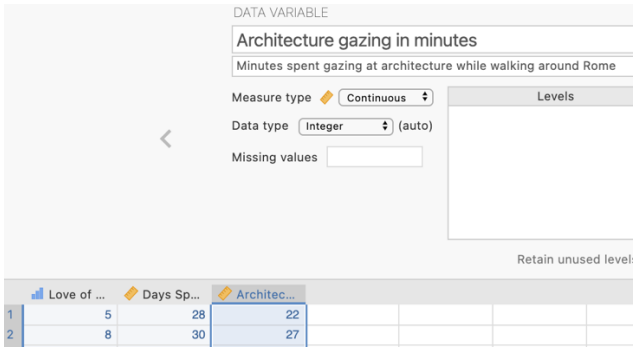
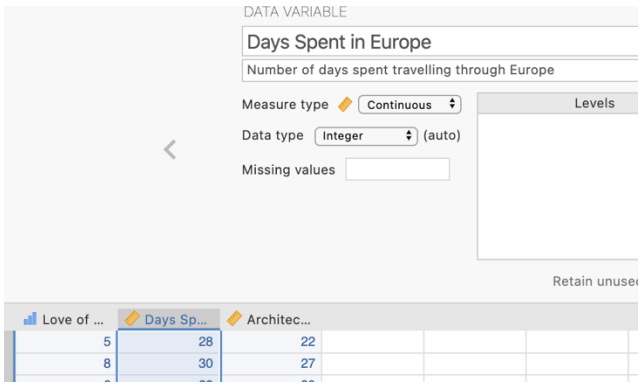
Conducting a Moderated Multiple Regression in *jamovi*

The European Tourism Board[1] were interested in exploring the extent to which time spent in Europe on holiday can predict the amount of time tourists spend paying attention to historical architecture. In particular they wondered whether a "ho hum" factor set in whereby the longer a tourist had spent in Europe the more normative the architecture would become and hence not draw as much attention. They further wanted to check whether a pre-existing love of/interest in architecture would influence this relationship. A team of field researchers were hired to intercept backpackers as they alighted from trains in Roma Termini train station and ask if they would be willing to participate in a study. Travellers were told the study was interested in how many kilometres of walking they would do in the day as a tourist and the kinds of places tourists are interested in. Those who agreed were given a survey to complete which included a measure of their love of historical architecture prior to the start of their European travels, and an indication of how many days they had spent on their holiday already. In addition each participant was fitted with a GPS unit that would monitor their movements for the day. At the end of the day the participants rendezvoused with the field researchers at a restaurant where they returned the GPS units and were given a voucher for a free dinner as a thank you. The data from the GPS units was analysed and the amount of time spent standing still in and around architecturally significant buildings was gathered as measure of time spent gazing at and admiring the architecture. Of interest from an analysis perspective is whether the number of days spent travelling already would predict the time spent gazing at architecture (with the feared result that this relationship would be negative), and a further check to see if a pre-existing love of architecture would change the strength or direction of this relationship.

Step 1 – Taking a look at the data.

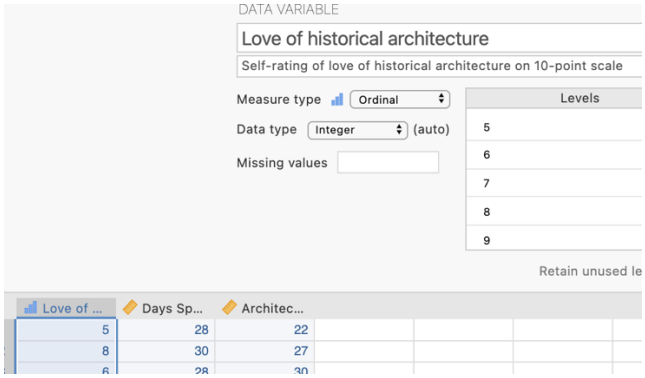


Our dependent variable is Architecture Gazing which is measured in minutes and hence is a continuous measure.



Our independent variable Days Spent in Europe which is also continuous in nature.

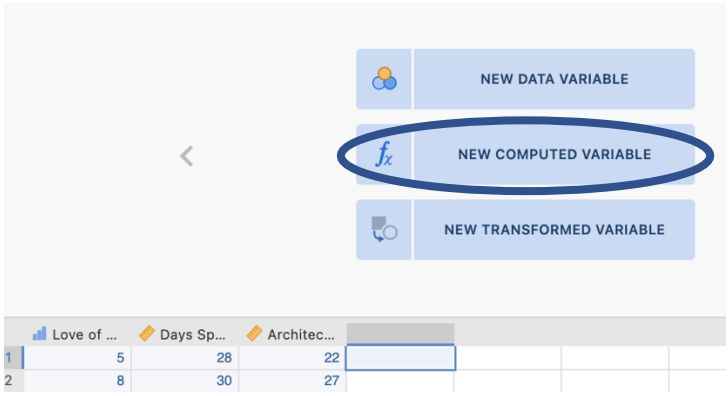
[1] Disclaimer: This board does not exist and this is a fictitious study and data set.



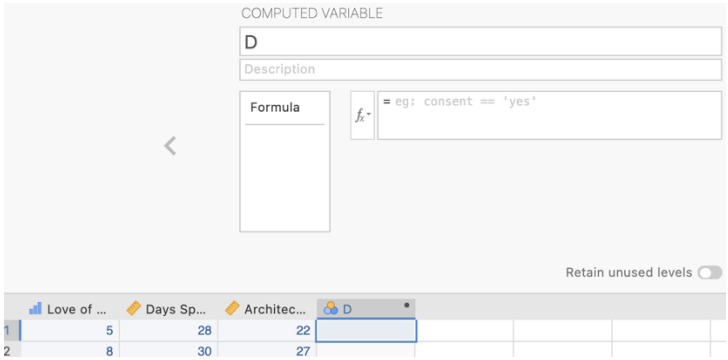
Finally our proposed moderator is a self-rated item on pre-existing Love of Historical Architecture rated on a 10-point Likert scale

Step 2 – Creating centred versions of our Independent Variable and Moderator Variable

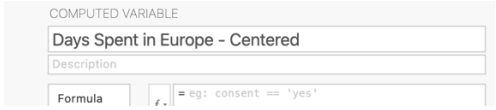
When conducting a moderated regression an interaction term is created by multiplying the independent variable and the moderator variable. The creation of this term creates a multicollinearity issue. It is not surprising that there would be multicollinearity between a variable and variables used to create it. This creates instability in the regression model, in particular for the standard errors (and hence significance tests) for the direct effects of the independent variable and the moderator in the presence of the interaction term. To minimise this we "centre" our independent variable and moderator. Centering is another way of saying that we convert the scores to deviation scores from the mean. So our first step will be to create these centered terms in our data set.



