exist between complaints and reindeer becoming venison?
b. Calculate the slope (a) and intercept (b) for this data.
c. Plug the calculated slope and intercept values into the least-squares regression line (Y = a + bX), and then plot it over the previously constructed scattergram.
d. Calculate the Pearson's-r correlation coefficient for this data set.
e. Using the appropriate t test, determine if the relationship between complaints and number of reindeer becoming venison is statistically significant.
f. Having made all of these calculations, what would one conclude about the relationship between complaints and the number of reindeer becoming venison?

Table 10.9 Complaints and Venison Data

Month	*Number of Complaints Received Per Month (X)*	*Number Turned Into Venison (Y)*
January	2	1
February	1	2
March	3	3
April	4	2
May	3	5
June	5	8
July	4	7
August	7	10
September	8	9
October	10	13
November	11	16
December	14	13

2. Once again, in reference to this cartoon, we are given the data set in Table 10.10. The independent variable is the age of the reindeer while the dependent variable represents the number of complaints received in a year's time. The data set obviously assumes that these 10 reindeer did not become venison in the (year period) of measurement.
a. Construct a scattergram with this data set. What type of relationship appears to exist between age and number of complaints received in a year?
b. Calculate the slope (a) and intercept (b) for this data set.
c. Plug the calculated slope and intercept values into the least-squares regression line (Y = a + bX), and then plot it over the previously constructed scattergram.

d. Calculate the Pearson's-r correlation coefficient for this data set.
e. Using the appropriate t test, determine if the relationship between complaints and number of reindeer becoming venison is statistically significant.
f. Having made all of these calculations, what can you conclude about the relationship between age and the number of complaints received?

Table 10.10 Age and Complaints Data

Reindeer's Name	*Age of Reindeer (X)*	*Number of Complaints Received for the Year (Y)*
Rudolph	1	20
Donner Jr.	1	15
Dasher II	3	17
Blixen	3	14
Dancer	4	15
Prancer	5	11
Comet	5	8
Cupid	5	10
Donner Sr.	7	9
Dasher III	10	8

3. Once again, in reference to this cartoon, we are given the data set in Table 10.11. In this study, we are interested in whether Santa displaces any of his anger at the reindeer's complaints toward elves. Here, the independent variable is the number of complaints from the reindeer, while the dependent variable represents the number of elves that Santa kicks in a given month.
a. Construct a scattergram with this data set. What type of relationship appears to exist between the number of reindeer complaints received in a month at the number of elves kicked?
b. Calculate the slope (a) and intercept (b) for this data set.
c. Plug the calculated slope and intercept values into the least-squares regression line ($Y = a + bX$), and then plot it over the previously constructed scattergram.
d. Calculate the Pearson's-r correlation coefficient for this data set.
e. Using the appropriate t test, determine if the relationship between complaints and the number of elves getting kicked is statistically significant.
f. Having made all of these calculations, what can you conclude about the relationship between the number of complaints received from reindeer and the number of elves Santa kicked?

Table 10.11 Complaints and Elves Kicked Data

Month	*Number of Complaints Received Per Month (X)*	*Number of Elves Kicked (Y)*
January	2	6
February	1	2
March	3	5
April	4	1
May	3	7
June	5	4
July	4	3
August	7	8
September	8	3
October	10	12
November	11	3
December	14	5

4. For our final example, assume that we want to assess the relationship between the average number of toys that children request each month, and the number of elves that Santa kicks. Here, the independent variable

is the average number of toys requested by children each month, while the dependent variable represents the number of elves that Santa kicks in a given month.

a. Construct a scattergram with this data set. What type of relationship appears to exist between number of toys requested and the number of elves that Santa kicks in a month?
b. Calculate the slope (a) and intercept (b) for this data set.
c. Plug the calculated slope and intercept values into the least-squares regression line (Y = a + bX), and then plot it over the previously constructed scattergram.
d. Calculate the Pearson's-r correlation coefficient for this data set.
e. Using the appropriate t test, determine if the relationship between average number of toys children ask for each month and the number of elves getting kicked is statistically significant.
f. Having made all of these calculations, what can you conclude about the relationship between the number of toys children request each month and the number of elves Santa kicked?

Table 10.12 Toys Asked for and Elves Kicked Data

Month	*Average Number of Toys Asked for Each Month (X)*	*Number of Elves Kicked (Y)*
January	0	12
February	1	9
March	1	8
April	3	7
May	4	4
June	5	2
July	5	1
August	6	4
September	8	6
October	9	8
November	10	10
December	12	12

5. For any of the above exercises where the correlation was not significant, how many observations/cases (what size *n*) would need be added in order for the correlation to be significant. This of course assumes that adding new observations does not change the strength or direction of the original relationship identified. Report both the number of new observations/cases needed and the resulting total sample size. Note, remember that the degrees of freedom for correlations are *n* - 2.