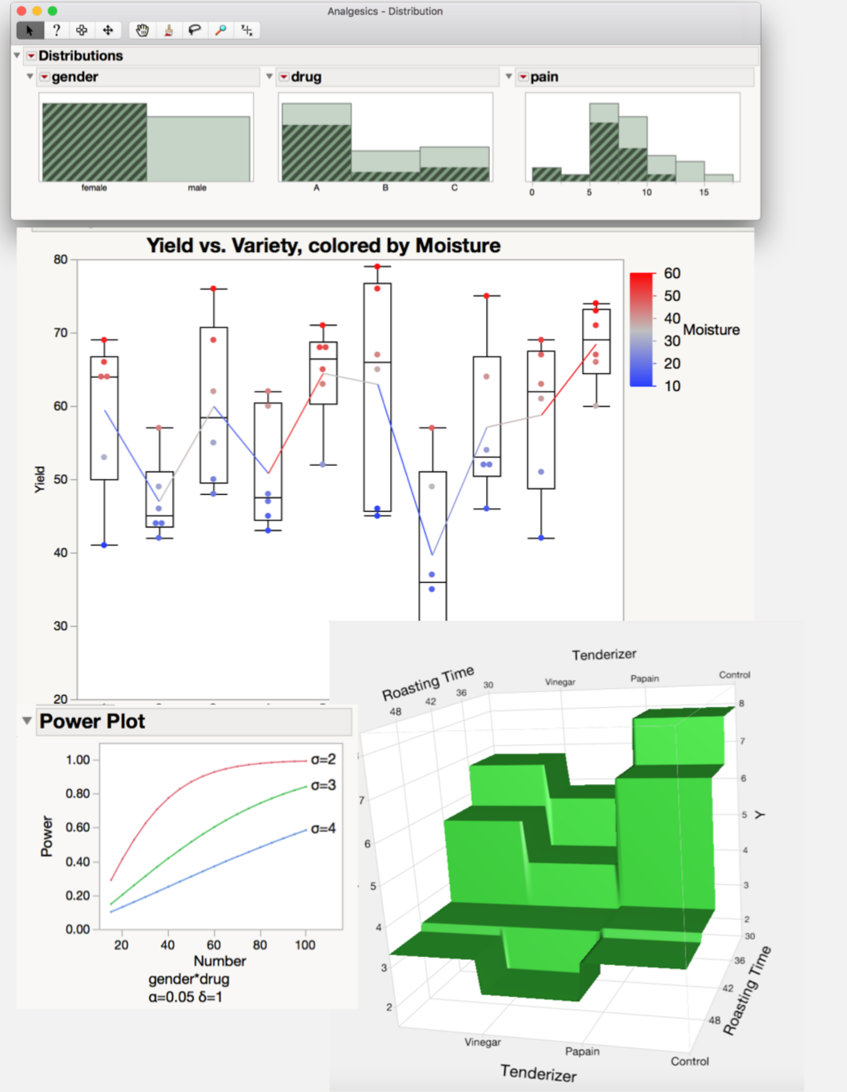
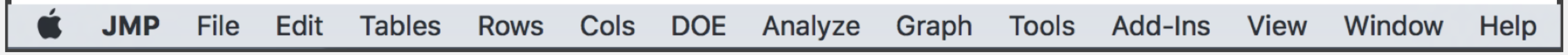
JMP FOR BIOSTATISTICS AND THE HEALTH AND LIFE SCIENCES

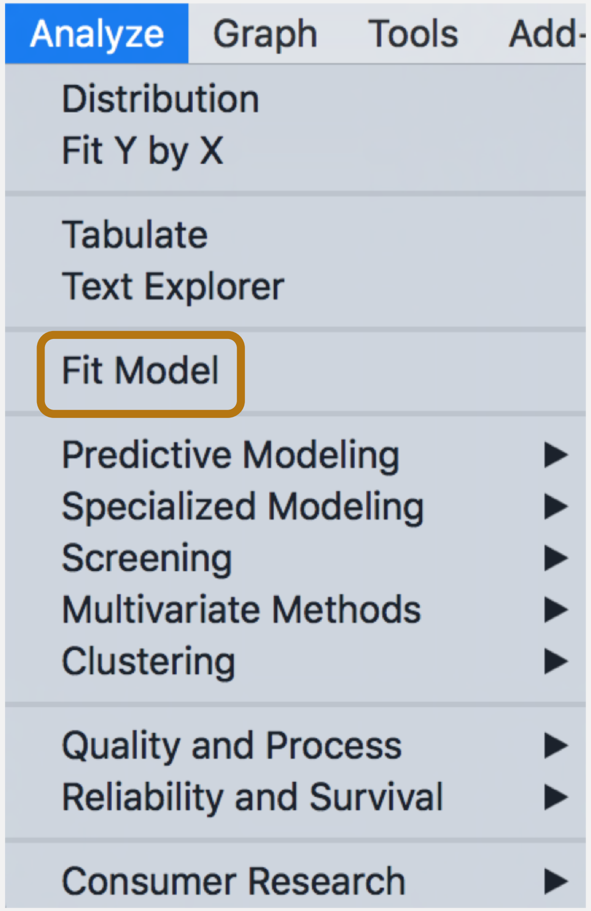
We are talking about: *how to use the tools and techniques commonly needed by researchers, practitioners, professors, and students in biostatistics and the health and life sciences fields.*

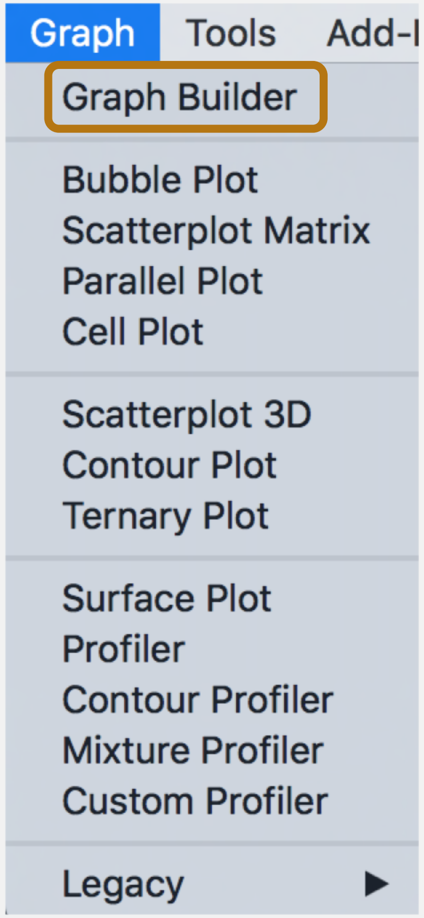
- **ANOVA and regression** (including variable selection and penalized regression)
- **mixed models** (hierarchical/split-plot and repeated measures)
- **survival analysis**
- **designed experiments**

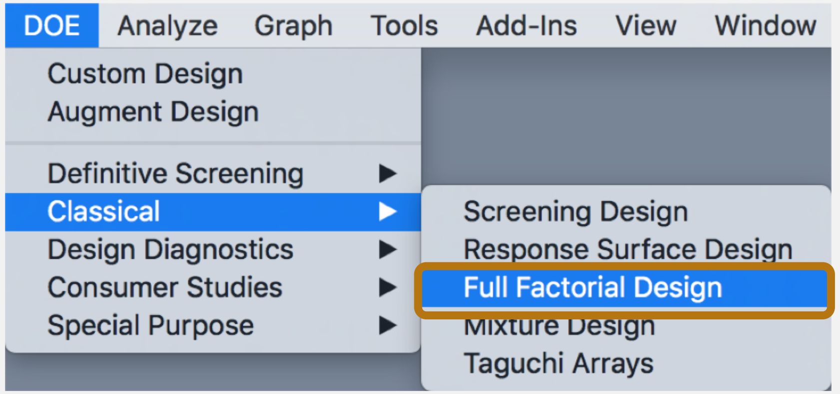


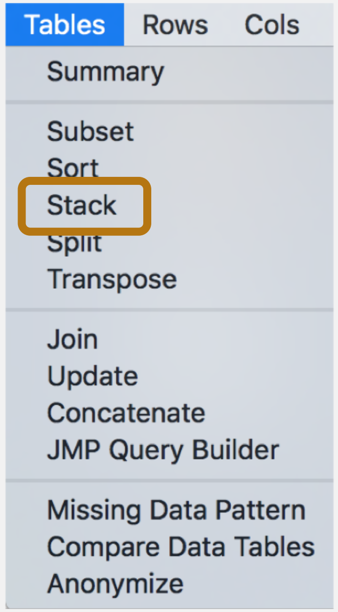
USING JMP

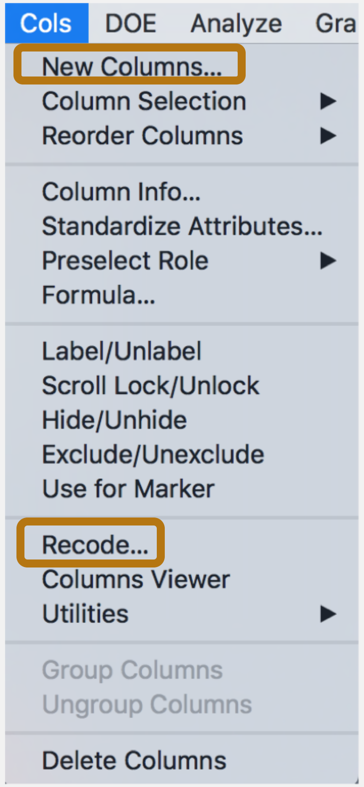

JMP
File
Edit
Tables
Rows
Cols
DOE
Analyze
Graph
Tools
Add-Ins
View
Window
Help


Analyze
Graph
Tools
Add-
Distribution
Fit Y by X
Tabulate
Text Explorer
Fit Model
Predictive Modeling
Specialized Modeling
Screening
Multivariate Methods
Clustering
Quality and Process
Reliability and Survival
Consumer Research


Graph
Tools
Add-I
Graph Builder
Bubble Plot
Scatterplot Matrix
Parallel Plot
Cell Plot
Scatterplot 3D
Contour Plot
Ternary Plot
Surface Plot
Profiler
Contour Profiler
Mixture Profiler
Custom Profiler
Legacy

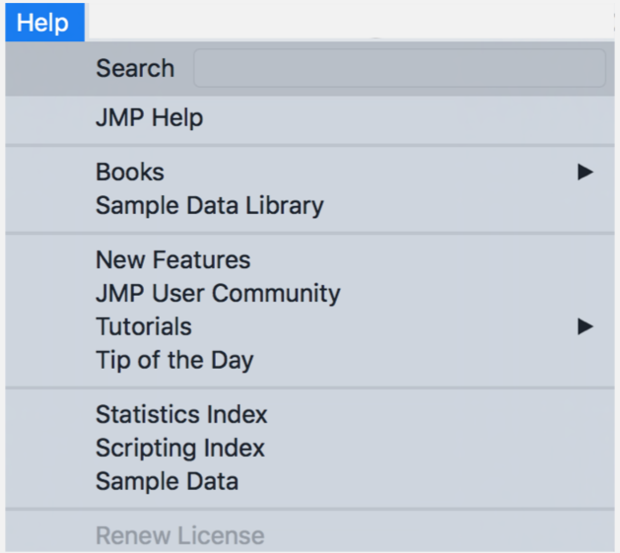

DOE
Analyze
Graph
Tools
Add-Ins
View
Window
Custom Design
Augment Design
Definitive Screening
Classical
Design Diagnostics
Consumer Studies
Special Purpose
Screening Design
Response Surface Design
Full Factorial Design
Mixture Design
Taguchi Arrays


Tables
Rows
Cols
Summary
Subset
Sort
Stack
Split
Transpose
Join
Update
Concatenate
JMP Query Builder
Missing Data Pattern
Compare Data Tables
Anonymize


Cols
DOE
Analyze
Gra
New Columns...
Column Selection
Reorder Columns
Column Info...
Standardize Attributes...
Preselect Role
Formula...
Label/Unlabel
Scroll Lock/Unlock
Hide/Unhide
Exclude/Unexclude
Use for Marker
Recode...
Columns Viewer
Utilities
Group Columns
Ungroup Columns
Delete Columns

GETTING HELP IN JMP

- www.jmp.com/webinar → "Data Summary and Analysis"
- Other resources at www.jmp.com/teach
- Within JMP, the Help Menu (or the online help)



- The JMP User Community! community.jmp.com

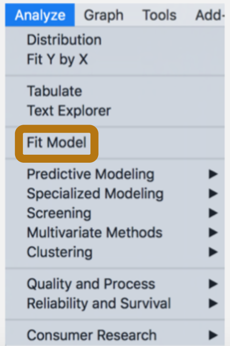
ANOVA AND REGRESSION

1. ANOVA to compare the pain ratings of men and women taking one of three different analgesics

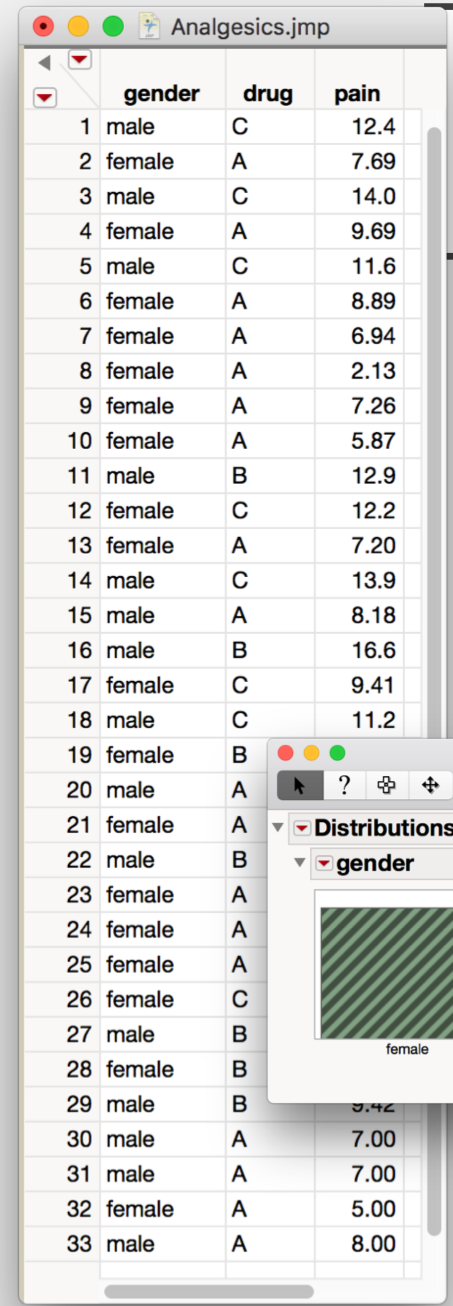
2. Regression to predict how long an animal sleeps per night based on other characteristics

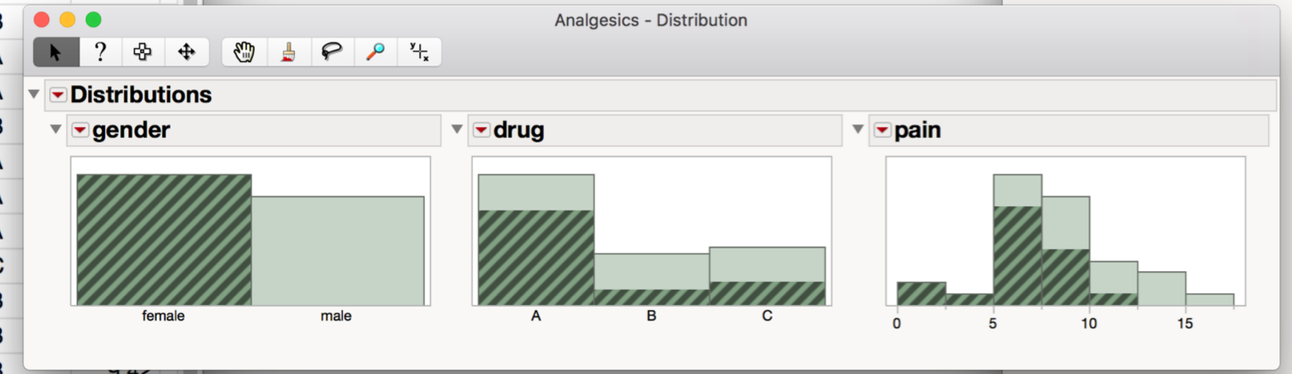
ANOVA

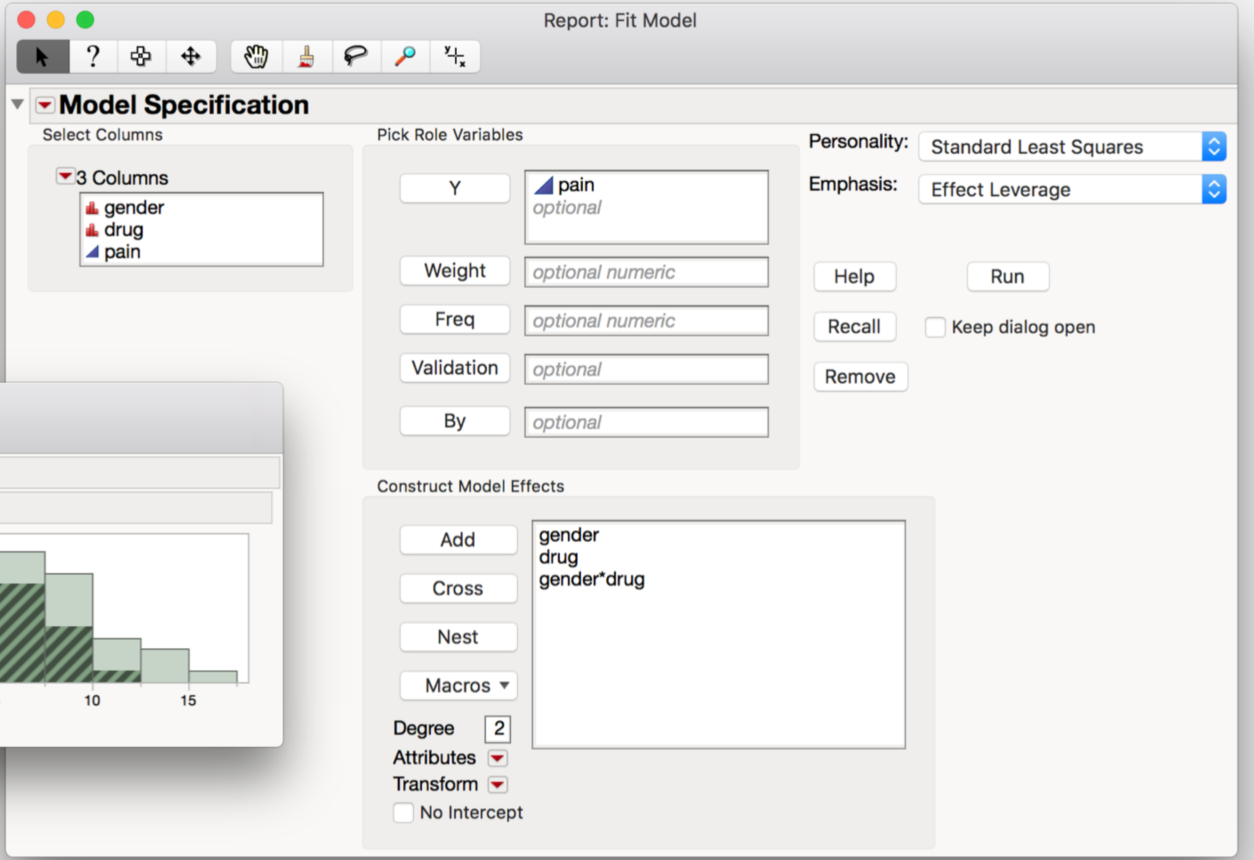
Go to **Analyze → Fit Model**



Patients experiencing pain are randomized to one of three different analgesics and asked to rate their pain.