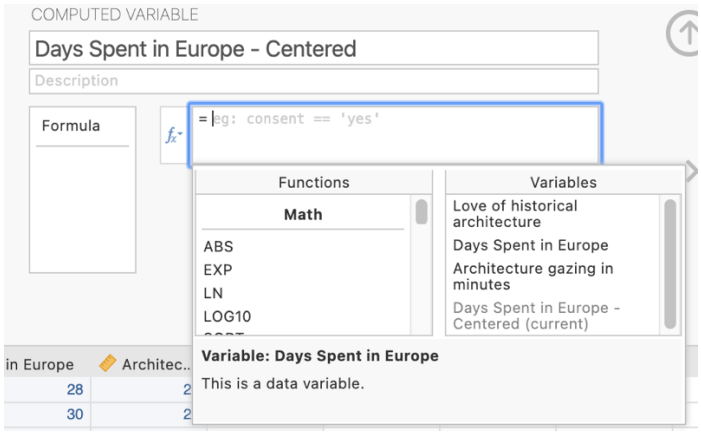
In the data spreadsheet if we click in the first empty column you'll see we get the choice to enter a new data variable, or compute or transform an existing one to create a new variable. We want to select to create a new computed variable.



A new variable has been created and given a letter as its default name until we name it. You'll see there is also a spot to enter a formula.

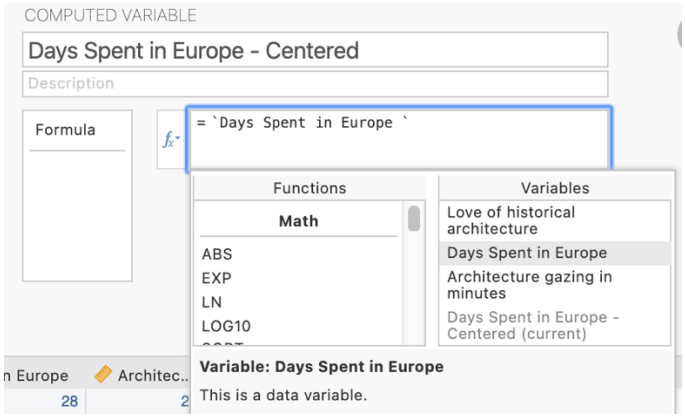


We'll create a centered version of our Days Spent in Europe independent variable first so let's name it accordingly.

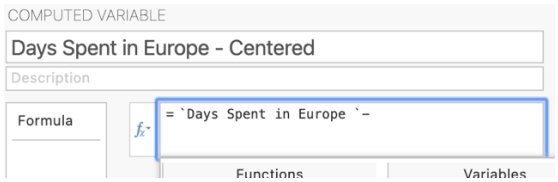


Clicking on the fx function symbol will bring up this listing of formula functions as well as a list of the variables existing in the data set.

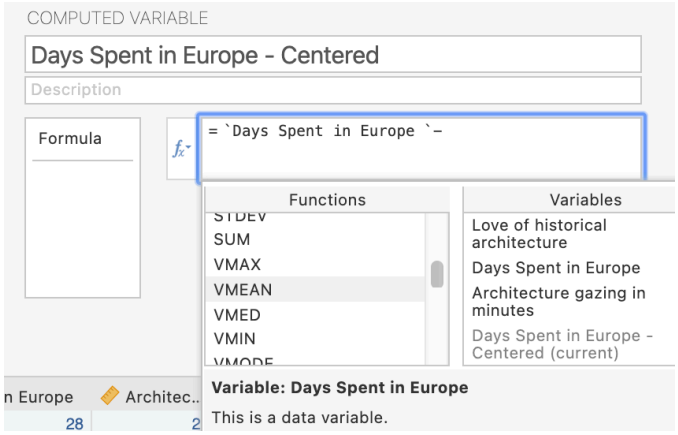
We want to create a formula in the formula box that will subtract the mean of this variable from each individual score. In other words we want to create the formula: $X - M_X$



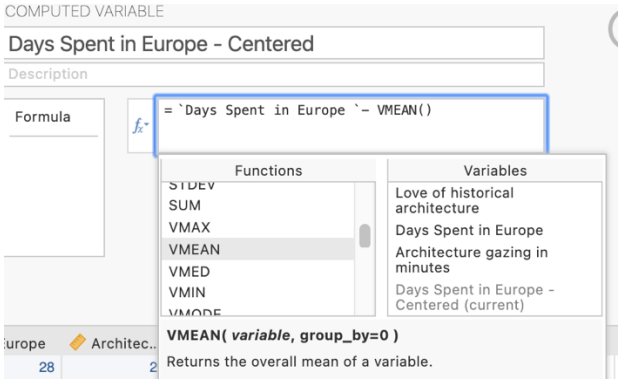
First double click on the Days in Spent in Europe variable in the Variables list. It will appear in the formula box with single quotation marks around it.



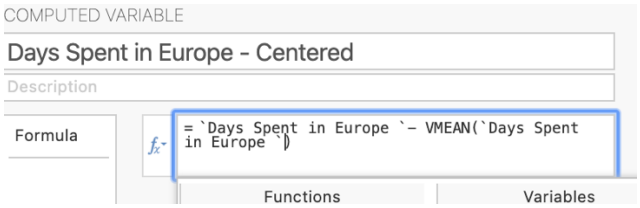
Next type a space and then a dash to serve as a subtraction sign and then another space.



Next scroll down the functions list to find the VMEAN function. This function will use the mean of whatever variable you select to use with it. Double click it to move it into the formula box.



You will see that now VMEAN( ) appears in the formula box with the cursor within the parentheses. Double Click on Days Spent in Europe again in the Variables list and it will appear in the parentheses next to VMEAN.



The formula should look like this now. Hit enter and this should process the formula and populate the variable with centered scores.

Love of ...	Days Spent in Europe	Architec...	Days Sp...
5	28	22	0.931
8	30	27	2.931
6	28	30	0.931
8	21	26	-6.069
8	29	34	1.931
6	26	23	-1.069
8	25	30	-2.069
5	29	22	1.931
6	23	24	-4.069
7	27	30	-0.069
6	24	27	-3.069
9	29	28	1.931
7	29	27	1.931

You should now have an extra column of data, the Days Spent in Europe – Centered scores. You'll see there is a black dot in the corner of the heading cell of this data column which denotes it is a computed variable with a formula attached.

Repeat this process to create the centered version of the Love of Historical Architecture moderator variable.

