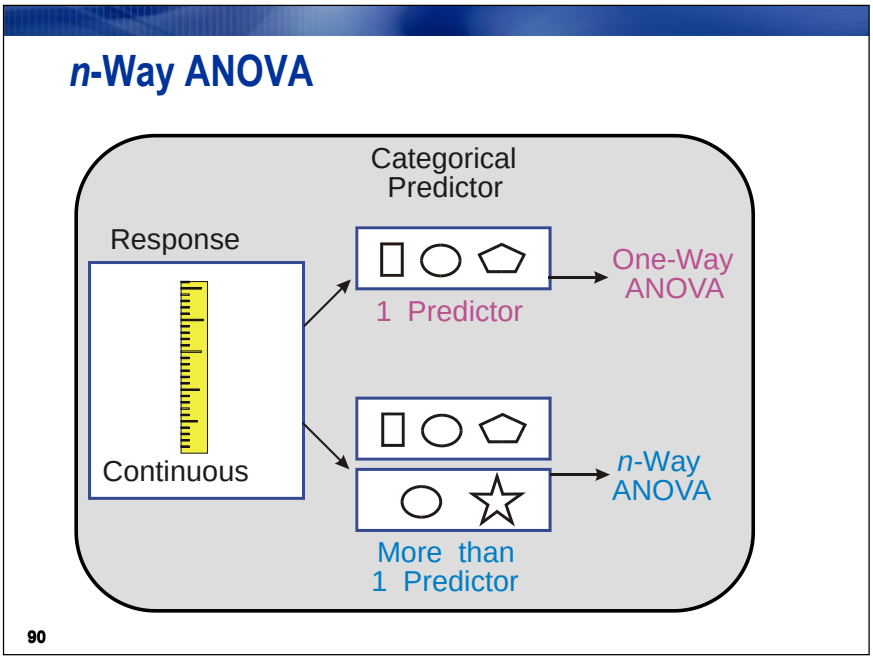
SASEG 6C – Two Way ANOVA

(Fall 2015)

Sources (adapted with permission)-

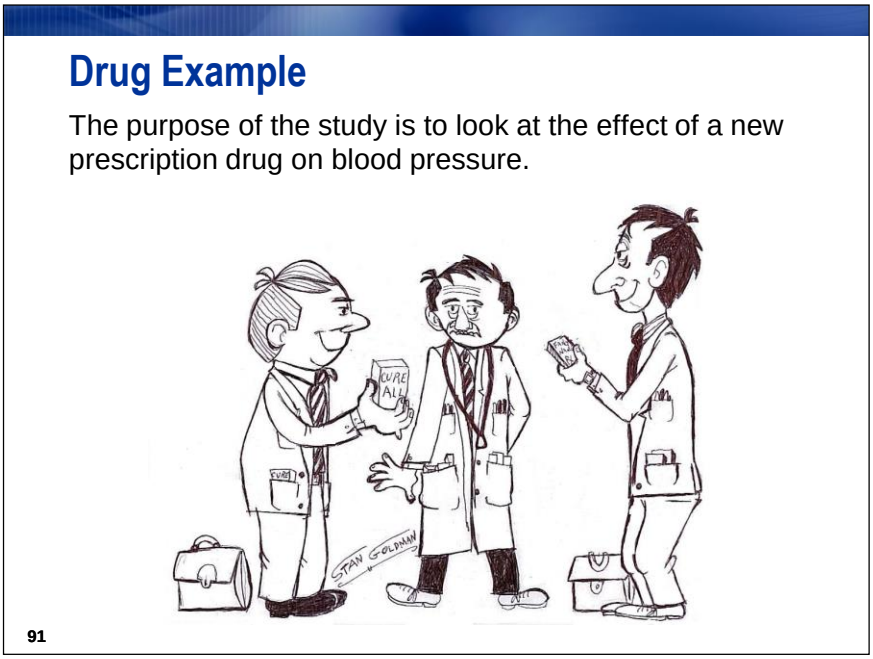
T. P. Cronan, Jeff Mullins, Ron Freeze and David E. Douglas Course and Classroom Notes
Enterprise Systems, Sam M. Walton College of Business, University of Arkansas, Fayetteville
Microsoft Enterprise Consortium
IBM Academic Initiative
SAS® Multivariate Statistics Course Notes & Workshop, 2010
SAS® Advanced Business Analytics Course Notes & Workshop, 2010
Microsoft® Notes
Teradata® University Network

