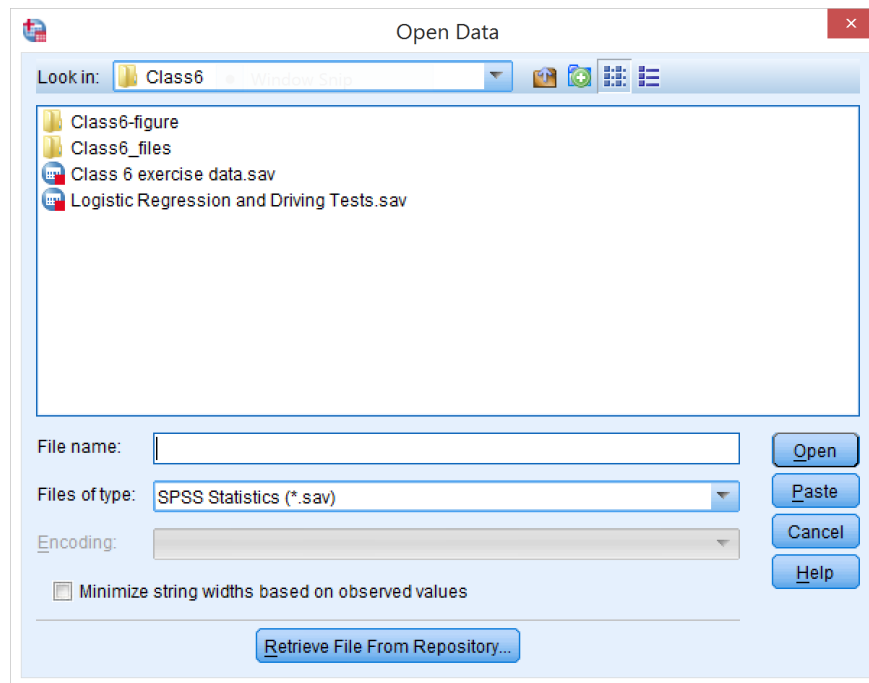


# Doing the Class 6 exercise in SPSS

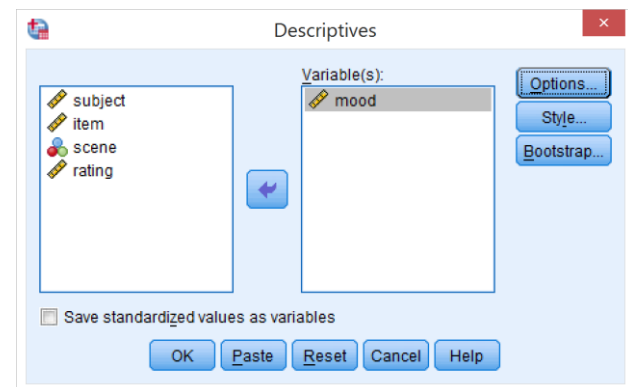
# Open the data file

- If you want to open the csv file directly, the instructions for that are in the “Homework 4 in SPSS” presentation on myBU.



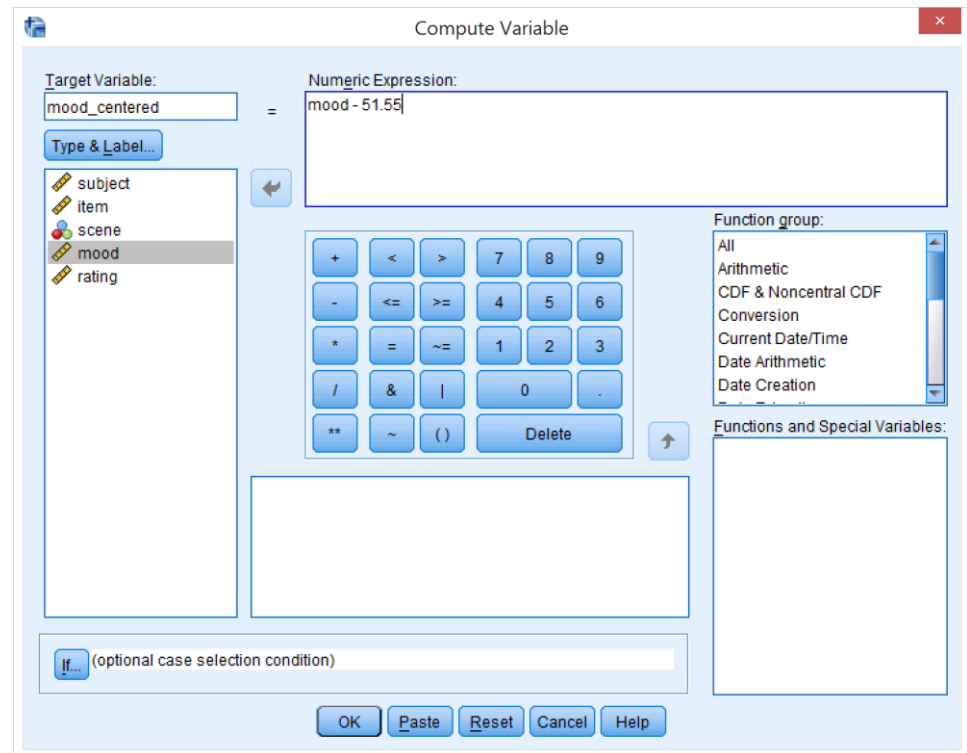
# Perform the linear regression

- We want to have two main effects and one interaction
- First thing to do: center the mood variable
- Need to know the mean first:
- Analyze
  - Descriptive Statistics
  - Descriptives
- Select the mood variable
- Click OK



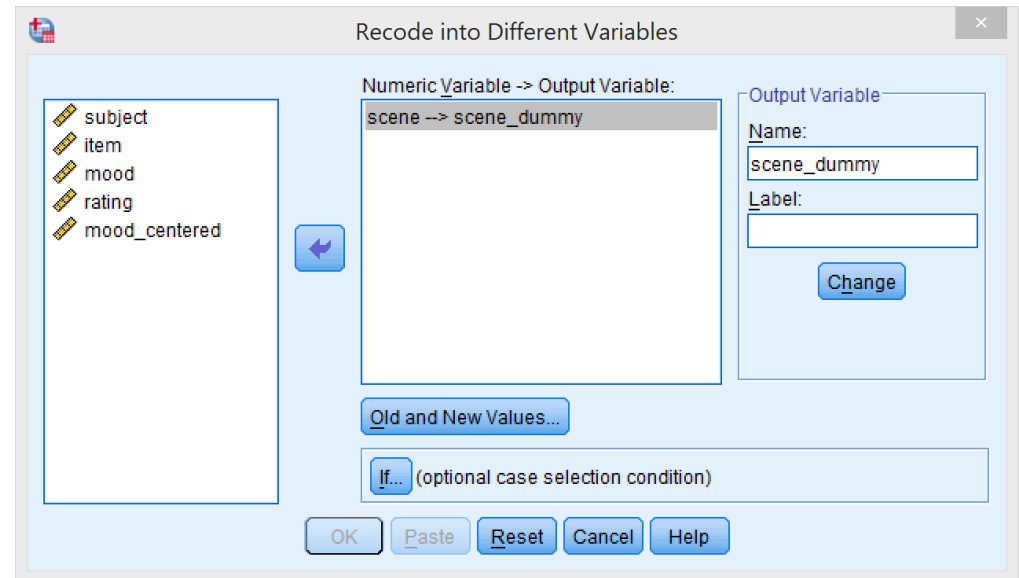
# Center the mood variable

- From the output we learn that the mean of the mood variable is 51.55
- Now make a new variable `mood_centered` by subtracting the mean from each value
- Transform → Compute Variable
- Enter the new variable name, the formula “`mood – 51.55`”
- Click OK