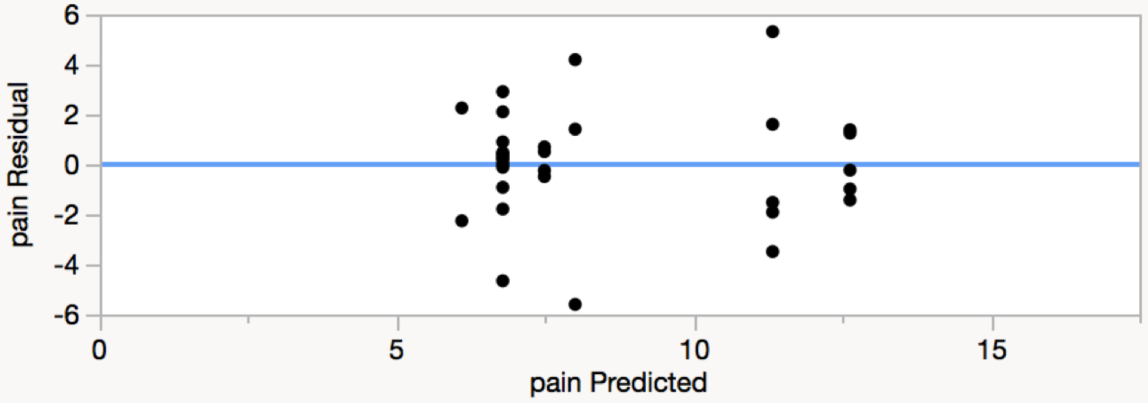


Effect Summary

Source	LogWorth	PValue
gender	2.848	0.00142
drug	1.645	0.02267
gender*drug	1.038	0.09161

Remove Add Edit FDR

Residual by Predicted Plot



Summary of Fit

RSquare	0.534258
RSquare Adj	0.448009
Root Mean Square Error	2.41667
Mean of Response	8.533515
Observations (or Sum Wgts)	33

Analysis of Variance

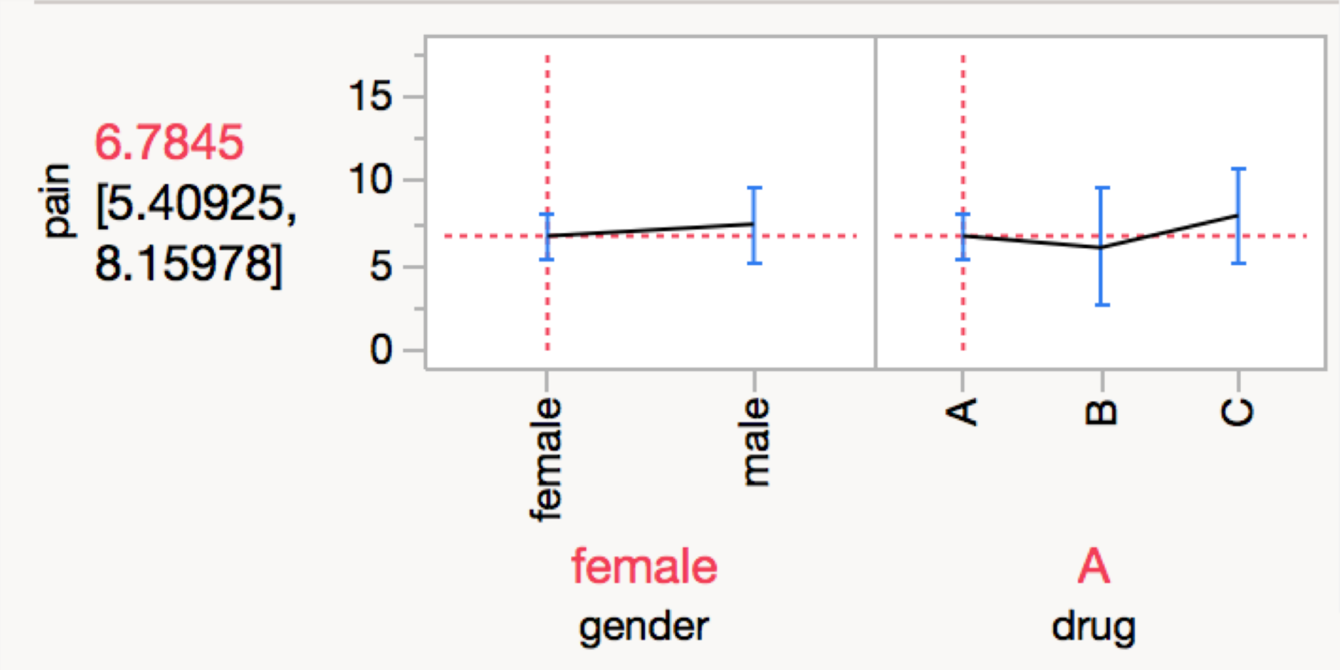
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	180.88537	36.1771	6.1944
Error	27	157.68799	5.8403	Prob > F
C. Total	32	338.57335		0.0006*

Parameter Estimates

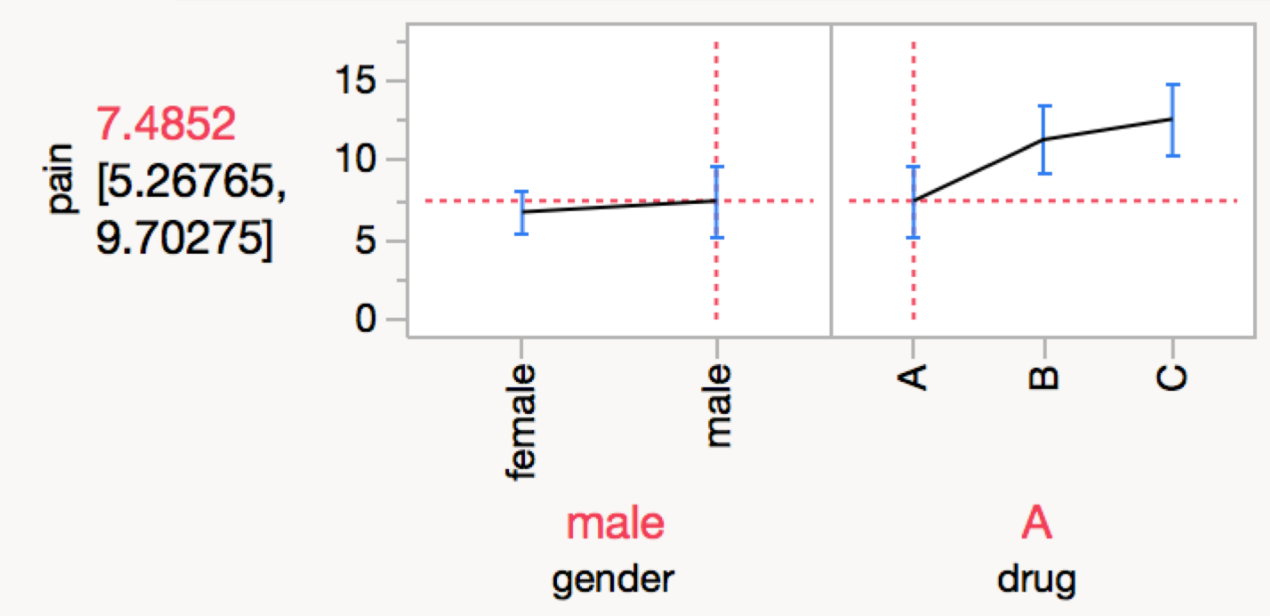
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	8.7207484	0.494984	17.62	<.0001*
gender[female]	-1.759652	0.494984	-3.55	0.0014*
drug[A]	-1.585892	0.616267	-2.57	0.0159*
drug[B]	-0.009982	0.765306	-0.01	0.9897
gender[female]*drug[A]	1.409308	0.616267	2.29	0.0303*
gender[female]*drug[B]	-0.856115	0.765306	-1.12	0.2731

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
gender	1	1	73.808295	12.6378	0.0014*
drug	2	2	51.059196	4.3713	0.0227*
gender*drug	2	2	30.542763	2.6148	0.0916



6.7845
pain
[5.40925,
8.15978]
15
10
5
0
female
male
A
B
C
female
gender
A
drug



7.4852
pain
[5.26765,
9.70275]
15
10
5
0
female
male
A
B
C
male
gender
A
drug

You can also:

- Save residuals, predicted, confidence intervals, hats, Cook's D, leverage
- Look at profilers, like the interaction profiler or the surface profiler
- Get the indicator parameterization estimates
- Get sequential tests, multiple comparisons adjustments, estimates, contrasts, slices
- Look at retrospective power and sample size analyses

Least Squares Means Table

Level	Least Sq Mean	Std Error	Mean
female	6.961097	0.76856047	6.9111
male	10.480400	0.62398161	10.4804

LSMeans Differences Student's t

α= 0.050 t= 2.05183

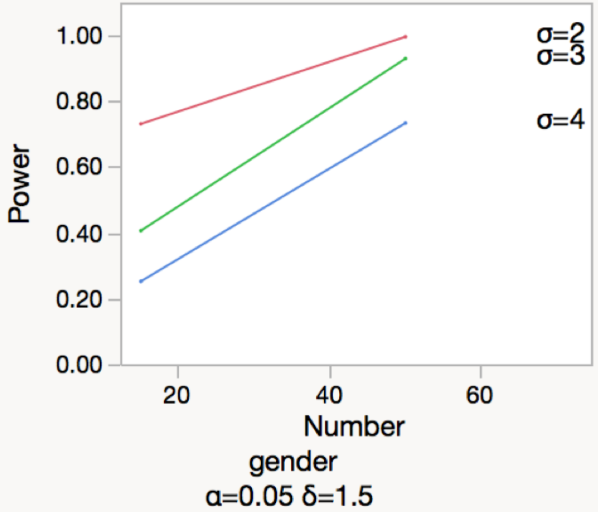
Level			Least Sq Mean
male	A		10.480400
female		B	6.961097

Levels not connected by same letter are significantly different.

Power

α	σ	δ	Number	Power
0.0500	2	1.5	15	0.7344
0.0500	2	1.5	50	0.9994
0.0500	3	1.5	15	0.4093
0.0500	3	1.5	50	0.9329
0.0500	4	1.5	15	0.2552
0.0500	4	1.5	50	0.7368

Power Plot



Least Squares Means Table

Level	Least Sq Mean	Std Error	Mean
A	7.134856	0.6358680	6.9791
B	8.710767	1.0109658	9.8318
C	10.316622	0.8824433	10.8948

LSMeans Differences Tukey HSD

α= 0.050 Q= 2.47942

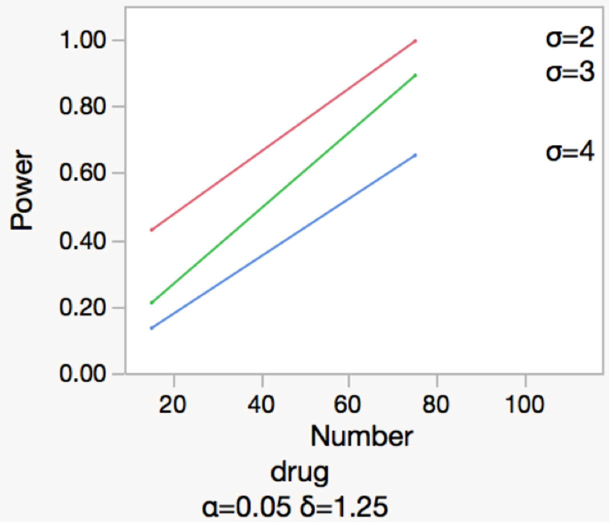
Level			Least Sq Mean
C	A		10.316622
B	A	B	8.710767
A		B	7.134856

Levels not connected by same letter are significantly different.

Power

α	σ	δ	Number	Power
0.0500	2	1.25	15	0.4322
0.0500	2	1.25	75	0.9986
0.0500	3	1.25	15	0.2139
0.0500	3	1.25	75	0.8953
0.0500	4	1.25	15	0.1382
0.0500	4	1.25	75	0.6562

Power Plot



Least Squares Means Table

Level	Least Sq Mean	Std Error
female,A	6.784513	0.6702638
female,B	6.095000	1.7088440
female,C	8.003778	1.3952653
male,A	7.485200	1.0807679
male,B	11.326533	1.0807679
male,C	12.629467	1.0807679

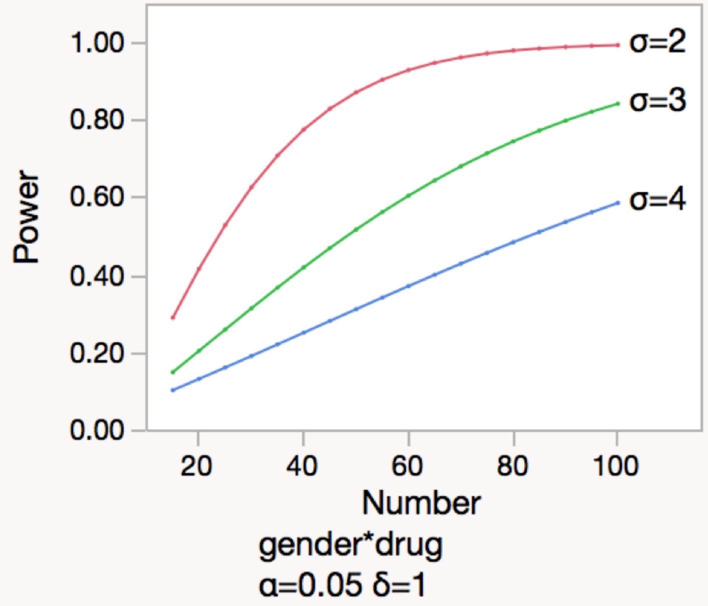
LSMeans Differences Tukey HSD

α= 0.050 Q= 3.06385

Level				Least Sq Mean
male,C	A			12.629467
male,B	A	B		11.326533
female,C	A	B	C	8.003778
male,A		B	C	7.485200
female,A			C	6.784513
female,B		B	C	6.095000

Levels not connected by same letter are significantly different.

Power Plot



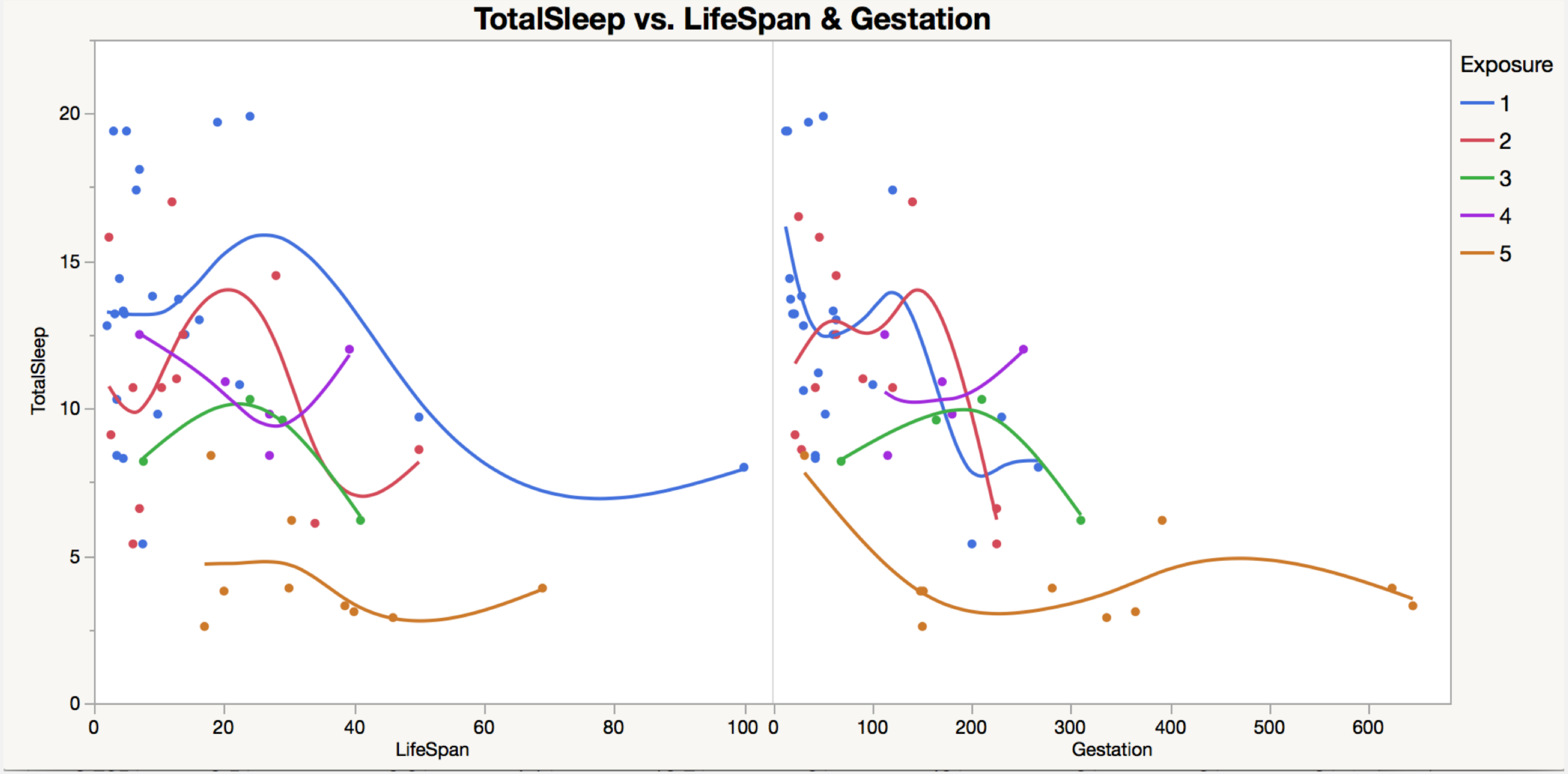
REGRESSION

- Includes brain and body weight, life span, gestation time, time sleeping, and predation and danger indices for 62 species of mammals.
- Of interest is to predict the time spent sleeping.

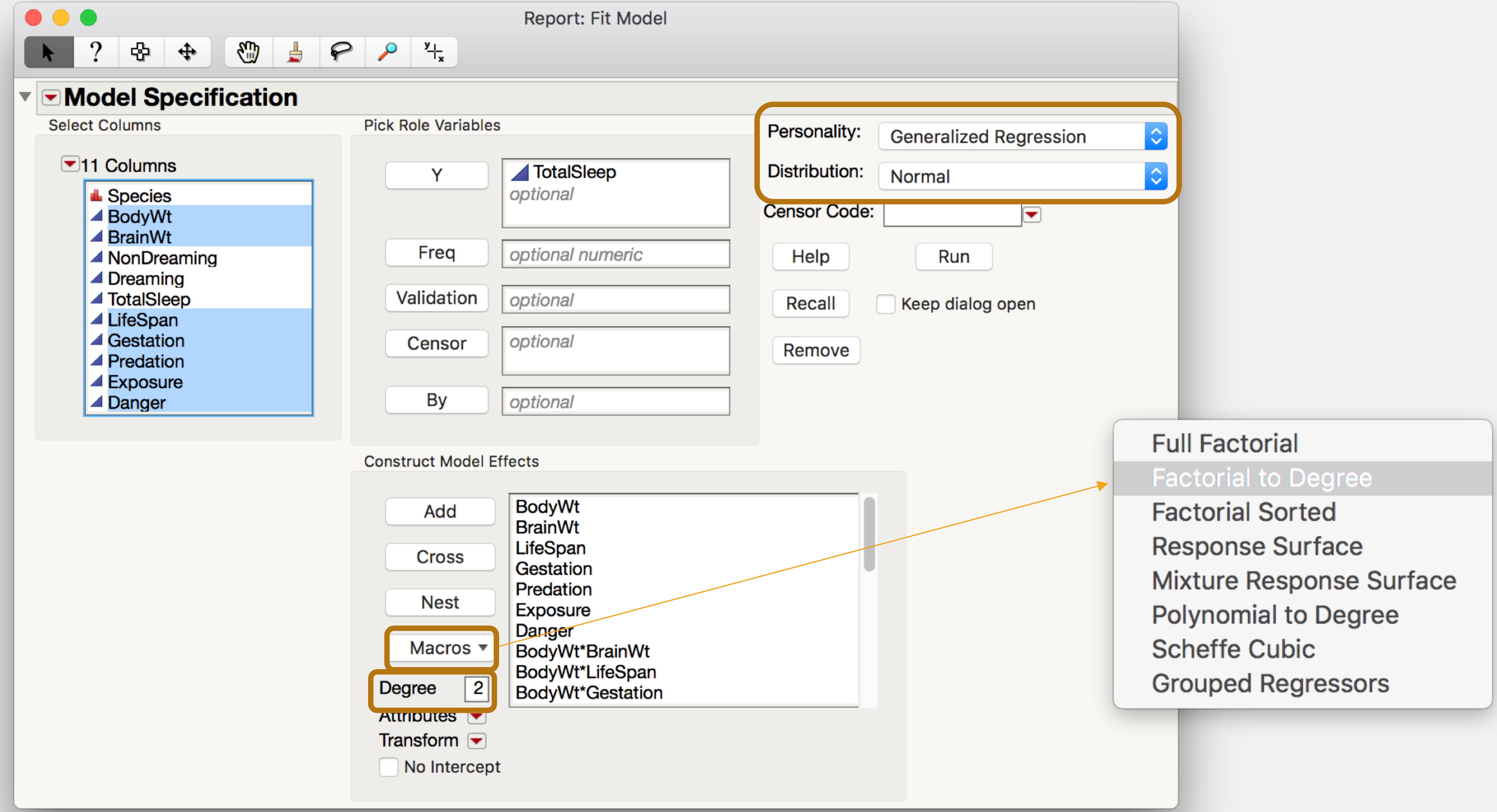
Sleeping Animals.jmp

	Species	BodyWt	BrainWt	NonDreaming	Dreaming	TotalSleep	LifeSpan	Gestation	Predation	Exposure	Danger
1	African elephant	6654	5712	•	•	3.3	38.6	645	3	5	3
2	African giant pouched rat	1	6.6	6.3	2	8.3	4.5	42	3	1	3
3	Arctic Fox	3.385	44.5	•	•	12.5	14	60	1	1	1
4	Arctic ground squirrel	0.92	5.7	•	•	16.5	•	25	5	2	3
5	Asian elephant	2547	4603	2.1	1.8	3.9	69	624	3	5	4
6	Baboon	10.55	179.5	9.1	0.7	9.8	27	180	4	4	4
7	Big brown bat	0.023	0.3	15.8	3.9	19.7	19	35	1	1	1
8	Brazilian tapir	160	169	5.2	1	6.2	30.4	392	4	5	4
9	Cat	3.3	25.6	10.9	3.6	14.5	28	63	1	2	1
10	Chimpanzee	52.16	440	8.3	1.4	9.7	50	230	1	1	1
11	Chinchilla	0.425	6.4	11	1.5	12.5	7	112	5	4	4
12	Cow	465	423	3.2	0.7	3.9	30	281	5	5	5
13	Desert hedgehog	0.55	2.4	7.6	2.7	10.3	•	•	2	1	2
14	Donkey	187.1	419	•	•	3.1	40	365	5	5	5
15	Eastern American mole	0.075	1.2	6.3	2.1	8.4	3.5	42	1	1	1
16	Echidna	3	25	8.6	0	8.6	50	28	2	2	2
17	European hedgehog	0.785	3.5	6.6	4.1	10.7	6	42	2	2	2
18	Galago	0.2	5	9.5	1.2	10.7	10.4	120	2	2	2
19	Genet	1.41	17.5	4.8	1.3	6.1	34	•	1	2	1