Two-Way ANOVA with Interactions

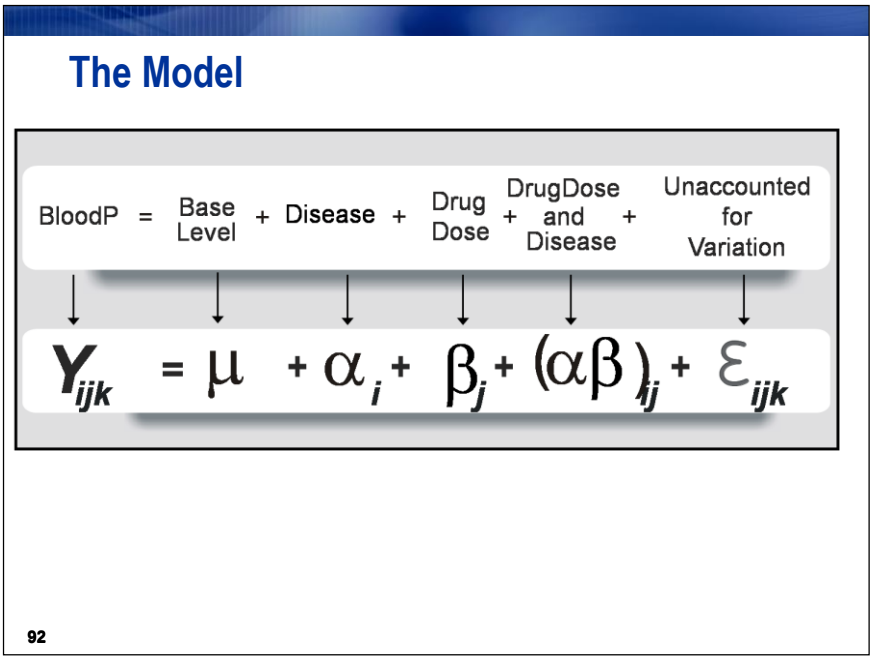


In the previous section, you considered the case where you had one categorical predictor and a blocking variable. In this section, consider a case with two categorical predictors. In general, any time you have more than one categorical predictor variable and a continuous response variable, it is called *n*-way ANOVA. The *n* can be replaced with the number of categorical predictor variables.

The analysis for a randomized block design is actually a special type of *n*-way ANOVA.



Data was collected in an effort to determine whether different dose levels of a given drug have an effect on blood pressure for people with one of three types of heart disease.



Y_{ijk} the observed **`BloodP`** for each subject

μ the overall base level of the response, **`BloodP`**

α_i the effect of the i^{th} **`Disease`**

β_j the effect of the j^{th} **`DrugDose`**

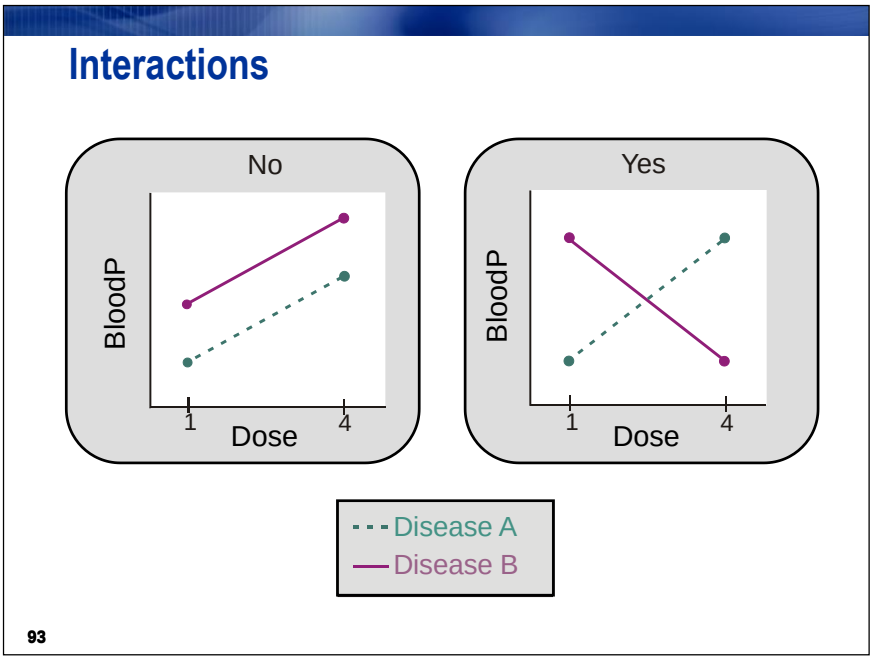
$(\alpha\beta)_{ij}$ the effect of the interaction between the i^{th} **`Disease`** and the j^{th} **`DrugDose`**

