





Chapter 10: Moderation, mediation and more regression

Smart Alex's Solutions

Task 1

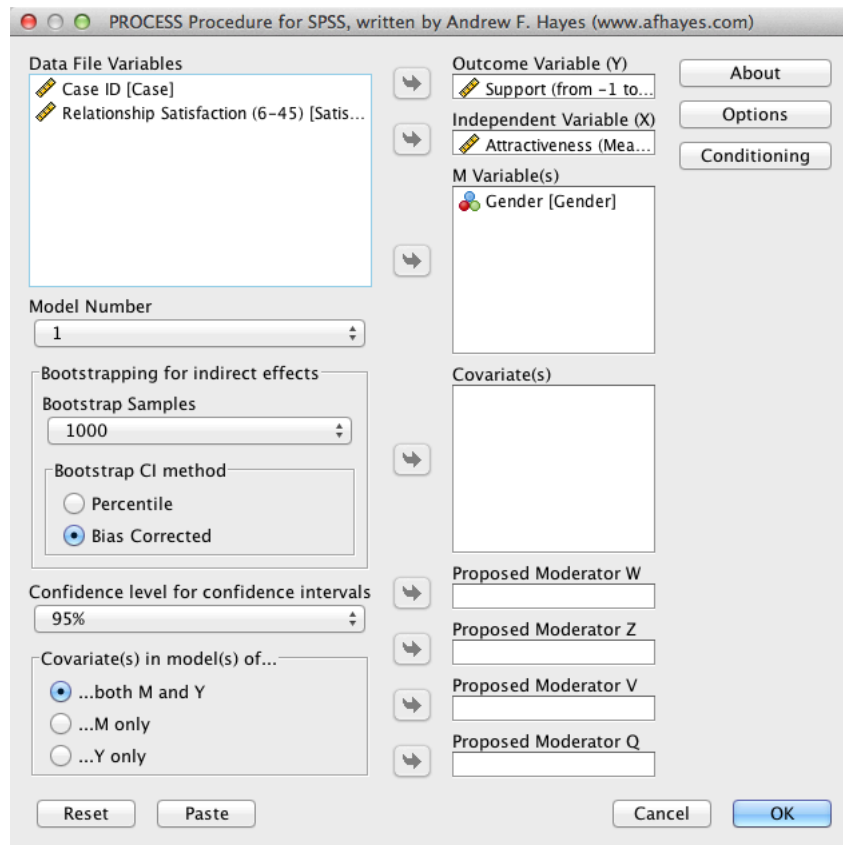
*McNulty et al. (2008) found a relationship between a person's **Attractiveness** and how much **Support** they give their partner in newlyweds. Is this relationship moderated by gender (i.e., whether the data were from the husband or wife)? The data are in **McNulty et al. (2008).sav**.*

To access the *PROCESS* dialog box in (see below) select **Analyze Regression**  **PROCESS, by Andrew F. Hayes (<http://www.afhayes.com>)**. The variables in your data file will be listed in the box labelled *Data File Variables*. Select the outcome variable (in this case **Support**) and drag it to the box labelled *Outcome Variable (Y)*, or click on . Similarly, select the predictor variable (in this case **Attractiveness**) and drag it to the box labelled *Independent Variable (X)*. Finally, select the moderator variable (in this case **Gender**) and drag it to the box labelled *M Variable(s)*, or click on . This box is where you specify any moderators (you can have more than one).

PROCESS can test 74 different types of model, and these models are listed in the drop-down box labelled *Model Number*. Simple moderation analysis is represented by model 1, but the default model is 4 (mediation, which we'll look at next). Therefore, activate this drop-down list and select **1** . The rest of the options in this dialog box are for models other than simple moderation, so we'll ignore them.

If you click on **Options...** another dialog box will appear containing four useful options for moderation. Selecting (1) *Mean center for products* centres the predictor and moderator for you; (2) *Heteroscedasticity-consistent SEs* means we need not worry about having heteroscedasticity in the model; (3) *OLS/ML confidence intervals* produces confidence intervals for the model, and I've tried to emphasize the importance of these throughout the book; and (4) *Generate data for plotting* is helpful for interpreting and visualizing the simple slopes analysis. Talking of simple slopes analysis, if you click on **Conditioning**, you can change whether you

want simple slopes at ± 1 standard deviation of the mean of the moderator (the default, which is fine) or at percentile points (it uses the 10th, 25th, 50th, 75th and 90th percentiles). Back in the main dialog box, click **OK** to run the analysis.



```

*****
Model = 1
  Y = Support
  X = Attracti
  M = Gender

Sample size
  164

*****
Outcome: Support

Model Summary
      R      R-sq      F      df1      df2      p
      .2921   .0853   4.8729   3.0000  160.0000   .0029

Model
      coeff      se      t      p      LLCI      ULCI
constant  .2335   .0161  14.5416   .0000   .2018   .2652
Gender    .0239   .0321   .7440   .4580  -.0395   .0873
    
```

| | | | | | | |
|----------|--------|-------|--------|-------|--------|-------|
| Attracti | -.0072 | .0148 | -.4852 | .6282 | -.0363 | .0220 |
| int_1 | .1053 | .0295 | 3.5677 | .0005 | .0470 | .1636 |

Interactions:

| | | | |
|-------|----------|---|--------|
| int_1 | Attracti | X | Gender |
|-------|----------|---|--------|

Output 1

Output 1 is the main moderation analysis. We're told the b -value for each predictor, the associated standard errors (which have been adjusted for heteroscedasticity because we asked for them to be). Each b is compared to zero using a t -test, which is computed from the beta divided by its standard error. The confidence interval for the b is also produced (because we asked for it). Moderation is shown up by a significant interaction effect, and in this case the interaction is highly significant, $b = 0.105$, 95% CI [0.047, 0.164], $t = 3.57$, $p < .01$, indicating that the relationship between attractiveness and support is moderated by gender.

Conditional effect of X on Y at values of the moderator(s)

| Gender | Effect | se | t | p | LLCI | ULCI |
|--------|--------|-------|---------|-------|--------|--------|
| -.5000 | -.0598 | .0202 | -2.9549 | .0036 | -.0998 | -.0198 |
| .5000 | .0455 | .0215 | 2.1177 | .0357 | .0031 | .0879 |

Values for quantitative moderators are the mean and plus/minus one SD from mean

Output 2

To interpret the moderation effect we can examine the simple slopes, which are shown in Output . Essentially, the table shows us the results of two different regressions: the regression for attractiveness as a predictor of support (1) when the value for gender is $-.5$ (i.e., low). Because husbands were coded as zero, this represents the value for males; and (2) when the value for gender is $.5$ (i.e., high). Because wives were coded as 1, this represents the female end of the gender spectrum. We can interpret these three regressions as we would any other: we're interested the value of b (called *Effect* in the output), and its significance. From what we have already learnt about regression we can interpret the two models as follows:

1. When gender is low (male), there is a significant negative relationship between attractiveness and support, $b = -0.060$, 95% CI $[-0.100, -0.020]$, $t = -2.95$, $p < .01$.
2. When gender is high (female), there is a significant positive relationship between attractiveness and support, $b = 0.05$, 95% CI $[0.003, 0.088]$, $t = 2.12$, $p < .05$.

These results tell us that the relationship between attractiveness of a person and amount of support given to their spouse is different for men and women. Specifically, for women, as attractiveness increases the level of support that they give to their husbands increases,

whereas for men, as attractiveness increases the amount of support they give to their wives decreases.