Use the Drag-and-Drop interface in **Graph → Graph Builder** to explore the data



Go to **Analyze → Fit Model**



Standard Least Squares

Model Summary

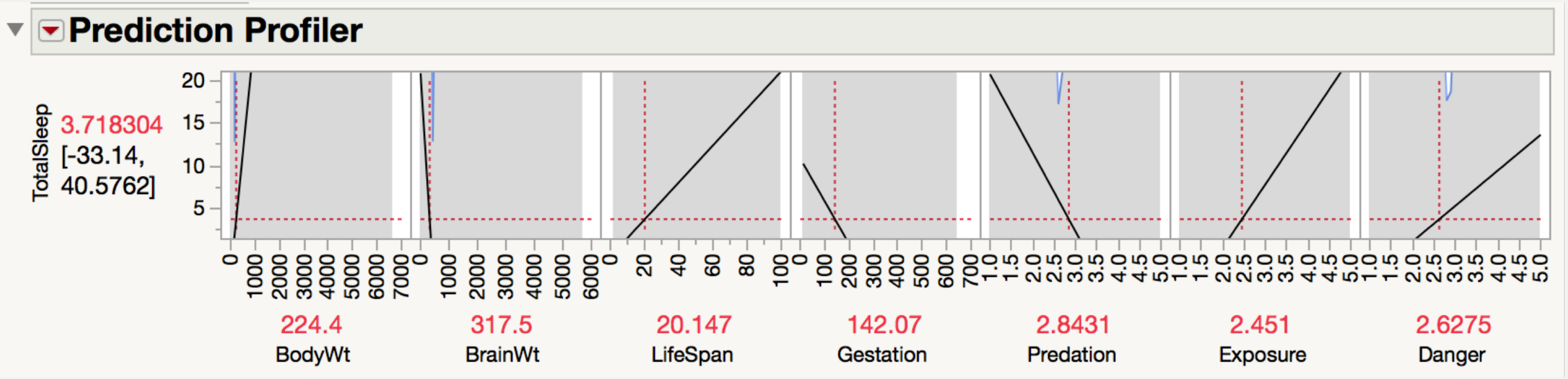
Response	TotalSleep
Distribution	Normal
Estimation Method	Standard Least Squares
Validation Method	None
Mean Model Link	Identity
Scale Model Link	Identity

Measure	
Number of rows	62
Sum of Frequencies	51
-LogLikelihood	107.24504
Number of Parameters	30
BIC	332.44485
AICc	367.49008
RSquare	0.8170884
RSquare Adj	0.5842918
RMSE	1.9816016

Parameter Estimates for Original Predictors

Term	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare
Intercept	14.222181	18.543901	0.5882071	0.4431
BodyWt	0.0287448	0.323292	0.0079055	0.9292
BrainWt	-0.054196	0.1807508	0.0899013	0.7643
LifeSpan	0.2181922	0.2705285	0.6505076	0.4199
Gestation	-0.050316	0.0412041	1.4911651	0.2220
Predation	-9.262813	81.82251	0.0128156	0.9099
Exposure	7.4745626	3.7790178	3.9121338	0.0479*
Danger	4.1941935	81.884386	0.0026236	0.9591
(BodyWt-224.363)*(BrainWt-317.502)	2.0964e-6	1.6069e-5	0.0170197	0.8962
(BodyWt-224.363)*(LifeSpan-20.1471)	-0.000873	0.0014008	0.3886397	0.5330
(BodyWt-224.363)*(Gestation-142.069)	9.7234e-5	0.0002753	0.1247439	0.7239
(BodyWt-224.363)*(Predation-2.84314)	0.0834584	1.4908411	0.0031338	0.9554
(BodyWt-224.363)*(Exposure-2.45098)	0.0217658	0.0326038	0.4456676	0.5044
(BodyWt-224.363)*(Danger-2.62745)	-0.112957	1.4913449	0.0057368	0.9396
(BrainWt-317.502)*(LifeSpan-20.1471)	0.0004776	0.0004081	1.3694114	0.2419
(BrainWt-317.502)*(Gestation-142.069)	-0.000285	0.0002072	1.8849887	0.1698
(BrainWt-317.502)*(Predation-2.84314)	-0.101604	0.8338899	0.0148459	0.9030
(BrainWt-317.502)*(Exposure-2.45098)	0.0171796	0.0221224	0.6030602	0.4374
(BrainWt-317.502)*(Danger-2.62745)	0.1115525	0.8348508	0.0178542	0.8937
(LifeSpan-20.1471)*(Gestation-142.069)	0.0037244	0.0016205	5.2818749	0.0215*
(LifeSpan-20.1471)*(Predation-2.84314)	-0.107643	0.6017127	0.032003	0.8580
(LifeSpan-20.1471)*(Exposure-2.45098)	-0.211865	0.1535277	1.9043442	0.1676
(LifeSpan-20.1471)*(Danger-2.62745)	0.127296	0.6568542	0.0375571	0.8463
(Gestation-142.069)*(Predation-2.84314)	0.0345868	0.0454583	0.5788875	0.4467
(Gestation-142.069)*(Exposure-2.45098)	-0.013856	0.0134418	1.0626404	0.3026
(Gestation-142.069)*(Danger-2.62745)	-0.030575	0.0425874	0.5154178	0.4728
(Predation-2.84314)*(Exposure-2.45098)	-0.522191	2.1683587	0.0579957	0.8097
(Predation-2.84314)*(Danger-2.62745)	-1.072457	0.8276892	1.6789	0.1951
(Exposure-2.45098)*(Danger-2.62745)	1.2165131	1.8223287	0.4456359	0.5044

Rather useless model…



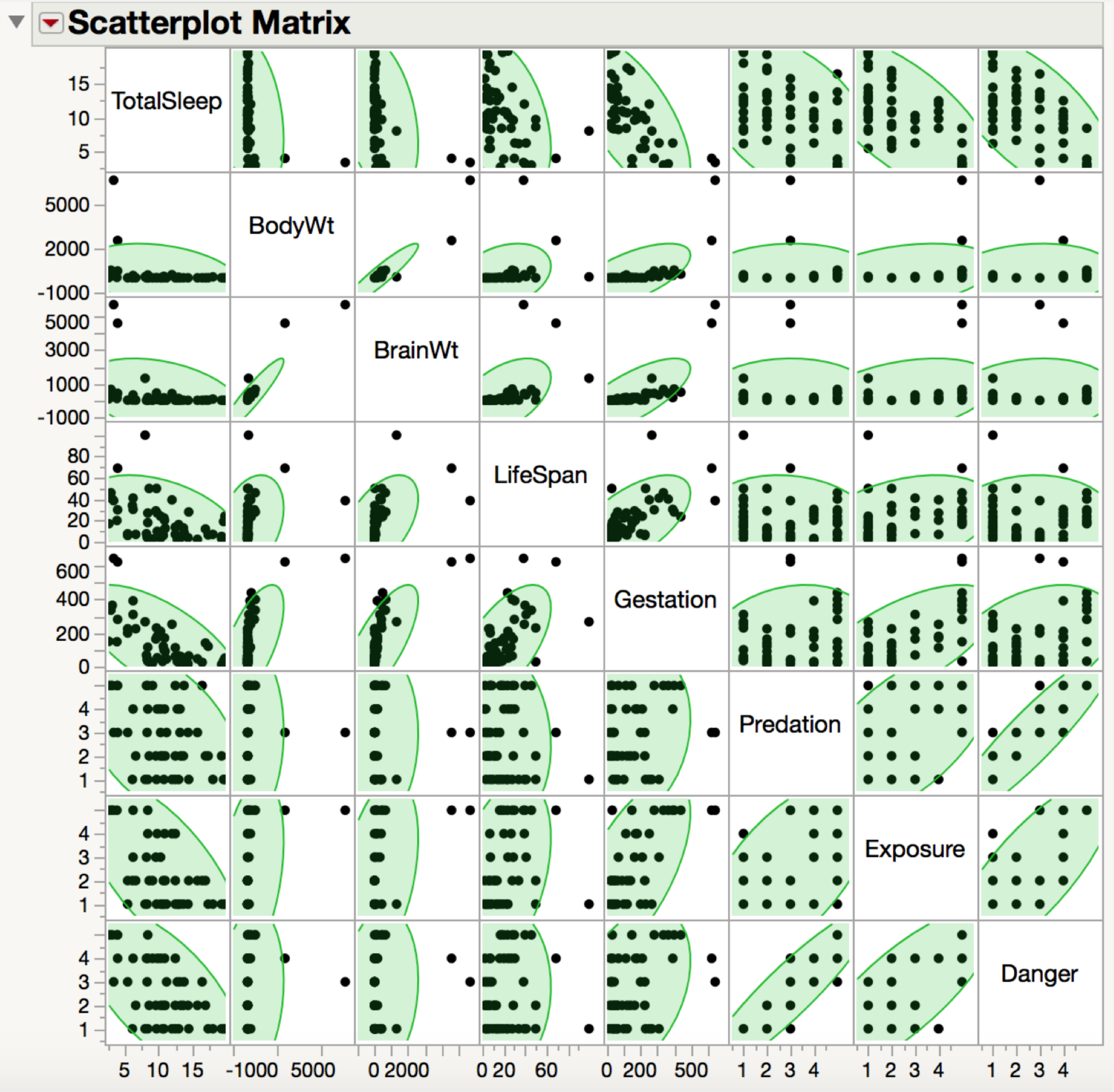
Use

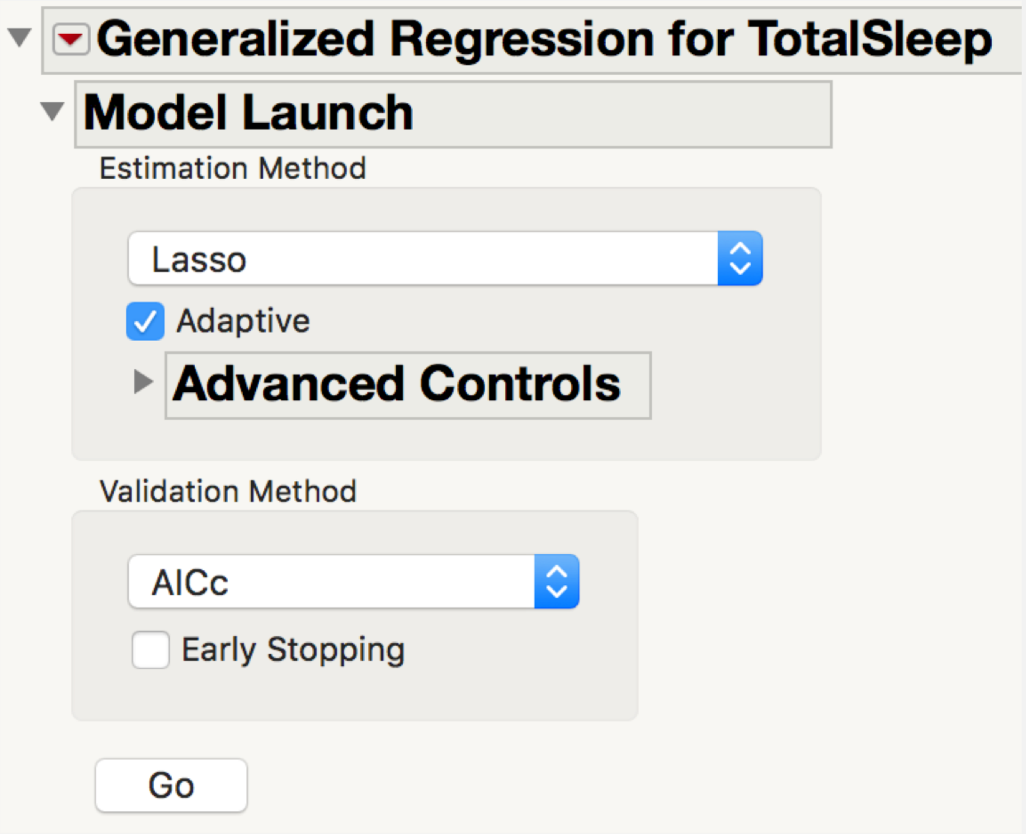
Analyze →

Multivariate Methods →

Multivariate

to get correlations:





Options such as:

- Force certain terms to be in the model

- Force "Effect heredity" (so that if an interaction effect will be in the model, then its main effects also need to be in the model)

Adaptive Lasso with AICc Validation

Model Summary

Response	TotalSleep
Distribution	Normal
Estimation Method	Adaptive Lasso
Validation Method	AICc
Mean Model Link	Identity
Scale Model Link	Identity

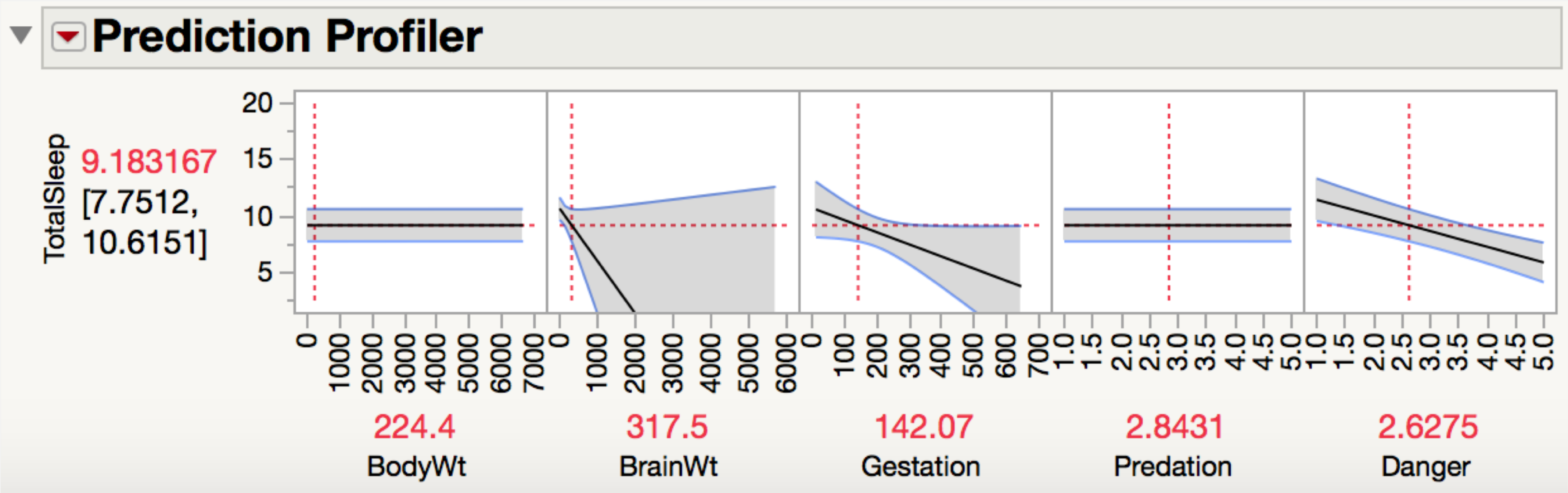
Measure	
Number of rows	62
Sum of Frequencies	51
-LogLikelihood	122.09385
Number of Parameters	9
BIC	279.57412
AICc	266.57794
ERIC	254.93618
RSquare	0.6725586
RMSE	2.0516221
Lambda Penalty	77.203421

Estimation Details

Number of Grid Points	150
Minimum Penalty Fraction	0
Grid Scale	Square Root

Parameter Estimates for Original Predictors

Term	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare
Intercept	15.859904	0.9230672	295.21231	<.0001*
BodyWt	0	0	0	1.0000
BrainWt	-0.004671	0.0025969	3.235129	0.0721
LifeSpan	0	0	0	1.0000
Gestation	-0.010792	0.005891	3.3558682	0.0670
Predation	0	0	0	1.0000
Exposure	0	0	0	1.0000
Danger	-1.393155	0.2987683	21.743512	<.0001*
(BodyWt-224.363)*(BrainWt-317.502)	0	0	0	1.0000
(BodyWt-224.363)*(LifeSpan-20.1471)	0	0	0	1.0000
(BodyWt-224.363)*(Gestation-142.069)	4.6541e-6	2.5904e-6	3.2279754	0.0724
(BodyWt-224.363)*(Predation-2.84314)	-0.003993	0.0015299	6.810896	0.0091*
(BodyWt-224.363)*(Exposure-2.45098)	0	0	0	1.0000
(BodyWt-224.363)*(Danger-2.62745)	0.0060396	0.0025199	5.7444161	0.0165*
(BrainWt-317.502)*(LifeSpan-20.1471)	0	0	0	1.0000
(BrainWt-317.502)*(Gestation-142.069)	0	0	0	1.0000
(BrainWt-317.502)*(Predation-2.84314)	-0.000745	0.0012981	0.3290839	0.5662
(BrainWt-317.502)*(Exposure-2.45098)	0	0	0	1.0000
(BrainWt-317.502)*(Danger-2.62745)	0	0	0	1.0000
(LifeSpan-20.1471)*(Gestation-142.069)	0	0	0	1.0000
(LifeSpan-20.1471)*(Predation-2.84314)	0	0	0	1.0000
(LifeSpan-20.1471)*(Exposure-2.45098)	0	0	0	1.0000
(LifeSpan-20.1471)*(Danger-2.62745)	0	0	0	1.0000
(Gestation-142.069)*(Predation-2.84314)	0	0	0	1.0000
(Gestation-142.069)*(Exposure-2.45098)	0	0	0	1.0000
(Gestation-142.069)*(Danger-2.62745)	0	0	0	1.0000
(Predation-2.84314)*(Exposure-2.45098)	0	0	0	1.0000
(Predation-2.84314)*(Danger-2.62745)	0	0	0	1.0000
(Exposure-2.45098)*(Danger-2.62745)	0	0	0	1.0000



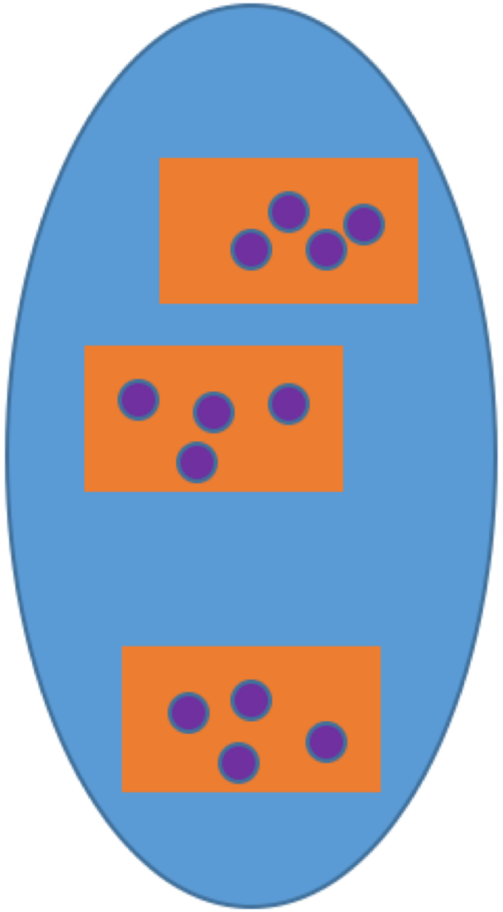
Prediction Profiler
TotalSleep
9.183167
[7.7512,
10.6151]
20
15
10
5
0 1000 2000 3000 4000 5000 6000 7000
0 1000 2000 3000 4000 5000 6000
0 100 200 300 400 500 600 700
1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
224.4
BodyWt
317.5
BrainWt
142.07
Gestation
2.8431
Predation
2.6275
Danger

MIXED MODELS

1. A split-plot (or hierarchical) model

2. Repeated measurements taken on the same subject over time

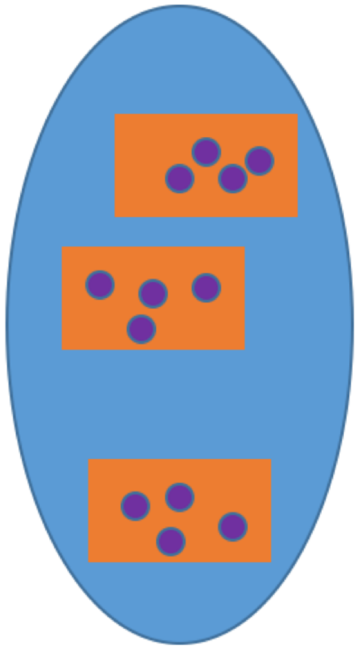
MIXED MODELS: SPLIT-PLOT / HIERARCHICAL



Blue oval: Block
Orange rectangle: Whole Plot
Purple dot: Subplot

1. Randomly assign the levels of Factor A to the Whole Plots within each Block.

2. Randomly assign the levels of Factor B to the Subplots within each Whole Plot.



Blue oval: Block
Orange rectangle: Whole Plot
Purple dot: Subplot

1. Randomly assign the levels of Factor A to the Whole Plots within each Block.

2. Randomly assign the levels of Factor B to the Subplots within each Whole Plot.

Typical Split-Plot Design
(pooling Blocks*B and Blocks*A*B to create the subplot error):

Source:	***df:***	***Effect Type:***
Blocks	r-1	Random
A	a-1	Fixed
Blocks*A	(r-1)(a-1)	Random
B	b-1	Fixed
A*B	(a-1)(b-1)	Fixed
Error		