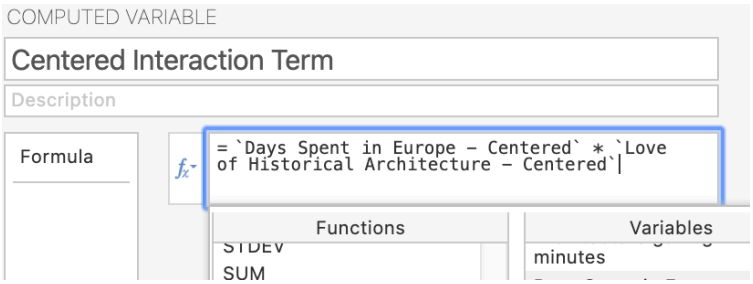
Love of ...	Days Spent in Europe	Architec...	Days Sp...	Love of ...
5	28	22	0.931	-2.444
8	30	27	2.931	0.556
6	28	30	0.931	-1.444
8	21	26	-6.069	0.556
8	29	34	1.931	0.556
6	26	23	-1.069	-1.444
8	25	30	-2.069	0.556
5	29	22	1.931	-2.444
6	23	24	-4.069	-1.444
7	27	30	-0.069	-0.444

Following the same process for the Love of Historical Architecture variable should result in the creation of two new centered variables in the data set as can be seen here.

Step 3 – Creating the interaction term with the centered variables

Finally we need to create the interaction term to add to the regression model. This needs to be created from the two centered variables that we just created.

We'll use the same compute variable process, but specify a different formula.



We need to create a variable that is the centered independent variable multiplied by the centered moderator. The asterisk is used to denote multiplication.

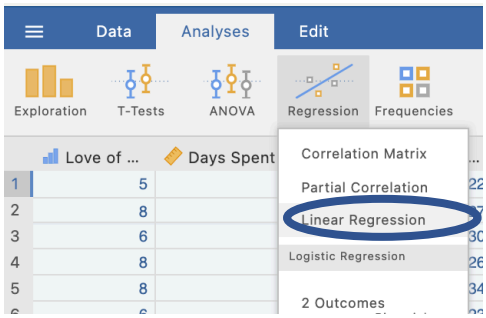
We should now have six variables all together in our data set. Our original dependent, independent and moderator variables, the centered versions of the independent and moderator variables and the interaction variable.

Retain unused levels

Love of ...	Days Spent in Europe	Architec...	Days Sp...	Love of ...	Centere...
5	28	22	0.931	-2.444	-2.275
8	30	27	2.931	0.556	1.628
6	28	30	0.931	-1.444	-1.344
8	21	26	-6.069	0.556	-3.372
8	29	34	1.931	0.556	1.073
6	26	23	-1.069	-1.444	1.545

We are now ready to run our moderated multiple regression to see if we have evidence for moderation.

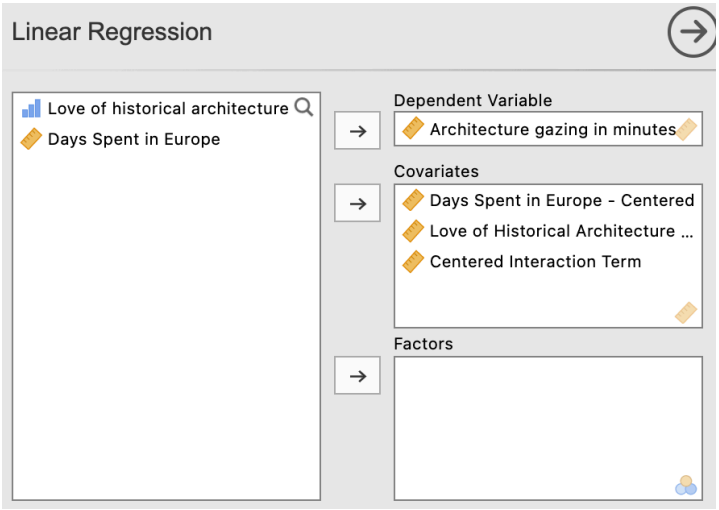
Step 4a – Running our moderated multiple regression.

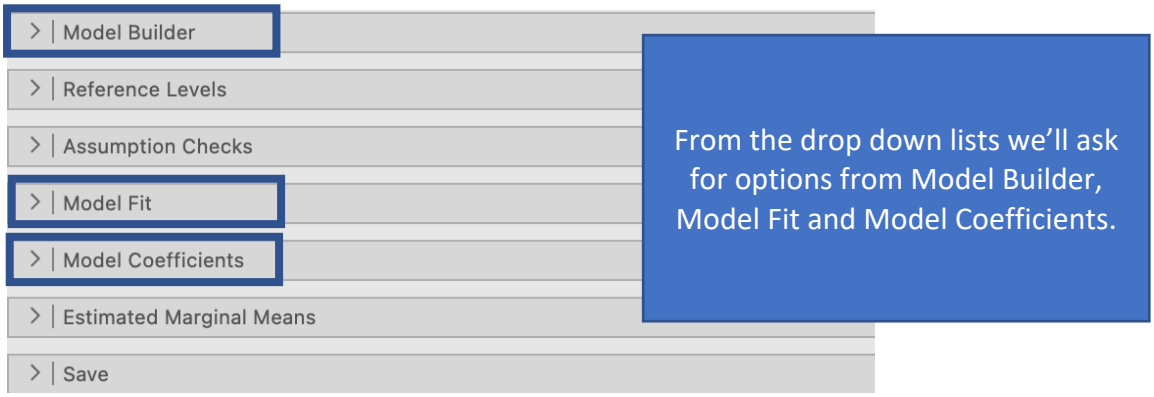


On the Analyses tab select the Regression menu, then select Linear Regression.

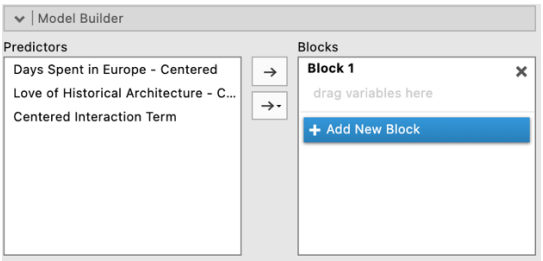
Step 4b – Selecting analysis options

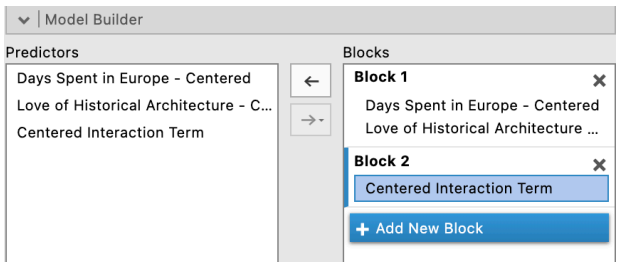
The first thing we need to do is tell *jamovi* what variable is our criterion/dependent variable and which are our predictors. It is important to note that jamovi refers to continuous predictors as covariates and categorical predictors as factors. In our case all our predictors are continuous so we will need to move them across to the covariates box. We'll move the centered versions of our independent and moderator variables as well as the interaction term all to the covariates box.