Reporting the results

We can report the results in a table as follows:

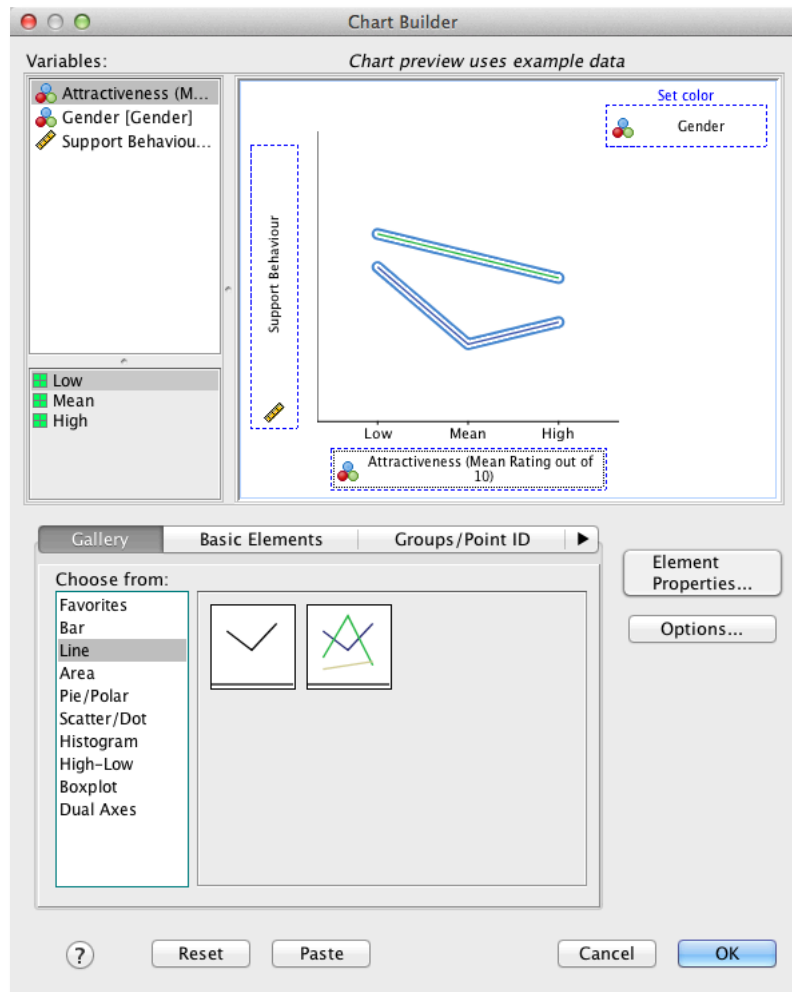
| | <i>b</i> | <i>SE B</i> | <i>t</i> | <i>p</i> |
|--------------------------|------------------------|-------------|----------|------------|
| Constant | 0.23 [0.20, 0.27] | 0.026 | 14.54 | $p < .001$ |
| Gender (centred) | 0.02 [-0.04, 0.09] | 0.032 | 0.74 | $p = .458$ |
| Attractiveness (centred) | -0.01 [-0.04, 0.02] | 0.015 | -0.49 | $p = .628$ |
| Gender × Attractiveness | 0.11 [0.05, 0.16] | 0.030 | 3.57 | $p < .01$ |

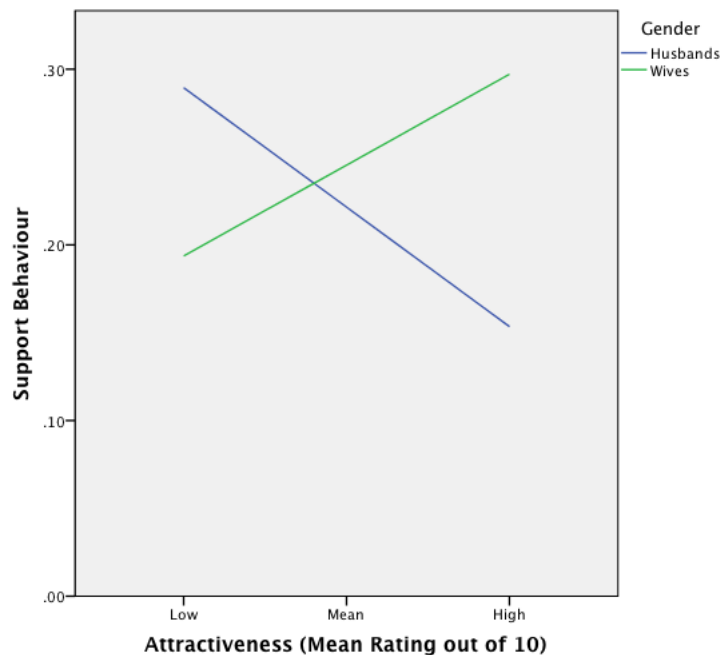
Note. $R^2 = .09$.

Task 2

Produce the simple slopes graphs for the above example.

| | Attractiveness | Gender | Support | var | var |
|---|----------------|--------|---------|-----|-----|
| 1 | Low | Low | .29 | | |
| 2 | Mean | Low | .22 | | |
| 3 | High | Low | .15 | | |
| 4 | Low | High | .19 | | |
| 5 | Mean | High | .25 | | |
| 6 | High | High | .30 | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |





Output 3



The graph in Output 3 above confirms our results from the simple slopes analysis in the previous task (Output 2). You can see that the direction of the relationship between attractiveness and support is different for men and women by the fact that the two regression lines slope in different directions. Specifically, for husbands (blue line) the relationship is negative (the regression line slopes downwards), whereas for wives (green line) the relationship is positive (the regression line slopes upwards). Additionally, the fact that the lines cross indicates a significant interaction effect (moderation). So basically, we can conclude that the relationship between attractiveness and support is positive for wives (more attractive wives give their husbands more support), but negative for husbands (more attractive husbands give their wives less support than unattractive ones). Although they didn't test moderation, this mimics the findings of McNulty et al. (2008).


Task 3




*McNulty et al. (2008) also found a relationship between a person's **Attractiveness** and their relationship **Satisfaction** as newlyweds. Using the same data as the previous examples, is this relationship moderated by gender?*

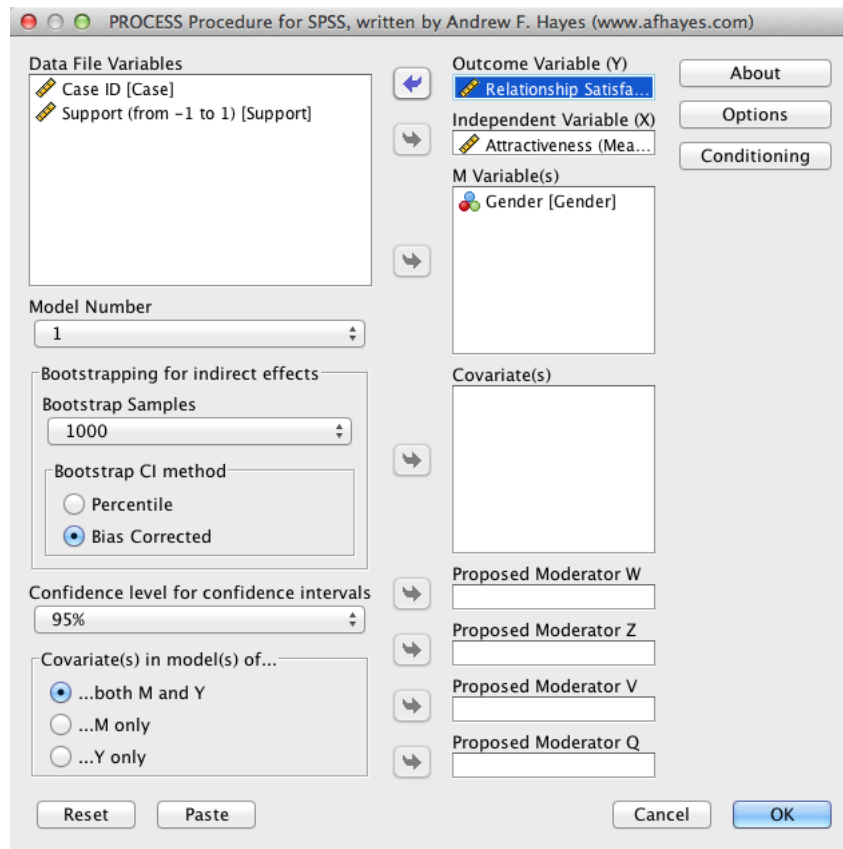
To access the dialog box shown below select **Analyze > Regression**

PROCESS, by Andrew F. Hayes (<http://www.afhayes.com>). The variables in your data file will be listed in

the box labelled *Data File Variables*. Select the outcome variable (in this case **Relationship Satisfaction**) and drag it to the box labelled *Outcome Variable (Y)*, or click on . Similarly, select the predictor variable (in this case **Attractiveness**) and drag it to the box labelled *Independent Variable (X)*. Finally, select the moderator variable (in this case **Gender**) and drag it to the box labelled *M Variable(s)*, or click on . This box is where you specify any moderators (you can have more than one).