ε_{ijk} error term, or residual

In the model it is assumed that the:

- observations are independent
- data are normally distributed for each treatment
- variances are equal across treatments

Verifying ANOVA assumptions with more than two variables is covered in Statistics II: ANOVA and Regression.



An interaction occurs when a change in the level of one factor results in a change in the difference between levels of the other factor.

The average blood pressure change over different doses were plotted in mean plots and then connected for disease A and B.

In the left plot above, different types of disease show the same change across different levels of dose.

In the right plot, however, as the dose increases, average blood pressure **increases** in those with disease A, but **decreases** for those with disease B. This indicates an interaction between the variables **`DrugDose`** and **`Disease`**.

When you analyze an *n*-way ANOVA with interactions you should first look at any tests for interaction among factors.

If there is no interaction between the factors, the tests for the individual factor effects can be interpreted to determine their significance/non-significance.

If an interaction exists between any factors, the tests for the individual factor effects might be misleading, due to masking of these effects by the interaction. This is especially true for unbalanced data.

In the previous section, you used a blocking variable and a categorical predictor as effects in the model. It is generally assumed that blocks do not interact with other factors. In this section, there are just two separate factors. An interaction between the two might be hypothesized and tested.

Nonsignificant Interaction

Analyze the main effects with the interaction in the model.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

…or…

Delete the interaction from the model, and then analyze the main effects.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$$

94

When the interaction is not statistically significant, the main effects can be analyzed with the model as originally written. This is generally the method used when analyzing designed experiments.

However, even when analyzing designed experiments, some statisticians suggest that if the interaction is nonsignificant, then the interaction effect can be deleted from the model and then the main effects are analyzed. This increases the power of the main effects tests.

Neter, Kutner, Wasserman, and Nachtsheim (1996) suggest the following guidelines for when to delete the interaction from the model:

- there are fewer than 5 degrees of freedom for the error, **and**
- the F Value for the interaction term is < 2.

When you analyze data from an observational study, it is more common to delete the non-significant interaction from the model and then analyze the main effects.

Two-Way ANOVA with Interactions

Perform a two-way ANOVA using the DRUG data set.

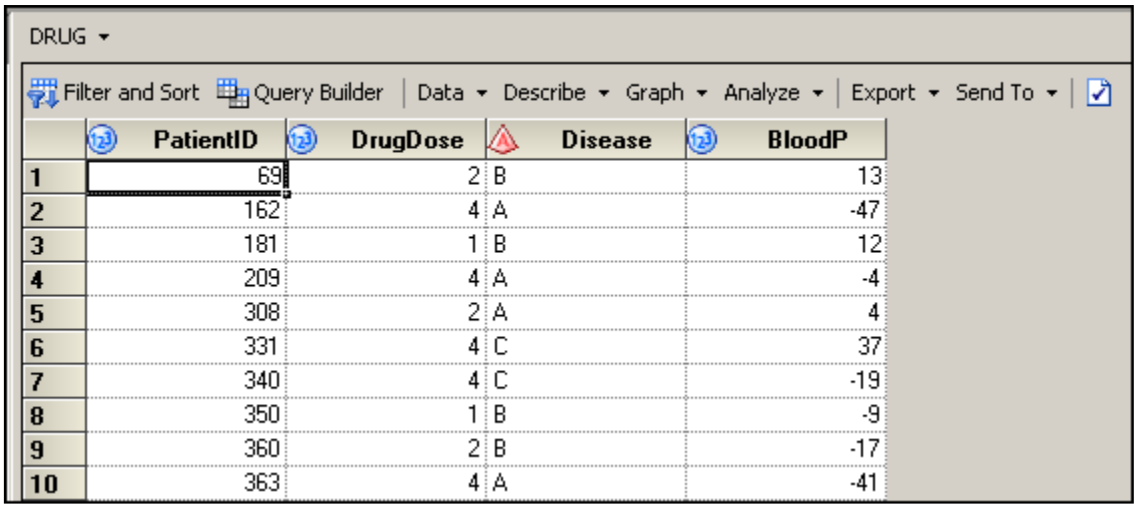
The **`Drug`** data set contains the following variables:

`DrugDose`	dosage level of drug (Placebo, 100 mg, 200 mg, 500 mg)
`Disease`	heart disease category
`BloodP`	change in diastolic blood pressure after 2 weeks treatment

Before conducting an analysis of variance, you should explore the data.

Presume that the initial data exploration was completed (output not shown here) and that no particular concerns were noted about unusual data values or the distribution of the data. During this exploration, you determine that the sample sizes for all treatments are not equal. The researchers recruited 240 patients (80 per heart disease category), but only 170 were randomized into the trial.

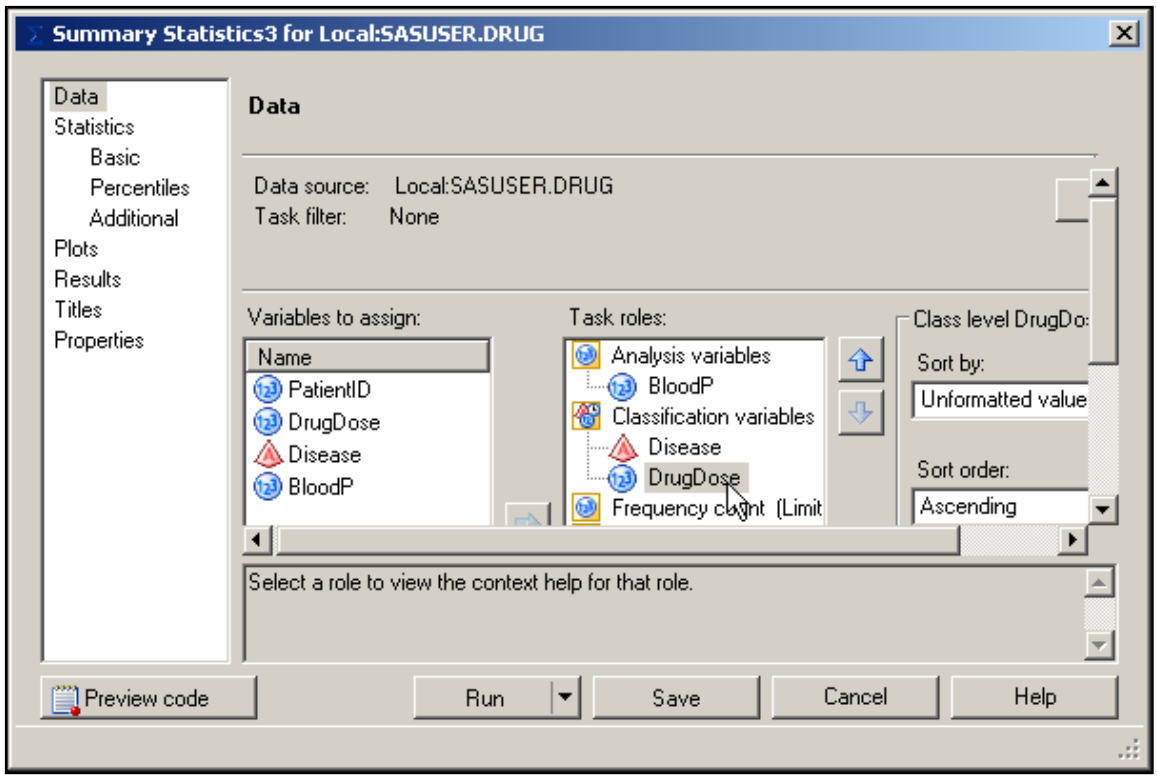
1. Open the **DRUG** data set.



