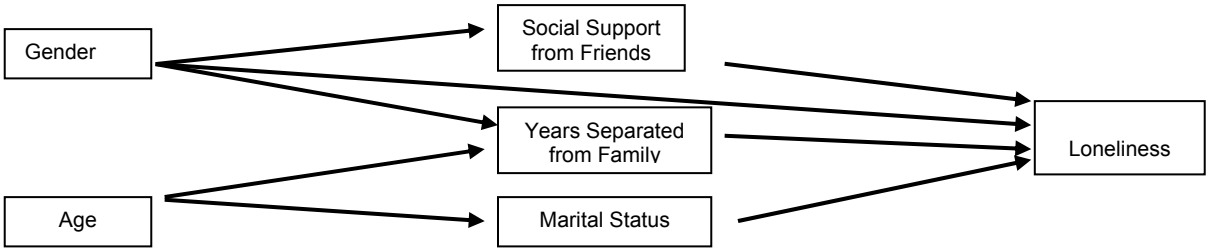
A Slightly More Interesting Path Model

Here's the proposed structural model for how these variables relate to reports of Loneliness (RULS) of young adults.

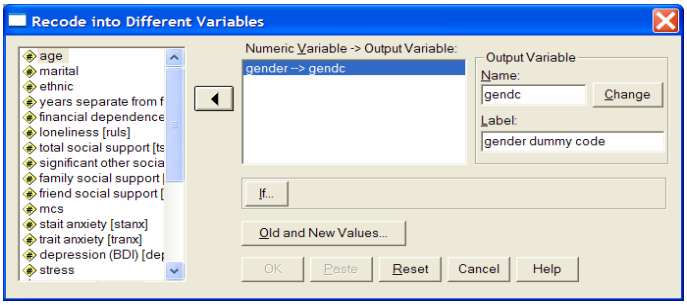
Remember, a deleted path doesn't mean no correlation is expected between those variables! Rather, it means that variable is not expected to make a unique contribution to the relevant model. For example, the deleted path between Gender and Marital Status doesn't mean that the two are expected to not be correlated, only that in a model predicting Marital Status from Gender and Age, Gender has no unique contribution. Notice that both Gender and Age are expected to contribute to a model predicting Years Separated from Family and that Gender, but not Age, is expected to contribute to a model predicting Social Support from Friends.

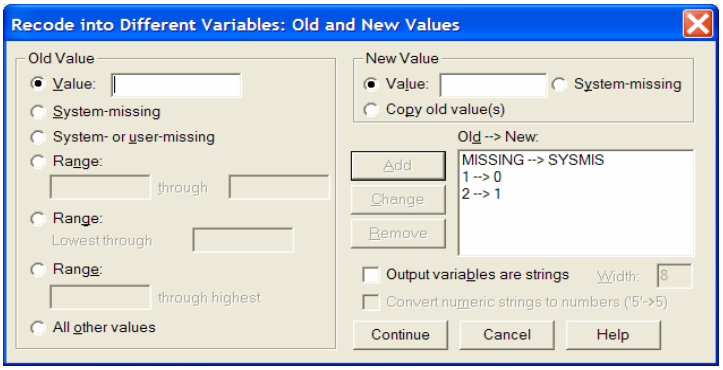


"Preparing" Variables for the Analysis

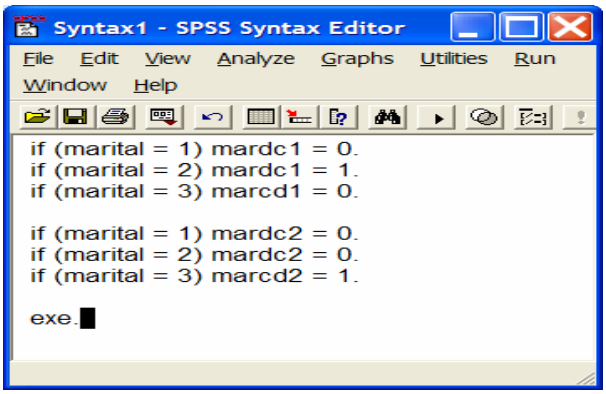
Age, Years Separated from Family, Social Support from Friends are all quantitative and require no "preparation" (other than the usual distributional cleansing).