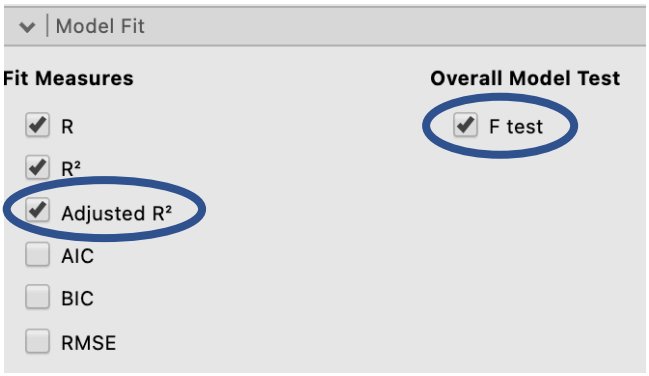


In Model Builder we need to build a hierarchical regression with our centered independent and moderator variables in Block 1 and our interaction term in Block 2. Drag the variables across to create Block 1. Click to add the second block and drag the interaction term across so that your model specification looks like the screen shot on the right side below.





Under Model Fit, in addition to the *R* and R^2 already set as a default, we'll ask for adjusted R^2 and the Overall Model *F* test.



We are given the R and R2 for the whole model as well as unstandardised regression coefficients (labelled Estimate) and associated significance testing.

However, there are a few more elements and tests we require for completeness.

Under the Model Coefficients drop down menu, we'll ask for confidence intervals for our unstandardised regression coefficients, *B*s, (under Estimate) as well as asking for standardised regression coefficients, βs, (under Standardised Estimate).

Our regression output so far looks like this:

Linear Regression

Model Fit Measures

Model	R	R^2	Adjusted R^2	Overall Model Test F	df1	df2	p
1	0.48488	0.23510	0.21293	10.60422	2	69	0.00010
2	0.52744	0.27819	0.24634	8.73578	3	68	0.00006

Model Comparisons

Comparison Model		Model	ΔR^2	F	df1	df2	p
1	-	2	0.04308	4.05873	1	68	0.04790

Note that in our Model Coefficients table we only see one model result set at a time. We can't view the coefficients for Model 1 (Block 1) and Model 2 (Block 2) at the same time. However we can toggle between them by clicking on the toggle next to the Model Specific Results heading.

Model Specific Results Model 2

Model Coefficients - Architecture gazing in minutes

Predictor	Estimate	SE	95% Confidence Interval Lower	Upper	t	p	Stand. Estimate
Intercept	26.16837	0.36277	25.44448	26.89226	72.13539	<.00001	
Days Spent in Europe - Centered	0.35442	0.10253	0.14982	0.55902	3.45669	0.00095	0.37115
Love of Historical Architecture - Centered	0.77889	0.26015	0.25976	1.29802	2.99395	0.00384	0.31237
Centered Interaction Term	0.14775	0.07334	0.00141	0.29409	2.01463	0.04790	0.21584

Model Specific Results Model 1

Model Coefficients - Architecture gazing in minutes

Predictor	Estimate	SE	95% Confidence Interval Lower	Upper	t	p	Stand. Estimate
Intercept	26.25000	0.36840	25.51506	26.98494	71.25380	<.00001	
Days Spent in Europe - Centered	0.30036	0.10113	0.09862	0.50211	2.97013	0.00409	0.31454
Love of Historical Architecture - Centered	0.83971	0.26406	0.31292	1.36650	3.17998	0.00221	0.33676

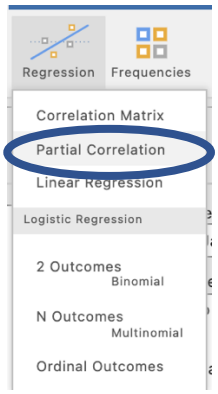
Whether or not we can argue for the presence of moderation hinges on the significance of the interaction term. This test can be found as a test of R^2_{change} for the second block where we added the interaction term in the Model Comparisons table. It can be found mirrored in the Model Coefficients table as well for the test of the coefficient for the interaction term.

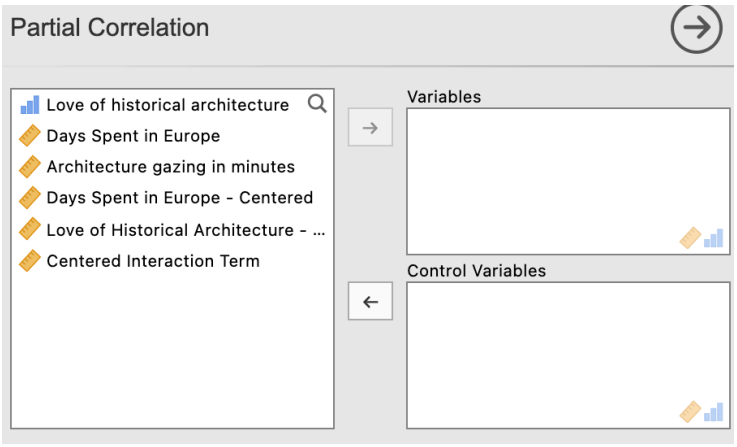
The *p* value associated with this test is .048 which, being less than .05, denotes the interaction term adds significantly to the model and hence suggests the presence of moderation. Given this we will need to go on to conduct post hoc analysis in the form of simple slopes analysis.

Step 4c – Obtaining $r^2_{a(b,c)}$, for our model

One key statistic that is conventionally reported with multiple regression results, namely squared semi-partial correlations, or $r^2_{a(b,c)}$, are not available via the Linear Regression menu. To obtain these we need to run some separate analysis.

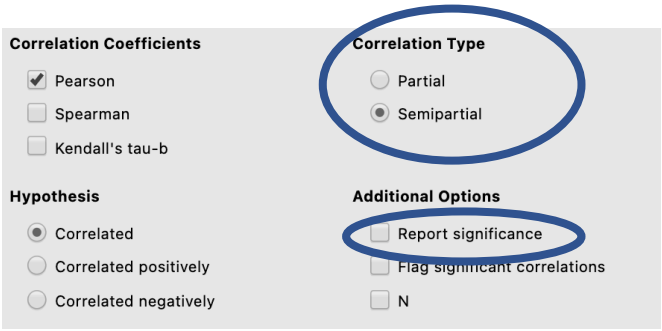
Head to the Analyses ribbon, select Regression and then Partial Correlations



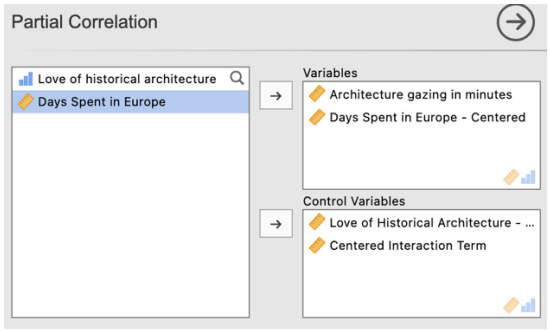


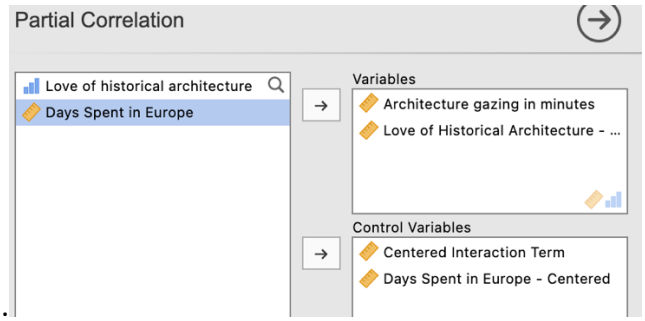
In the current *jamovi* set up for obtaining partial and semi-partial correlations they need to be specified one at a time (as at version 1.8.4). In other words we need to specify a pair of variables to be correlated and which variables to use as control variables, or variables that will be partialled out of the calculations.