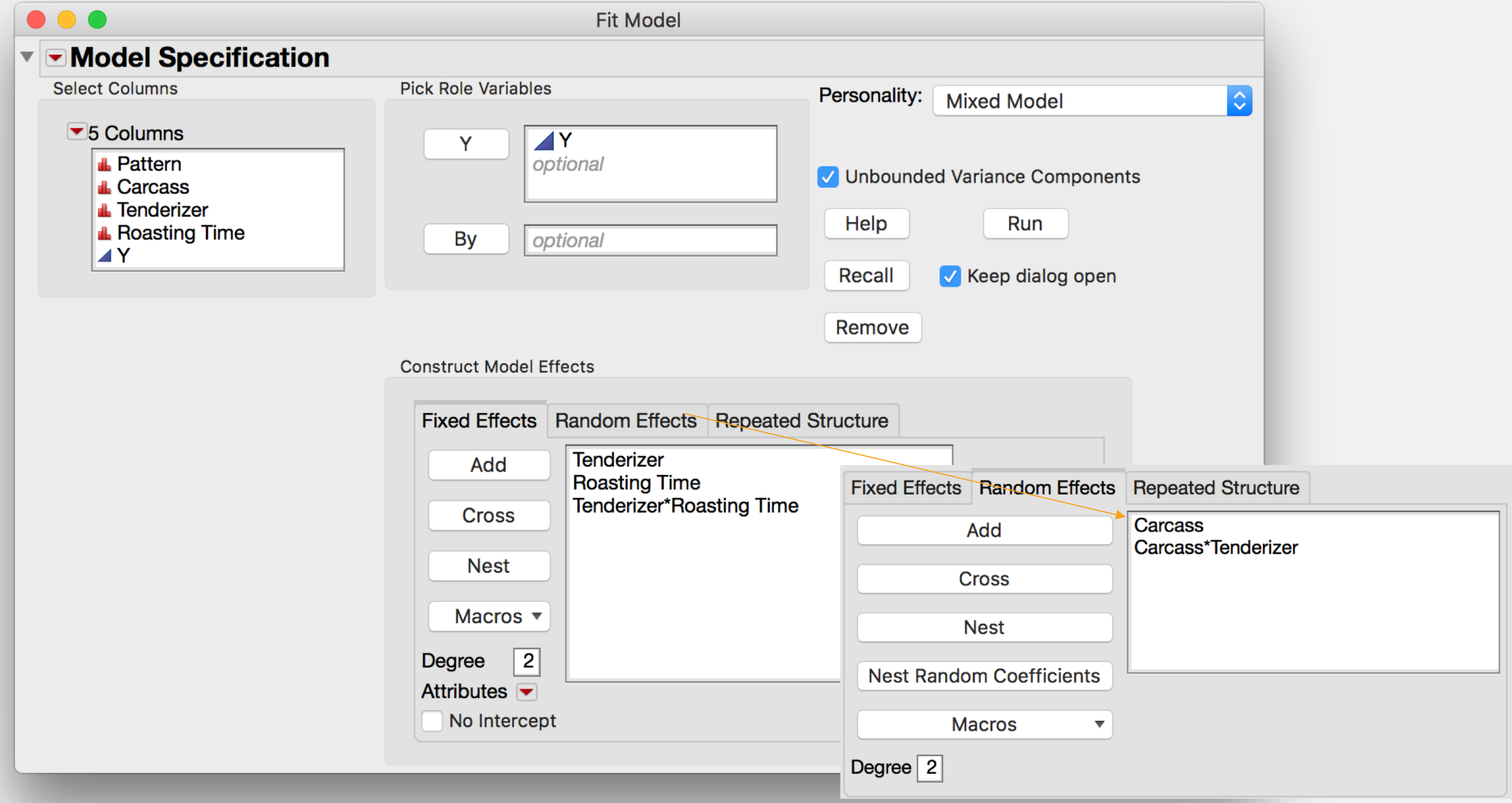
Could be:

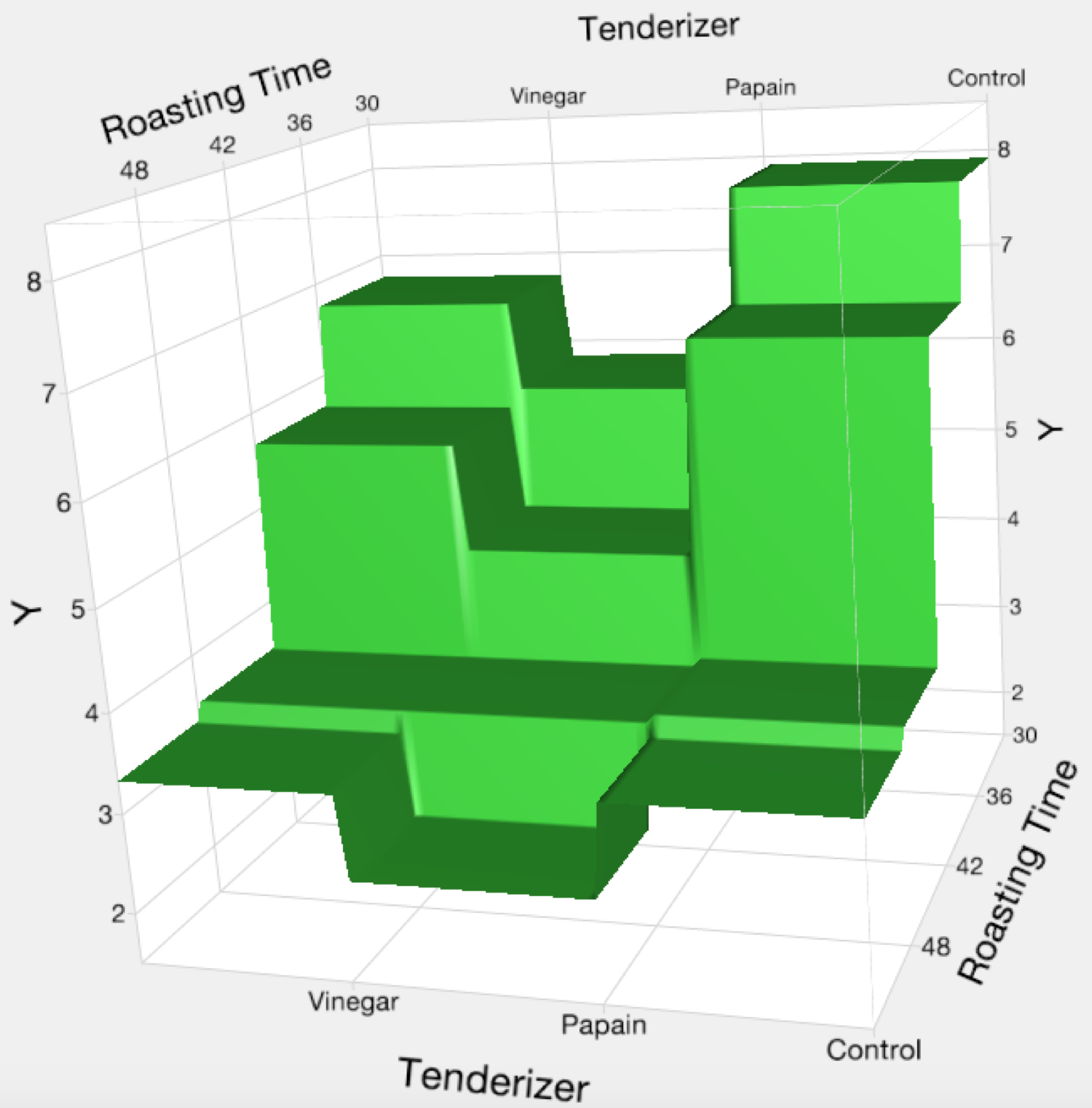
- Different **Nursing Methods** (Factor A) are applied to randomly selected **Units** (Whole Plots) within **Hospitals** (Blocks), but three different **Rating Scales** (Factor B) are randomized to the individual **Patients** (Subplots) in the hospitals. (Response is Patient Satisfaction.)

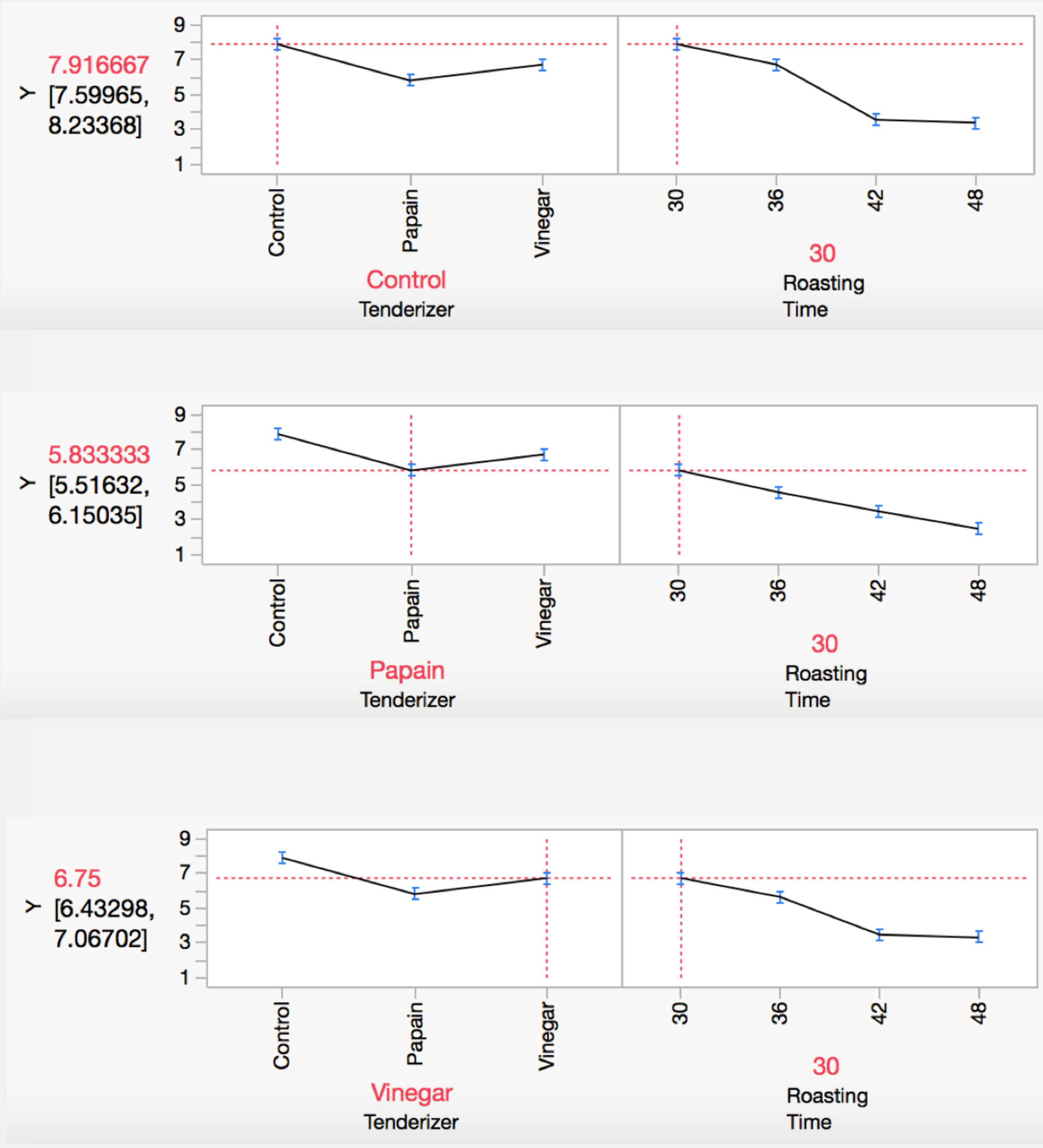
- Different **Watering Methods** (Factor A) are applied to randomly selected **Large Fields** (Whole Plots) within **Different Farms** (Blocks), but three different **Fertilizers** (Factor B) are randomized to the individual **Small Fields** (Subplots) into which the large fields are subdivided. (Response is Yield.)

- Different **Tenderizers** (Factor A) are applied to randomly selected **cuts of meat** (Whole Plots) within **different carcasses** (Blocks), but different **Roasting Times** (Factor B) are randomized to the individual **Cores** (Subplots) in the cuts of meat. (Response is Toughness.)

Go to **Analyze → Fit Model**



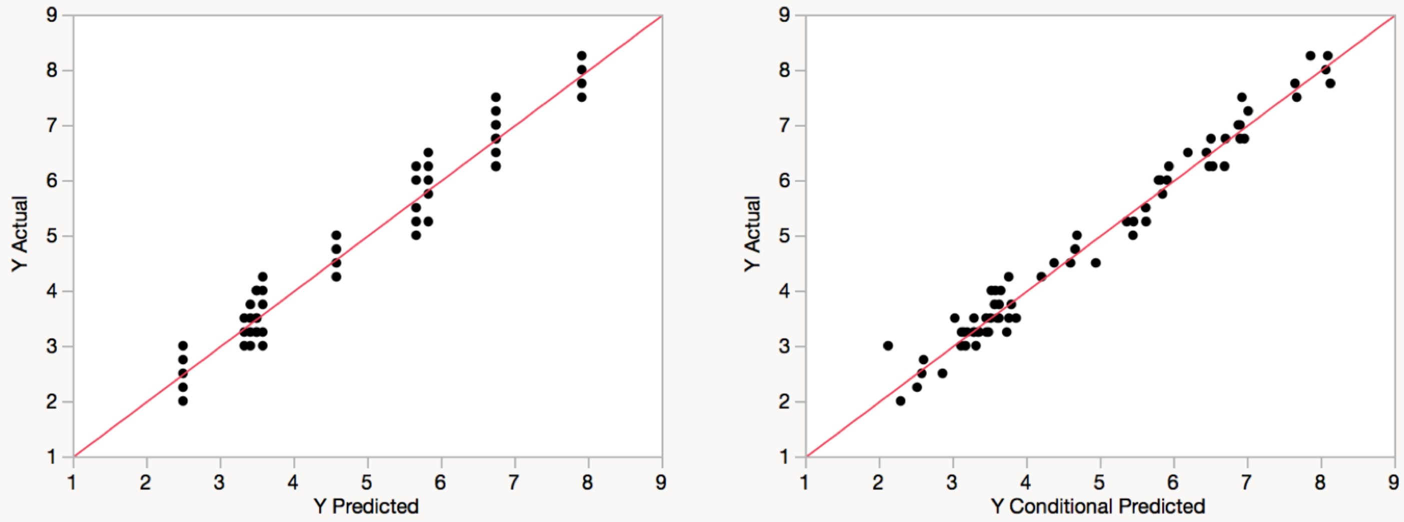

Tenderizer
Roasting Time
30
36
42
48
Vinegar
Papain
Control
Y
8
7
6
5
4
3
2
Vinegar
Papain
Control
Tenderizer
30
36
42
48
Roasting Time


Y
7.916667
[7.59965,
8.23368]
9
7
5
3
1
Control
Papain
Vinegar
Control
Tenderizer
30
36
42
48
30
Roasting
Time
Y
5.833333
[5.51632,
6.15035]
9
7
5
3
1
Control
Papain
Vinegar
Papain
Tenderizer
30
36
42
48
30
Roasting
Time
Y
6.75
[6.43298,
7.06702]
9
7
5
3
1
Control
Papain
Vinegar
Vinegar
Tenderizer
30
36
42
48
30
Roasting
Time

Fit Mixed

Actual by Predicted Plot

Actual by Conditional Predicted Plot



Fit Statistics

-2 Residual Log Likelihood	68.324206
-2 Log Likelihood	23.659713
AICc	62.231142
BIC	87.809705

Random Effects Covariance Parameter Estimates

Variance Component	Estimate	Std Error	95% Lower	95% Upper	Wald p-Value
Carcass	0.0605042	0.0411897	-0.020226	0.1412346	0.1419
Carcass*Tenderizer	-0.007958	0.0076058	-0.022865	0.0069492	0.2954
Residual	0.0863433	0.0182028	0.0594013	0.1369749	
Total	0.1468476	0.0450325	0.0871689	0.2983101	

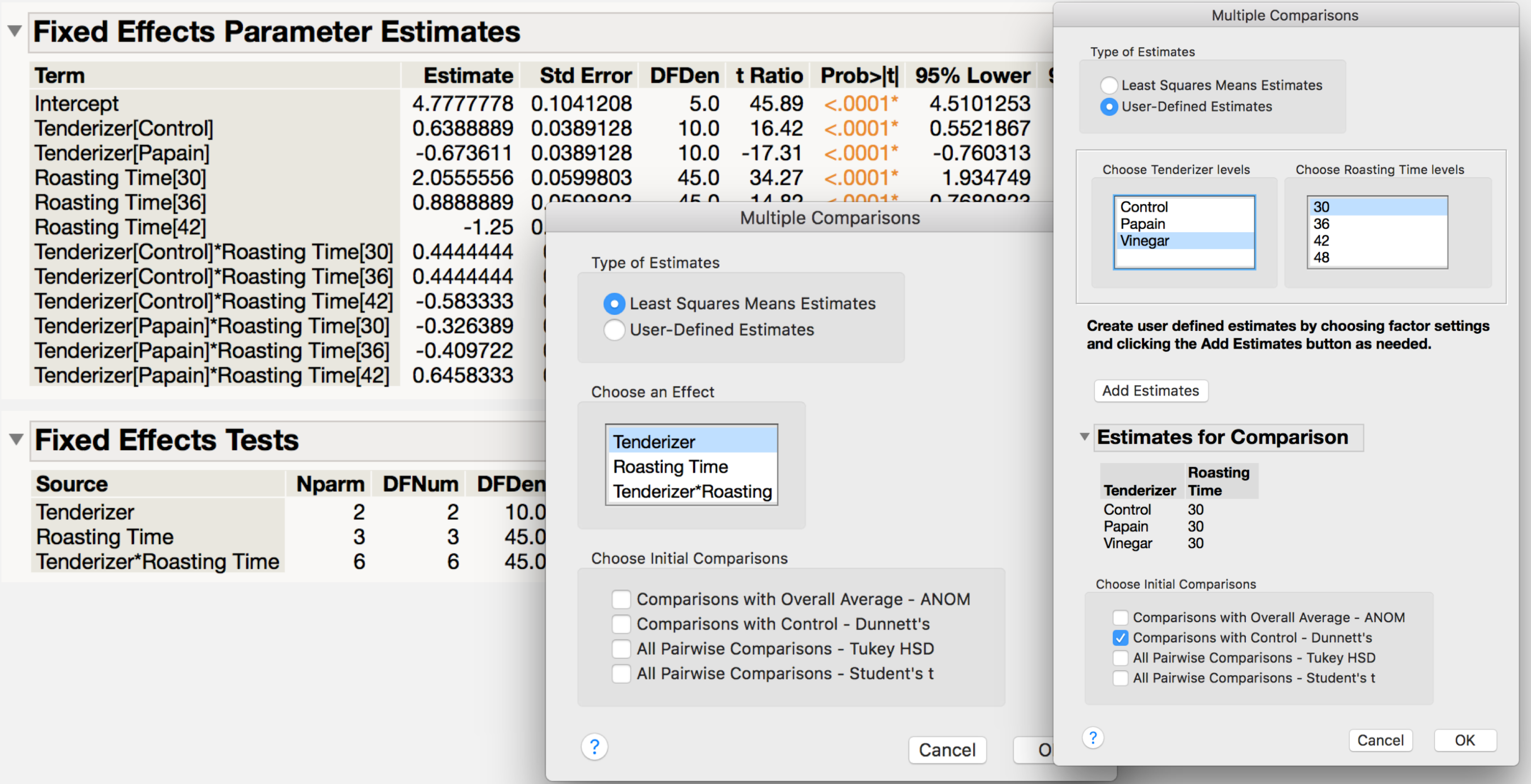
Random Coefficients

Carcass

Carcass	Intercept
1	0.2842167
2	0.1291894
3	0.1679462
4	-0.025838
5	-0.258379
6	-0.297136

Carcass*Tenderizer

Carcass	Tenderizer	Intercept
1	Control	-0.101675
1	Papain	0.0808084
1	Vinegar	-0.016516
2	Control	0.0267775
2	Papain	-0.046216
2	Vinegar	0.0024464
3	Control	0.0494094
3	Papain	-0.060081
3	Vinegar	-0.011418
4	Control	-0.027253
4	Papain	0.0457398
4	Vinegar	-0.015088
5	Control	0.0194383
5	Papain	0.0559349
5	Vinegar	-0.041389
6	Control	0.033303
6	Papain	-0.076187
6	Vinegar	0.0819652


Fixed Effects Parameter Estimates
Term | Estimate | Std Error | DFDen | t Ratio | Prob>|t| | 95% Lower
Intercept 4.7777778 0.1041208 5.0 45.89 <.0001* 4.5101253
Tenderizer[Control] 0.6388889 0.0389128 10.0 16.42 <.0001* 0.5521867
Tenderizer[Papain] -0.673611 0.0389128 10.0 -17.31 <.0001* -0.760313
Roasting Time[30] 2.0555556 0.0599803 45.0 34.27 <.0001* 1.934749
Roasting Time[36] 0.8888889
Roasting Time[42] -1.25
Tenderizer[Control]*Roasting Time[30] 0.4444444
Tenderizer[Control]*Roasting Time[36] 0.4444444
Tenderizer[Control]*Roasting Time[42] -0.583333
Tenderizer[Papain]*Roasting Time[30] -0.326389
Tenderizer[Papain]*Roasting Time[36] -0.409722
Tenderizer[Papain]*Roasting Time[42] 0.6458333
Fixed Effects Tests
Source | Nparm | DFNum | DFDen
Tenderizer 2 2 10.0
Roasting Time 3 3 45.0
Tenderizer*Roasting Time 6 6 45.0
Multiple Comparisons
Type of Estimates
Least Squares Means Estimates
User-Defined Estimates
Choose an Effect
Tenderizer
Roasting Time
Tenderizer*Roasting
Choose Initial Comparisons
Comparisons with Overall Average - ANOM
Comparisons with Control - Dunnett's
All Pairwise Comparisons - Tukey HSD
All Pairwise Comparisons - Student's t
?
Cancel
Multiple Comparisons
Type of Estimates
Least Squares Means Estimates
User-Defined Estimates
Choose Tenderizer levels
Control
Papain
Vinegar
Choose Roasting Time levels
30
36
42
48
Create user defined estimates by choosing factor settings and clicking the Add Estimates button as needed.
Add Estimates
Estimates for Comparison
Tenderizer | Roasting Time
Control 30
Papain 30
Vinegar 30
Choose Initial Comparisons
Comparisons with Overall Average - ANOM
Comparisons with Control - Dunnett's
All Pairwise Comparisons - Tukey HSD
All Pairwise Comparisons - Student's t
?
Cancel
OK

MIXED MODELS: REPEATED MEASURES

- A study was performed to test two new cholesterol drugs against a control drug.
- Five patients with high cholesterol were randomly assigned to each of four treatments (the two experimental drugs, the control, and a placebo).
- Each patient's total **cholesterol** was measured at six times during the study: the first day in April, May, and June in the morning and afternoon.
- You are interested in whether either of the new drugs is effective at lowering cholesterol and in whether time and treatment interact.

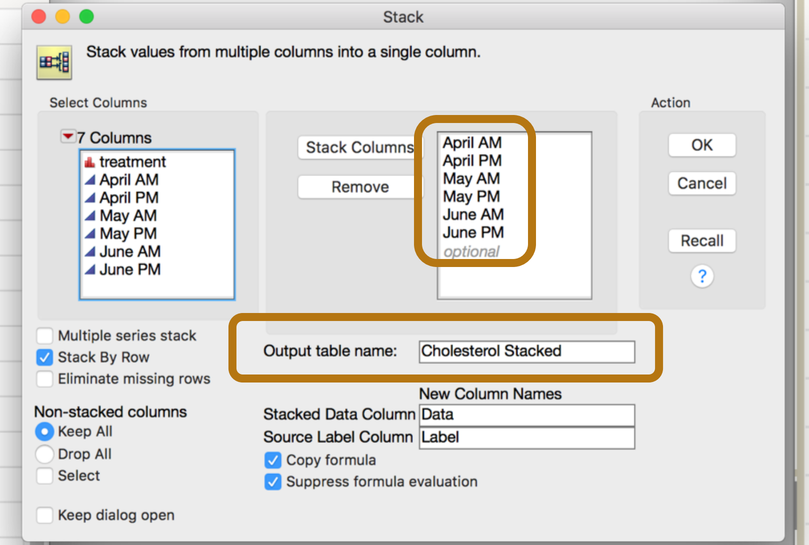
➢ Unique feature: The measurements are not totally independent, as we would assume for an ANOVA or Regression, but instead are correlated within subject.