Gender is binary, and while we could leave it in its current form (1=male, 2=female), it is a good idea to convert it to a dummy code if there is any chance that you'll be considering interactions including this variable. The following will create a dummy coded version of gender with 0=male and 1=female.





Marital Status has 3 conditions (1=single, 2=married, 3=divorced, separated or widowed) and must be converted to dummy codes before it can be used in the analysis. This syntax will create 2 dummy codes. When used together in a multiple regression the first compares single with married and the second compares single with divorced (i.e., single is the comparison group and is coded "0" for both dummy codes, with each of the other groups being the target group for one dummy code).



Getting the Correlations (the bivariate description of how these variables related to each other)

The correlations among the criterion and the quantitative and binary predictors are no problem!

Correlations

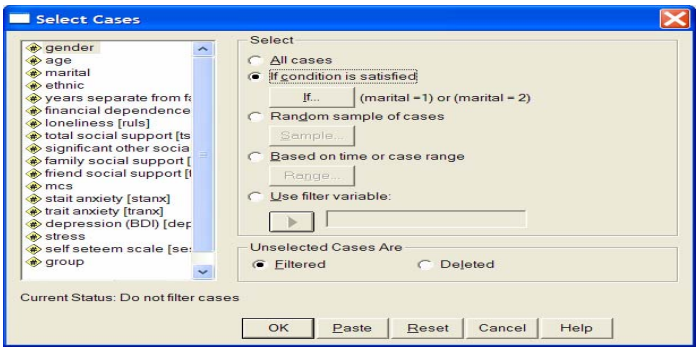
		loneliness	GENDER	AGE	years separate from family	friend social support
loneliness	Pearson Correlation	1	-.021	.225	.214	-.577
	Sig. (2-tailed)	.	.678	.000	.000	.000
	N	405	405	405	405	405
GENDER	Pearson Correlation	-.021	1	.146	.186	.112
	Sig. (2-tailed)	.678	.	.003	.000	.024
	N	405	405	405	405	405
AGE	Pearson Correlation	.225	.146	1	.971	-.256
	Sig. (2-tailed)	.000	.003	.	.000	.000
	N	405	405	405	405	405
years separate from family	Pearson Correlation	.214	.186	.971	1	-.240
	Sig. (2-tailed)	.000	.000	.000	.	.000
	N	405	405	405	405	405
friend social support	Pearson Correlation	-.577	.112	-.256	-.240	1
	Sig. (2-tailed)	.000	.024	.000	.000	.
	N	405	405	405	405	405

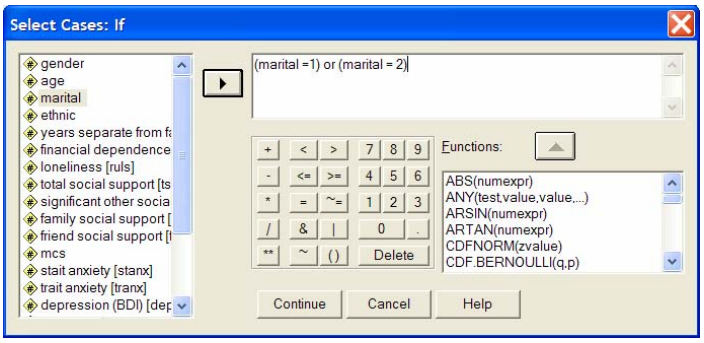
However, getting the correlations for Marital Status, which we've coded into two dummy codes, will take a bit of work,

We can't use the original 3-category variable to get the correlations (the 1-2-3 values don't representative a quantitative variable). We can't use the dummy codes to get the correlations (each used in isolation compares that target group for that code with the amalgam of the other two groups, which isn't comparable with the interpretation of the regression weight from the multiple regression). What to do?