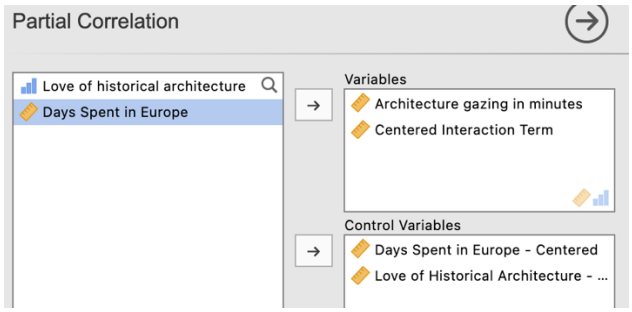
Before we start let's make sure the options we need have been selected. The most important selection is to change the selection to "Semipartial" under Correlation Type as Partial is the default. Also untick "Report significance" under Additional Options to keep the output minimal and easy to interpret.



There are three semi-partial correlations we need to request. These are between our dependent variable Architecture Gazing with each of the three predictor variables, controlling for the respective remaining two each time. Below is how you will specify each semi-partial correlation that you need for this regression model.







Note that after running each semi-partial correlation you will need to click underneath the output so that jamovi knows to create an additional output, rather than editing the one you just created. Jamovi will edit whichever output is currently highlighted in white. To ensure you get additional output rather than overwritten output ensure you have clicked underneath the output created and it has turned grey.

Our three semi-partial correlation runs will provide the following output:

Partial Correlation

Semipartial Correlation - Pearson's r

	Architecture gazing in minutes	Days Spent in Europe - Centered
Architecture gazing in minutes	—	0.35614
Days Spent in Europe - Centered	0.37096	—

Note. controlling for 'Love of Historical Architecture - Centered' and 'Centered Interaction Term'
Note. variation from the control variables is only removed from the variables in the columns

Partial Correlation

Semipartial Correlation - Pearson's r

	Architecture gazing in minutes	Love of Historical Architecture - Centered
Architecture gazing in minutes	—	0.30846
Love of Historical Architecture - Centered	0.33701	—

Note. controlling for 'Days Spent in Europe - Centered' and 'Centered Interaction Term'
Note. variation from the control variables is only removed from the variables in the columns

Partial Correlation

Semipartial Correlation - Pearson's r

	Architecture gazing in minutes	Centered Interaction Term
Architecture gazing in minutes	—	0.20756
Centered Interaction Term	0.22823	—

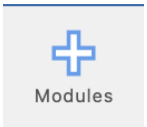
Note. controlling for 'Days Spent in Europe - Centered' and 'Love of Historical Architecture - Centered'
Note. variation from the control variables is only removed from the variables in the columns

In each of these semi-partial correlation tables the coefficient we want is listed in the column heading with the predictor name of interest as circled on each.

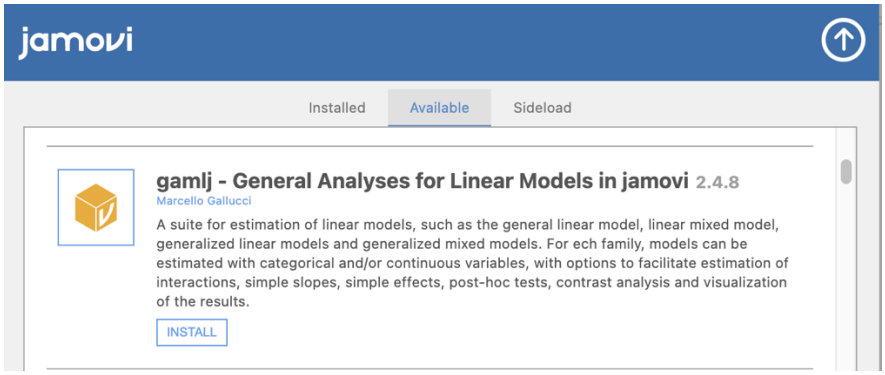
Step 5 – Running our Simple Slopes Analysis

Within the regression menus that are part of the base *jamovi* program, simple slopes are not currently an option that can be requested (as at v. 1.8.4). However *jamovi* has a range of add on modules that can be activated which expand the analysis options available. One of these, the *General Analyses for Linear Models in jamovi*, or *gamlj* module, will allow us to obtain our simple slopes.

In order to install an add on module you need to click on the addition icon in the top right hand corner of your *jamovi* window.

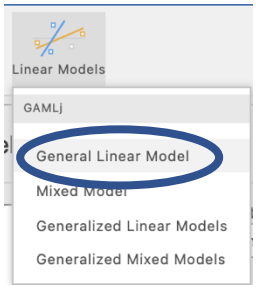


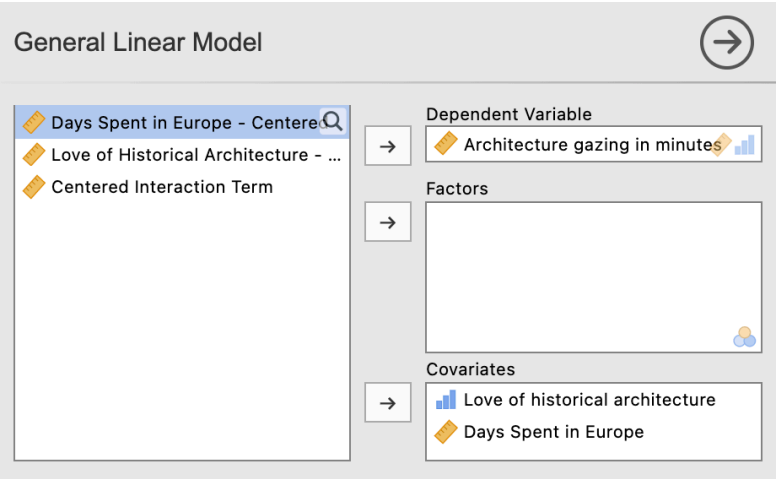
Go to the library tab, and scroll to find the *gamlj* module and click to install it.



You will now see that you have an additional option on your Analyses toolbar called Linear Models.