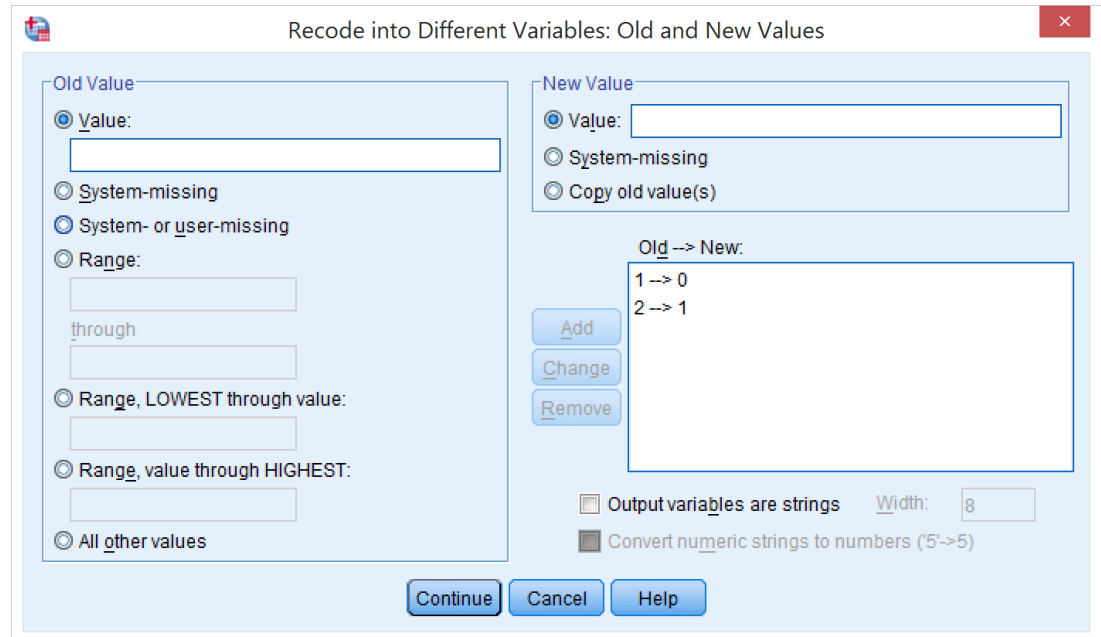
# Adding the interaction term

- SPSS forces you to compute the interaction term by hand if you want to add it to a regression model
- To start, we need to dummy-code “scene” to make sure that the correct contrasts are used
- Transform → Recode into different variables
- Select “scene”
- Choose a name for the output variable (“scene\_dummy”)
- Click Old and New Values



# Recoding scene

- Old and New Values
- Replace 1 with 0 and 2 with 1 by entering each pair of numbers in “Old value” and “New value” and pressing “Add”
- Click Continue
- Click OK



The image shows the "Recode into Different Variables: Old and New Values" dialog box in SPSS. The "Old Value" section on the left has the "Value:" radio button selected, with an empty text box below it. The "New Value" section on the right also has the "Value:" radio button selected, with an empty text box below it. In the center, the "Old --> New:" list contains two entries: "1 --> 0" and "2 --> 1". Below this list are three buttons: "Add", "Change", and "Remove". At the bottom right, there are two checkboxes: "Output variables are strings" (unchecked) and "Convert numeric strings to numbers ('5'-->5)" (checked). The "Width:" field next to the first checkbox is set to 8. At the bottom of the dialog are three buttons: "Continue", "Cancel", and "Help".

Recode into Different Variables: Old and New Values

**Old Value**

☒ Value:

☐ System-missing

☐ System- or user-missing

☐ Range:

through

☐ Range, LOWEST through value:

☐ Range, value through HIGHEST:

☐ All other values

**New Value**

☒ Value:

☐ System-missing

☐ Copy old value(s)

Old --> New:

1 --> 0

2 --> 1

Add

Change

Remove

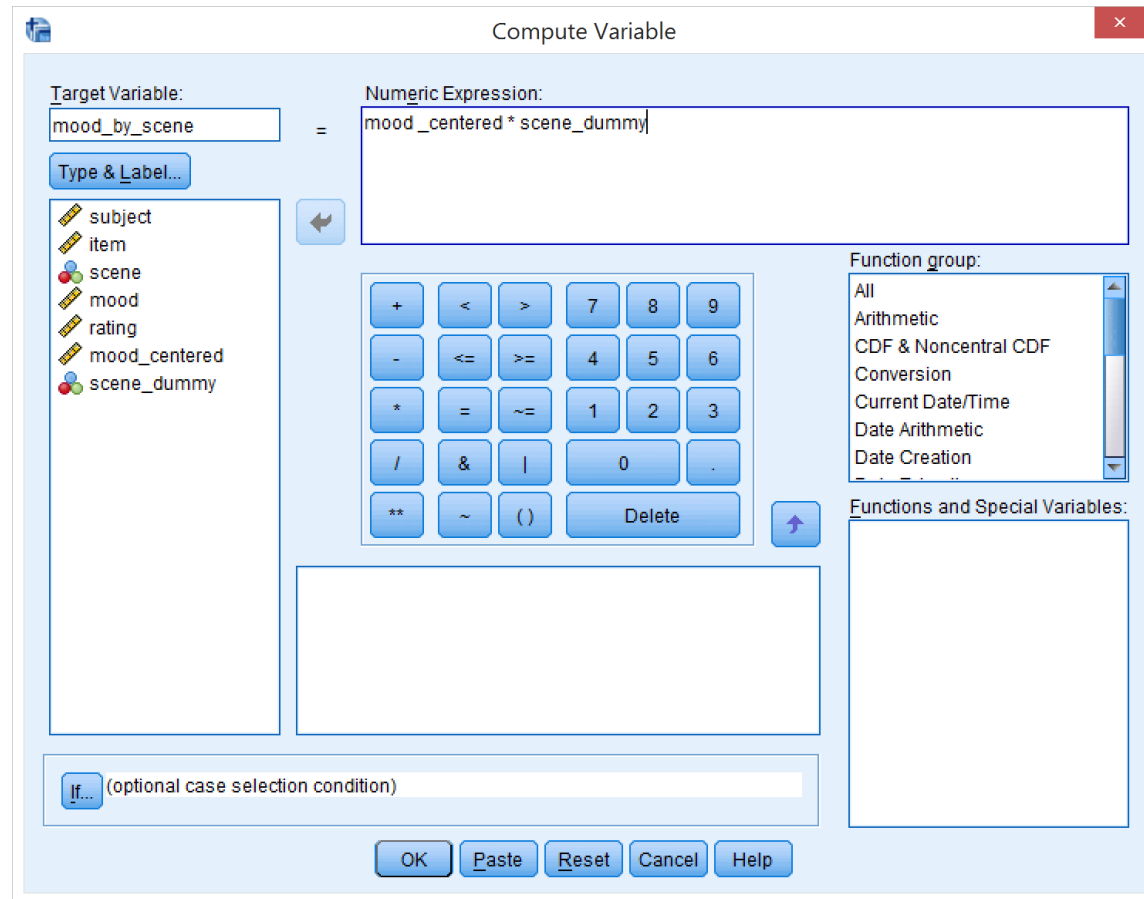
☐ Output variables are strings Width:

☒ Convert numeric strings to numbers ('5'-->5)

Continue Cancel Help

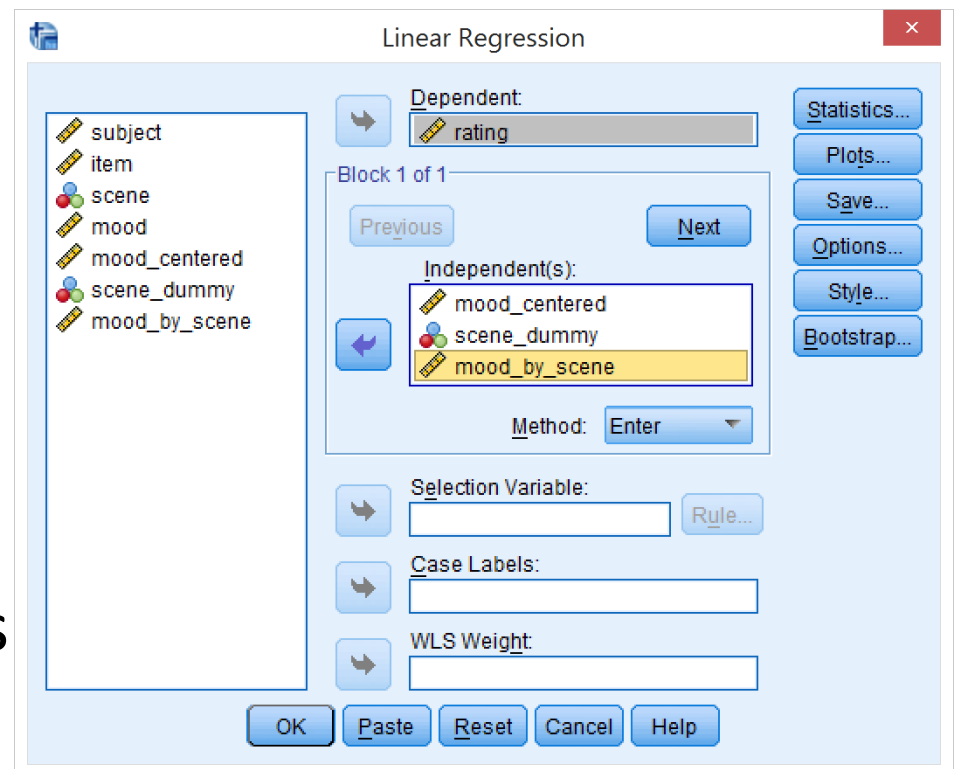
# Computing the interaction term

- Transform → Compute Variable
- Give it a good name, e.g. mood\_by\_scene
- Enter the formula:  
mood\_centered  
\* scene\_dummy
- Click OK



# Running the regression

- Analyze → Regression  
→ Linear
- Enter rating as the dependent variable
- Enter the centered mood variable, the recoded scene variable, and the interaction term as independent variables





# Running the regression (2)

- Click Statistics...
- Select Descriptives
- Select Collinearity diagnostics
- Click Continue

Linear Regression: Statistics

**Regression Coefficients**

- ☒ Estimates
- ☐ Confidence intervals  
Level(%): 95
- ☐ Covariance matrix

**Model fit**

- ☒ Model fit
- ☐ R squared change
- ☒ Descriptives
- ☐ Part and partial correlations
- ☒ Collinearity diagnostics

**Residuals**

- ☐ Durbin-Watson
- ☐ Casewise diagnostics
  - ☒ Outliers outside: 3 standard deviations
  - ☐ All cases

Continue Cancel Help

# Running the regression (3)

- Click Save...
- Select Standardized Residuals
- Select Cook's distance
- Click Continue
- Click OK to run the regression

Linear Regression: Save

**Predicted Values**

- ☐ Unstandardized
- ☐ Standardized
- ☐ Adjusted
- ☐ S.E. of mean predictions

**Residuals**

- ☐ Unstandardized
- ☒ Standardized
- ☐ Studentized
- ☐ Deleted
- ☐ Studentized deleted

**Distances**

- ☐ Mahalanobis
- ☒ Cook's
- ☐ Leverage values

**Prediction Intervals**

- ☐ Mean ☐ Individual
- Confidence Interval: 95 %

**Influence Statistics**

- ☐ DfBeta(s)
- ☐ Standardized DfBeta(s)
- ☐ DfFit
- ☐ Standardized DfFit
- ☐ Covariance ratio

**Coefficient statistics**

- ☐ Create coefficient statistics
- ☒ Create a new dataset  
Dataset name:
- ☐ Write a new data file

**Export model information to XML file**

☒ Include the covariance matrix

# Regression results

- Look at the output: Nothing is significant
  - No multicollinearity issues (max VIF = 2)

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	40.283	3	13.428	1.926	.123 <sup>b</sup>
	Residual	11128.235	1596	6.973		
	Total	11168.518	1599			