Some notes:

- We need each cholesterol measurement needs to be in its own row.
  - ➢ We can use the Tables → Stack to reformat the data. For example, take Cholesterol.jmp and transform it into Cholesterol Stacked.jmp

Cholesterol.jmp

	treatment	April AM	April PM	May AM	May PM	June AM	June PM
1	A	278	280	204	208	171	175
2	A	278	281	195	199	185	189
3	A	276	280	213	219	179	181
4	A	276	281	201	211	183	188
5	A	279	285	188	192	170	174
6	B	266	270	220	224	180	184
7	B	280	284	228	232	200	204
8	B	284	288	233	237	175	179
9	B	273	277	215	219	202	207
10	B	281	285	237	241	199	205
11	Control	278	282	273	277	281	285
12	Control	273	277	274	278	280	282
13	Control	282	285	276	281	279	282
14	Control	274	277	285	289	274	278
15	Control	277	281	279	284	284	285
16	Placebo	279	283	278	284	268	272
17	Placebo	277	279	291	291	280	285
18	Placebo	275	279	280	283	281	283
19	Placebo	276	282	277	282	274	279
20	Placebo	282	286	281	285	282	285

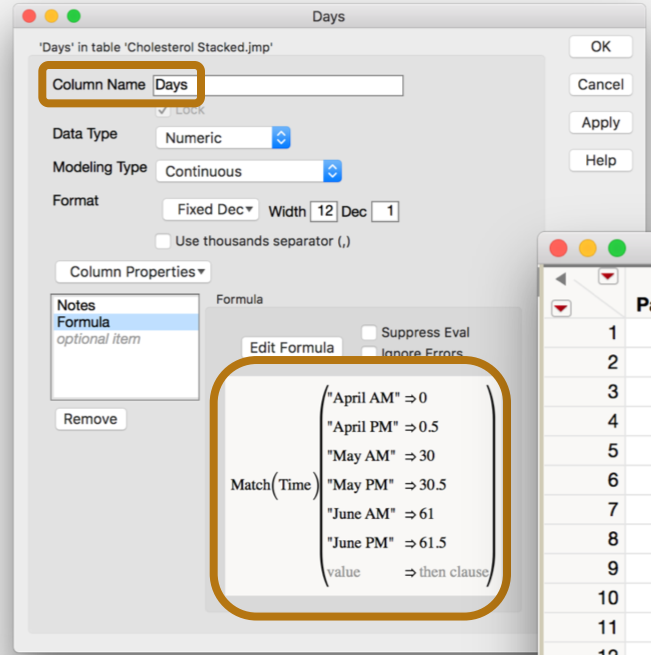


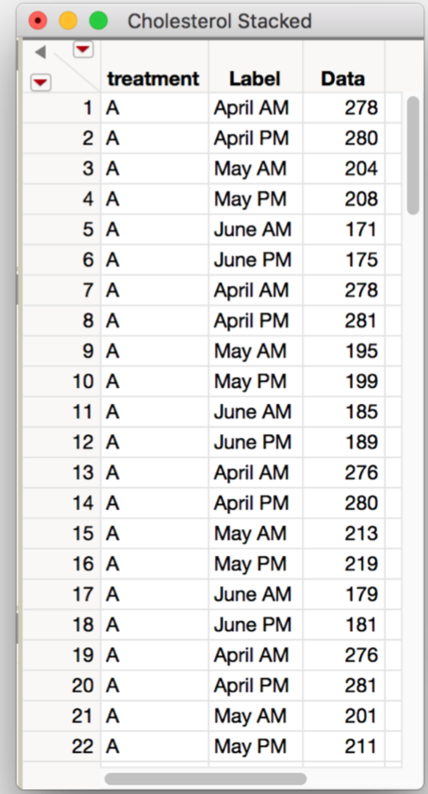
Cholesterol Stacked

	treatment	Label	Data
1	A	April AM	278
2	A	April PM	280
3	A	May AM	204
4	A	May PM	208
5	A	June AM	171
6	A	June PM	175
7	A	April AM	278
8	A	April PM	281
9	A	May AM	195
10	A	May PM	199
11	A	June AM	185
12	A	June PM	189
13	A	April AM	276
14	A	April PM	280
15	A	May AM	213
16	A	May PM	219
17	A	June AM	179
18	A	June PM	181
19	A	April AM	276
20	A	April PM	281
21	A	May AM	201
22	A	May PM	211

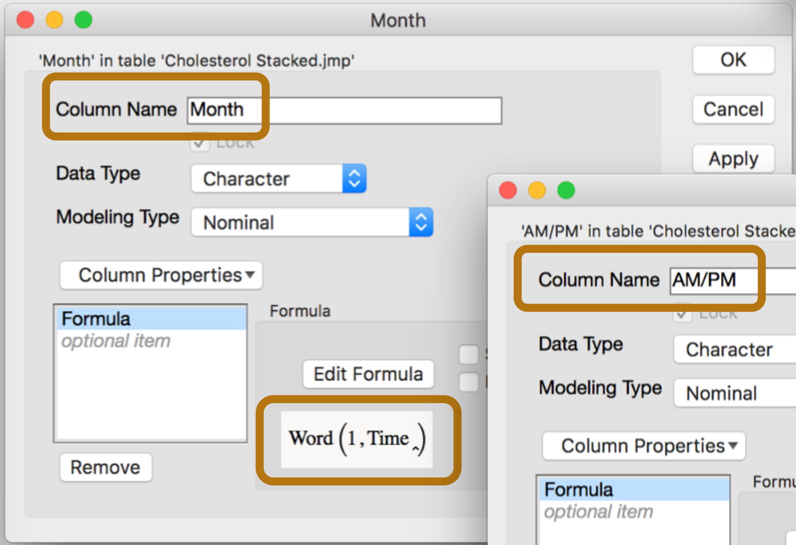
More notes:

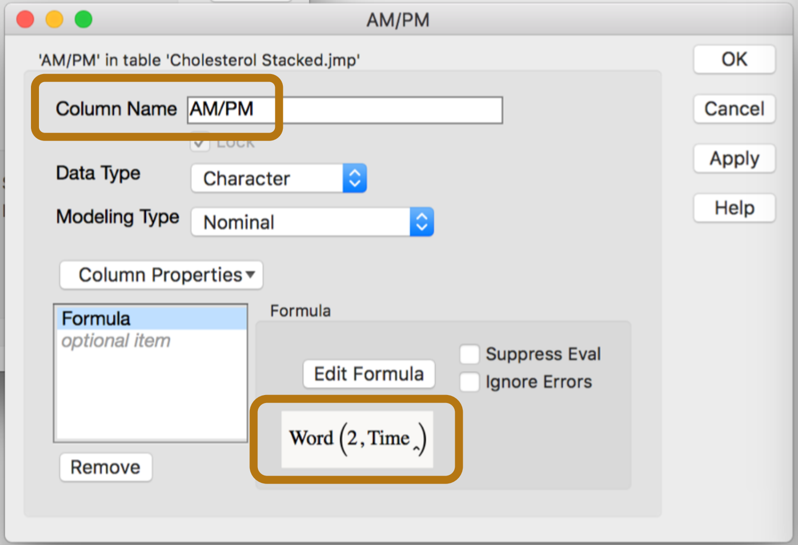
- Now let's clean up that data table and add columns for each factor we are interested in ("Month" and "AM/PM")
- We also need a continuous variable to indicate the time over which the repeated measures were taken. This will be necessary if we want to consider the AR(1) structure for our analysis.

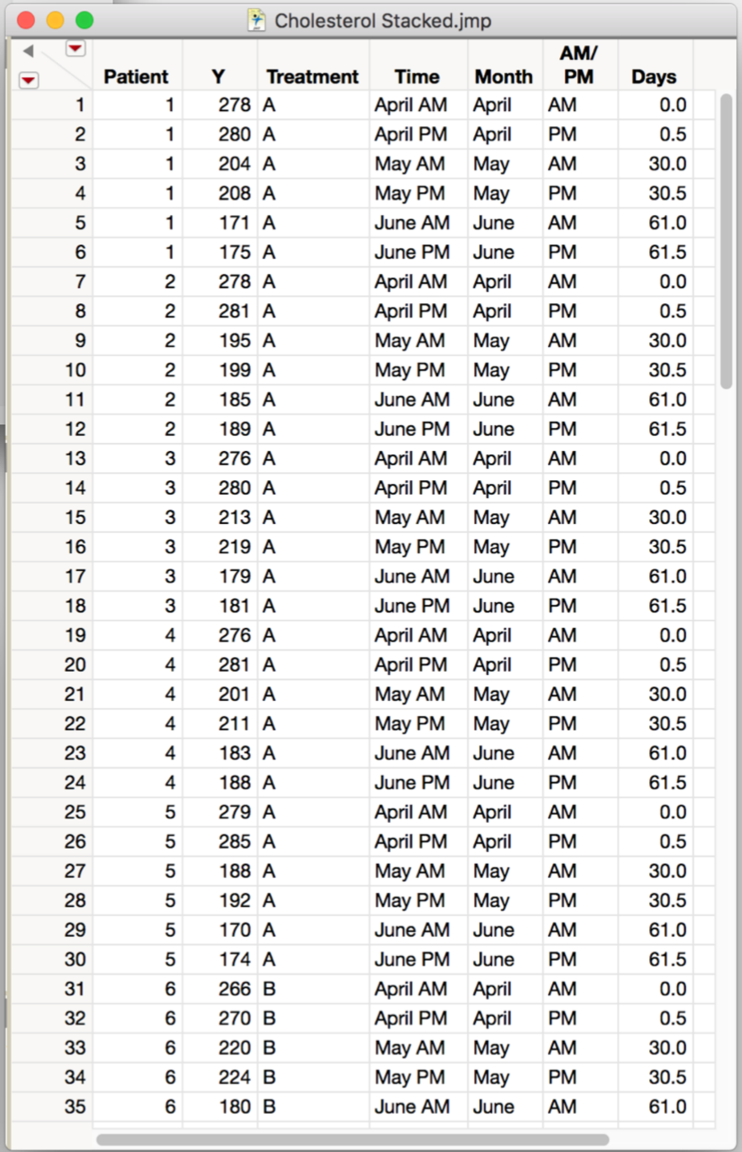



Cholesterol Stacked

	treatment	Label	Data
1	A	April AM	278
2	A	April PM	280
3	A	May AM	204
4	A	May PM	208
5	A	June AM	171
6	A	June PM	175
7	A	April AM	278
8	A	April PM	281
9	A	May AM	195
10	A	May PM	199
11	A	June AM	185
12	A	June PM	189
13	A	April AM	276
14	A	April PM	280
15	A	May AM	213
16	A	May PM	219
17	A	June AM	179
18	A	June PM	181
19	A	April AM	276
20	A	April PM	281
21	A	May AM	201
22	A	May PM	211

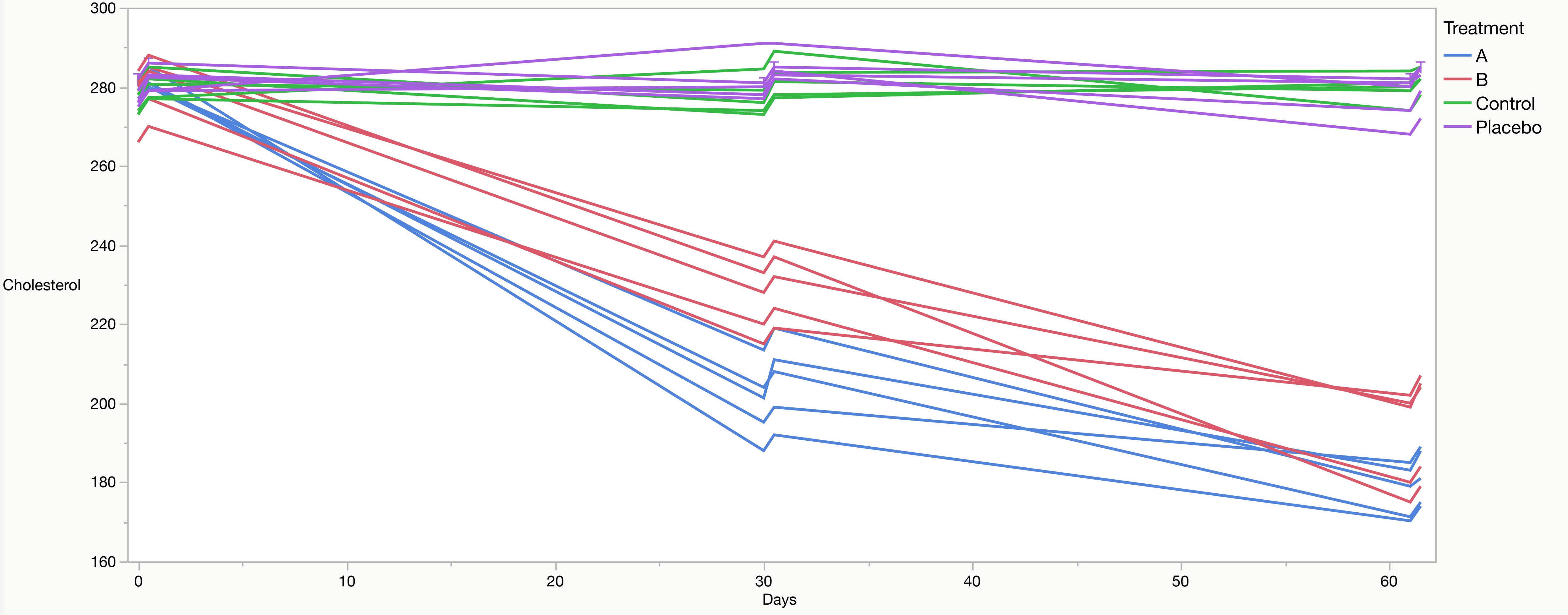





Cholesterol Stacked.jmp

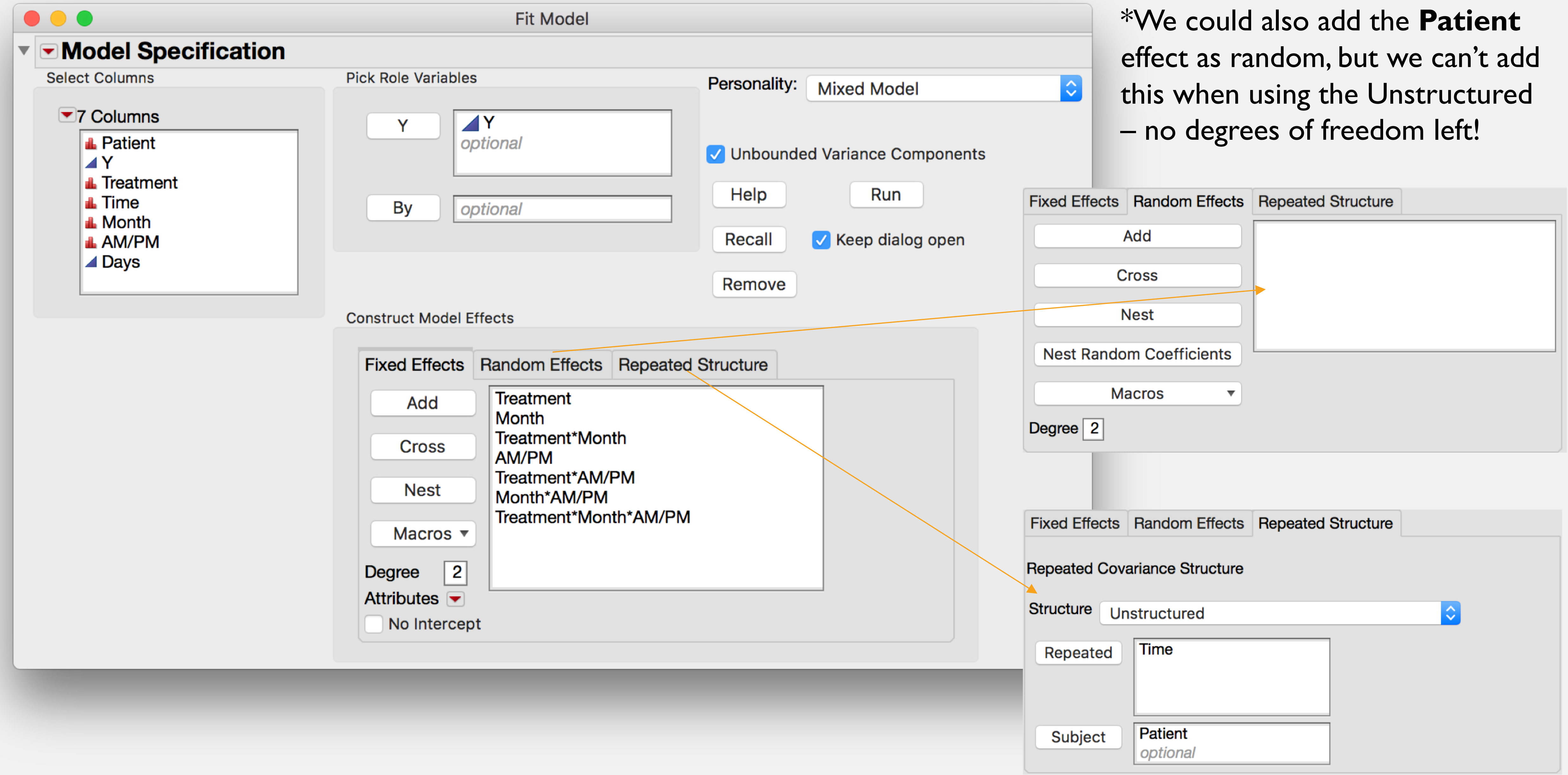
	Patient	Y	Treatment	Time	Month	AM/PM	Days
1	1	278	A	April AM	April	AM	0.0
2	1	280	A	April PM	April	PM	0.5
3	1	204	A	May AM	May	AM	30.0
4	1	208	A	May PM	May	PM	30.5
5	1	171	A	June AM	June	AM	61.0
6	1	175	A	June PM	June	PM	61.5
7	2	278	A	April AM	April	AM	0.0
8	2	281	A	April PM	April	PM	0.5
9	2	195	A	May AM	May	AM	30.0
10	2	199	A	May PM	May	PM	30.5
11	2	185	A	June AM	June	AM	61.0
12	2	189	A	June PM	June	PM	61.5
13	3	276	A	April AM	April	AM	0.0
14	3	280	A	April PM	April	PM	0.5
15	3	213	A	May AM	May	AM	30.0
16	3	219	A	May PM	May	PM	30.5
17	3	179	A	June AM	June	AM	61.0
18	3	181	A	June PM	June	PM	61.5
19	4	276	A	April AM	April	AM	0.0
20	4	281	A	April PM	April	PM	0.5
21	4	201	A	May AM	May	AM	30.0
22	4	211	A	May PM	May	PM	30.5
23	4	183	A	June AM	June	AM	61.0
24	4	188	A	June PM	June	PM	61.5
25	5	279	A	April AM	April	AM	0.0
26	5	285	A	April PM	April	PM	0.5
27	5	188	A	May AM	May	AM	30.0
28	5	192	A	May PM	May	PM	30.5
29	5	170	A	June AM	June	AM	61.0
30	5	174	A	June PM	June	PM	61.5
31	6	266	B	April AM	April	AM	0.0
32	6	270	B	April PM	April	PM	0.5
33	6	220	B	May AM	May	AM	30.0
34	6	224	B	May PM	May	PM	30.5
35	6	180	B	June AM	June	AM	61.0

CHOLESTEROL VS. TIME IN DAYS



First, fit the model we want (full factorial of Treatment, Month, and AM/PM), and use the Unstructured in order to see (1) how the observations are correlated with each other over time and (2) to decide if the variances are the same or different over time.

*We could also add the **Patient** effect as random, but we can't add this when using the Unstructured – no degrees of freedom left!