PROCESS can test 74 different types of model, and these models are listed in the drop-down box labelled *Model Number*. Simple moderation analysis is represented by model 1, but the default model is 4 (mediation, which we'll look at next). Therefore, activate this drop-down list and select . The rest of the options in this dialog box are for models other than simple moderation, so we'll ignore them.

If you click on  another dialog box will appear containing four useful options for moderation. Selecting (1) *Mean center for products* centres the predictor and moderator for you; (2) *Heteroscedasticity-consistent SEs* means we need not worry about having heteroscedasticity in the model; (3) *OLS/ML confidence intervals* produces confidence intervals for the model, and I've tried to emphasize the importance of these throughout the book; and (4) *Generate data for plotting* is helpful for interpreting and visualizing the simple slopes analysis. Talking of simple slopes analysis, if you click on , you can change whether you want simple slopes at ± 1 standard deviation of the mean of the moderator (the default, which is fine) or at percentile points (it uses the 10th, 25th, 50th, 75th and 90th percentiles). Back in the main dialog box, click  to run the analysis.



```

*****
Model = 1
  Y = Satisfac
  X = Attracti
  M = Gender

Sample size
  164

*****
Outcome: Satisfac

Model Summary
      R      R-sq      F      df1      df2      p
      .1679    .0282    1.9911    3.0000    160.0000    .1175

Model
      coeff      se      t      p      LLCI      ULCI
constant    33.7160    .3512    96.0087    .0000    33.0224    34.4095
Gender      -.0236    .7024    -.0336    .9733    -1.4107    1.3635
Attracti    -.6104    .2887    -2.1145    .0360    -1.1805    -.0403
int_1       .5467    .5773    .9470    .3451    -.5935    1.6869

Interactions:

int_1    Attracti    X    Gender
*****

```


| Conditional effect of X on Y at values of the moderator(s) | | | | | | |
|--|--------|-------|---------|-------|---------|--------|
| Gender | Effect | se | t | p | LLCI | ULCI |
| -.5000 | -.8837 | .3827 | -2.3090 | .0222 | -1.6396 | -.1279 |
| .5000 | -.3370 | .4322 | -.7797 | .4367 | -1.1906 | .5166 |

Values for quantitative moderators are the mean and plus/minus one SD from mean

Output 4

Output is the results of the moderation analysis. We're told the b -value for each predictor, the associated standard errors (which have been adjusted for heteroscedasticity because we asked for them to be). Each b is compared to zero using a t -test, which is computed from the beta divided by its standard error. The confidence interval for the b is also produced (because we asked for it). Moderation is shown up by a significant interaction effect; however, in this case the interaction is non-significant, $b = 0.547$, 95% CI [-0.594, 1.687], $t = 0.95$, $p > .05$, indicating that the relationship between attractiveness and relationship satisfaction is not moderated by gender.

Task 4

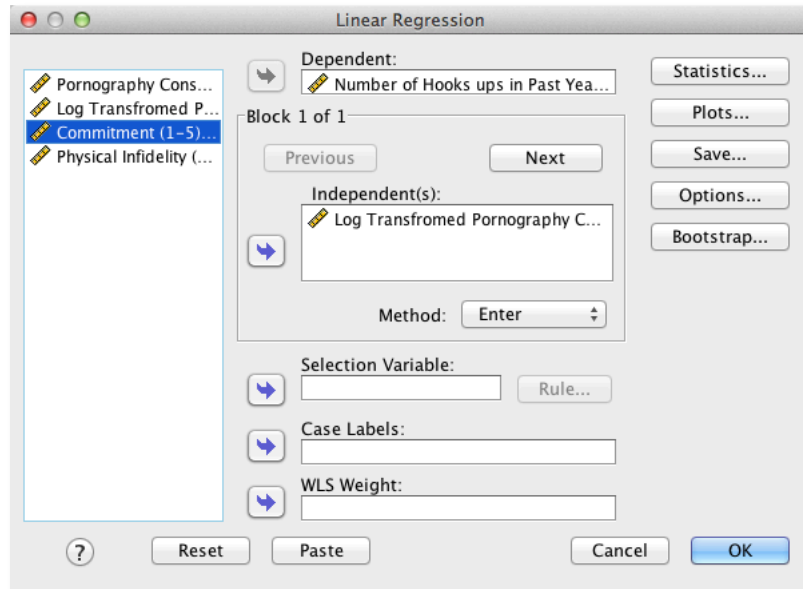
*In the chapter we tested a mediation model of infidelity for Lambert et al.'s data using Baron and Kenny's regressions. Repeat this analysis but using **Hook_Ups** as the measure of infidelity.*

Baron and Kenny suggested that mediation is tested through three regression models:

1. A regression predicting the outcome (**Hook_Ups**) from the predictor variable (**Consumption**).
2. A regression predicting the mediator (**Commitment**) from the predictor variable (**Consumption**).
3. A regression predicting the outcome (**Hook_Ups**) from both the predictor variable (**Consumption**) and the mediator (**Commitment**).

These models test the four conditions of mediation: (1) the predictor variable (**Consumption**) must significantly predict the outcome variable (**Hook_Ups**) in model 1; (2) the predictor variable (**Consumption**) must significantly predict the mediator (**Commitment**) in model 2; (3) the mediator (**Commitment**) must significantly predict the outcome (**Hook_Ups**) variable in model 3; and (4) the predictor variable (**Consumption**) must predict the outcome variable (**Hook_Ups**) less strongly in model 3 than in model 1.

Model 1: Predicting Infidelity from Consumption



Dialog box for model 1

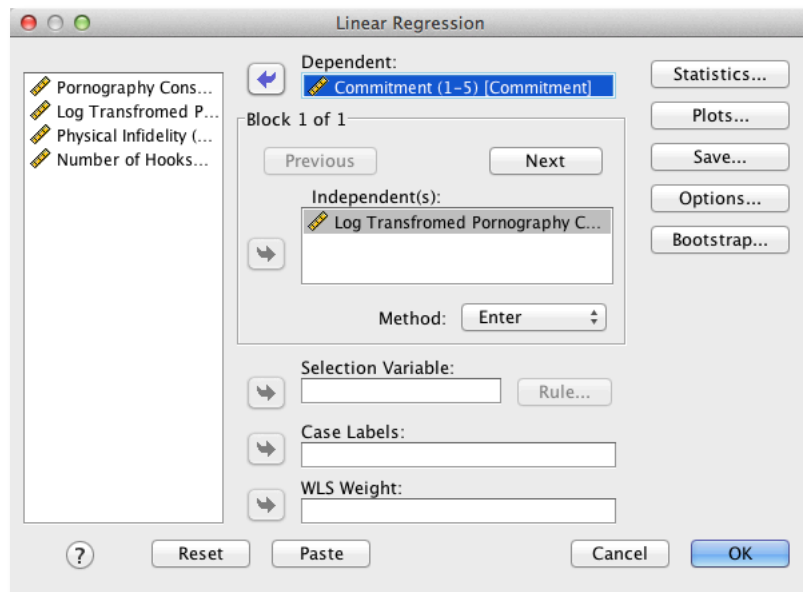
Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | 95.0% Confidence Interval for B | |
|-------|---|-----------------------------|------------|---------------------------|--------|------|---------------------------------|-------------|
| | | B | Std. Error | Beta | | | Lower Bound | Upper Bound |
| 1 | (Constant) | 1.626 | .111 | | 14.667 | .000 | 1.408 | 1.845 |
| | Log Transformed Pornography Consumption | 1.580 | .435 | .229 | 3.636 | .000 | .724 | 2.436 |