a. Dependent Variable: rating

b. Predictors: (Constant), mood\_by\_scene, scene\_dummy, mood\_centered

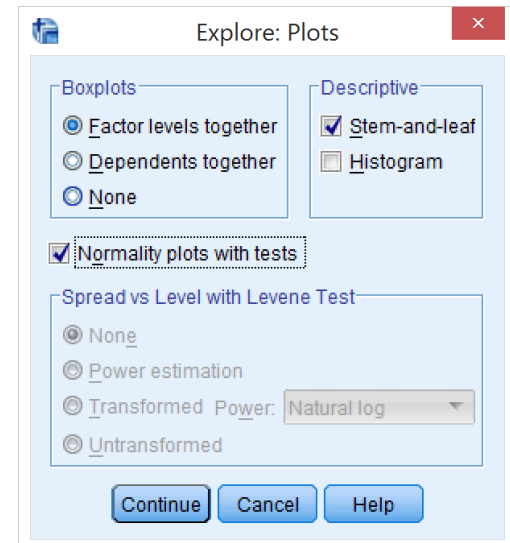
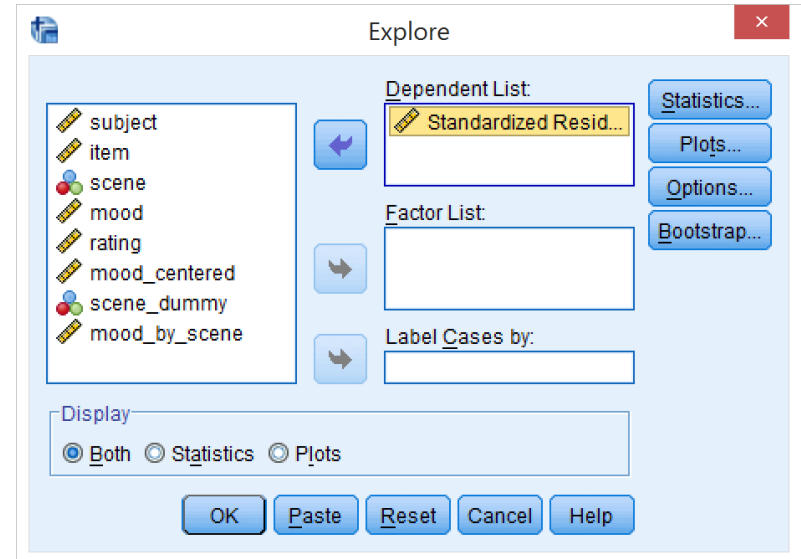
**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	11.328	.093		121.342	.000		
	mood_centered	-.007	.004	-.058	-1.650	.099	.500	2.000
	scene_dummy	-.208	.132	-.039	-1.579	.114	1.000	1.000
	mood_by_scene	.011	.006	.060	1.696	.090	.500	2.000

a. Dependent Variable: rating

# Regression results

- Check normality of residuals
  - Analyze → Descriptives → Explore
- Select the new standardized residuals variable
- Click “Plots...”
- Select Normality plots with tests
- Click Continue
- Click OK



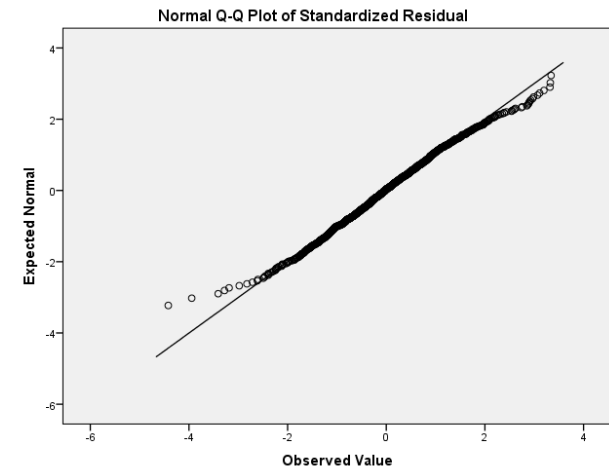
# Testing normality of the residuals

Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.020	1600	.129	.995	1600	.000

a. Lilliefors Significance Correction

- Shapiro-Wilk shows a significant deviation from normality (although K-F doesn't)
- Visual inspection of the normal Quantile-Quantile plot also shows deviations from normality

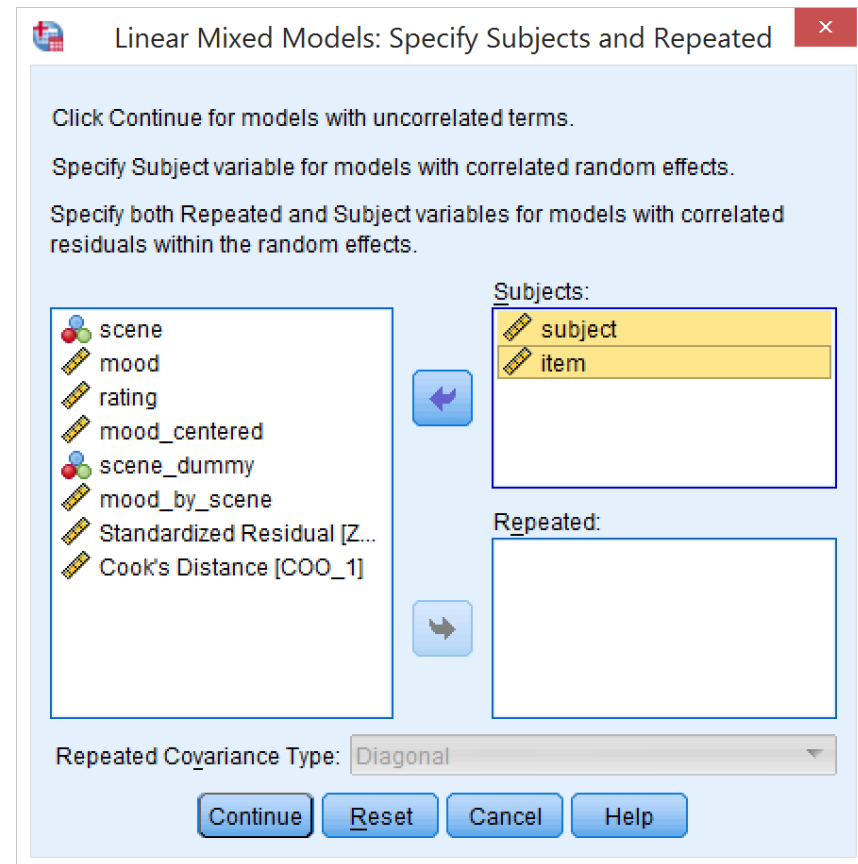


# Regression results

- Outliers: in the new Cook's D column (COO\_1), check if any values are  $> 1$

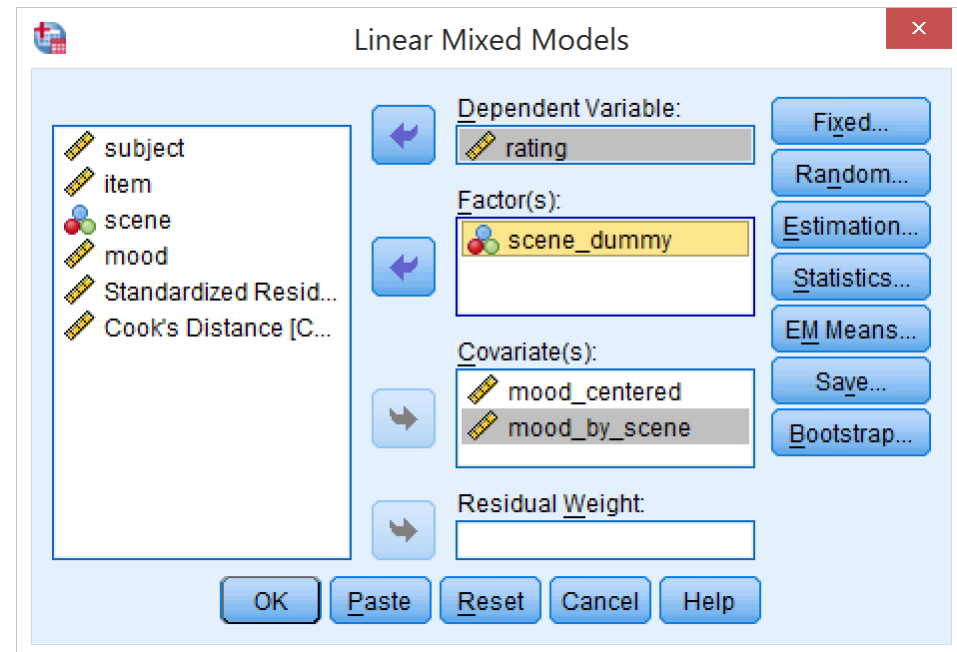
# Linear mixed model

- Analyze → Mixed → Linear
- Enter “subject” and “item” as Subjects
- Click Continue