Repeated Effects Covariance Parameter Estimates

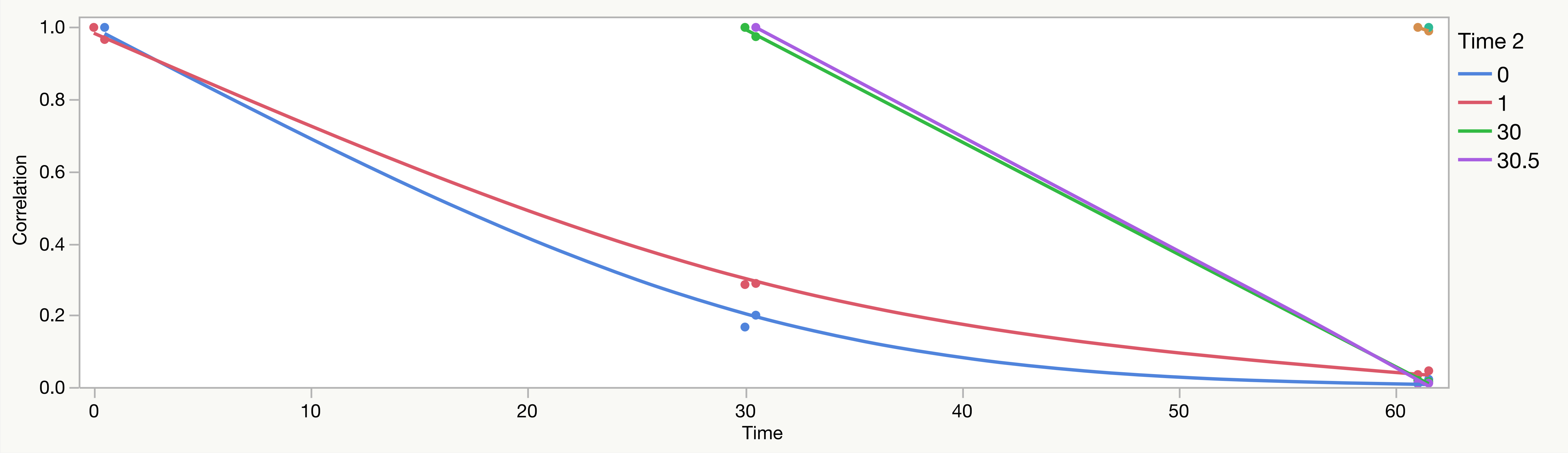
Repeated Effect: Time

Subject: Patient

Covariance Parameter	Estimate	Std Error	95% Lower	95% Upper
Var(April AM)	18.725	6.6202872	5.7494754	31.700525
Cov(April PM,April AM)	18.354884	6.6035621	5.4121399	31.297628
Var(April PM)	19.268932	6.8125963	5.9164888	32.621376
Cov(May AM,April AM)	9.2756709	8.4629182	-7.311344	25.862686
Cov(May AM,April PM)	5.5074058	8.3704001	-10.89828	21.913089
Var(May AM)	56.603347	20.012305	17.379949	95.826744
Cov(May PM,April AM)	9.4147226	8.5038584	-7.252534	26.081979
Cov(May PM,April PM)	6.6230805	8.4532303	-9.944946	23.191108
Cov(May PM,May AM)	55.365523	19.83529	16.489068	94.241978
Var(May PM)	57.058089	20.173081	17.519577	96.596602
Cov(June AM,April AM)	1.1945478	8.6266098	-15.7133	18.102392
Cov(June AM,April PM)	0.3183447	8.7461245	-16.82374	17.460434
Cov(June AM,May AM)	1.106725	14.992149	-28.27735	30.490796
Cov(June AM,May PM)	0.6455101	15.050552	-28.85303	30.14405
Var(June AM)	63.512277	22.454981	19.501323	107.52323
Cov(June PM,April AM)	1.5810194	8.7687993	-15.60551	18.76755
Cov(June PM,April PM)	0.7647113	8.8882627	-16.65596	18.185386
Cov(June PM,May AM)	0.9262595	15.232066	-28.92804	30.780561
Cov(June PM,May PM)	0.6543895	15.292238	-29.31785	30.626625
Cov(June PM,June AM)	63.878686	22.700344	19.386829	108.37054
Var(June PM)	65.568482	23.181959	20.132677	111.00429

- Table Style ▶
- Columns ▶
- Sort by Column...
- Make into Data Table
- Make Combined Data Table
- Make Into Matrix
- Format Column...
- Show Properties
- Copy Column
- Copy Table
- Simulate

CORRELATIONS VS. TIME



Conclusions:

-Correlations are high between AM/PM on the same day (Time 0 with Time 0 and Time 1, and Time 30 with Time 30 and Time 30.5)

-Correlations decay the farther apart in time the measurements are from each other.

SOME POSSIBLE STRUCTURES:

- **Unequal Variances** structure assumes unequal variances BUT DOES NOT ALLOW ANY NON-ZERO CORRELATIONS.
- **Compound Symmetry, CS,** structure assumes same correlation from any time point to any other time point. This is equivalent to adding a random effect for each patient.
- There is also **CS with Unequal Variances**.

Correlation matrix:

1	r	r	r
	1	r	r
		1	r
			1

Example correlation matrix:

1	.97	.97	.97
	1	.97	.97
		1	.97
			1

- **Auto-Regressive of Order 1, AR(1),** structure assumes that the correlations will degrade exponentially over time.

Correlation matrix:

1	r	r^2	r^3
	1	r	r^2
		1	r
			1

Example correlation matrix:

1	.90	.81	.73
	1	.90	.81
		1	.90
			1

Let’s compare:

- Residual Only
- Compound Symmetry (this is just using Patient as a random effect)
- AR(1)
- AR(1) plus Patient as a random effect
- Unstructured

(simpler to more complicated)

WHICH MODEL IS BEST?
(SMALLER FIT STATISTICS ARE BETTER)

Unstructured:

Fit Statistics

-2 Residual Log Likelihood	554.98818
-2 Log Likelihood	557.89101
AICc	703.83696
BIC	773.32814

AR(1) plus random patient effects:

Fit Statistics

-2 Residual Log Likelihood	576.71933
-2 Log Likelihood	585.05495
AICc	655.48974
BIC	714.31723

AR(1):

Fit Statistics

-2 Residual Log Likelihood	577.09954
-2 Log Likelihood	585.53021
AICc	652.62698
BIC	710.00499

Compound Symmetry:

Fit Statistics

-2 Residual Log Likelihood	721.03949
-2 Log Likelihood	765.45515
AICc	832.55192
BIC	889.92993

Residual Only (no accounting for dependence):

Fit Statistics

-2 Residual Log Likelihood	731.09452
-2 Log Likelihood	778.02393
AICc	841.85372
BIC	897.71123

AR(1):

Fixed Effects Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|	95% Lower	95% Upper
Intercept	253.67179	0.9973849	22.8	254.34	<.0001*	251.60745	255.73613
Treatment[A]	-32.98848	1.7275214	22.8	-19.10	<.0001*	-36.56402	-29.41294
Treatment[B]	-20.03845	1.7275214	22.8	-11.60	<.0001*	-23.61399	-16.46291
Treatment[Control]	26.051892	1.7275214	22.8	15.08	<.0001*	22.476351	29.627433
Month[April]	25.485843	1.1353123	73.2	22.45	<.0001*	23.223268	27.748418
Month[May]	-5.052009	1.0147916	95.9	-4.98	<.0001*	-7.066372	-3.037646

In this case, adding the **Patient** random effects does not help us.

AR(1) with random patient effects:

Fixed Effects Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|	95% Lower	95% Upper
Intercept	253.67179	0.9378746	16.5	270.48	<.0001*	251.68838	255.6552
Treatment[A]	-32.98848	1.6244464	16.5	-20.31	<.0001*	-36.42384	-29.55311
Treatment[B]	-20.03845	1.6244464	16.5	-12.34	<.0001*	-23.47382	-16.60309
Treatment[Control]	26.051892	1.6244464	16.5	16.04	<.0001*	22.616526	29.487258
Month[April]	25.485843	1.2059822	28.7	21.13	<.0001*	23.018127	27.953559
Month[May]	-5.052009	1.0417762	71.3	-4.85	<.0001*	-7.129116	-2.974902

Random Coefficients

Patient

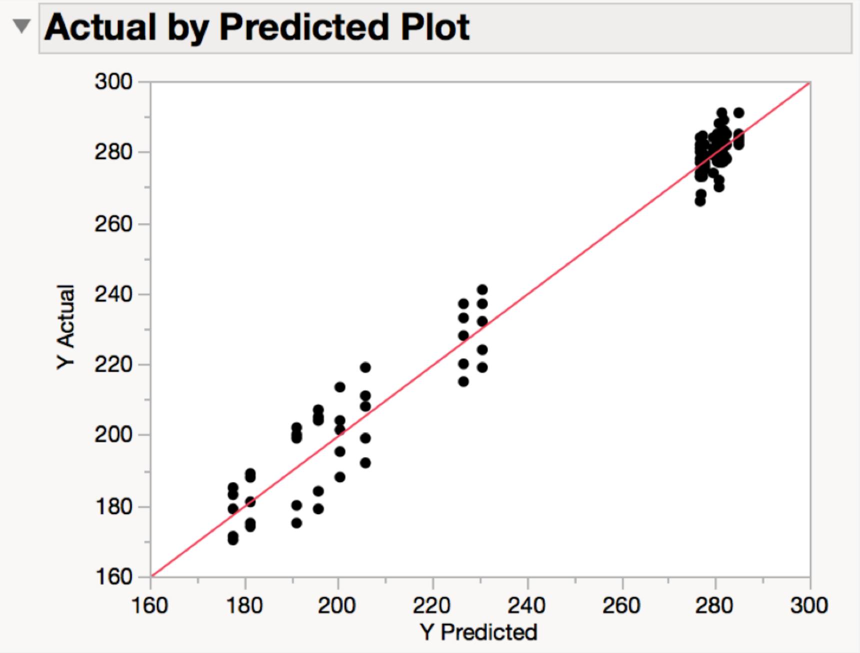
Patient	Intercept
1	0.8246733
2	-0.665516
3	-1.430139
4	-1.318725
5	2.5897066
6	4.9122149
7	-2.296658
8	0.9279618
9	0.1211668
10	-3.664686
11	-0.080837
12	1.1362682
13	-0.706763
14	0.5086569
15	-0.857325
16	1.686069
17	-1.395608
18	0.2097648
19	1.0523585
20	-1.552584

Remember, Std Errors are made up of the related **error** components divided by the related **degrees of freedom**.

→ If adding model complexity can reduce the unexplained variance, it will be easier to detect significant results.

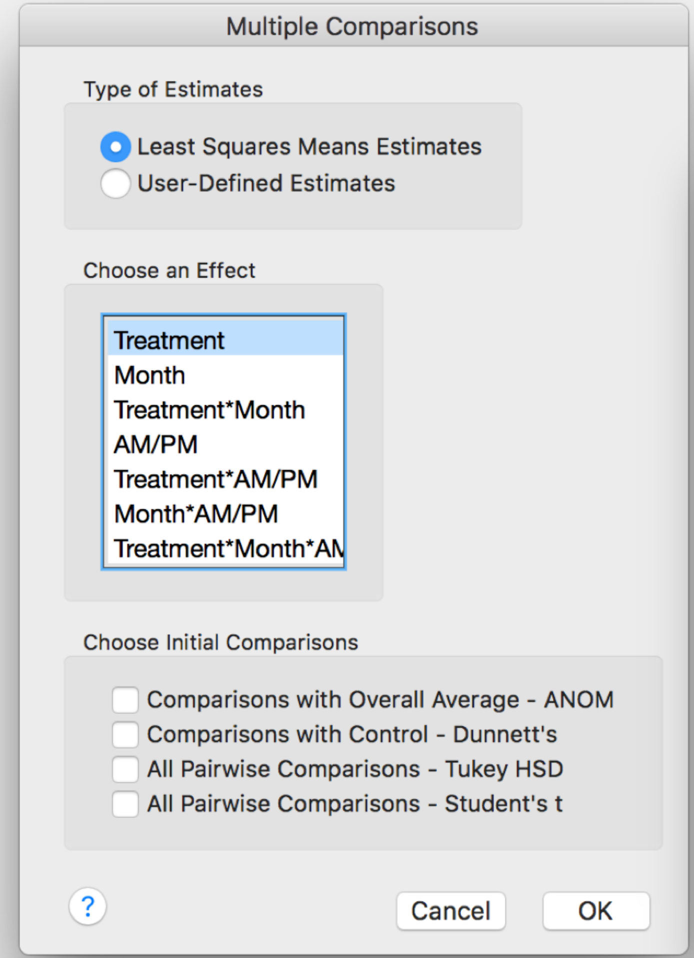
→ But if you don't gain enough in explaining the variance, the loss of degrees of freedom might actually make it HARDER to detect significant differences.

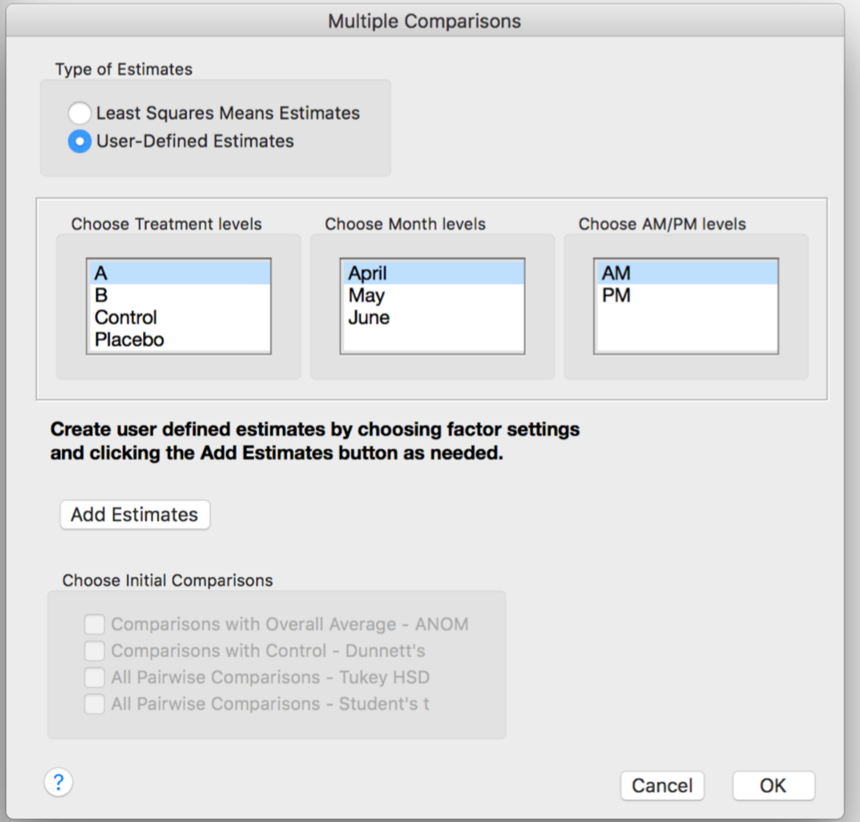
MORE FROM THE FINAL MODEL:

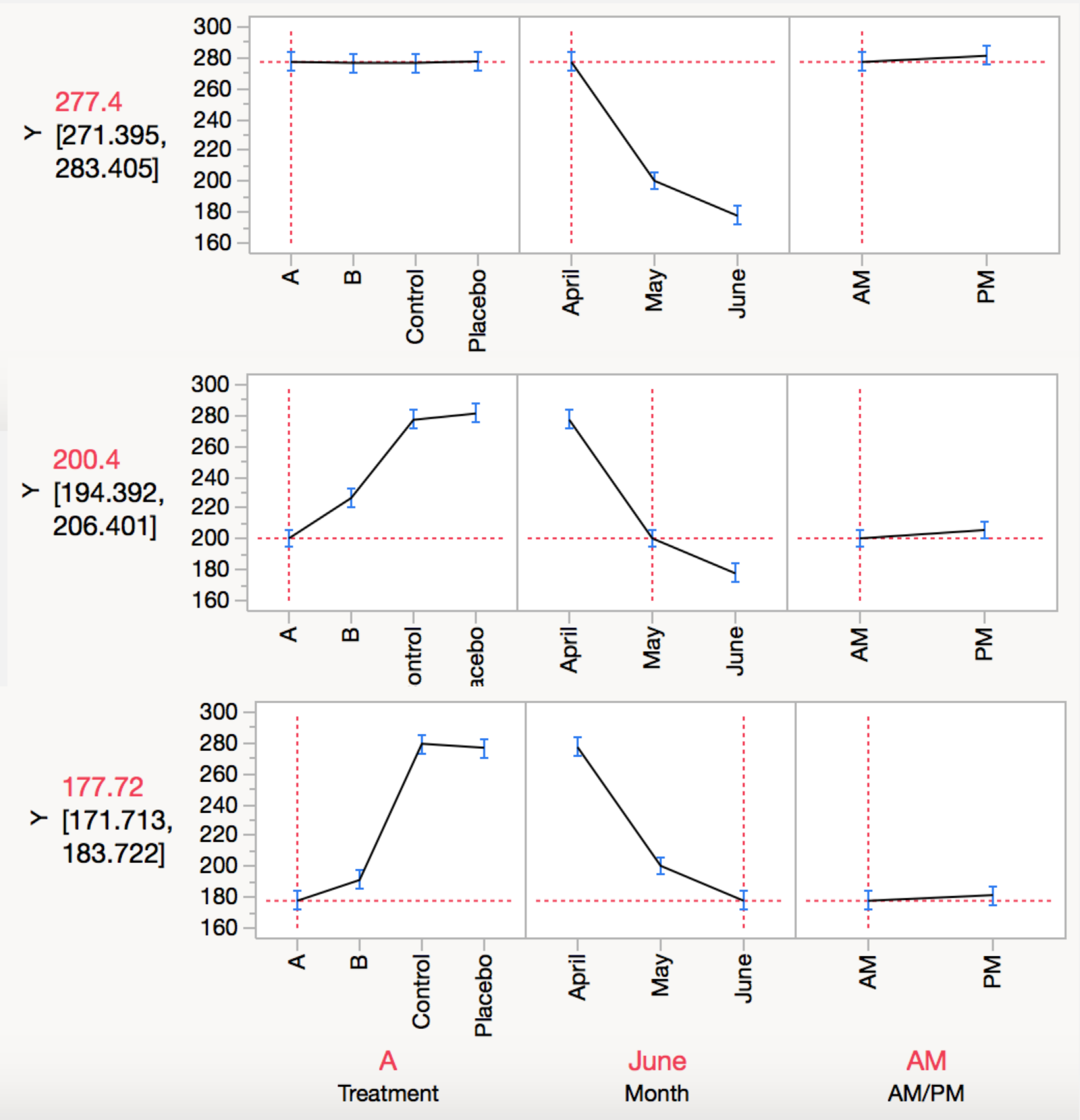


Assess residuals; save out predicted values or standard errors

Make estimates, comparisons, or contrasts







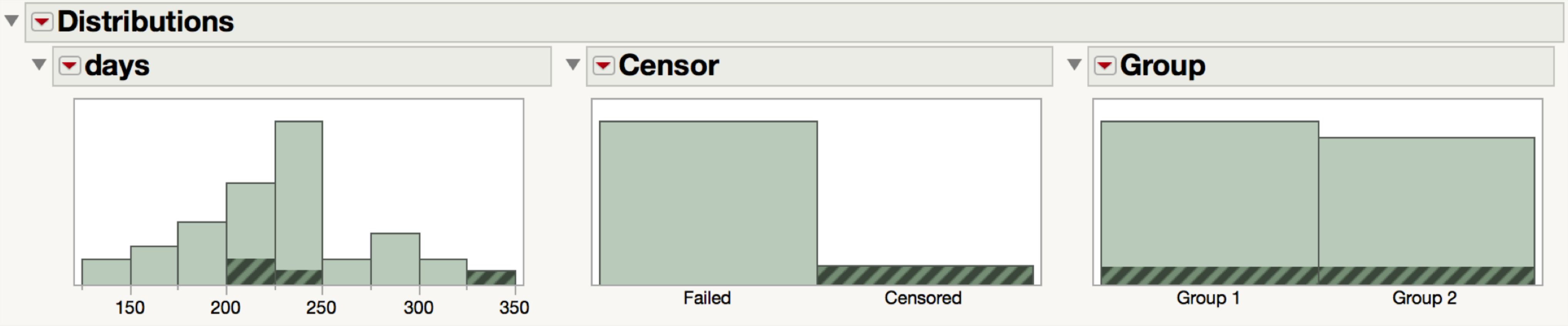
- If we do not appropriately account for the *lack of independence* between repeated measures, as well as any *random effects* that can explain some of the variance, we might not have the power to **detect significant effects**, or we might have **wide CIs**.

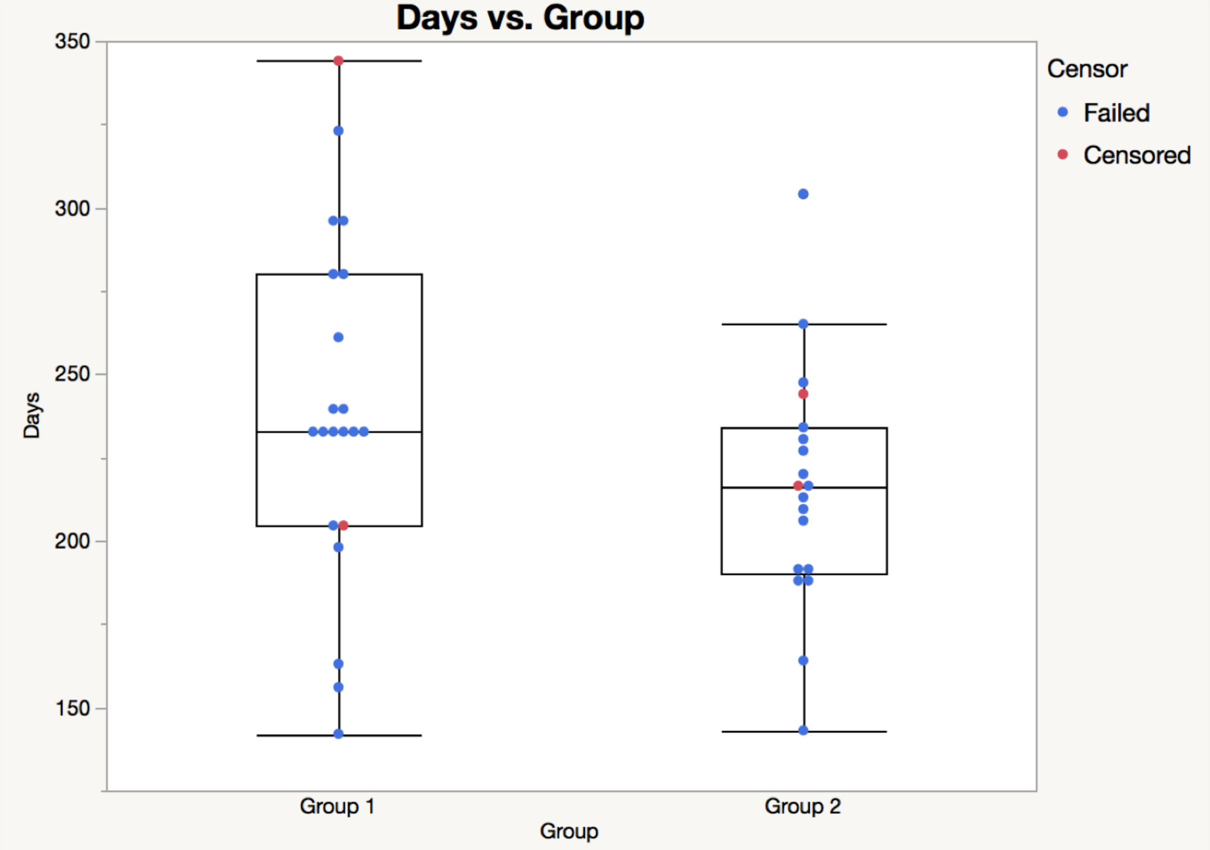
SURVIVAL ANALYSIS

- An experiment was undertaken to characterize the survival time of rats exposed to a carcinogen in two treatment groups.
- The event in this example is death.
- The objective is to see whether rats in one treatment group live longer (more days) than rats in the other treatment group.