a. Dependent Variable: Number of Hooks ups in Past Year

Output 5

Model 2: Predicting Commitment from Consumption



Dialog box for model 2

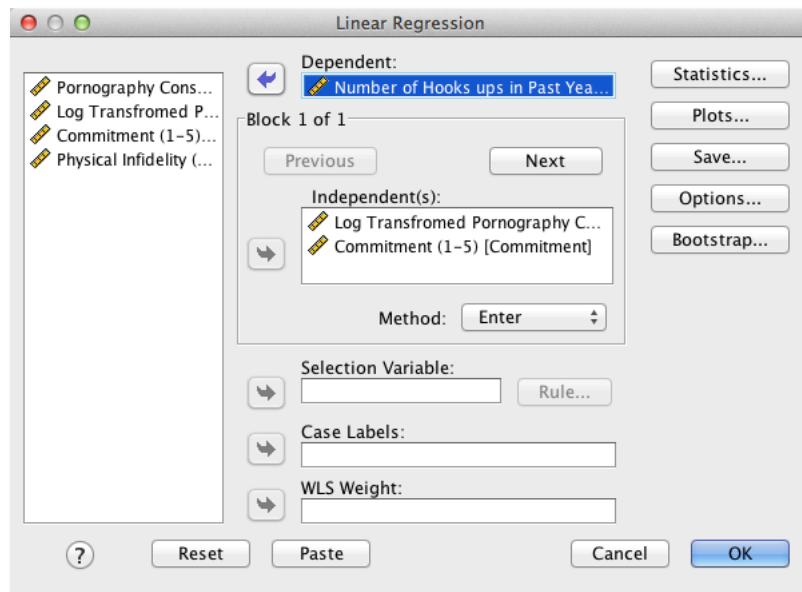
Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | 95.0% Confidence Interval for B | |
|-------|---|-----------------------------|------------|---------------------------|--------|------|---------------------------------|-------------|
| | | B | Std. Error | Beta | | | Lower Bound | Upper Bound |
| 1 | (Constant) | 4.203 | .054 | | 77.178 | .000 | 4.095 | 4.310 |
| | Log Transformed Pornography Consumption | -.470 | .213 | -.142 | -2.205 | .028 | -.889 | -.050 |

a. Dependent Variable: Commitment (1-5)

Output 6

Model 3: Predicting Infidelity from Consumption and Commitment



Dialog box for model 3

Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | 95.0% Confidence Interval for B | |
|-------|---|-----------------------------|------------|---------------------------|--------|------|---------------------------------|-------------|
| | | B | Std. Error | Beta | | | Lower Bound | Upper Bound |
| 1 | (Constant) | 4.243 | .544 | | 7.806 | .000 | 3.172 | 5.313 |
| | Log Transformed Pornography Consumption | 1.281 | .420 | .186 | 3.050 | .003 | .454 | 2.109 |
| | Commitment (1-5) | -.622 | .127 | -.299 | -4.903 | .000 | -.872 | -.372 |

a. Dependent Variable: Number of Hooks ups in Past Year

Output 7

Is there evidence for mediation?



- Output shows that pornography consumption significantly predicts hook-ups, $b = 1.58$, 95% CI [0.72, 2.45], $t = 3.64$, $p < .001$. As pornography consumption increases, the number of hook-ups increases also.
- Output shows that pornography consumption significantly predicts relationship commitment, $b = -0.47$, 95% CI [-0.89, -0.05], $t = -2.21$, $p = .028$. As pornography consumption increases commitment declines.
- Output 7 shows that relationship commitment significantly predicts hook-ups, $b = -0.62$, 95% CI [-0.87, -0.37], $t = -4.90$, $p < .001$. As relationship commitment increases the number of hook-ups decreases.


- The relationship between pornography consumption and infidelity is stronger in model 1, $b = 1.58$ (Output) than in model 3, $b = 1.28$ (Output).




As such, the four conditions of mediation have been met.

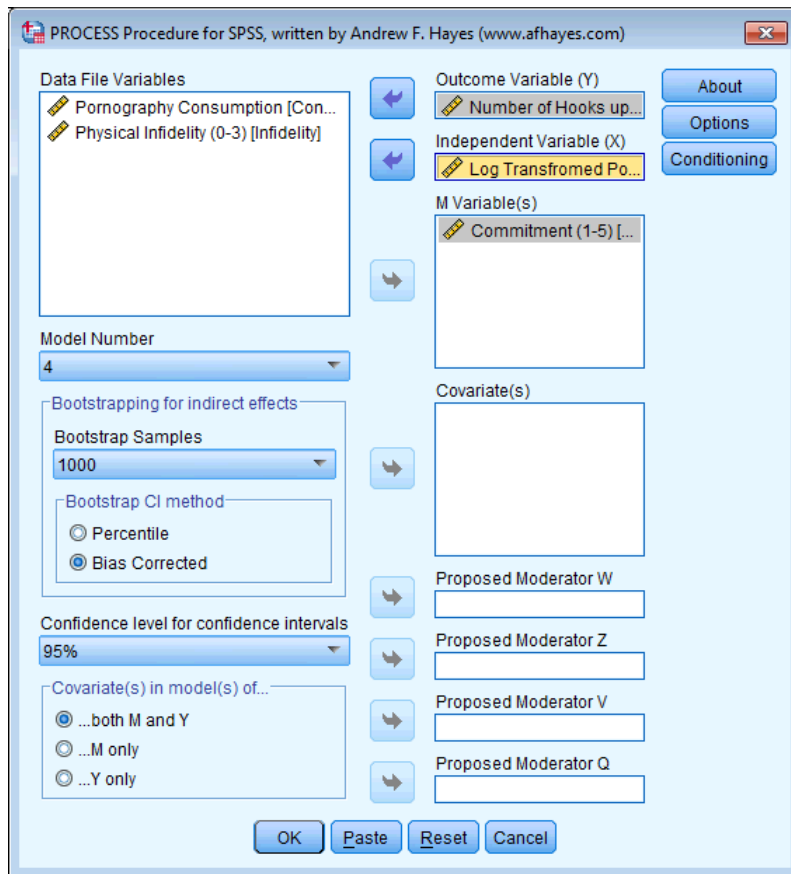
Task 5

Repeat the above analysis but using the PROCESS tool to estimate the indirect effect and its confidence interval.

To access the dialog box shown below select **Analyze > Regression > PROCESS**, by Andrew F. Hayes (<http://www.afhayes.com>). The variables in your data file will be listed in the box labelled *Data File Variables*. Select the outcome variable (in this case **Hook_Ups**) and drag it to the box labelled *Outcome Variable (Y)*, or click on . Similarly, select the predictor variable (in this case **LnConsumption**) and drag it to the box labelled *Independent Variable (X)*. Finally, select the mediator variable (in this case **Commitment**) and drag it to the box labelled *M Variable(s)*, or click on . This box is where you specify any mediators (you can have more than one).

As I mentioned before, *PROCESS* can test many different types of model, and simple mediation analysis is represented by model 4 (this model is selected by default). Therefore, make sure that  is selected in the drop-down list under *Model Number*. Unlike moderation, there are other options in this dialog box that are useful: for example, to test the indirect effects we will use bootstrapping to generate a confidence interval around the indirect effect. By default *PROCESS* uses 1000 bootstrap samples, and will compute bias corrected and accelerated confidence intervals. These default options are fine, but just be aware that you can ask for percentile bootstrap confidence intervals instead.