# Linear mixed model

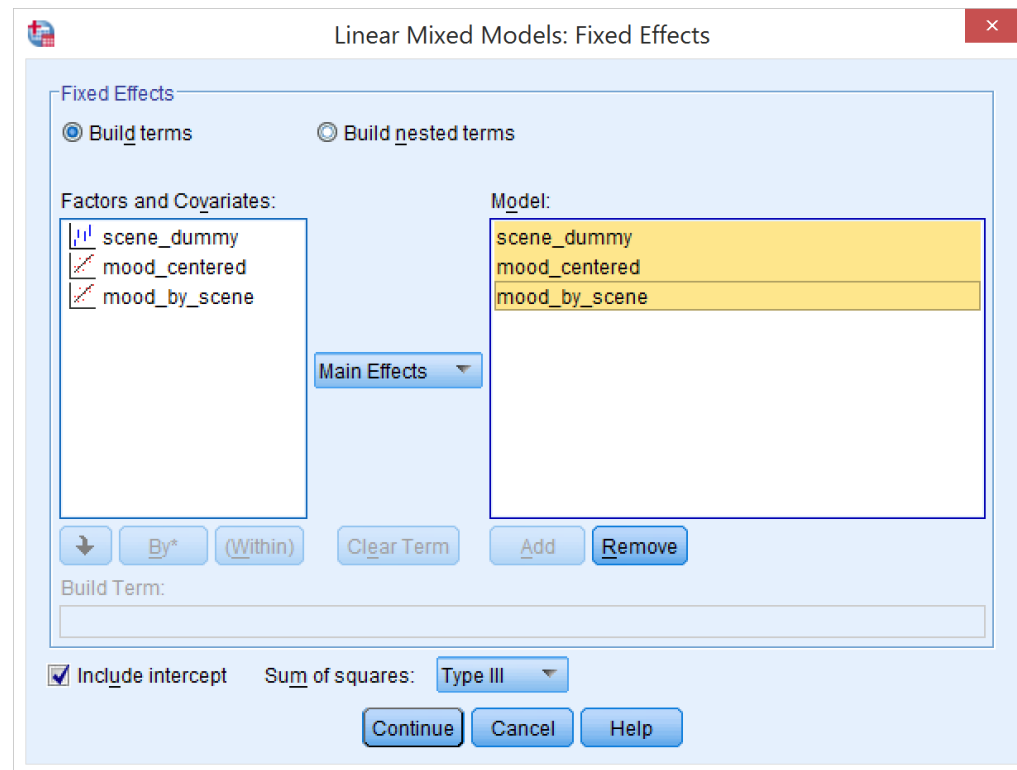
- Enter “rating” as the dependent variable
- Enter the recoded scene variable as a factor
- Enter the centered mood variable as a covariate
- Enter the interaction term as a covariate
- Click “Fixed”





# Linear mixed model

- Select all three factors/covariates
- In the drop-down menu, select “Main effects” (we coded the interaction term by hand, so to SPSS it looks like a main effect)
- Click Add
- Click Continue



# Linear mixed model

- Click Random...
- Define Subject random effects:
  - Enter “subject” into the “Combinations” filed
  - Select “Include intercept”
  - Click “Next”

Linear Mixed Models: Random Effects

Random Effect 1 of 1

Previous Next

Covariance Type: Variance Components

Random Effects

☒ Build terms ☐ Build nested terms ☒ Include intercept

Factors and Covariates: scene\_dummy mood\_centered mood\_by\_scene

Model:

Factorial

Build Term:

Subject Groupings

Subjects: subject item

Combinations: subject

Continue Cancel Help

# Linear mixed model

- Click Random...
- Define Item random effects:
  - Enter “item” into the “Combinations” field
  - Select “Include intercept”
  - Select “mood\_centered”
  - Click “Add”
  - In the Covariance Type dropdown menu, select “Unstructured”
  - Click “Next”

Note: you don't need any random slopes for covariates in the assignment – the intercept is sufficient

Linear Mixed Models: Random Effects

Random Effect 2 of 2

Previous Next

Covariance Type: Unstructured

Random Effects

☒ Build terms ☐ Build nested terms ☒ Include intercept

Factors and Covariates:

- scene\_dummy
- mood\_centered
- mood\_by\_scene

Model:

- mood\_centered

Factorial

Build Term:

Subject Groupings

Subjects:

- subject
- item

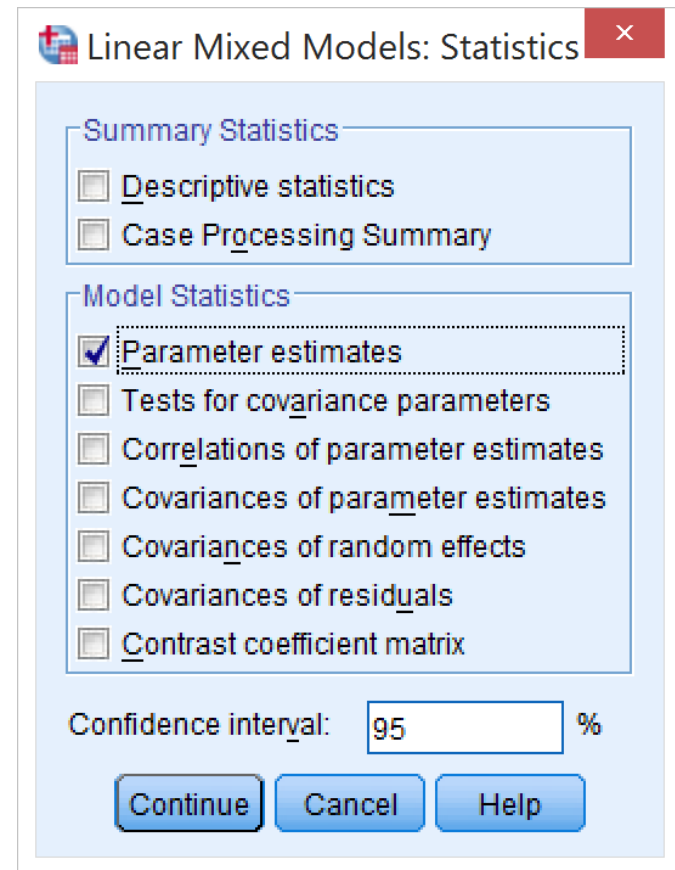
Combinations:

- item

Continue Cancel Help

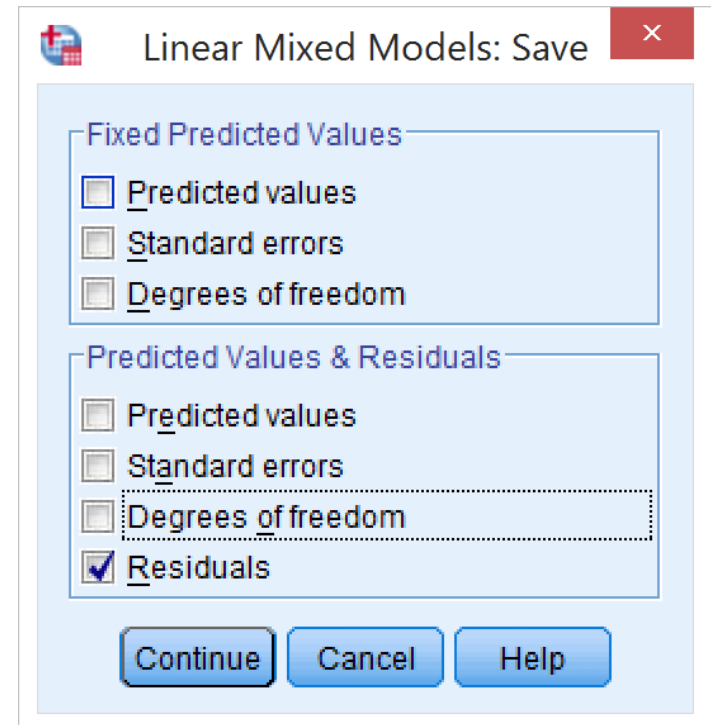
# Linear mixed model

- Final touches
- Click “Statistics”
- Select “Parameter estimates”
- Click Continue



# Linear mixed model

- Click Save
- Select Residuals
- Click continue
- Click OK to run the model



# Linear mixed model results

- Both scene and mood by scene are significant
  - Report the coefficient, t, and p values of the significant predictors

Estimates of Fixed Effects<sup>a</sup>

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	11.092576	.118148	77.302	93.887	.000	10.857327	11.327825
[scene_dummy=.00]	.262848	.097945	1482.420	2.684	.007	.070723	.454973
[scene_dummy=1.00]	0 <sup>b</sup>	0	.	.	.	.	.
mood_centered	-.007034	.013662	45.596	-.515	.609	-.034542	.020474
mood_by_scene	.010055	.004662	1501.837	2.157	.031	.000910	.019199

a. Dependent Variable: rating.

b. This parameter is set to zero because it is redundant.

# Linear mixed model residuals

- Analyze → Descriptive Statistics → Explore
- Same procedure as with the linear regression, just choose RESID\_1 this time
- Now the Shapiro-Wilk test is no longer significant

**Tests of Normality**

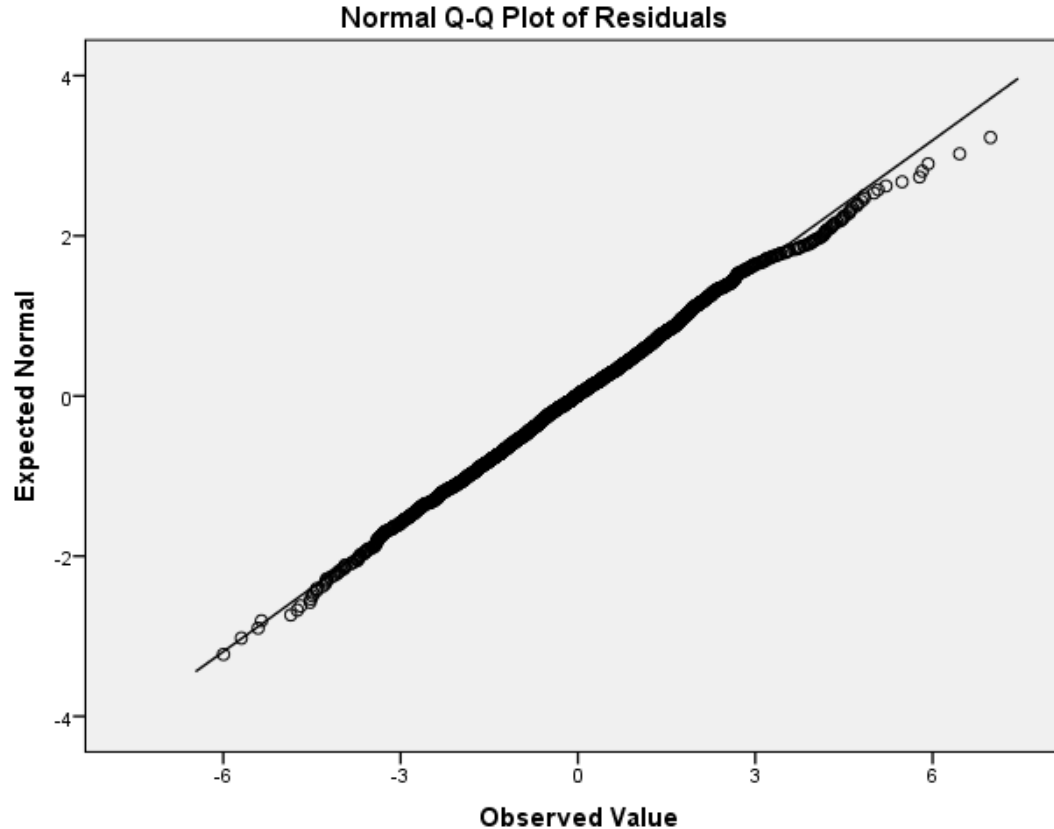
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residuals	.016	1600	.200 <sup>*</sup>	.998	1600	.075