	PatientID	DrugDose	Disease	BloodP
1	69	2	B	13
2	162	4	A	-47
3	181	1	B	12
4	209	4	A	-4
5	308	2	A	4
6	331	4	C	37
7	340	4	C	-19
8	350	1	B	-9
9	360	2	B	-17
10	363	4	A	-41

Negative values for **BloodP** mean that diastolic blood pressure was reduced by that amount. Positive values mean that the **BloodP** increased.

2. Use the summary statistics task, choosing **BloodP** for the analysis variables role and **Disease** and **DrugDose** for the classification variables role.



Analysis Variable : BloodP							
Disease	DrugDose	N Obs	Mean	Std Dev	Minimum	Maximum	N
A	1	12	1.3333333	13.5333483	-22.0000000	25.0000000	12
	2	16	-9.6875000	18.8881577	-37.0000000	19.0000000	16
	3	13	-26.2307692	18.1390640	-51.0000000	11.0000000	13
	4	18	-22.5555556	21.0970369	-61.0000000	12.0000000	18
B	1	15	-8.1333333	16.9109714	-39.0000000	22.0000000	15
	2	15	5.4000000	21.8886794	-45.0000000	35.0000000	15
	3	14	24.7857143	23.7427838	-24.0000000	60.0000000	14
	4	13	23.2307692	23.5872630	-22.0000000	55.0000000	13
C	1	14	0.4285714	20.2929100	-38.0000000	45.0000000	14
	2	13	-4.8461538	24.0341637	-36.0000000	50.0000000	13
	3	14	-5.1428571	13.9827209	-26.0000000	27.0000000	14
	4	13	1.3076923	28.7847894	-57.0000000	42.0000000	13

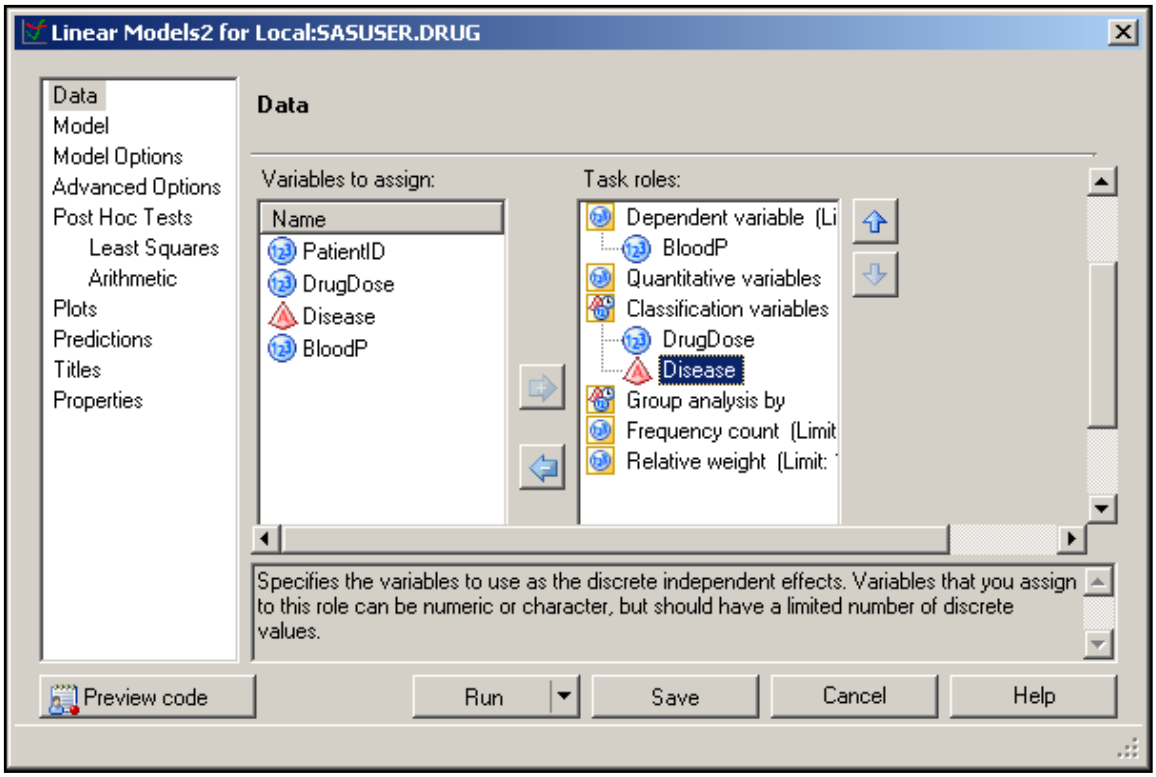
The means do seem to vary from group to group. However, it is not obvious if there is a consistent trend across different levels of ***DrugDose****. It seems that the trend of* ***BloodP*** *across levels of* ***DrugDose*** *may be different from disease to disease.*