If you click on  another dialog box will appear containing four useful options for mediation. Selecting (1) *Effect Size* produces the estimates of the size of the indirect effect; (2) *Sobel Test* produces a significance test of the indirect effect devised by Sobel; (3) *Total effect model* produces the direct effect of the predictor on the outcome (in this case the regression of infidelity predicted from pornography consumption). None of the options activated by clicking on  apply to simple mediation models, so we can ignore this button and click  to run the analysis.



```
*****
Model = 4
Y = Hook_Ups
X = LnConsum
M = Commitme
```

```
Sample size
    239
```

```
*****
Outcome: Commitme
```

```
Model Summary
      R      R-sq      F      df1      df2      p
    .1418    .0201    4.8633    1.0000    237.0000    .0284
```

```
Model
      coeff      se      t      p
constant  4.2027    .0545   77.1777   .0000
LnConsum  -.4697    .2130   -2.2053   .0284
```

Output 8

Output shows the first part of the output, which initially tells us the name of the outcome (Y), the predictor (X) and the mediator (M) variables, which have been shortened to eight letters. This is useful for double-checking we have entered the variables in the correct place: the outcome is number of hook-ups, the predictor is consumption, and the mediator is commitment. The next part of the output shows us the results of the simple regression of commitment predicted from pornography consumption. This output is interpreted just as we would interpret any regression: we can see that pornography consumption significantly predicts relationship commitment, $b = -0.47$, $t = -2.21$, $p = .028$. The R^2 value tells us that pornography consumption explains 2% of the variance in relationship commitment, and the fact that the b is negative tells us that the relationship is negative also: as consumption increases, commitment declines (and vice versa).

```
*****
Outcome: Hook_Ups

Model Summary
      R      R-sq      F      df1      df2      p
      .3739      .1398      19.1747      2.0000      236.0000      .0000

Model
      coeff      se      t      p
constant      4.2427      .5435      7.8062      .0000
Commitme      -.6218      .1268      -4.9027      .0000
LnConsum      1.2811      .4201      3.0498      .0026

Output 9
```

Output shows the results of the regression of number of hook-ups predicted from both pornography consumption and commitment. We can see that pornography consumption significantly predicts number of hook-ups even with relationship commitment in the model, $b = 1.28$, $t = 3.05$, $p = .003$; relationship commitment also significantly predicts number of hook-ups, $b = -0.62$, $t = -4.90$, $p < .001$. The R^2 value tells us that the model explains 14.0% of the variance in number of hook-ups. The negative b for commitment tells us that as commitment increases, number of hook-ups declines (and vice versa), but the positive b for consumptions indicates that as pornography consumption increases, the number of hook-ups increases also. These relationships are in the predicted direction.

```
***** TOTAL EFFECT MODEL *****
Outcome: Hook_Ups

Model Summary
      R      R-sq      F      df1      df2      p
      .2284      .0522      13.0449      1.0000      237.0000      .0004

Model
      coeff      se      t      p
constant      1.6296      .1114      14.6324      .0000
LnConsum      1.5731      .4356      3.6118      .0004

Output 10
```

Output 10 shows the total effect of pornography consumption on number of hook-ups (outcome). You will get this bit of the output only if you selected *Total effect model*. The total effect is the effect of the predictor on the outcome when the mediator is not present in the model. When relationship commitment is not in the model, pornography consumption significantly predicts the number of hook-ups, $b = 1.57$, $t = 3.61$, $p < .001$. The R^2 value tells us that the model explains 5.22% of the variance in number of hook-ups. As is the case when we include relationship commitment in the model, pornography consumption has a positive relationship with number of hook-ups (as shown by the positive b -value).

***** TOTAL, DIRECT, AND INDIRECT EFFECTS *****

| Total effect of X on Y | | | | |
|---|--------|---------|----------|----------|
| | Effect | SE | t | p |
| | 1.5731 | .4356 | 3.6118 | .0004 |
| Direct effect of X on Y | | | | |
| | Effect | SE | t | p |
| | 1.2811 | .4201 | 3.0498 | .0026 |
| Indirect effect of X on Y | | | | |
| | Effect | Boot SE | BootLLCI | BootULCI |
| Commitme | .2920 | .1572 | .0310 | .6591 |
| Partially standardized indirect effect of X on Y | | | | |
| | Effect | Boot SE | BootLLCI | BootULCI |
| Commitme | .1904 | .0962 | .0189 | .3894 |
| Completely standardized indirect effect of X on Y | | | | |
| | Effect | Boot SE | BootLLCI | BootULCI |
| Commitme | .0424 | .0215 | .0043 | .0889 |
| Ratio of indirect to total effect of X on Y | | | | |
| | Effect | Boot SE | BootLLCI | BootULCI |
| Commitme | .1856 | .2086 | .0128 | .5298 |
| Ratio of indirect to direct effect of X on Y | | | | |
| | Effect | Boot SE | BootLLCI | BootULCI |
| Commitme | .2280 | 1.0581 | .0052 | 1.0776 |
| R-squared mediation effect size (R-sq_med) | | | | |
| | Effect | Boot SE | BootLLCI | BootULCI |
| Commitme | .0183 | .0123 | .0014 | .0548 |
| Preacher and Kelley (2011) Kappa-squared | | | | |
| | Effect | Boot SE | BootLLCI | BootULCI |
| Commitme | .0433 | .0210 | .0059 | .0904 |

Output 11

Output 11 is the most important part of the output because it displays the results for the indirect effect of pornography consumption on number of hook-ups (i.e. the effect via relationship commitment). First, we're told the effect of pornography consumption on the number of hook-ups in isolation (the total effect). Next, we're told the effect of pornography

consumption on the number of hook-ups when relationship commitment is included as a predictor as well (the direct effect). The first bit of new information is the *Indirect Effect of X on Y*, which in this case is the indirect effect of pornography consumption on the number of hook-ups. We're given an estimate of this effect ($b = 0.292$) as well as a bootstrapped standard error and confidence interval. As we have seen many times before, 95% confidence intervals contain the true value of a parameter in 95% of samples. Therefore, we tend to assume that our sample isn't one of the 5% that does not contain the true value and use them to infer the population value of an effect. In this case, assuming our sample is one of the 95% that 'hits' the true value, we know that the true b -value for the indirect effect falls between 0.031 and 0.659. This range does not include zero, and remember that $b = 0$ would mean 'no effect whatsoever'; therefore, the fact that the confidence interval does not contain zero means that there is likely to be a genuine indirect effect. Put another way, relationship commitment is a mediator of the relationship between pornography consumption and the number of hook-ups.

The rest of Output you will see only if you selected *Effect size*; it contains various standardized forms of the indirect effect. In each case they are accompanied by a bootstrapped confidence interval. As with the unstandardized indirect effect, if the confidence intervals don't contain zero then we can be confident that the true effect size is different from 'no effect'. In other words, there is mediation. All of the effect size measures have confidence intervals that don't include zero, so whichever one we look at we can be fairly confident that the indirect effect is greater than 'no effect'. Focusing on the most useful of these effect sizes, the standardized b for the indirect effect, its value is $b = .042$, 95% BCa CI [.004, .089], and similarly, $\kappa^2 = .043$, 95% BCa CI [.006, .090]. κ^2 is bounded to fall between 0 and 1, so we can interpret this as the indirect effect being about 4.3% of the maximum value that it could have been, which is a fairly small effect. We might, therefore, want to look for other potential mediators to include in the model in addition to relationship commitment.

```
Normal theory tests for indirect effect
  Effect      se      Z      p
   .2920    .1477    1.9773   .0480
```

Output 12

Output shows the results of the Sobel test. As I have mentioned before, it is better to interpret the bootstrap confidence intervals than formal tests of significance; however, if you selected *Sobel test* this is what you will see. Again, we're given the size of the indirect effect ($b = 0.292$), the standard error, associated z-score ($z = 1.98$) and p -value ($p = .048$). The p -value is just under the not-at-all magic .05 threshold so we'd conclude that there is a significant indirect effect, so there is compelling evidence that there is a small, but meaningful mediation effect.

Reporting the Results of the Mediation Analysis

Some people report only the indirect effect in mediation analysis, and possibly the Sobel test. However, I have repeatedly favoured using bootstrap confidence intervals, so you should report these, and preferably the effect size κ^2 and its confidence interval:

- ✓ There was a significant indirect effect of pornography consumption on the number of hook-ups through relationship commitment, $b = 0.292$, BCa CI [0.031, 0.659]. This represents a relatively small effect, $\kappa^2 = .043$, 95% BCa CI [.006, .090].