We can get the appropriate correlations, but its takes a couple of steps. Recall that the intent of mardc1 is to compare married (= 1) and single (= 0) folks. So, if we select only these folks and correlate the dummy code with the other variables, we'll get the correlation we want! Here's how to do that.

Data → Select Cases





Then we get the correlations between Marital Status and the other variables for these folks.

These are the correlations between whether a person is single (= 0) or married (= 1) and each of the other variables.

Correlations

		loneliness	GENDER	AGE	years separate from family	friend social support
MARDC1	Pearson Correlation	.091	.109	.783	.788	-.195
	Sig. (2-tailed)	.082	.037	.000	.000	.000
	N	363	363	363	363	363

Substitute "(marital = 1) or (marital = 3)" into the Select Cases window to select only those involved in the mardc2, and then correlate dc2 with the other variables.

These are the correlations between whether a person is single (= 0) or divorced, separated or widowed (= 1) and each of the other variables.

For both variables, I tossed the other correlations in the matrix because they are the correlations among the other variables for just these two groups, and so, don't give us any useful information.

Correlations

		loneliness	GENDER	AGE	years separate from family	friend social support
MARDC2	Pearson Correlation	.167	.172	.757	.765	-.182
	Sig. (2-tailed)	.005	.004	.000	.000	.002
	N	284	284	284	284	284

Now we have the information to look at how all of these variables are related to each other, "bivariately." Path analysis goes beyond these correlations in three ways: 1) looking at a specific structural model (arrangement of boxes), 2) looking "multivariately" (unique contributions), and 3) providing tests of research hypotheses about the unique contributions

Building and Testing Hypotheses About the Structural Model

Basically we have to build two models and then compare the: 1) the full model including all the possible right-pointing arrows and 2) the hypothesized model with the selection of arrows we think are "sufficient".

The Full Model

We need to obtain a multiple regression "for each box that has arrows pointing to it." These include:

1. Social Support from Friends as the criterion, with Age and Gender as predictors
2. Years Separated from Family as the criterion, with Age and Gender as predictors
3. Marital Status as the criterion, with Age and Gender as predictors
4. Loneliness as the criterion, with Marital Status (using mardc1 & mardc2) , Age, Gender, Years Separated from Family, and Social Support from Friends as the predictors

The first three of these require multiple regression analyses, but the 4th, with Marital Status as the criterion is a little different. In fact, you can find at least two suggested ways of doing this analysis. One approach is to use a MANOVA or an linear discriminant function to ask how to estimate the unique contribution of Age and Gender to account for Marital Status (the original 3-group variable). Another approach is based on our earlier decision to represent Marital Status as two 2-group comparisons (single vs. married & single vs. divorced, separated or widowed). Using this approach we separate the Marital Status box into two boxes, using each as a criterion when analyzing the 3rd above, and both as predictors when analyzing the 4th above. I prefer the second, because it gives more specific information about how the 2-group comparisons we have chosen to represent Marital Status relate to the other variables in the model.

Using this approach, we have five multiple regressions to complete:

1. Social Support from Friends as the criterion, with Age and Gender as predictors
2. Years Separated from Family as the criterion, with Age and Gender as predictors
3. mardc1 (single=0 vs. married=1) as the criterion, with Age and Gender as predictors
4. mardc2 (single=0 vs. divorced, separated or widowed=1) as the criterion, with Age and Gender as predictors
5. Loneliness as the criterion, with Marital Status (using mardc1 & mardc2) , Age, Gender, Years Separated from Family, and Social Support from Friends as the predictors

For each multiple regression we will obtain the related path coefficients as the β weights of the predictors and we will compute the e value as $\sqrt{(1-R^2)}$. When using mardc1 and mardc2 as criterion, we have to remember to select only the participants in the two related conditions of this variable, as we did for the correlations above. Here are the results of those analyses… (I've adjusted the SPSS output to show the R^2 values with greater precision).