To test the hypothesis that effect of **DrugDose** differs across diseases, perform a 2-way ANOVA with an interaction.

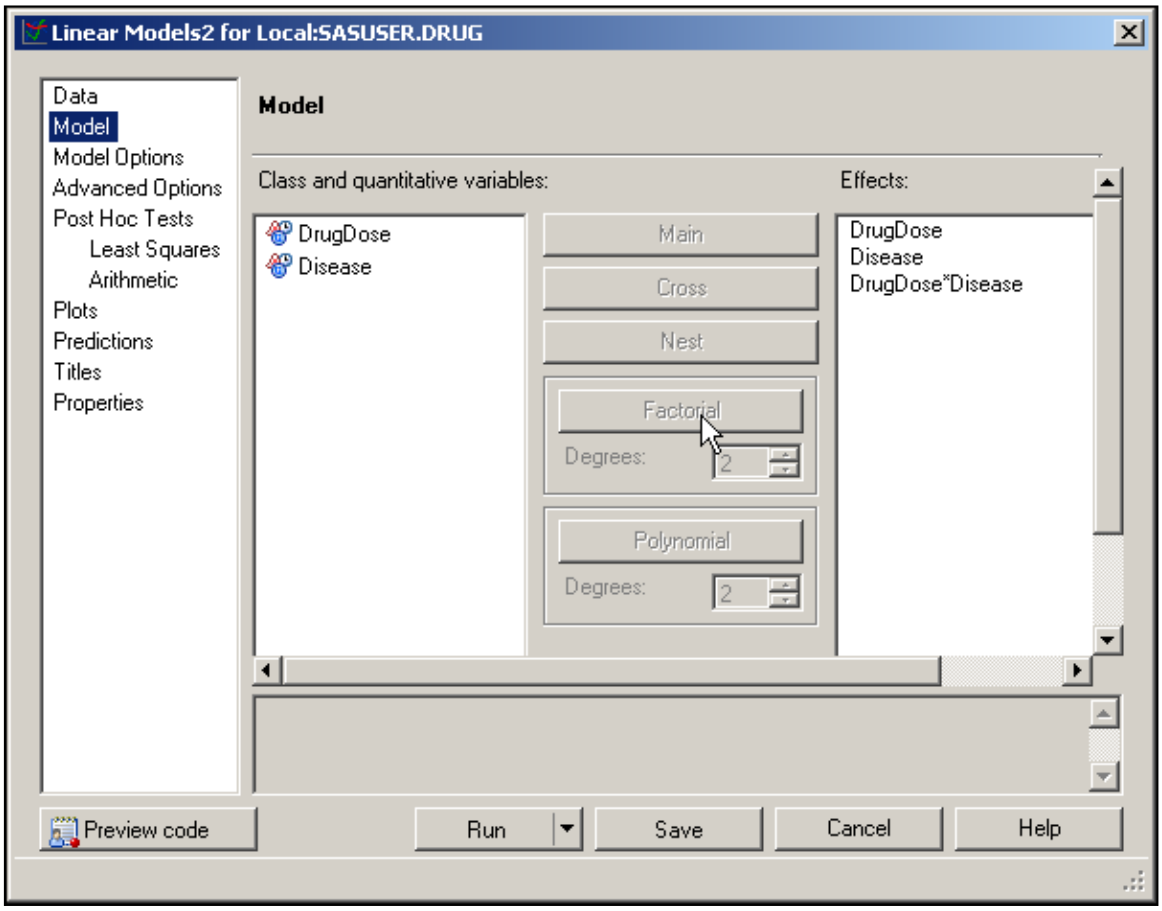
1. Select **Tasks** (or **Analyze**) ⇨ **ANOVA** ⇨ **Linear Models…**.
2. Under Data, assign **Bloodp** to the dependent variable task role and **Disease** and **DrugDose** to the classification variables role.