Task 6

*In Chapter 3 (Task 5) we looked at data from people who had been forced to marry goats and dogs and measured their life satisfaction and, also, how much they like animals (**Goat or Dog.sav**). Run a regression predicting life satisfaction from the type of animal to which a person was married. Write out the final model.*

Coefficients^a

| Model | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---------------------|-----------------------------|------------|---------------------------|-------|------|
| | B | Std. Error | Beta | | |
| 1 (Constant) | 16.208 | 9.452 | | 1.715 | .104 |
| Type of Animal Wife | 21.958 | 6.372 | .630 | 3.446 | .003 |

a. Dependent Variable: Life Satisfaction (%)

Output 13

Looking at the coefficients above, we can see that type of animal wife significantly predicted life satisfaction because the p -value is less than .05 (.003). The positive standardized beta value (.630) indicates a positive relationship between type of animal wife and life satisfaction. Remember that **goat** was coded as 0 and **dog** was coded as 1, therefore as type of animal wife increased from goat to dog, life satisfaction also increased. In other words, men who were married to dogs were more satisfied than those who were married to goats. By replacing the b -values from Output 13 in the equation from the book chapter, we can define our specific model as:

$$\text{Life satisfaction}_i = b_0 + b_1 \text{ type of animal wife}_i = 16.21 + (21.96 \text{ type of animal wife}_i)$$

Task 7

Repeat the analysis above but include animal liking in the first block, and type of animal wife in the second block. Do your conclusions about the relationship between type of animal and life satisfaction change?

Model Summary

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|-------------------|----------|-------------------|----------------------------|
| 1 | .512 ^a | .262 | .221 | 15.447 |
| 2 | .791 ^b | .625 | .581 | 11.332 |

a. Predictors: (Constant), Love of Animals

b. Predictors: (Constant), Love of Animals, Type of Animal Wife

ANOVA^c

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|-------|------------|----------------|----|-------------|--------|-------------------|
| 1 | Regression | 1527.711 | 1 | 1527.711 | 6.402 | .021 ^a |
| | Residual | 4295.239 | 18 | 238.624 | | |
| | Total | 5822.950 | 19 | | | |
| 2 | Regression | 3639.810 | 2 | 1819.905 | 14.172 | .000 ^b |
| | Residual | 2183.140 | 17 | 128.420 | | |
| | Total | 5822.950 | 19 | | | |

a. Predictors: (Constant), Love of Animals

b. Predictors: (Constant), Love of Animals, Type of Animal Wife

c. Dependent Variable: Life Satisfaction (%)

Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|-------|---------------------|-----------------------------|------------|---------------------------|-------|------|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 25.940 | 8.993 | | 2.884 | .010 |
| | Love of Animals | .580 | .229 | .512 | | |
| 2 | (Constant) | -2.067 | 9.551 | | -.216 | .831 |
| | Love of Animals | .541 | .169 | .478 | | |
| | Type of Animal Wife | 21.011 | 5.181 | .603 | | |

a. Dependent Variable: Life Satisfaction (%)

Output 14

Looking at the coefficients from the final model in Output 14 (model 2), we can see that both **Love of animals**, $t(17) = 3.21$, $p < .01$, and **Type of animal wife**, $t(17) = 4.06$, $p < .01$, significantly predicted life satisfaction. This means that type of animal wife predicted life satisfaction over and above the effect of love of animals. In other words, even after controlling for the effect of love of animals, type of animal wife still significantly predicted life satisfaction.

R^2 is the squared correlation between the observed values of life satisfaction and the values of life satisfaction predicted by the model. The values in this output tell us that love of animals explains 26.2% of the variance in life satisfaction. When type of animal wife is factored in as well, 62.5% of variance in life satisfaction is explained (i.e., an additional 36.3%).

Task 8

Using the *GlastonburyDummy.sav* data, which you should've already analysed, comment on whether you think the model is reliable and generalizable.

Model Summary^b

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | | Durbin-Watson |
|-------|-------------------|----------|-------------------|----------------------------|-------------------|----------|-----|-----|---------------|---------------|
| | | | | | R Square Change | F Change | df1 | df2 | Sig. F Change | |
| 1 | .276 ^a | .076 | .053 | .68818 | .076 | 3.270 | 3 | 119 | .024 | 1.893 |

a. Predictors: (Constant), No Affiliation vs. Indie Kid, No Affiliation vs. Crusty, No Affiliation vs. Metallor

b. Dependent Variable: Change in Hygiene Over The Festival

Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | 95% Confidence Interval for B | | Collinearity Statistics | |
|-------|------------------------------|-----------------------------|------------|---------------------------|--------|------|-------------------------------|-------------|-------------------------|-------|
| | | B | Std. Error | Beta | | | Lower Bound | Upper Bound | Tolerance | VIF |
| 1 | (Constant) | -.554 | .090 | | -6.134 | .000 | -.733 | -.375 | | |
| | No Affiliation vs. Crusty | -.412 | .167 | -.232 | -2.464 | .015 | -.742 | -.081 | .879 | 1.138 |
| | No Affiliation vs. Metallor | .028 | .160 | .017 | .177 | .860 | -.289 | .346 | .874 | 1.144 |
| | No Affiliation vs. Indie Kid | -.410 | .205 | -.185 | -2.001 | .048 | -.816 | -.004 | .909 | 1.100 |