Social Support from Friends (N=405)

Model Summary

Model	R	R Square
1	.297[a]	.08828

a. Predictors: (Constant), GENDC, AGE

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	6.299	.180		34.938	.000
	AGE	-.033	.006	-.278	-5.780	.000
	GENDC	.401	.126	.153	3.169	.002

a. Dependent Variable: friend social support

Years Separated from Family (N=405)

Model Summary

Model	R	R Square
1	.972[a]	.94538

a. Predictors: (Constant), GENDC, AGE

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-17.274	.353		-48.955	.000
	AGE	.926	.011	.965	81.864	.000
	GENDC	.953	.247	.045	3.852	.000

a. Dependent Variable: years separate from family

Single vs. Married (N=363 -- must select only participants in these two Marital Status groups**)**

Model Summary

Model	R	R Square
1	.783[a]	.61333

a. Predictors: (Constant), GENDC, AGE

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-.648	.045		-14.334	.000
	AGE	.036	.002	.780	23.662	.000
	GENDC	.021	.031	.023	.689	.491

a. Dependent Variable: MARDC1

Single vs. Divorced, Married or Widowed (N=284) -- must select only participants in these two Marital Status groups**)**

Model Summary

Model	R	R Square
1	.760[a]	.57821

a. Predictors: (Constant), GENDC, AGE

Social Support from Friends

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-.564	.039		-14.337	.000
	AGE	.028	.001	.748	19.116	.000
	GENDC	.048	.028	.068	1.733	.084

a. Dependent Variable: MARDC2

Loneliness (N=405)

Model Summary

Model	R	R Square
1	.593[a]	.35150

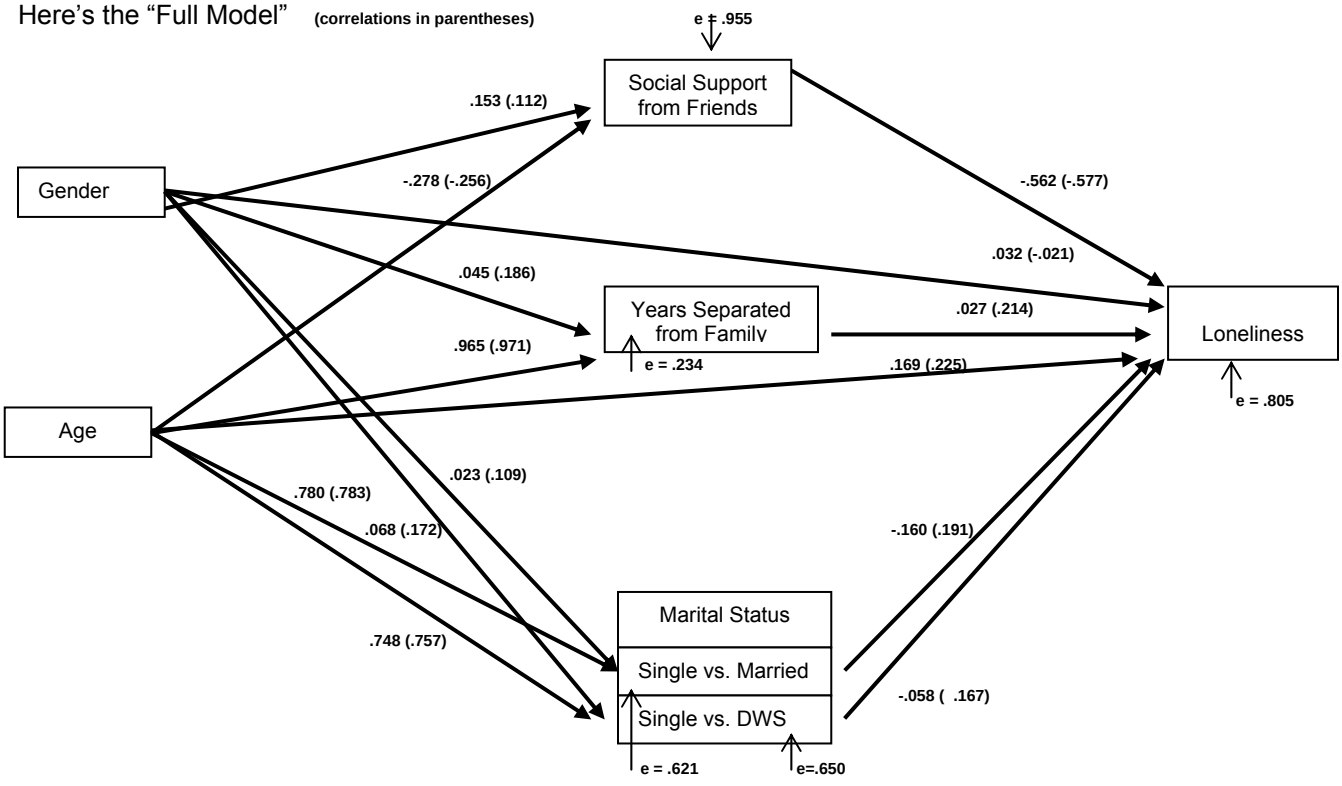
a. Predictors: (Constant), MARDC2, friend social support, GENDC, MARDC1, years separate from family, AGE