   DrugDose is saved in SAS as numeric and could have been assigned to the quantitative variables role. However, the scale of measurement of the variable is not continuous, but rather ordinal, so it should be added as a classification variable.

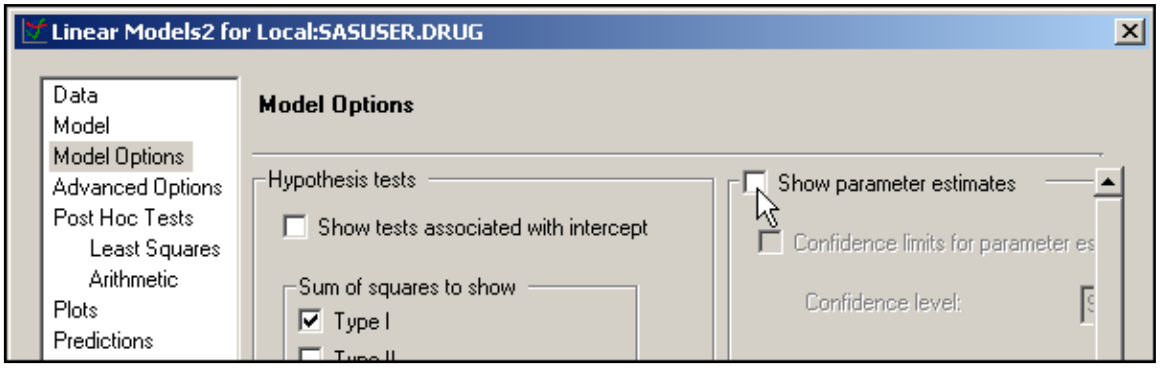
3. Select **Model** at the left and highlight both **Disease** and **DrugDose** in the `Class and quantitative variables` pane (by clicking on one and then holding down the CTRL key on the keyboard and then clicking on the other variable).


Linear Models2 for Local:SASUSER.DRUG
Data
Model
Model Options
Advanced Options
Post Hoc Tests
Least Squares
Arithmetic
Plots
Predictions
Titles
Properties
Data
Variables to assign:
Name
PatientID
DrugDose
Disease
BloodP
Task roles:
Dependent variable (Li
BloodP
Quantitative variables
Classification variables
DrugDose
Disease
Group analysis by
Frequency count (Limit
Relative weight (Limit:
Specifies the variables to use as the discrete independent effects. Variables that you assign to this role can be numeric or character, but should have a limited number of discrete values.
Preview code
Run
Save
Cancel
Help

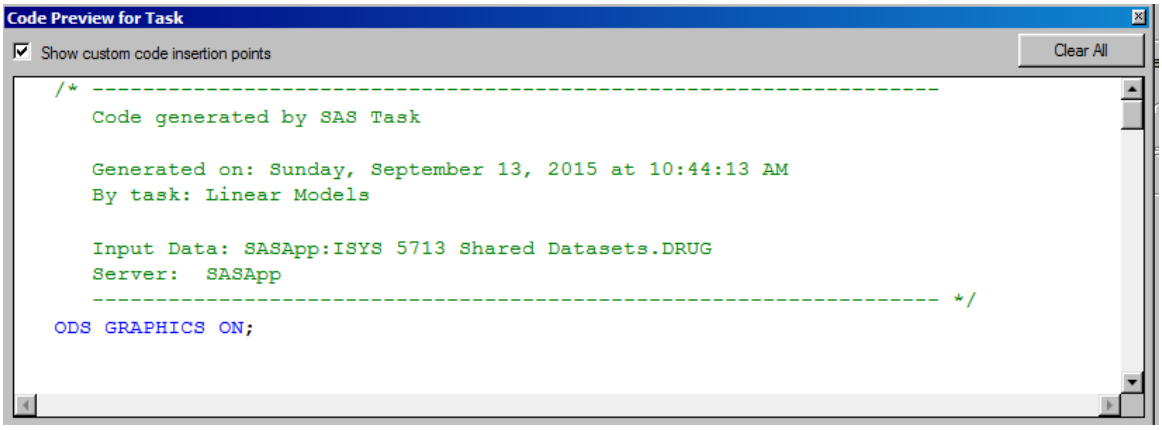
4. Click Factorial and see the terms, **DrugDose**, **Disease** and **DrugDose*Disease** displayed in the Effects pane.