a. Dependent Variable: Change in Hygiene Over The Festival

Collinearity Diagnostics^a

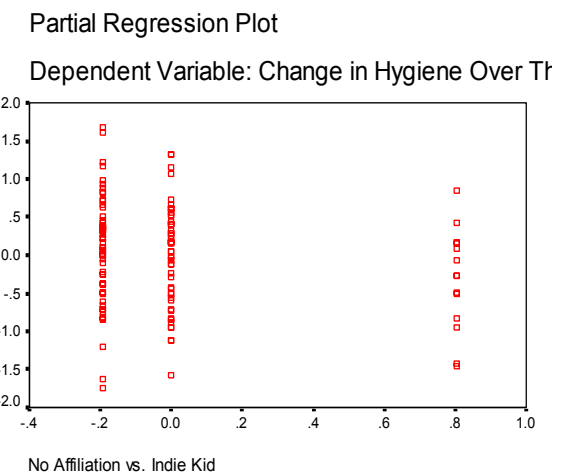
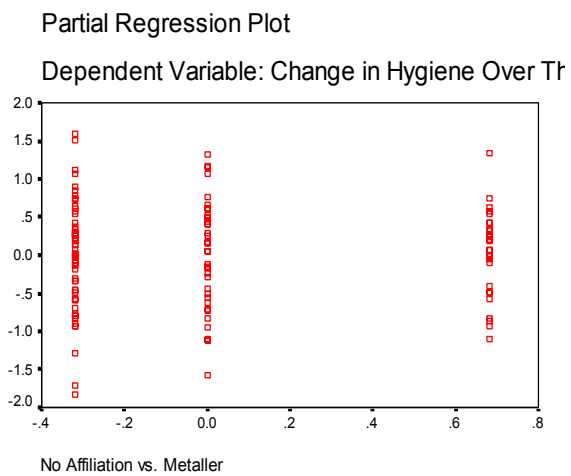
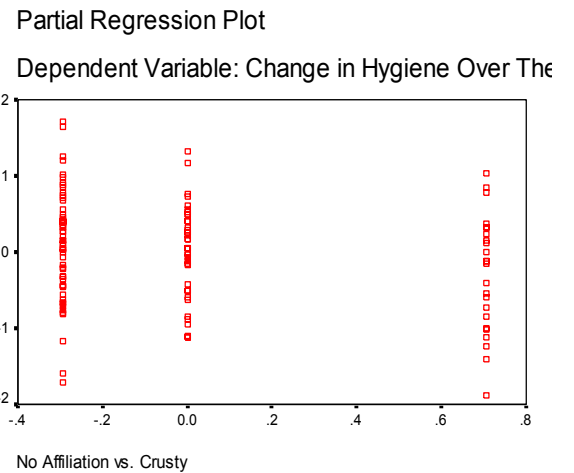
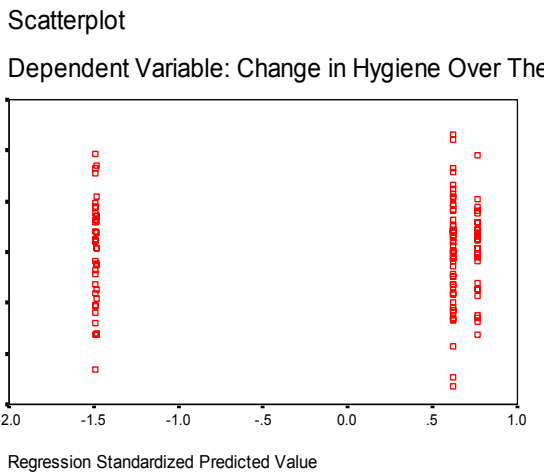
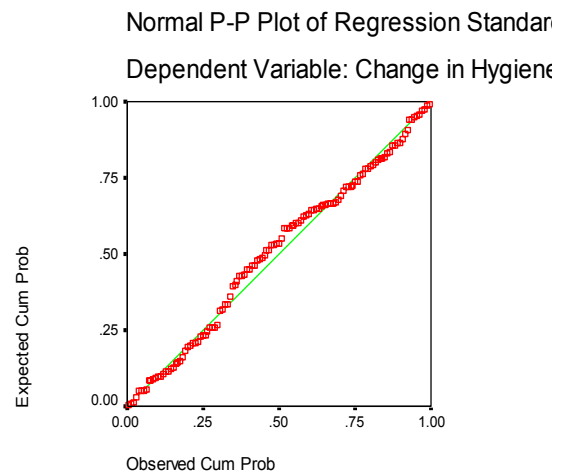
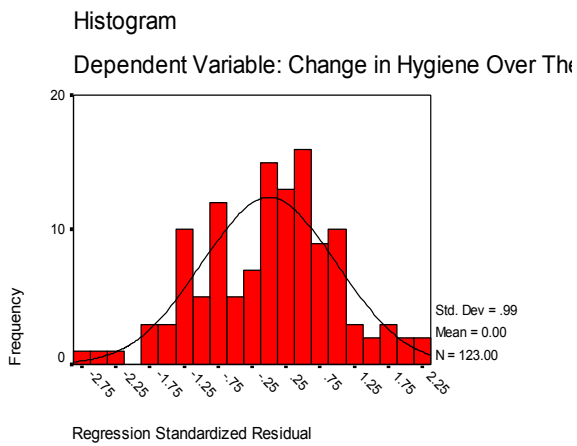
| Model | Dimension | Eigenvalue | Condition Index | Variance Proportions | | | |
|-------|-----------|------------|-----------------|----------------------|---------------------------|-----------------------------|------------------------------|
| | | | | (Constant) | No Affiliation vs. Crusty | No Affiliation vs. Metallor | No Affiliation vs. Indie Kid |
| 1 | 1 | 1.727 | 1.000 | .14 | .08 | .08 | .05 |
| | 2 | 1.000 | 1.314 | .00 | .37 | .32 | .00 |
| | 3 | 1.000 | 1.314 | .00 | .07 | .08 | .63 |
| | 4 | .273 | 2.515 | .86 | .48 | .52 | .32 |

a. Dependent Variable: Change in Hygiene Over The Festival

Casewise Diagnostics^a

| Case Number | Std. Residual | Change in Hygiene Over The Festival | Predicted Value | Residual |
|-------------|---------------|-------------------------------------|-----------------|----------|
| 31 | -2.302 | -2.55 | -.9658 | -1.5842 |
| 153 | 2.317 | 1.04 | -.5543 | 1.5943 |
| 202 | -2.653 | -2.38 | -.5543 | -1.8257 |
| 346 | -2.479 | -2.26 | -.5543 | -1.7057 |
| 479 | 2.215 | .97 | -.5543 | 1.5243 |

a. Dependent Variable: Change in Hygiene Over The Festival



Output 15

This question asks whether this model is valid.

Residuals: There are no cases that have a standardized residual greater than 3. If you look at the casewise diagnostics table in Output 15, you can see that there were 5 cases out of a total of 123 (for day 3) with standardized residuals above 2. As a percentage this would be $5/123 \times 100 = 4.07\%$, so that's as we would expect. There was only 1 case out of 123 with residuals above 2.5, which as a percentage would be $1/123 \times 100 = 0.81\%$ (and we'd expect 1%), which indicates the data are consistent with what we'd expect.

Normality of errors: The histogram looks reasonably normally distributed, indicating that the normality of errors assumption has probably been met. The normal P–P plot verifies this because the dashed line doesn't deviate much from the straight line (which indicates what you'd get from normally distributed errors).

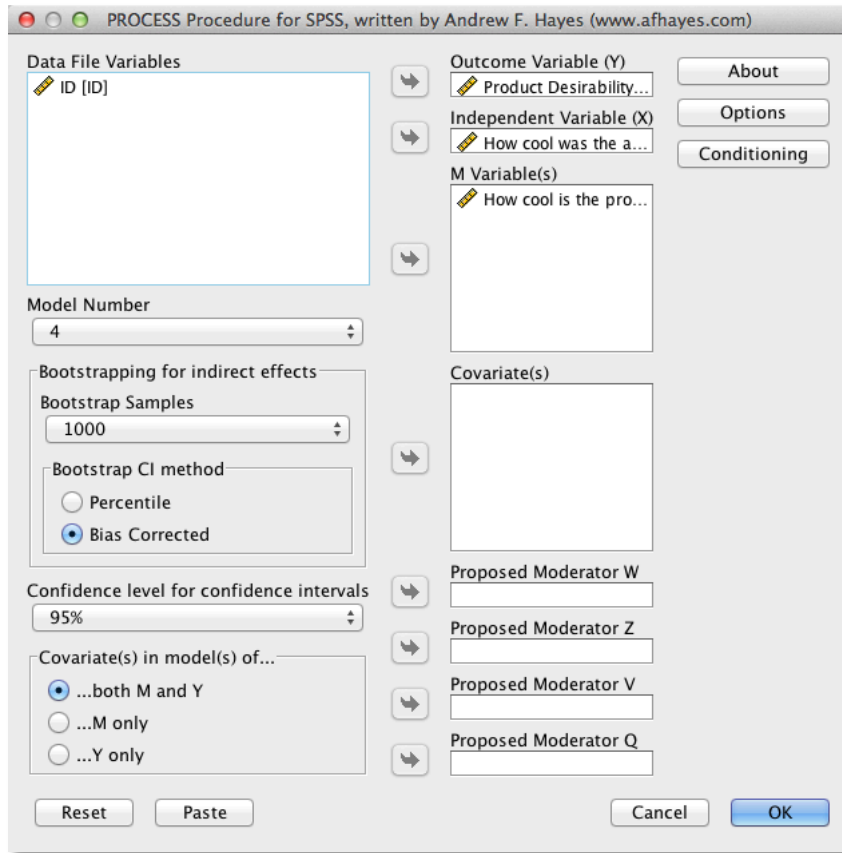
Homoscedasticity and independence of errors: The scatterplot of ZPRED vs. ZRESID does look a bit odd with categorical predictors, but essentially we're looking for the height of the lines to be about the same (indicating the variability at each of the three levels is the same). This is true, indicating homoscedasticity. The Durbin–Watson statistic also falls within the recommended boundaries of 1–3, which suggests that errors are reasonably independent.

Multicollinearity: For all variables in the model, VIF values are below 10 (or alternatively, tolerance values are all well above 0.2) indicating no multicollinearity in the data.

All in all, the model looks fairly reliable (but you should check for influential cases).