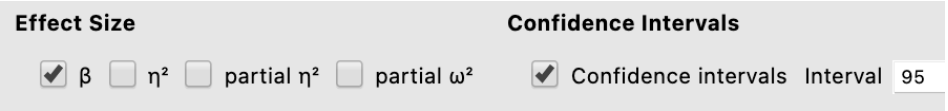
Click on Linear Models and then select General Linear Model.



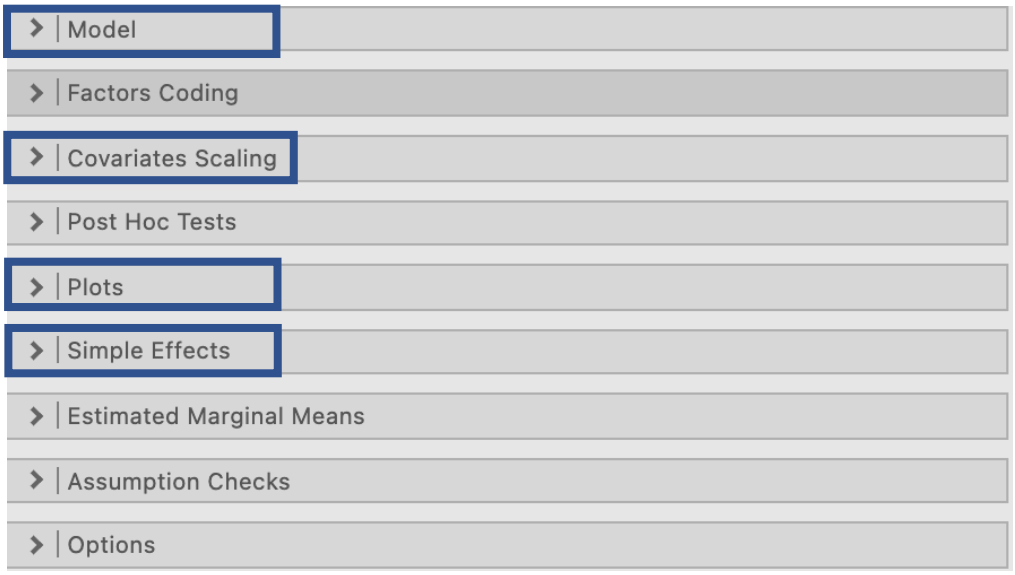


We'll use our raw variable versions for our independent variable and moderator here. Shift Architecture Gazing in Minutes to the Dependent Variable slot and the Love of Historical Architecture and Days Spent in Europe uncentered versions to the Covariates box (this is what *jamovi* refers to continuous predictors as).

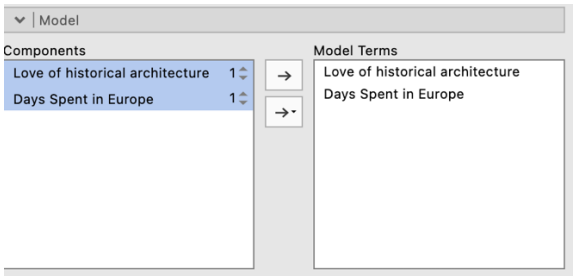
In the options immediately under the variable specification ensure just β and confidence intervals are selected to keep our output minimal and focussed on our key needs.

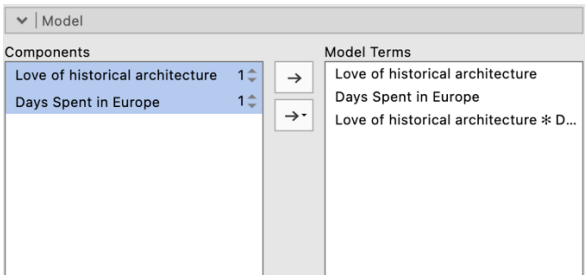


There are quite a few drop down menu options. We'll need to go into Model, Covariates Scaling, Plots and Simple Effects. We'll walk through them one at a time.



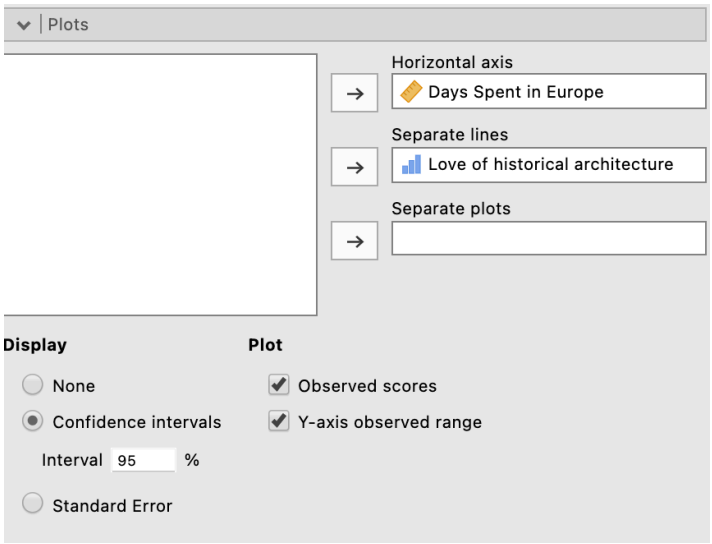
Under Model we need to specify our moderation model in a similar way to the regression model we ran. In other words, we need to add in an interaction term. If you highlight both predictor variables and click the arrow to move them to the right hand side you get the option to add them as an interaction term. If this has been done correctly there should now be a third variable in the Model Terms section with the independent variable and moderator with an asterisk between them to denote they are multiplied as a product term.



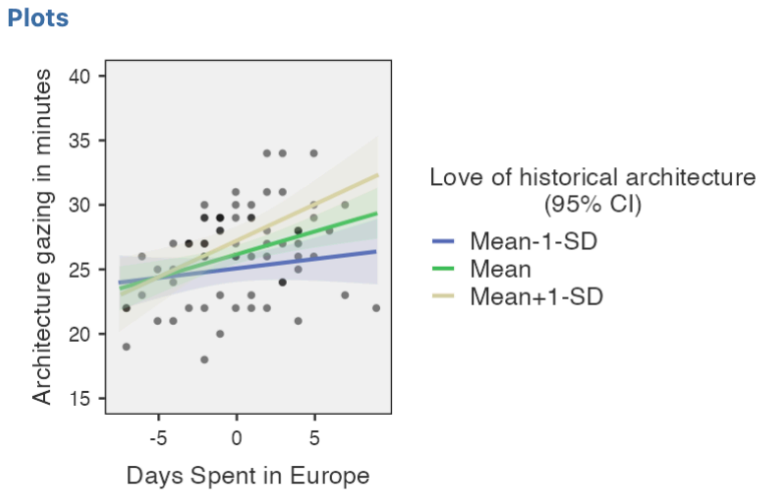


Under Covariates Scaling you'll note that as a default both the predictors have been set to be centered. Running the moderation model via this add on module means that we can get the analysis to centre for us and we don't have to do it manually.