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	60.170	4.258		14.130	.000
	GENDC	.721	.964	.032	.748	.455
	AGE	.176	.183	.169	.964	.336
	friend social support	-4.891	.368	-.562	-13.285	.000
	years separate from family	.029	.190	.027	.154	.878
	MARDC1	-3.967	1.605	-.160	-2.472	.014
	MARDC2	-2.163	2.122	-.058	-1.019	.309

a. Dependent Variable: loneliness

Here's the "Full Model" **(correlations in parentheses)**



The Hypothesized (Reduced) Model

We will obtain the same set of five regressions, but now deleting from each those predictors that correspond to the paths dropped from the models.

Social Support from Friends – dropping Age (N=405)

Model Summary

Model	R	R Square
1	.112[a]	.01252

a. Predictors: (Constant), GENDC

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	5.407	.097		55.796	.000
	GENDC	.294	.130	.112	2.260	.024

a. Dependent Variable: friend social support

Years Separated from Family -- stays the same (N=405)

Model Summary

Model	R	R Square
1	.972[a]	.94538

a. Predictors: (Constant), GENDC, AGE

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-17.274	.353		-48.955	.000
	AGE	.926	.011	.965	81.864	.000
	GENDC	.953	.247	.045	3.852	.000

a. Dependent Variable: years separate from family

Single vs. Married -- **dropping GENDC (N=363** -- must select only participants in these two Marital Status groups**)**

Model Summary

Model	R	R Square
1	.783[a]	.61282

a. Predictors: (Constant), AGE

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-.640	.044		-14.696	.000
	AGE	.036	.002	.783	23.903	.000

a. Dependent Variable: MARDC1

Single vs. D, S or W -- **dropping GENDC (N=284)** -- must select only participants in these two Marital Status groups**)**

Model Summary

Model	R	R Square
1	.757[a]	.57370

a. Predictors: (Constant), AGE

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-.548	.038		-14.306	.000
	AGE	.029	.001	.757	19.481	.000

a. Dependent Variable: MARDC2

Loneliness – dropping Age (N=405)

Model Summary

Model	R	R Square
1	.592[a]	.34998

a. Predictors: (Constant), MARDC2, friend social support, GENDC, MARDC1, years separate from family