Task 9

*Tablets like the iPad are very popular. A company owner was interested in how to make his brand of tablets more desirable. He collected data on how cool people perceived a products advertising to be (**Advert_Cool**), how cool they thought the product was (**Product_Cool**), and how desirable they found the product (**Desirability**). The data are in the file **Tablets.sav**. Test his theory that the relationship between cool advertising and product desirability is mediated by how cool people think the product is. Am I showing my age by using the word 'cool'?*



```

*****
Model = 4
  Y = Desirabi
  X = Advert_C
  M = Product_

Sample size
  240

*****
Outcome: Product_

Model Summary
      R      R-sq      F      df1      df2      p
      .1894   .0359   8.8507   1.0000   238.0000   .0032

Model
      coeff      se      t      p
constant   .4289   .1548   2.7710   .0060
Advert_C   .1990   .0669   2.9750   .0032

Output 15
    
```


Output shows the first part of the output, which initially tells us the name of the outcome (Y), the predictor (X) and the mediator (M) variables, which have been shortened to eight letters. This is useful for double-checking we have entered the variables in the correct place: the outcome is **Desirability**, the predictor is **Advert_Cool**, and the mediator is **Product_Cool**. The next part of the output shows us the results of the simple regression of **Product_Cool** predicted from cool advertising. This output is interpreted just as we would interpret any regression: we can see that how cool people perceive the advertising to be significantly predicts how cool they think the product is, $b = 0.20$, $t = 2.98$, $p = .003$. The R^2 value tells us that cool advertising explains 3.59% of the variance in how cool they think the product is, and the fact that the b is positive tells us that the relationship is positive also: the more 'cool' people think the advertising is, the more 'cool' they think the product is (and vice versa).

```
*****
Outcome: Desirabi

Model Summary
      R      R-sq      F      df1      df2      p
      .3601    .1297    17.6543    2.0000    237.0000    .0000

Model
      coeff      se      t      p
constant    3.1431    .1376    22.8486    .0000
Product_    .2479    .0567    4.3720    .0000
Advert_C    .1861    .0596    3.1235    .0020
```

Output 16

Output shows the results of the regression of **Desirability** predicted from both how cool people think the product is and how cool people think the advertising is. We can see that cool advertising significantly predicts product desirability even with **Product_Cool** in the model, $b = 0.19$, $t = 3.12$, $p = .002$; **Product_Cool** also significantly predicts product desirability, $b = 0.25$, $t = 4.37$, $p < .001$. The R^2 values tells us that the model explains 12.97% of the variance in product desirability. The positive bs for **Product_Cool** and **Advert_Cool** tells us that as adverts and products increase in how cool they are perceived to be, product desirability increases also (and vice versa). These relationships are in the predicted direction.

```
***** TOTAL EFFECT MODEL *****
Outcome: Desirabi

Model Summary
      R      R-sq      F      df1      df2      p
      .2439    .0595    15.0488    1.0000    238.0000    .0001

Model
      coeff      se      t      p
constant    3.2494    .1405    23.1352    .0000
Advert_C    .2355    .0607    3.8793    .0001
```

Output 17

Output shows the total effect of cool advertising on product desirability (outcome). You will get this bit of the output only if you selected *Total effect model*. The total effect is the effect of the predictor on the outcome when the mediator is not present in the model. When **Product_Cool** is not in the model, cool advertising significantly predicts product desirability, $b = .24$, $t = 3.88$, $p < .001$. The R^2 values tells us that the model explains 5.96% of the variance in product desirability. As is the case when we include **Product_Cool** in the model, **Advert_Cool** has a positive relationship with product desirability (as shown by the positive b -value).

***** TOTAL, DIRECT, AND INDIRECT EFFECTS *****

```

Total effect of X on Y
  Effect      SE      t      p
  .2355      .0607    3.8793  .0001

Direct effect of X on Y
  Effect      SE      t      p
  .1861      .0596    3.1235  .0020

Indirect effect of X on Y
  Effect      Boot SE    BootLLCI    BootULCI
Product_     .0493      .0219      .0135      .1023

Partially standardized indirect effect of X on Y
  Effect      Boot SE    BootLLCI    BootULCI
Product_     .0655      .0283      .0169      .1321

Completely standardized indirect effect of X on Y
  Effect      Boot SE    BootLLCI    BootULCI
Product_     .0511      .0226      .0139      .1057

Ratio of indirect to total effect of X on Y
  Effect      Boot SE    BootLLCI    BootULCI
Product_     .2095      .1373      .0638      .5553

Ratio of indirect to direct effect of X on Y
  Effect      Boot SE    BootLLCI    BootULCI
Product_     .2650      8.3929      .0666      1.2284

R-squared mediation effect size (R-sq_med)
  Effect      Boot SE    BootLLCI    BootULCI
Product_     .0236      .0136      .0049      .0621

Preacher and Kelley (2011) Kappa-squared
  Effect      Boot SE    BootLLCI    BootULCI
Product_     .0518      .0225      .0138      .1074

```

Output 18

Output is the most important part of the output because it displays the results for the indirect effect of cool advertising on product desirability (i.e. the effect via **Product_Cool**). First, we're again told the effect of cool advertising on the product desirability in isolation (the total effect). Next, we're told the effect of cool advertising on the product desirability when **Product_Cool** is included as a predictor as well (the direct effect). The first bit of new information is the *Indirect Effect of X on Y*, which in this case is the indirect effect of cool advertising on the product desirability. We're given an estimate of this effect ($b = 0.049$) as

well as a bootstrapped standard error and confidence interval. As we have seen many times before, 95% confidence intervals contain the true value of a parameter in 95% of samples. Therefore, we tend to assume that our sample isn't one of the 5% that does not contain the true value and use them to infer the population value of an effect. In this case, assuming our sample is one of the 95% that 'hits' the true value, we know that the true b -value for the indirect effect falls between .0135 and .1023. This range does not include zero, and remember that $b = 0$ would mean 'no effect whatsoever'; therefore, the fact that the confidence interval does not contain zero means that there is likely to be a genuine indirect effect. Put another way, **Product_Cool** is a mediator of the relationship between cool advertising and product desirability.

The rest of Output you will see only if you selected *Effect size*; it contains various standardized forms of the indirect effect. In each case they are accompanied by a bootstrapped confidence interval. As with the unstandardized indirect effect, if the confidence intervals don't contain zero then we can be confident that the true effect size is different from 'no effect'. In other words, there is mediation. All of the effect size measures have confidence intervals that don't include zero, so whatever one we look at we can be fairly confident that the indirect effect is greater than 'no effect'. Focusing on the most useful of these effect sizes, the standardized b for the indirect effect, its value is $b = .051$, 95% BCa CI [.014, .106], and similarly, $\kappa^2 = .052$, 95% BCa CI [.014, .107]. κ^2 is bounded to fall between 0 and 1, so we can interpret this as the indirect effect being about 5.2% of the maximum value that it could have been, which is a fairly small effect. We might, therefore, want to look for other potential mediators to include in the model in addition to how cool people think the product is.

```
Normal theory tests for indirect effect
      Effect      se      z      p
      .0493      .0204      2.4168      .0157
```

Output 19

Output shows the results of the Sobel test. As I have mentioned before, it is better to interpret the bootstrap confidence intervals than formal tests of significance; however, if you selected *Sobel test* this is what you will see. Again, we're given the size of the indirect effect ($b = 0.05$), the standard error, associated z-score ($z = 2.42$) and p -value ($p = .016$). The p -value is under the not-at-all magic .05 threshold so we'd conclude that there is a significant indirect effect, so there is compelling evidence that there is a meaningful mediation effect.

Reporting the results of the mediation analysis

Some people report only the indirect effect in mediation analysis, and possibly the Sobel test. However, I have repeatedly favoured using bootstrap confidence intervals, so you should report these, and preferably the effect size κ^2 and its confidence interval:

- ✓ There was a significant indirect effect of how cool people think a products' advertising is on the desirability of the product though how cool they think the product is, $b = 0.049$, BCa CI [0.014, 0.102]. This represent a relatively small effect, $\kappa^2 = .052$, 95% BCa CI [.014, .107].