Workshop 5

Correlation Analysis

Name: _____

Date Completed: _____

Provide all solutions, answers and requested outputs for each question.

Questions 1 & 2: (a) Compute the Pearson correlation for the following set of data:

X	*0*	*1*	*2*	*3*	*4*
Y	*3*	*1*	*5*	*9*	*7*

(b) Add 5 points to each *X* value, and compute the Person correlation again.
(c) When you add a constant to each score, what happens to *SS* for *X* and *Y*? What happens to the correlation between *X* and *Y*?
(d) Now multiply each *X* in the original data by 3, and compute the Pearson correlation again.
(e) When you multiply by a constant, what happens to *SS* for *X* and *Y*? What happens to the correlation between *X* and *Y*?

Question 3 & 4: Correlation studies are often used to help determine whether certain characteristics are controlled more by genetic influences or by environmental influences. These studies often examine adopted children and compare their behavior with the behavior**s** of their birth parents and their adoptive parents. One study examined how much time individuals spend watching TV (Plomin, Corley, Defries, & Fulker, 1990). The following data are similar to results obtained in the study.

Amount of Time Spent Watching TV		
Adopted Children	Birth Parents	Adoptive Parents
2	0	1
3	3	4
6	4	2
1	1	0
3	1	0
0	2	3
5	3	2
2	1	3
5	3	3

(a) What do scatter plots tell us about the relationships of TV watching between children and their birth and children and their adoptive parents?
(b) Compute the Spearman correlation between the children and their birth parents.
(c) Compute the Spearman correlation between the children and their adoptive parents.
(d) Based on the two correlations, does TV watching appear to be inherited from the birth parents or is it learned from the adoptive parents?

Question 5 & 6: (a) Compute both the Pearson and Spearman correlation for the Father's Education (FAED) and Mother's Education (MAED) variables for the HSB500.csv data set.
(b) What do the correlation tells us about the relationship between these two variables?

Question 7 & 8: A researcher has developed a new test of self-esteem. To evaluate the reliability of the test, the researcher obtains a sample of $n = 8$ participants. each individual takes the test on a Monday morning, then returns 2 weeks later to take the test again. The two scores for the individual are reported in the following table.

First Test	13	5	12	11	9	14	8	8
Second Test	15	4	13	11	10	13	8	6

(a) Which of the two correlation methods (Pearson or Spearman) is the best approach for this analysis and why?
(b) Using an appropriate correlation computational formula, what is the correlation between the first and second measure for this sample?
(c) Is the correlation between the two tests small, medium or large? Give justifications for your conclusion.

Question 9 & 10: It is well known that similarity in attitudes, beliefs, and interests plays an important role interpersonal attraction (see Byrne, 1971, for example). Therefore, correlations for attitudes between married couples should be strong. Suppose a researcher developed a questionnaire that measures how liberal or conservative one's attitudes are. Low scores indicate that the person has liberal attitudes, whereas high scores indicate conservatism. The following hypothetical data are scores for married couples.

Couple	Wife	Husband
A	11	14
D	6	7
C	16	15
D	4	7
E	1	3
F	10	9
G	5	9
H	3	8

(a) Compute Pearson correlation using a formula
(b) Compute Pearson using statistical packages
(c) Is there a significant relationship between these attitudes for husbands and wives? Why do you think so?