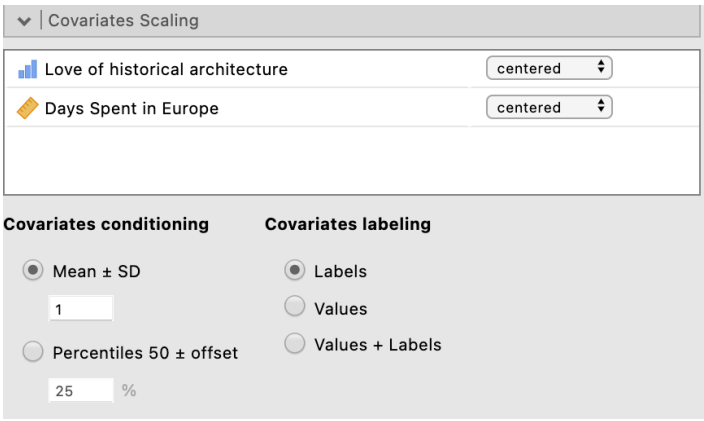
Under Plots we can ask for a simple slopes plot to help illustrate our moderation effect. Shift the independent variable Days Spent in Europe to the Horizontal axis and the moderator variable Love of Historical Architecture to the Separate Lines slot. We'll also ask for observed scores to be plotted as well as the *Y*-axis to be adjusted to reflect the observed range of scores. 95% *CI*s will be plotted as a default as well.



The chosen plot options will create a plot like the one below.



Finally under Simple Effects Under Simple Effects, we want to specify our independent variable Days Spent in Europe as the variable we want Simple Effects for, and specify our moderator Love of Historical Architecture.



Note that under "Covariates conditioning" we can choose to obtain simple slopes at the *M* and +/- 1 *SD* values of our moderator, or we can choose for them to be based on the median (50th percentile) +/- 25 percentiles. We can also choose to alter the defaults for the low and high slopes in either *SD*s or percentiles if we wish. We'll leave this at the default of *M* +/- 1 *SD* as this is what is standardly reported.

This provides us with a simple slopes/simple effects table with results for simple slopes when our moderator is at its mean as well as plus and minus one standard deviation.

Simple effects of Days Spent in Europe : Parameter estimates

Moderator levels			95% Confidence Interval				
Love of historical architecture	Estimate	SE	Lower	Upper	df	t	p
Mean-1·SD	0.14564	0.12527	−0.10433	0.39560	68.00000	1.16263	0.24904
Mean	0.35442	0.10253	0.14982	0.55902	68.00000	3.45669	0.00095
Mean+1·SD	0.56321	0.16375	0.23645	0.88997	68.00000	3.43940	0.00100

We have everything we need to pull our write up together now so let's do it!

Step 6a – Finding the components for reporting the overall model.

There are two parts to our moderated multiple regression write up. The overall model results and then results pertaining to the simple slopes.

The components we need for the overall model results are:

1. The *F* statistic, *df*s and *p* value – the omnibus ANOVA result for the full model and model 2.
2. Effect sizes in the form of **R^2, adjusted R^2** and **R^2_{change}**

Linear Regression

Model Fit Measures

Model	R	R^2	Adjusted R	Overall Model Test F	df1	df2	p
1	0.48488	0.23510	0.21293	10.60422	2	69	0.00010
2	0.52744	0.27819	0.24634	8.73578	3	68	0.00006

Model Comparisons

Comparison			ΔR^2	F	df1	df2	p
Model		Model					
1	-	2	0.04308	4.05873	1	68	0.04790

The Write Up (Part 1):

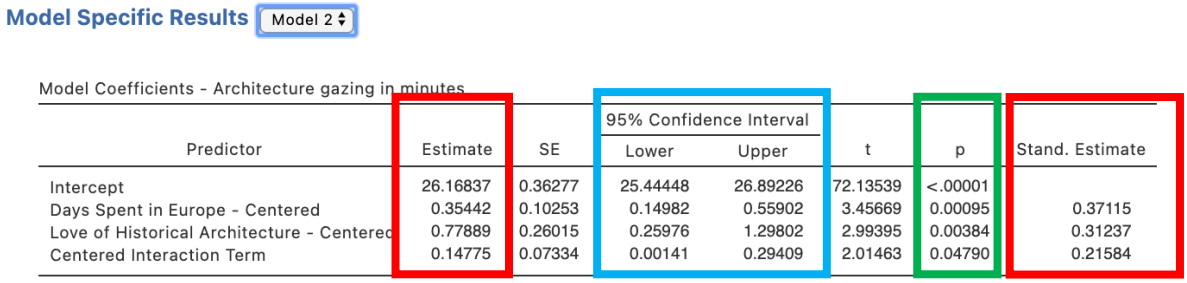
Data from 72 travellers were analysed to test whether the relationship between the number of days spent in Europe on holiday and the number of minutes spent gazing in architecture was moderated by the pre-existing level of love of architectural history. In order to test for the presence of an interactive effect, a hierarchical regression was conducted with Number of Days Spent in Europe and Love of Architectural History added in step one of the model and an interaction term, Number of Days Spent in Europe X Love of Architectural History, added in step two of the model. Predictors were centered to eliminate multicollinearity issues created by the presence of the interaction term when assessing direct effects. When the two predictors were considered additively, 23.5% of the variance in minutes spent gazing at architecture was explained, **R^2 = .24. Adj. R^2 = .21**, which represents a statistically significant effect, **F (2,69) = 10.60, p = .001**. The inclusion of the interaction term in step two of the model contributed a statistically significant addition to the model, **R^2_{change} = .04**, F_{change} (1, 68) = 4.06, p = .047, and provides statistical support for the presence of moderation. The total model explained **27.8%** of the variance in architecture gazing minutes **R^2 = .28. Adj. R^2 = .25**, **F (3, 68) = 8.74, p < .001**.

Getting the parts for the Write Up (Part 2):

We'll create a table to contain the results about the individual predictors. This will require us to consolidate the information from the Model Coefficients table from the regression output as well as the semi-partial correlations we ran separately. We'll put the regression output elements in the table first.

The elements needed for the regression table are:

1. *p* values for each predictor – to determine the significance of each predictor's contribution to the model
2. Effect sizes in the form of *B*s, βs, $r^2_{a(b,c)}$ (Squared semi-partial correlations) and R^2_{change}
3. Confidence intervals – around the unstandardised regression coefficients (*B*s).



Model Specific Results — Model 2

Model Coefficients - Architecture gazing in minutes

Predictor	Estimate	SE	95% Confidence Interval Lower	95% Confidence Interval Upper	t	p	Stand. Estimate
Intercept	26.16837	0.36277	25.44448	26.89226	72.13539	<.00001	
Days Spent in Europe - Centered	0.35442	0.10253	0.14982	0.55902	3.45669	0.00095	0.37115
Love of Historical Architecture - Centered	0.77889	0.26015	0.25976	1.29802	2.99395	0.00384	0.31237
Centered Interaction Term	0.14775	0.07334	0.00141	0.29409	2.01463	0.04790	0.21584

Table 1

Moderated Multiple Regression results for the prediction of Architectural Gazing in Minutes from Days Spent in Europe as Moderated by Love of Historical Architecture.