Rats.jmp

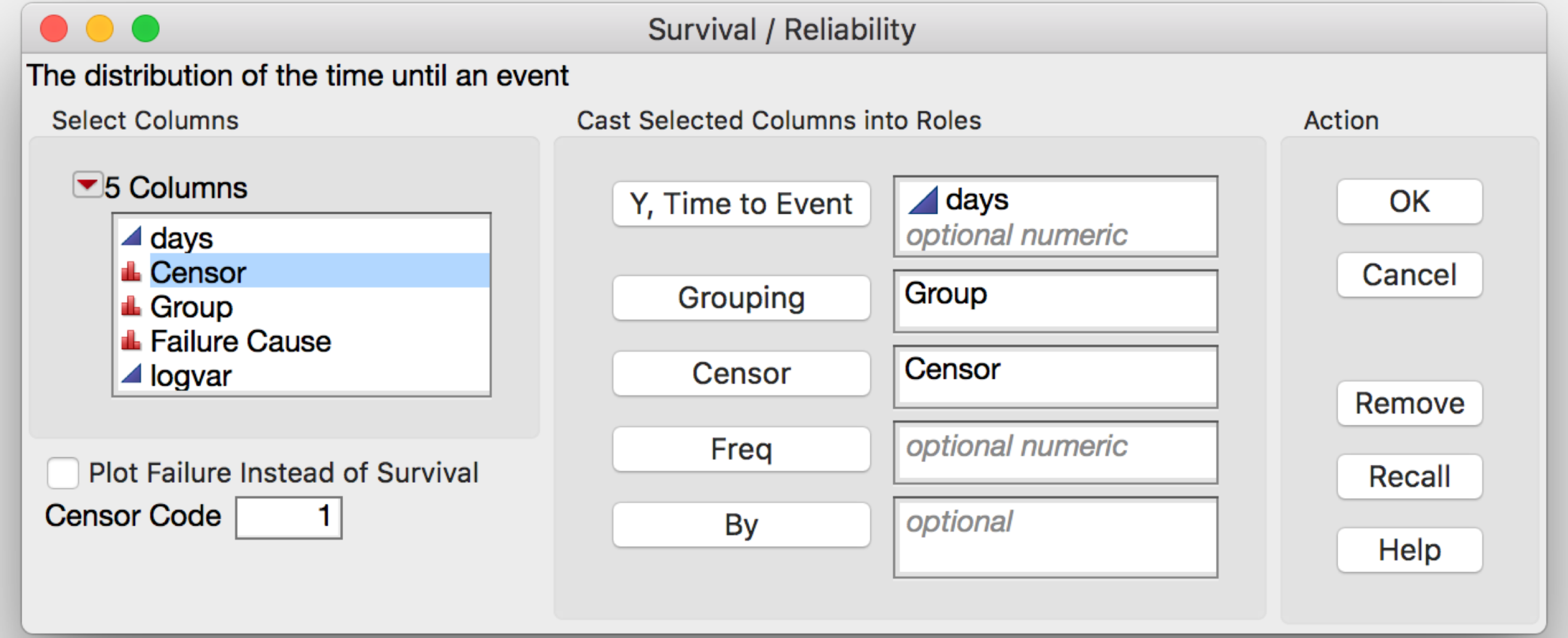
	days	Censor	Group	Failure Cause
1	142	Failed	Group 1	overdose
2	143	Failed	Group 2	overdose
3	156	Failed	Group 1	overdose
4	163	Failed	Group 1	overdose
5	164	Failed	Group 2	overdose
6	188	Failed	Group 2	accident
7	188	Failed	Group 2	accident
8	190	Failed	Group 2	accident
9	192	Failed	Group 2	accident
10	198	Failed	Group 1	accident
11	204	Censor...	Group 1	
12	205	Failed	Group 1	overdose
13	206	Failed	Group 2	overdose
14	209	Failed	Group 2	overdose
15	213	Failed	Group 2	overdose
16	216	Failed	Group 2	overdose
17	216	Censor...	Group 2	
18	220	Failed	Group 2	overdose



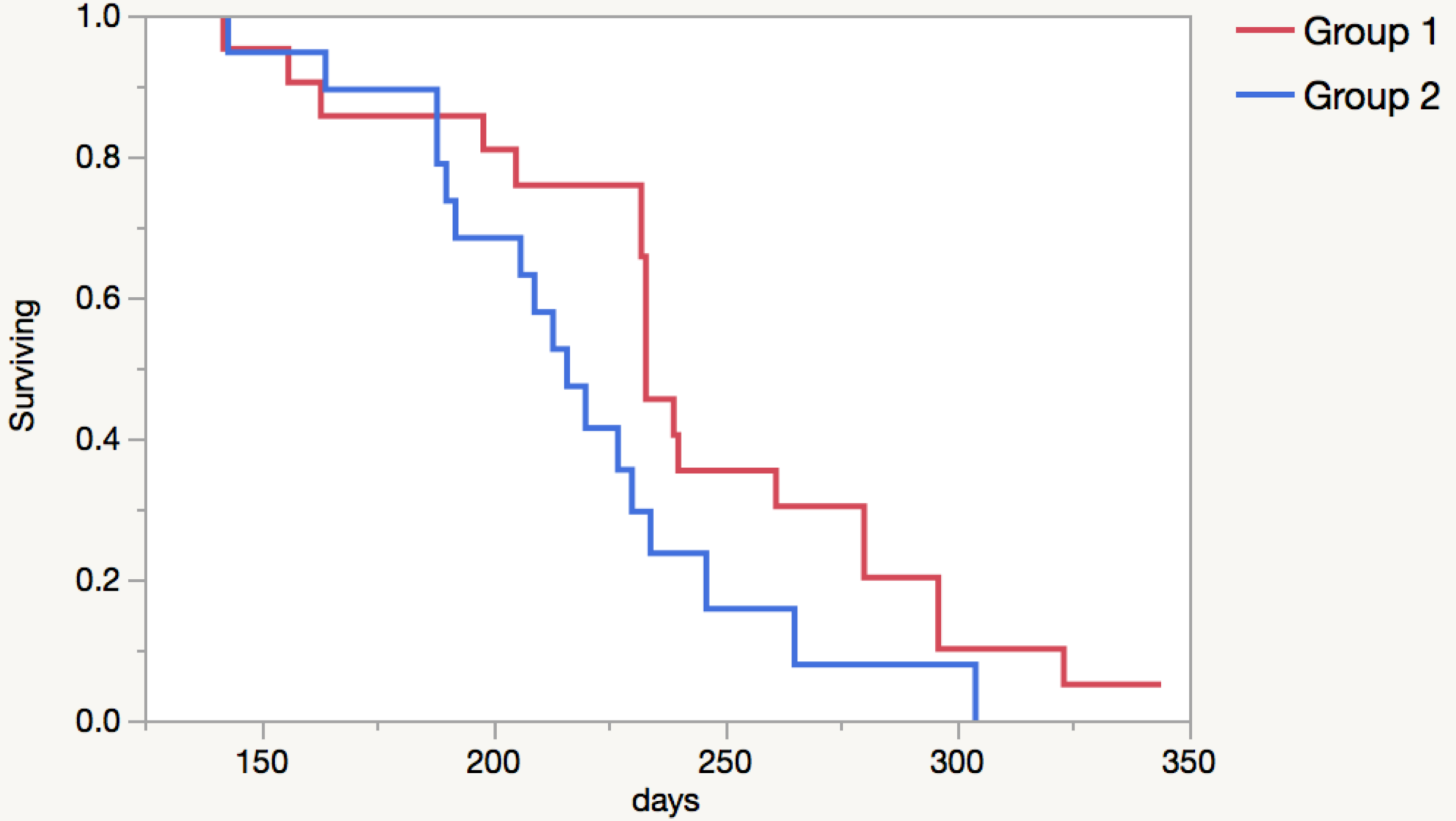


		days			
Group	**Censor**	**N**	**Min**	**Mean**	**Max**
Group 1	Failed	19.00	142.00	235.53	323.00
	Censored	2.00	204.00	274.00	344.00
Group 2	Failed	17.00	143.00	213.82	304.00
	Censored	2.00	216.00	230.00	244.00

Go to **Analyze → Reliability and Survival → Survival**



Survival Plot



Time to event: days
Censored by Censor
Censor Code 1
Grouped by Group

Group 1

days	Survival	Failure	SurvStdErr	Number failed	Number censored	At Risk
0.000	1.0000	0.0000	0.0000	0	0	21
142.000	0.9524	0.0476	0.0465	1	0	21
156.000	0.9048	0.0952	0.0641	1	0	20
163.000	0.8571	0.1429	0.0764	1	0	19
198.000	0.8095	0.1905	0.0857	1	0	18
204.000	0.8095	0.1905	0.0857	0	1	17
205.000	0.7589	0.2411	0.0941	1	0	16
232.000	0.6577	0.3423	0.1053	2	0	15
233.000	0.4554	0.5446	0.1114	4	0	13
239.000	0.4048	0.5952	0.1099	1	0	9
240.000	0.3542	0.6458	0.1072	1	0	8
261.000	0.3036	0.6964	0.1031	1	0	7
280.000	0.2024	0.7976	0.0902	2	0	6
296.000	0.1012	0.8988	0.0678	2	0	4
323.000	0.0506	0.9494	0.0493	1	0	2
344.000	0.0506	0.9494	0.0493	0	1	1

Group 2

days	Survival	Failure	SurvStdErr	Number failed	Number censored	At Risk
0.000	1.0000	0.0000	0.0000	0	0	19
143.000	0.9474	0.0526	0.0512	1	0	19
164.000	0.8947	0.1053	0.0704	1	0	18
188.000	0.7895	0.2105	0.0935	2	0	17
190.000	0.7368	0.2632	0.1010	1	0	15
192.000	0.6842	0.3158	0.1066	1	0	14
206.000	0.6316	0.3684	0.1107	1	0	13
209.000	0.5789	0.4211	0.1133	1	0	12
213.000	0.5263	0.4737	0.1145	1	0	11
216.000	0.4737	0.5263	0.1145	1	1	10
220.000	0.4145	0.5855	0.1145	1	0	8
227.000	0.3553	0.6447	0.1124	1	0	7
230.000	0.2961	0.7039	0.1082	1	0	6
234.000	0.2368	0.7632	0.1015	1	0	5
244.000	0.2368	0.7632	0.1015	0	1	4
246.000	0.1579	0.8421	0.0934	1	0	3
265.000	0.0789	0.9211	0.0728	1	0	2
304.000	0.0000	1.0000	0.0000	1	0	1

Combined

days	Survival	Failure	SurvStdErr	Number failed	Number censored	At Risk
0.000	1.0000	0.0000	0.0000	0	0	40
142.000	0.9750	0.0250	0.0247	1	0	40
143.000	0.9500	0.0500	0.0345	1	0	39
156.000	0.9250	0.0750	0.0416	1	0	38
163.000	0.9000	0.1000	0.0474	1	0	37
164.000	0.8750	0.1250	0.0523	1	0	36
188.000	0.8250	0.1750	0.0601	2	0	35
190.000	0.8000	0.2000	0.0632	1	0	33
192.000	0.7750	0.2250	0.0660	1	0	32
198.000	0.7500	0.2500	0.0685	1	0	31
204.000	0.7500	0.2500	0.0685	0	1	30
205.000	0.7241	0.2759	0.0708	1	0	29
206.000	0.6983	0.3017	0.0729	1	0	28
209.000	0.6724	0.3276	0.0746	1	0	27
213.000	0.6466	0.3534	0.0761	1	0	26
216.000	0.6207	0.3793	0.0773	1	1	25
220.000	0.5937	0.4063	0.0785	1	0	23
227.000	0.5667	0.4333	0.0795	1	0	22
230.000	0.5397	0.4603	0.0801	1	0	21
232.000	0.4858	0.5142	0.0807	2	0	20
233.000	0.3778	0.6222	0.0788	4	0	18
234.000	0.3508	0.6492	0.0776	1	0	14
239.000	0.3238	0.6762	0.0762	1	0	13
240.000	0.2969	0.7031	0.0745	1	0	12
244.000	0.2969	0.7031	0.0745	0	1	11
246.000	0.2672	0.7328	0.0727	1	0	10
261.000	0.2375	0.7625	0.0704	1	0	9
265.000	0.2078	0.7922	0.0676	1	0	8
280.000	0.1484	0.8516	0.0599	2	0	7
296.000	0.0891	0.9109	0.0485	2	0	5
304.000	0.0594	0.9406	0.0404	1	0	3
323.000	0.0297	0.9703	0.0291	1	0	2
344.000	0.0297	0.9703	0.0291	0	1	1

Summary

Group	Number failed	Number censored	Mean		Std Error
Group 1	19	2	240.795	Biased	11.206
Group 2	17	2	218.757		9.40318
Combined	36	4	230.729	Biased	7.57346

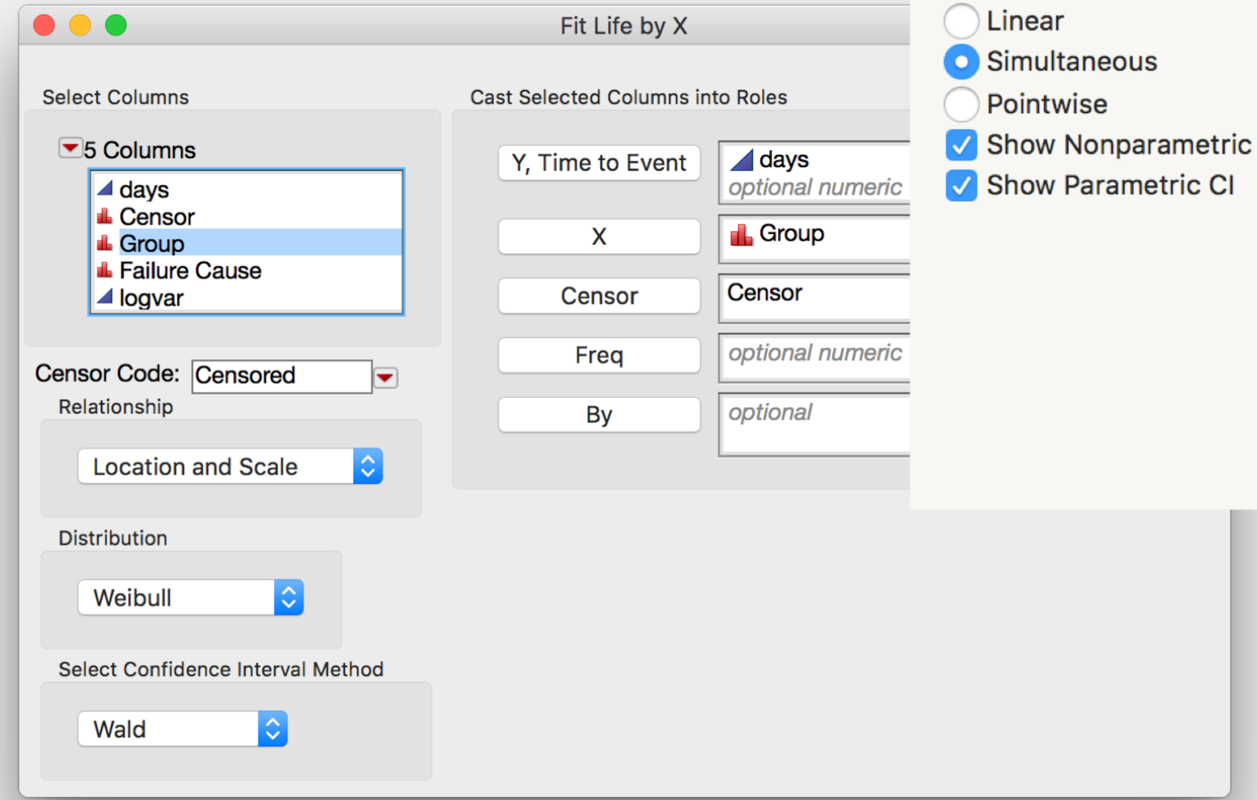
Quantiles

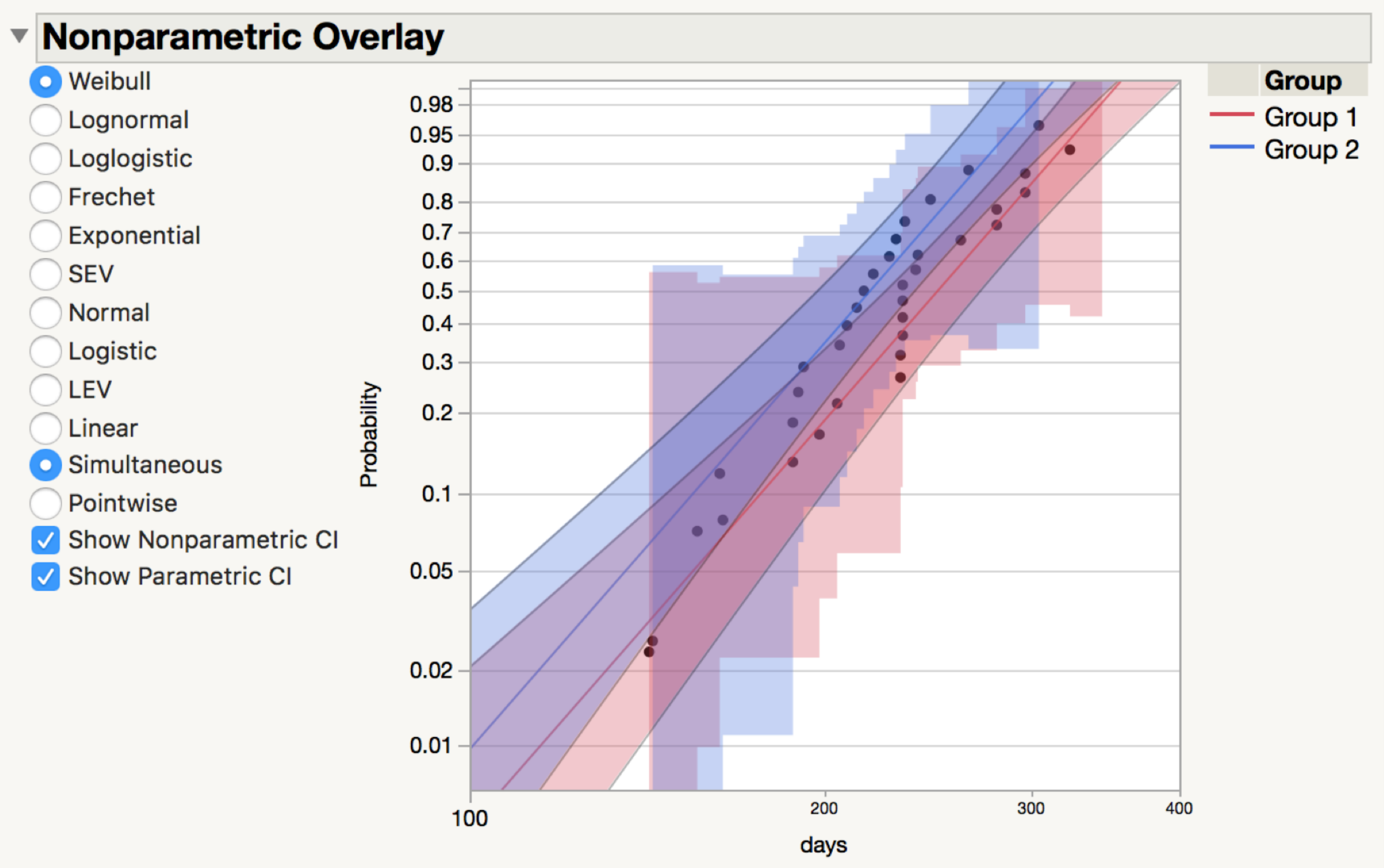
Group	Median Time	Lower 95%	Upper 95%	25% Failures	75% Failures
Group 1	233	232	280	232	280
Group 2	216	190	234	190	234
Combined	232	213	239	201.5	261

Tests Between Groups

Test	ChiSquare	DF	Prob>ChiSq
Log-Rank	3.1227	1	0.0772
Wilcoxon	2.6510	1	0.1035

If we want to make estimates of how long we expect the rat to survive, or estimate how many rats will die by a certain day, we can use **Analyze → Reliability and Survival → Life Distribution** or **Fit Life by X**





Custom Quantile Estimation

Failure Probability	Group	Life Time Quantile	Life Time Quantile Lower 95%	Life Time Quantile Upper 95%
0.10000000	Group 1	175.97737	152.23873	203.41758
0.10000000	Group 2	154.21881	133.64227	177.96345
0.50000000	Group 1	248.55432	227.34570	271.74146
0.50000000	Group 2	217.82205	198.73803	238.73863