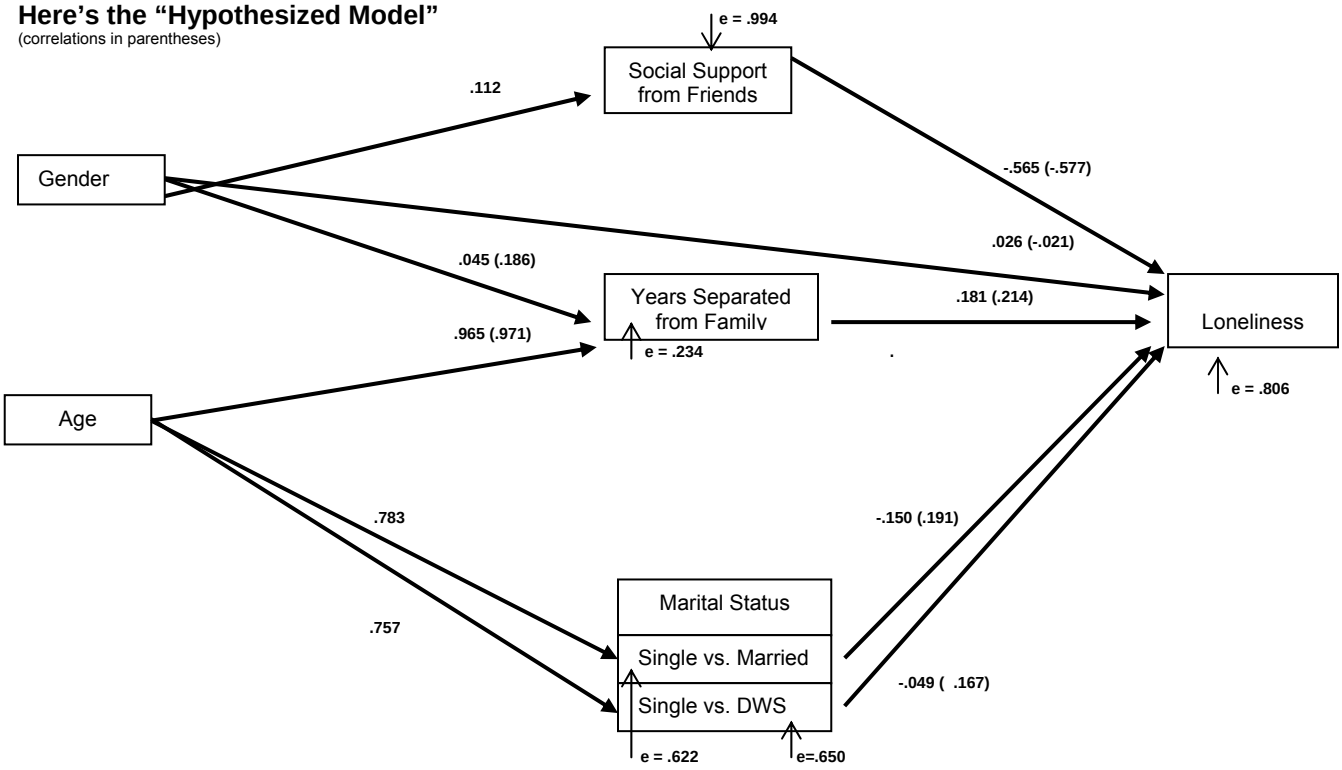
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	63.676	2.216		28.733	.000
	GENDC	.591	.954	.026	.620	.536
	friend social support	-4.917	.367	-.565	-13.393	.000
	years separate from family	.198	.075	.181	2.634	.009
	MARDC1	-3.713	1.583	-.150	-2.346	.019
	MARDC2	-1.842	2.096	-.049	-.879	.380

a. Dependent Variable: loneliness

Here's the "Hypothesized Model"

(correlations in parentheses)



Testing the Hypothesized Model

Testing the reduced model involves comparing how the hypothesized and full models fit the data.

$R^2_{Full} = 1 - \pi(e^2) = .9947$ $\quad$ $R^2_{Reduced} = 1 - \pi(e^2) = .9942$

`

The summary statistic showing the relative fit of the reduced model to the full model is

$$Q = \frac{1 - R^2_{Full}}{1 - R^2_{Reduced}} = \frac{1 - .9947}{1 - .9942} = .9102$$

The significance test comparing the fit of two models is (N = sample size d = number of dropped paths = 14 - 10)

$W = -(N - d) * \log_e Q = -(405 - 4) * \log_e .9102 = 37.74$ $\quad$ $X^2(df=4, p = .01) = 13.277$

We would conclude that the reduced model does not fit the data as well as the full model.