<sup>\*</sup>. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

# Linear mixed model residuals

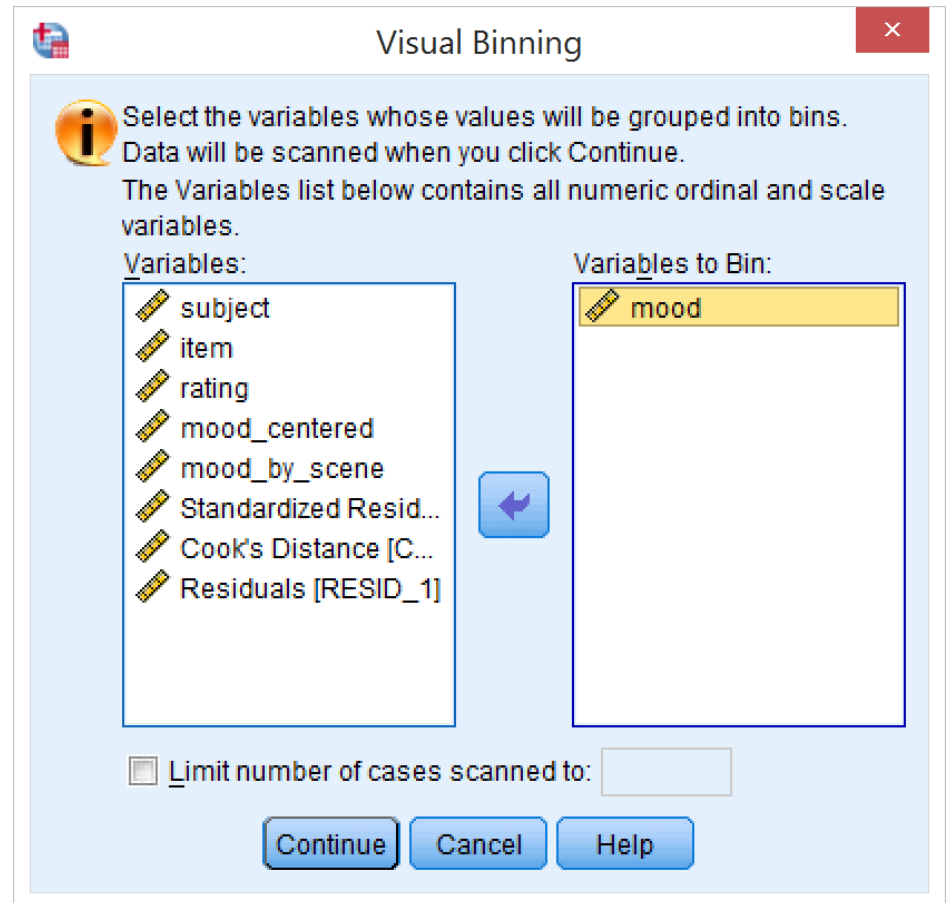
- The Quantile-Quantile plot also looks much better





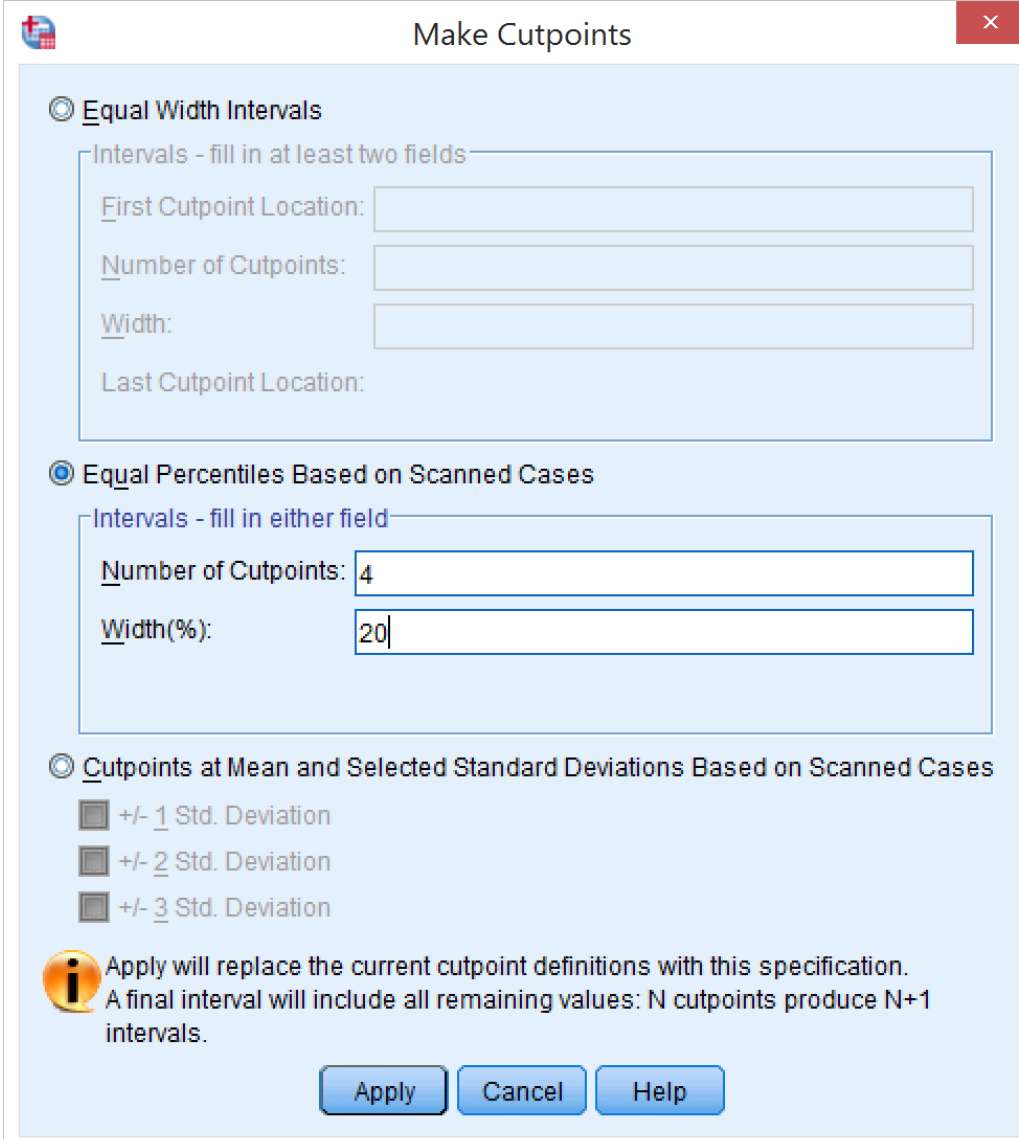
# Plotting the mood by scene interaction

- Need to split the mood variable into bins
- Transform → Visual binning
- Select “mood”
- Click Continue



# Plotting the mood by scene interaction

- Click Make Cutpoints
- Click “Equal Percentiles”
- Enter 20 in the Width(%) field for 5 equally sized bins
- Click Apply
- Click OK



**Make Cutpoints**

☐ Equal Width Intervals

Intervals - fill in at least two fields

First Cutpoint Location:

Number of Cutpoints:

Width:

Last Cutpoint Location:

☒ Equal Percentiles Based on Scanned Cases

Intervals - fill in either field

Number of Cutpoints:


Width(%):

☐ Cutpoints at Mean and Selected Standard Deviations Based on Scanned Cases

☐ +/- 1 Std. Deviation

☐ +/- 2 Std. Deviation

☐ +/- 3 Std. Deviation

 Apply will replace the current cutpoint definitions with this specification.  
A final interval will include all remaining values: N cutpoints produce N+1 intervals.

# Binning mood

- Choose a comprehensible name for the new binned variable and enter it in “Binned Variable”
- Press OK

Visual Binning

Scanned Variable List:

- mood

Current Variable: mood

Binned Variable: Mood\_binned

Label: mood (Binned)

Minimum: 5 Nonmissing Values Maximum: 94

Enter interval cutpoints or click Make Cutpoints for automatic intervals. A cutpoint value of 10, for example, defines an interval starting above the previous interval and ending at 10.

Grid:

	Value	Label
1	37.5	
2	44.0	
3	58.5	
4	69.5	
5	HIGH	
6		

Upper Endpoints

☒ Included (<=)

☐ Excluded (<)

Make Cutpoints...