	B	95% CI for *B* LL	95% CI for *B* UL	β	*p*	$r^2_{a(b,c)}$	R^2_{change}
Block 1							.24 **
Days Spent in Europe	0.35	0.15	0.56	.37	.001		
Love of Historical Architecture	0.78	0.26	1.30	.31	.004		
Block 2							.04 *
Days Spent in Europe X Love of Historical Architecture	0.15	0.00	0.29	.22	.048		

* $p < .05$, ** $p < .01$, *** $p < .001$

Getting our $r^2_{a(b,c)}$ from our semi-partial correlation output:

Partial Correlation

Semipartial Correlation - Pearson's r

	Architecture gazing in minutes	Days Spent in Europe - Centered
Architecture gazing in minutes	—	0.35614
Days Spent in Europe - Centered	0.37096	—

Note. controlling for 'Love of Historical Architecture - Centered' and 'Centered Interaction Term'
Note. variation from the control variables is only removed from the variables in the columns

Square this value for the $r^2_{a(b,c)}$ for the Days Spent in Europe predictor.

$0.35614^2 = 0.127$

Partial Correlation

Semipartial Correlation - Pearson's r

	Architecture gazing in minutes	Love of Historical Architecture - Centered
Architecture gazing in minutes	—	0.30846
Love of Historical Architecture - Centered	0.33701	—

Note. controlling for 'Days Spent in Europe - Centered' and 'Centered Interaction Term'
Note. variation from the control variables is only removed from the variables in the columns

Square this value for the $r^2_{a(b,c)}$ for the Love of Historical Architecture predictor.

$0.30846^2 = 0.095$

Partial Correlation

Semipartial Correlation - Pearson's r

	Architecture gazing in minutes	Centered Interaction Term
Architecture gazing in minutes	—	0.20756
Centered Interaction Term	0.22823	—

Note. controlling for 'Days Spent in Europe - Centered' and 'Love of Historical Architecture - Centered'
Note. variation from the control variables is only removed from the variables in the columns

Square this value for the $r^2_{a(b,c)}$ for interaction term.

$0.20756^2 = 0.043$

We can now complete our table below:

The Write Up (Part 2):

Full details of the model including regression coefficients can bee seen in Table 1.

Table 1

Moderated Multiple Regression results for the prediction of Architectural Gazing in Minutes from Days Spent in Europe as Moderated by Love of Historical Architecture.

	B	95% CI for *B*		*β*	*p*	$r^2_{a(b,c)}$	R^2_{change}
		LL	UL				
Block 1							.24 **
Days Spent in Europe	0.35	0.15	0.56	.37	.001	.127	
Love of Historical Architecture	0.78	0.26	1.30	.31	.004	.095	
Block 2							.04 *
Days Spent in Europe X Love of Historical Architecture	0.15	0.00	0.29	.22	.048	.043	

* $p < .05$, ** $p < .01$, *** $p < .001$

Step 6b – Finding the components for reporting the simple slopes analysis.

The elements needed for simple slopes reporting are:

1. *p* values for each simple slope – to determine the significance of each simple slope at levels of the moderator.
2. Effect sizes in the form of unstandardised regression coefficients (*Bs*)
3. Confidence intervals – around the unstandardised regression coefficients (*B*s).

Simple effects of Days Spent in Europe : Parameter estimates

Moderator levels			95% Confidence Interval				
Love of historical architecture	Estimate	SE	Lower	Upper	df	t	p
Mean-1·SD	0.14564	0.12527	−0.10433	0.39560	68.00000	1.16263	0.24904
Mean	0.35442	0.10253	0.14982	0.55902	68.00000	3.45669	0.00095
Mean+1·SD	0.56321	0.16375	0.23645	0.88997	68.00000	3.43940	0.00100

The Write Up (Part 3):

Simple slopes analysis was conducted to further explore the moderating impact of prior love of historical architecture on the relationship days spent in Europe and gazing at architecture. The relationship between days spent in Europe and minutes spent gazing at architecture is positive and is seen to increase with increasing scores on prior love of historical architecture. The simple slope is not significant when prior love of historical architecture is at one standard deviation below its mean, ***B* = 0.15**, ***p* = .249**, **95% *CI* [-0.10, 0.40]**, but is significant when prior love of historical architecture is at its mean, ***B* = 0.35**, ***p* = .001**, **95% *CI* [0.15, 0.56]**, and one standard deviation above its mean., ***B* = 0.56**, ***p* = .001**, **95% *CI* [0.24, 0.89]**.

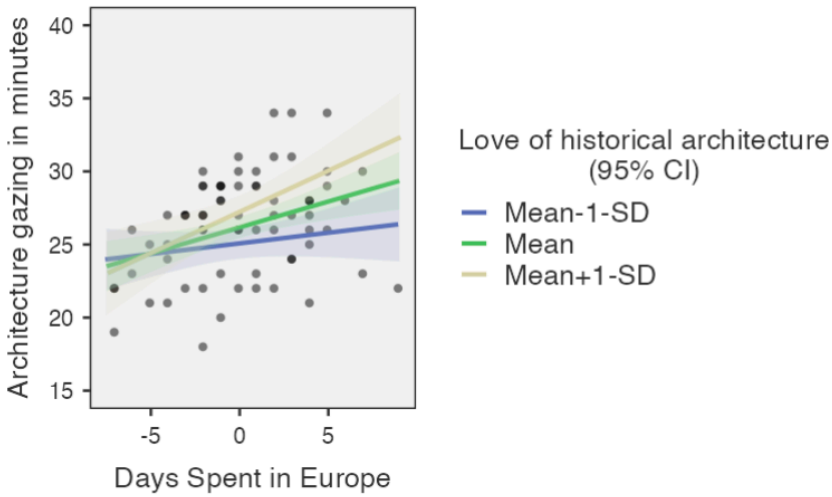
Step 6c – Including the Simple Slopes Plot

The Write Up (Part 4):

Figure 1 provides a visual representation of this moderated relationship.

Figure 1

Simple Slopes for the Relationship Between Days Spent in Europe and Architecture Gazing in Minutes as Moderated by Love of Historical Architecture



[2] The Statistics Working Group was formed in November 2020 to review the use of statistical packages in teaching across the core undergraduate statistics unit. The working group is led by Winnifred Louis and Philip Grove, with contributions from Timothy Ballard, Stefanie Becker, Jo Brown, Jenny Burt, Nathan Evans, Mark Horswill, David Sewell, Eric Vanman, Bill von Hippel, Courtney von Hippel, Zoe Walter, and Brendan Zietsch.

[3] The jamovi project (2021). jamovi (Version 1.8.4) [Computer Software]. Retrieved from https://www.jamovi.org