**Study**: Subjects took two questionnaires. First, they filled out a survey rating different words on their meaningfulness or pleasantness. Scores were rated on a Likert scale from 1 (not meaningful, not pleasant) to 5 (very meaningful, very pleasant). These words were grouped into sets based on previous research, and the data set contains the averages for the words by set. Then they completed a meaning in life questionnaire (scores on questions were totaled).

**Important:** You will need to analyze BOTH questions below (aka you should have LOTS of boxes). However, you only need to do ONE write up, whichever you like better. You will pick mediation OR moderation for the second question. These questions use very different variables, so you will need to data screen each one separately (i.e. start over with the original data set for question 2).

**Hierarchical Multiple Linear Regression**

IV X-Variables:

* Control variables: Age, gender, priming type (1=meaningful, 2=pleasantness)
* Education words (accomplish, College, Degree, Education, Grades, Graduate, School, Teacher, Undergrad, University, educ avg)
* Goals words (achieve, ambition, become, goals, progress, success, goals avg)
* Nouns words (everything, know, lot, many, mind, much, right, some, something, thing, time, what, when, noun avg)
* Religion words (serve, glorify, religion avg)

DV – Y-variable:

* PIL total – scores on the purpose in life questionnaire

Hypothesis – we thought that word ratings would predict scores on the PIL questionnaire in some format. First control for demographic and experiment variables in step 1, then use the average of their word ratings (step 2) to predict the PIL total (DV).

1. Data screening:
   1. Accuracy – you can assume the data are accurate.
   2. Missing data:
      1. Include a box with a missing data line to show the data has/has no missing data.
      2. Fix/list what you did with the missing data if necessary.
   3. Outliers
      1. Mahalanobis
         1. What is the cut off score (*df* and *X2*)?
         2. How many were Mahalanobis outliers?
      2. Leverage
         1. What is the cut off score?
         2. How many leverage outliers did you have (note this is the total number of people just for leverage)?
      3. Cooks
         1. What is the cut off score?
         2. How many Cooks outliers did you have (note this is the total number of people just for Cooks)?
      4. How many overall outlier problems did you have with two or more issues?
      5. Delete all outliers with two or more problems.
   4. Multicollinearity
      1. Include a correlation table of the IVs.
      2. Do you have any multicollinearity issues?
   5. Normality
      1. Include the multivariate normality chart.
      2. Is the data normal?
   6. Linearity
      1. Include the PP Plot.
      2. Is the data linear?
   7. Homogeneity/Homoscedasticity
      1. Include the residuals graph.
      2. Is the data homogeneic?
      3. Is the data homoscedastic?
2. Regression test:
   1. Include the model box (make sure this has R2 change):
   2. Include the coefficients box (make sure this has pr and sr):
3. Chart:
   1. Include a regression chart for the y and y-hat values.
   2. Be sure to add the line of best fit.
   3. Be sure to clean up the X and Y axis (not the range, but the labels).
4. Write up (remember there is only one write up of the options: MLR, Mediation, or Moderation):
   1. Include a brief description of the experiment, variables, and order entered into steps.
   2. Include a brief section on the data screening/assumptions.
   3. Include the all F-values for each step of the model.
   4. Include all the b or beta values for variables *in the step they were entered*. So, you will not have double b values for any predictor.
      1. You can put either model values or b values in a table (so you don’t have a paragraph full of numbers).
      2. Be sure each model value (F, p, R2) and b value (t, p, pr2) has the appropriate statistics with it (even if not significant).

**Mediation or Moderation (don’t do both):**

Mediation: Use education average as a mediator variable between religion average and happiness totals (PIL). Does education ratings of importance mediate the relationship between religion importance and happiness ratings? See below.

Moderation: Use goals and religion averages to predict happiness and determine if their interaction predicts different levels of happiness. See below.

**Note**:

* On either these analyses, work the whole assignment (i.e normally you would stop if one of the steps or the interaction was not significant, but examples are hard to make up, so practice ☺).
* Use the PIL total as the happiness scores for the DV.

1. Data screening:
   1. Accuracy – you can assume the data are accurate.
   2. Missing data:
      1. Include a box with a missing data line to show the data has/has no missing data.
      2. Fix/list what you did with the missing data if necessary.
   3. Outliers
      1. Mahalanobis
         1. What is the cut off score (*df* and *X2*)?
         2. How many were Mahalanobis outliers?
      2. Leverage
         1. What is the cut off score?
         2. How many leverage outliers did you have (note this is the total number of people just for leverage)?
      3. Cooks
         1. What is the cut off score?
         2. How many Cooks outliers did you have (note this is the total number of people just for Cooks)?
      4. How many overall outlier problems did you have with two or more issues?
      5. Delete all outliers with two or more problems.
   4. Multicollinearity
      1. Include a correlation table of the IVs.
      2. Do you have any multicollinearity issues?
   5. Normality
      1. Include the multivariate normality chart.
      2. Is the data normal?
   6. Linearity
      1. Include the PP Plot.
      2. Is the data linear?
   7. Homogeneity/Homoscedasticity
      1. Include the residuals graph.
      2. Is the data homogeneic?
      3. Is the data homoscedastic?
2. Regression test:
   1. Include the output from PROCESS here.
   2. Mediation only:
      1. Is the c path significant? List b, t, and p values in APA style.
      2. Is the a path significant? List b, t, and p values in APA style.
      3. Is the b path significant? List b, t, and p values in APA style.
      4. Is the c’ path non-significant? List b, t, and p values in APA style.
      5. What is the effect size?
      6. What is the indirect effect?
      7. Does the Sobel test show if there was significant mediation? List Z and p values in APA style.
   3. Moderation only:
      1. Is the overall model significant? List the *F* value in APA style.
      2. Are the main effects significant? List the coefficient values in APA style.
      3. Is the interaction significant? List the coefficient value in APA style.
      4. List the simple slopes in APA style:
         1. Low
         2. Average
         3. High
      5. Using the Johnson-Neyman output where are the zones of significance?
      6. Create a line graph of the interaction.
         1. Make sure the x and y axis are appropriately labeled.
         2. Make sure the graph has three distinct lines.
   4. Include a write up of the results from your analysis.
3. Write up (remember there is only one write up of the options: MLR, Mediation, or Moderation):
   1. Mediation:
      1. Include a short description of variables/analysis.
      2. Include a brief section on the data screening/assumptions.
      3. Include all the path coefficients you listed above.
      4. Include the results of the indirect test.
      5. Include the Sobel test to determine if significant mediation occurred.
      6. Include a picture of the mediation (triangle diagram).
      7. Include a table of the model F values for each model test (should be 3).
   2. Moderation:
      1. Include a short description of the variables/analysis.
      2. Include a brief section on the data screening/assumptions.
      3. Include the model *F* value and effect size.
      4. Include the main effects and interaction coefficients from above.
      5. Include the figure created above.
      6. Explain the interaction by listing the simple slopes for each of the low, average, and high groups for your moderator. Are they all significant? What do they mean (explain like you would a correlation)?