**DATA ANALYTICS FOR MULTIVARIATE DISTRIBUTIONS**

**MULTIPLE LINEAR REGRESSION & PARTIAL AND MULTIPLE CORRELATION**

***--Show All Work--***

The ISU Registrar’s lead administrator has asked you to conduct a preliminary scientific study of variables that appear to be associated with the GPA’s of undergraduates.

Suppose data have been collected on undergraduate GPA’s (Y), high school GPA’s (X), and SAT scores (Z) for a random sample of twenty students. Here are the data:

X Z Y Student

3.45 1232.00 3.38 1.00

2.78 1070.00 2.89 2.00

2.52 1086.00 2.76 3.00

3.67 1287.00 3.55 4.00

3.24 1130.00 3.19 5.00

3.44 1232.00 3.30 1.00

2.77 1070.00 2.81 2.00

2.55 1086.00 2.60 3.00

3.66 1287.00 3.54 4.00

3.22 1130.00 3.17 5.00

3.45 1232.00 3.38 1.00

2.78 1070.00 2.89 2.00

2.52 1086.00 2.76 3.00

3.67 1287.00 3.55 4.00

3.24 1130.00 3.19 5.00

3.44 1232.00 3.30 1.00

2.77 1070.00 2.81 2.00

2.55 1086.00 2.60 3.00

3.66 1287.00 3.54 4.00

3.22 1130.00 3.17 5.00

Your tasks as a sociological researcher are to:

1. Compute the zero order correlations among all **three** variables.

* Construct scattergrams to be sure you have met the linearity assumption.
* Interpret these outcomes.

1. Compute the partial correlations between all combinations of variables.

* Interpret these outcomes.

1. Compute the multiple correlation between undergraduate GPA’s (Y), high school GPA’s (X), and SAT scores (Z).

* Interpret the outcome.

1. Compute the multiple regression equation:

* standardized values/equation =
* unstandardized values/equation =

What are the advantages and disadvantages of each?

Note: Formulae for multiple and partial correlations: I suggest using the ones in the Multivariate chapter of ***Data Analytics*** and lecture notes.

5. Using the regression equation, make predictions on the Y variable for Students 10 & 15.

Interpret what this means.

What do your *sociological* and *statistical imaginations* conclude from the empirical study you conducted?