Custom Probability Estimation

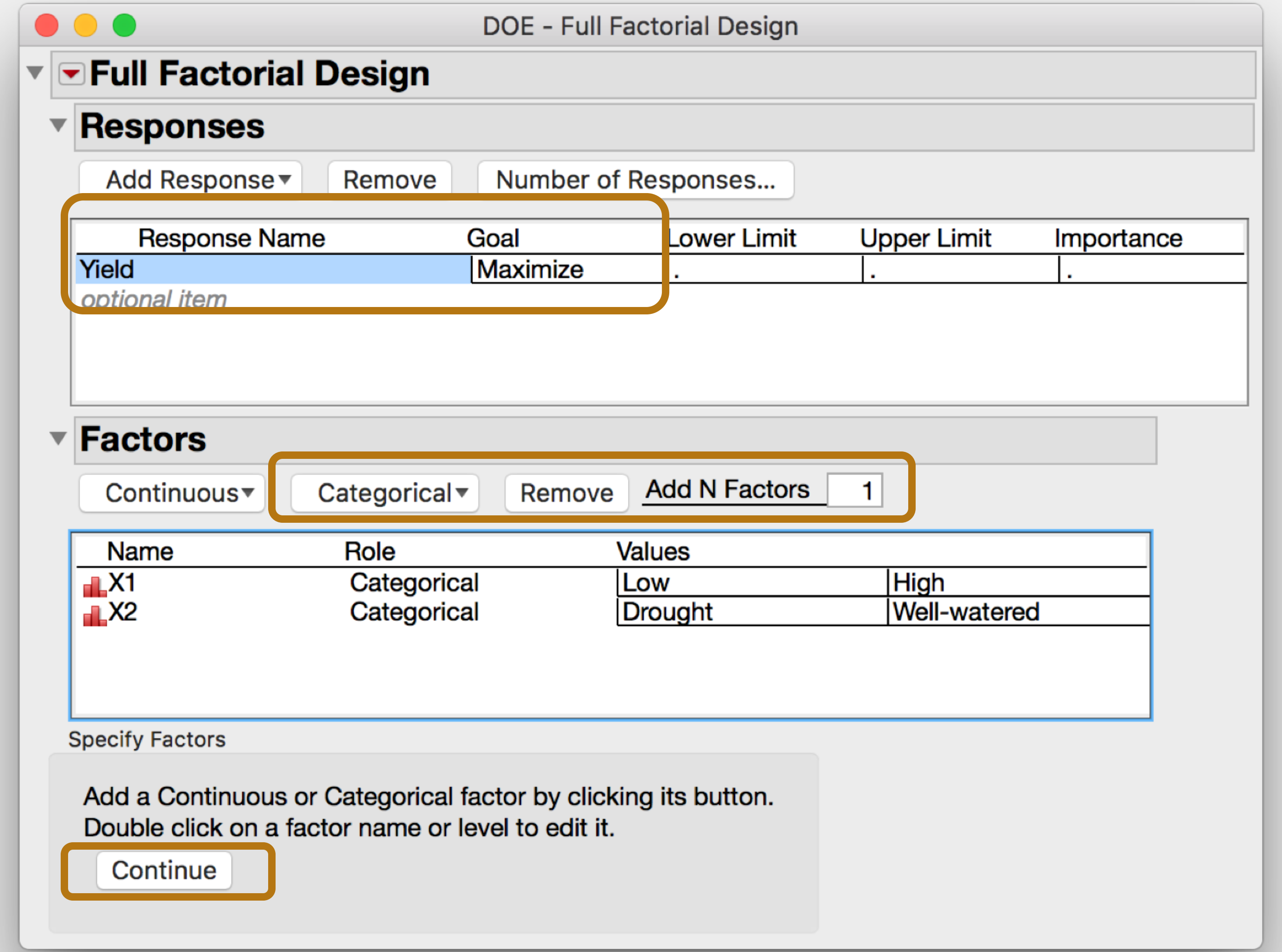
days	Group	Failure Probability	Failure Probability Lower 95%	Failure Probability Upper 95%
200.00000	Group 1	0.19084879	0.10403175	0.33518689
200.00000	Group 2	0.35279158	0.22259227	0.52849164
300.00000	Group 1	0.85549034	0.70705377	0.95253491
300.00000	Group 2	0.98120760	0.89927006	0.99897335

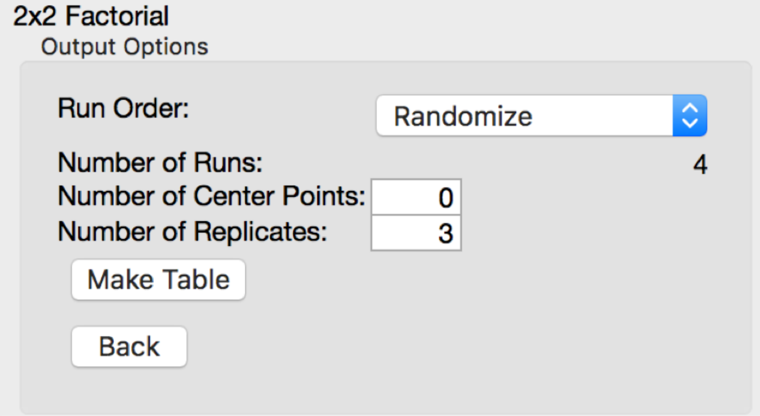
- If we want to add more predictor terms into the model, or do modern variable selection we can use **Analyze → Fit Model** and choose the
  - **Generalized Regression**,
  - **Parametric Survival**, or
  - **Proportional Hazards** personalities

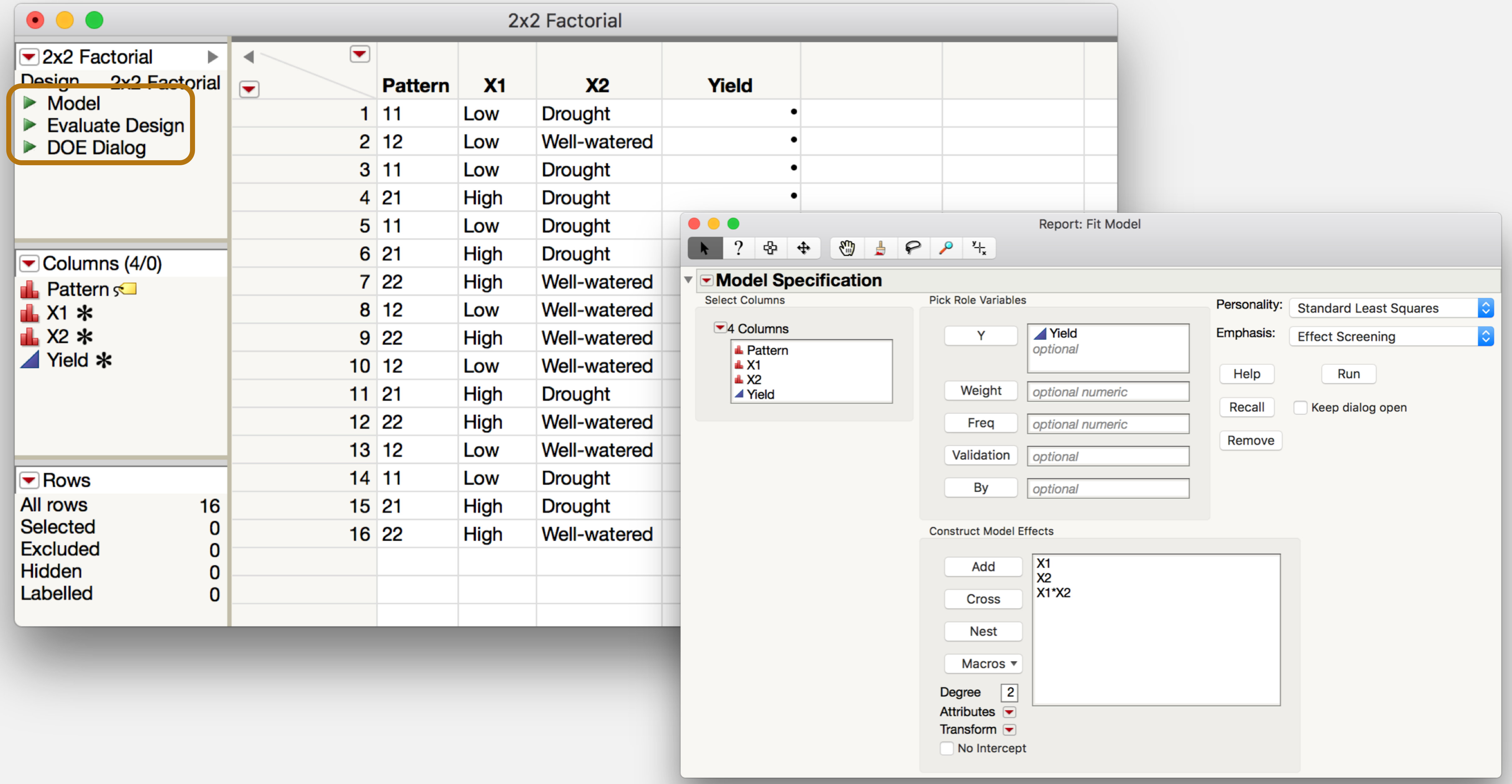
DESIGNED EXPERIMENTS

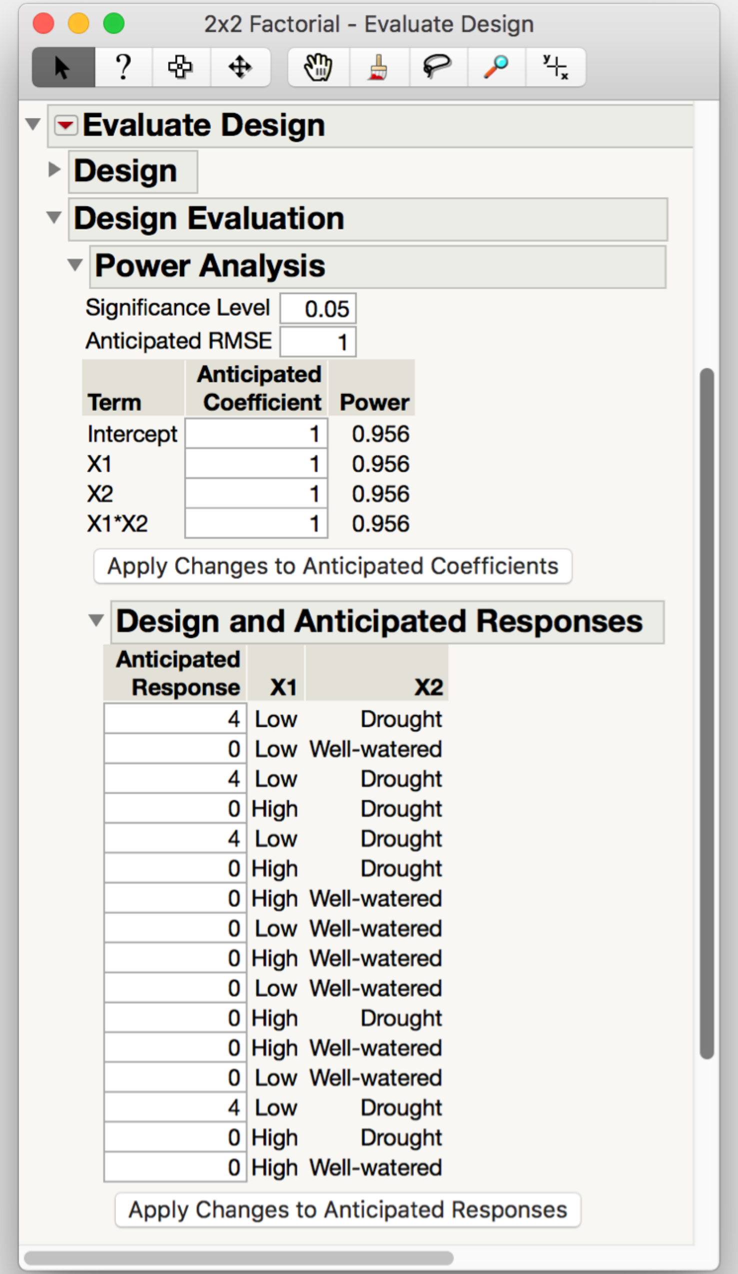
- Plan to measure Yield of plants that receive either the high or low level of a Fertilizer treatment and under either Drought or Well-watered conditions.
- You plan to run this as a balanced full factorial.
- You can run 4 experimental units in each treatment combination.

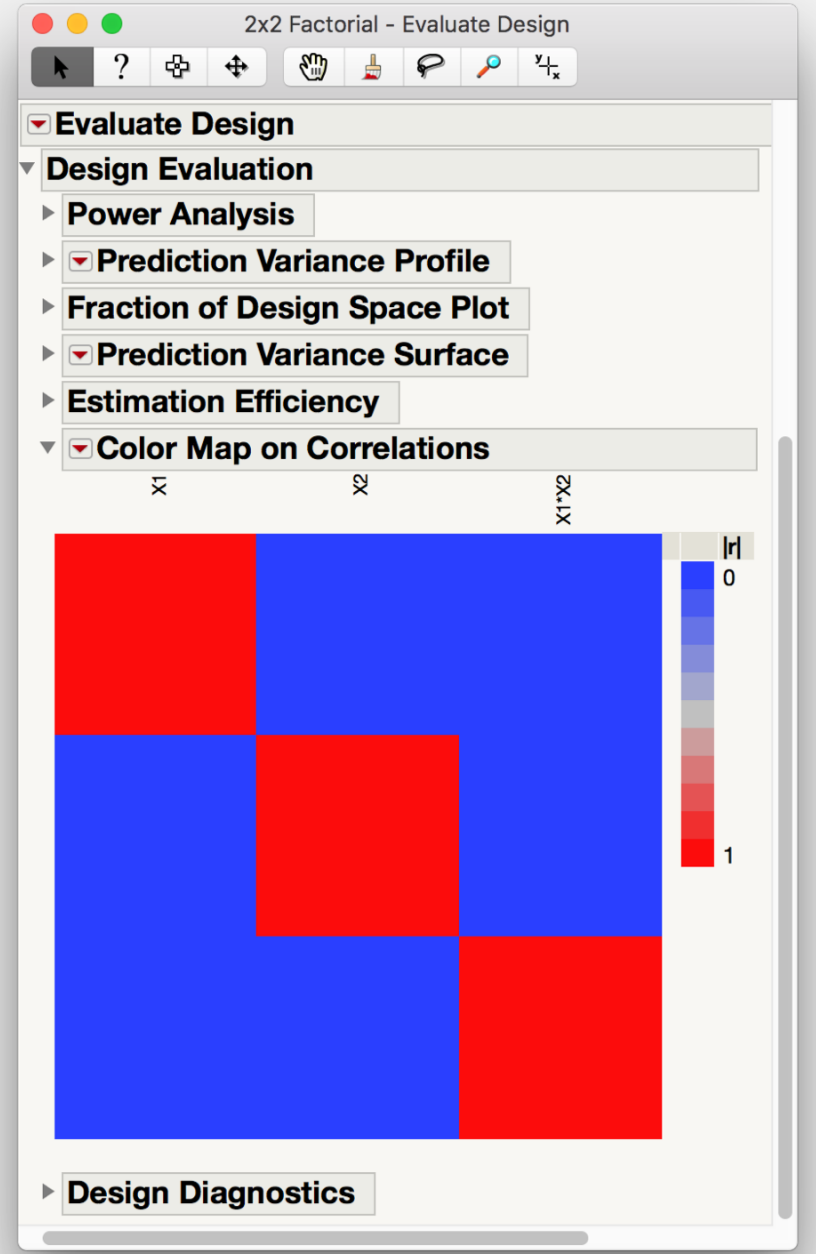
Go to **DOE → Classical → Full Factorial Design**

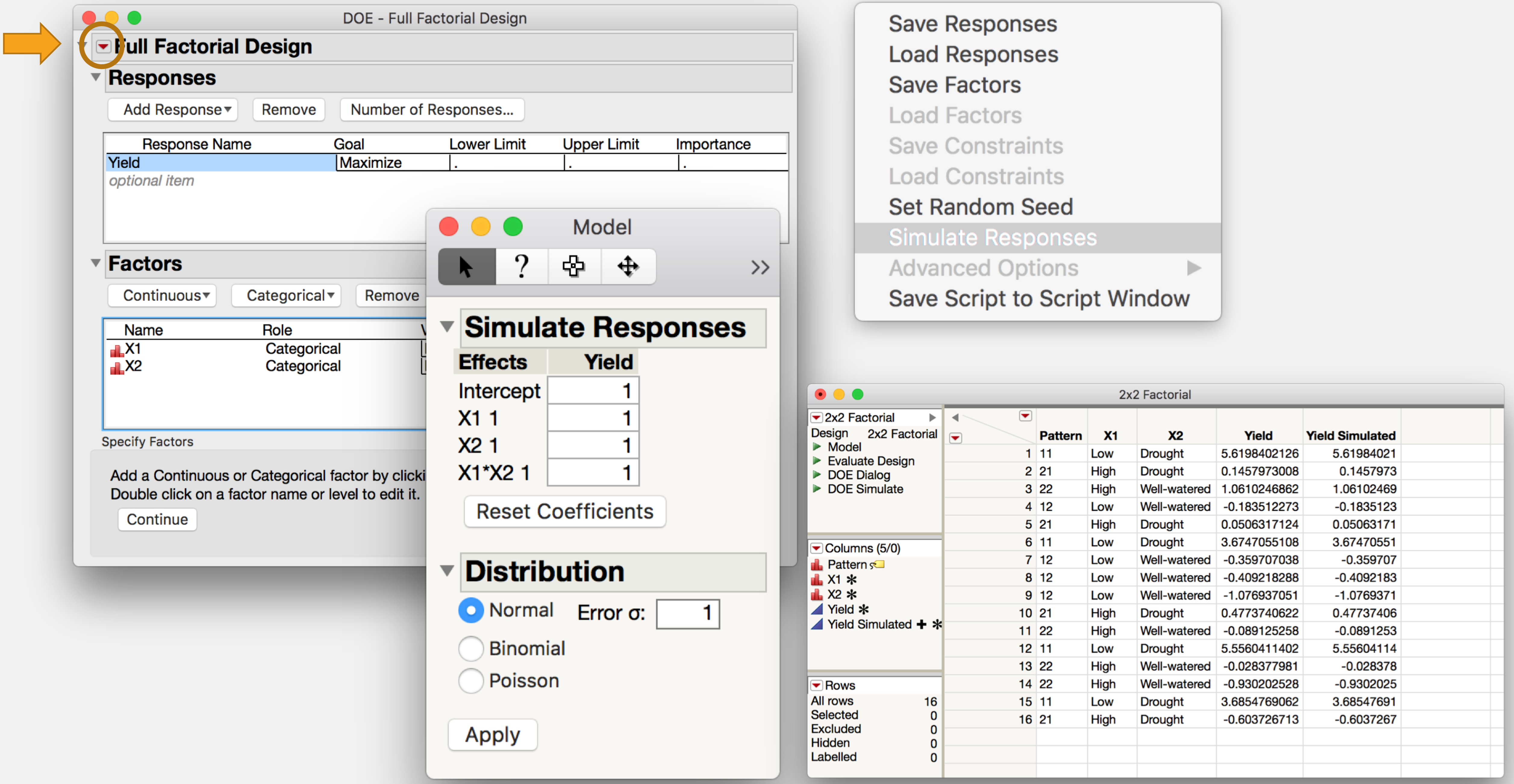





2x2 Factorial
2x2 Factorial
Design 2x2 Factorial
Model
Evaluate Design
DOE Dialog
Columns (4/0)
Pattern
X1
X2
Yield
Rows
All rows 16
Selected 0
Excluded 0
Hidden 0
Labelled 0
Pattern X1 X2 Yield
1 11 Low Drought
2 12 Low Well-watered
3 11 Low Drought
4 21 High Drought
5 11 Low Drought
6 21 High Drought
7 22 High Well-watered
8 12 Low Well-watered
9 22 High Well-watered
10 12 Low Well-watered
11 21 High Drought
12 22 High Well-watered
13 12 Low Well-watered
14 11 Low Drought
15 21 High Drought
16 22 High Well-watered
Report: Fit Model
Model Specification
Select Columns
4 Columns
Pattern
X1
X2
Yield
Pick Role Variables
Y
Yield
optional
Weight
optional numeric
Freq
optional numeric
Validation
optional
By
optional
Personality: Standard Least Squares
Emphasis: Effect Screening
Help
Run
Recall
Keep dialog open
Remove
Construct Model Effects
Add
Cross
Nest
Macros
Degree 2
Attributes
Transform
No Intercept
X1
X2
X1*X2


2x2 Factorial - Evaluate Design
Evaluate Design
Design
Design Evaluation
Power Analysis
Significance Level 0.05
Anticipated RMSE 1
Term | Anticipated Coefficient | Power
Intercept 1 0.956
X1 1 0.956
X2 1 0.956
X1*X2 1 0.956
Apply Changes to Anticipated Coefficients
Design and Anticipated Responses
Anticipated Response | X1 | X2
4 Low Drought
0 Low Well-watered
4 Low Drought
0 High Drought
4 Low Drought
0 High Drought
0 High Well-watered
0 Low Well-watered
0 High Well-watered
0 Low Well-watered
0 High Drought
0 High Well-watered
0 Low Well-watered
4 Low Drought
0 High Drought
0 High Well-watered
Apply Changes to Anticipated Responses


2x2 Factorial - Evaluate Design
Evaluate Design
Design Evaluation
Power Analysis
Prediction Variance Profile
Fraction of Design Space Plot
Prediction Variance Surface
Estimation Efficiency
Color Map on Correlations
X1
X2
X1*X2
|r|
0
1
Design Diagnostics



	Pattern	X1	X2	Yield	Yield Simulated
1	11	Low	Drought	5.6198402126	5.61984021
2	21	High	Drought	0.1457973008	0.1457973
3	22	High	Well-watered	1.0610246862	1.06102469
4	12	Low	Well-watered	-0.183512273	-0.1835123
5	21	High	Drought	0.0506317124	0.05063171
6	11	Low	Drought	3.6747055108	3.67470551
7	12	Low	Well-watered	-0.359707038	-0.359707
8	12	Low	Well-watered	-0.409218288	-0.4092183
9	12	Low	Well-watered	-1.076937051	-1.0769371
10	21	High	Drought	0.4773740622	0.47737406
11	22	High	Well-watered	-0.089125258	-0.0891253
12	11	Low	Drought	5.5560411402	5.55604114
13	22	High	Well-watered	-0.028377981	-0.028378
14	22	High	Well-watered	-0.930202528	-0.9302025
15	11	Low	Drought	3.6854769062	3.68547691
16	21	High	Drought	-0.603726713	-0.6037267

OTHER TOPICS

- Connecting JMP with SAS, R, Matlab, or Python, or capturing real-time data into JMP:

    Help → Books → Scripting Guide → Extending JMP

- More Power and Sample Size situations

    www.jmp.com/learn → Basic Proportions and Means

    Help → Books → Design of Experiments Guide → Prospective Sample Size and Power

- More examples on specific topics, search for "**JMP On-Demand Webcasts**" to find many!
    - Text Mining to find topics or insights from unstructured text data, such as patient charts
    - More generalized regression, more survival analysis
    - More Designed Experiments