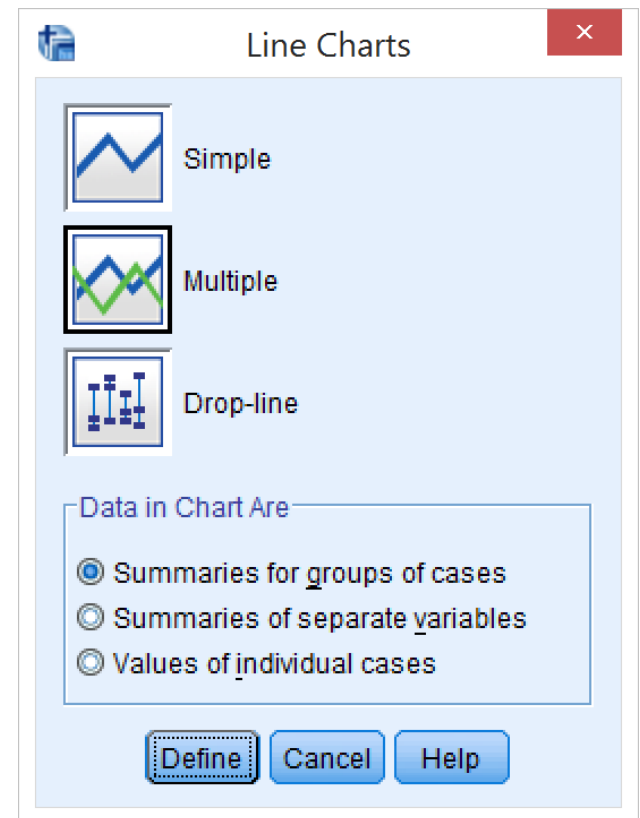
Make Labels

☐ Reverse scale

OK Paste Reset Cancel Help

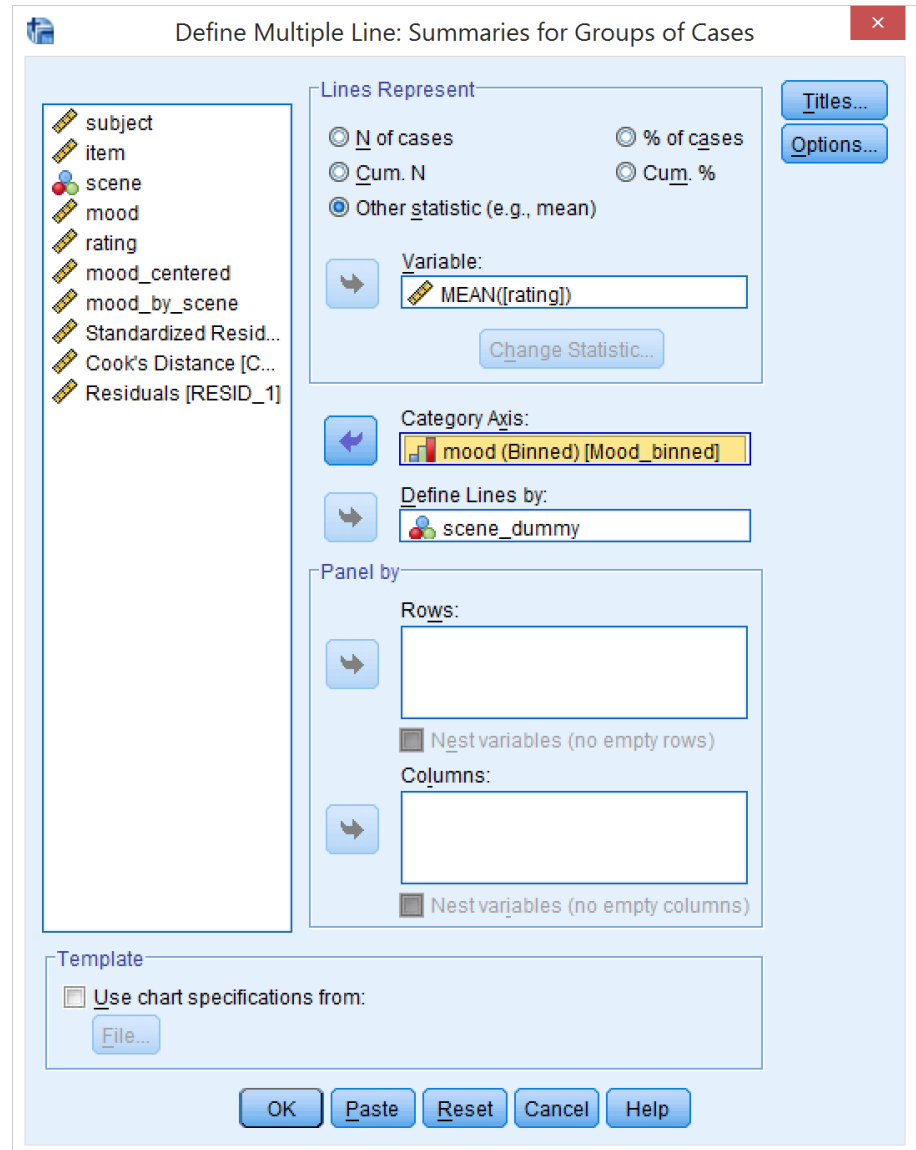
# Making the plot

- Graphs → Legacy Dialogs → Line
- Select “Multiple” and “Summaries for groups of cases”
- Click Define



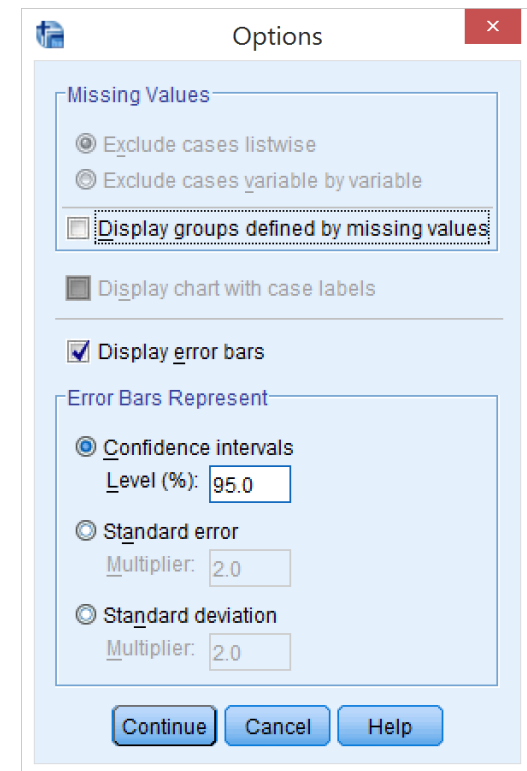
# Line plot

- Select “Other statistic”
- Select rating as Variable
- Select the binned mood variable as the Category Axis
- Select “scene” as “Define lines by”
- Click “Options”



# Line plot

- Select “Display error bars”
- Select 95% confidence intervals
- Click Continue
- Click OK



# Line plot

- There you go. Now just fix the axis labels and the legend