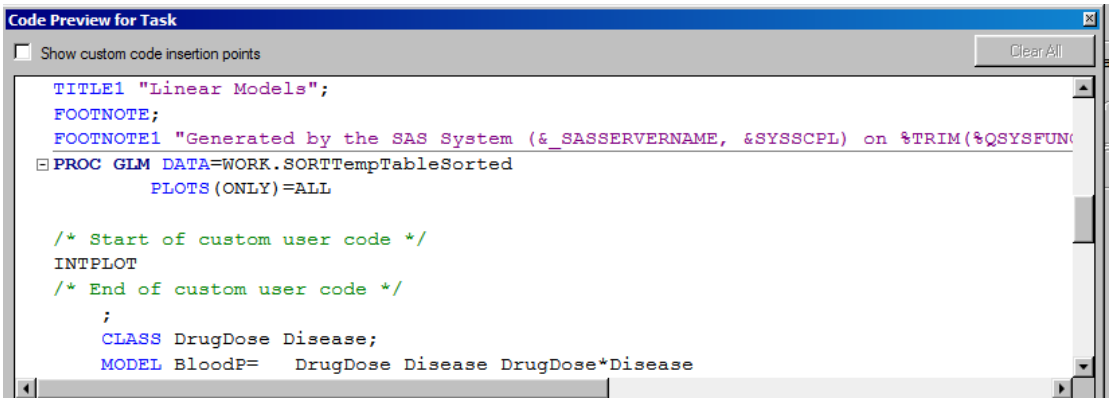
5. Under Model Options, deselect **Show parameter estimates**.



6. Click Preview code at the bottom of the window.
7. You will now see a window showing the SAS code created by the t Test task. This window is where you can directly edit the Code generated by SAS Task.
8. Select the Show custom code insertion points and…



9. Scroll down in the code to just under the words **PLOTS(ONLY)=ALL**.
10. Click on the <insert custom code here> in the area just after **PLOTS(ONLY)=ALL**
11. Type **INTPLOT**. *This is the option to produce an interaction plot.*
12. Uncheck the Show custom code insertion points. The code should look like the following.



13. Close the window by clicking ☒ in the upper right corner.
14. Click Run.

Class Level Information

Class	Levels	Values
DrugDose	4	1 2 3 4
Disease	3	A B C

Number of Observations Read	170
Number of Observations Used	170

The next part of the output, below, shows the source table with the *F* test for the overall model. This tests the null hypothesis that none of the effects in the model are statistically different, in other words, that there are no differences among the 12 group means (one for each **`DrugDose`** * **`Disease`** combination).

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	36476.8353	3316.0759	7.66	<.0001
Error	158	68366.4589	432.6991		
Corrected Total	169	104843.2941			

R-Square	Coeff Var	Root MSE	BloodP Mean
0.347918	-906.7286	20.80142	-2.294118

The descriptive statistics indicate that the average blood pressure change for all observations is -2.294118. The R^2 for this model is 0.347918.

The *p*-value given is <0.0001. Presuming an alpha equal to 0.05, you reject the null hypothesis and conclude that at least one treatment mean is different from one other treatment mean. Which factor(s) explain this difference?

The next part of the output shows tests of the main effects and the interaction.

Source	DF	Type I SS	Mean Square	F Value	Pr > F
DrugDose	3	54.03137	18.01046	0.04	0.9886
Disease	2	19276.48690	9638.24345	22.27	<.0001
DrugDose*Disease	6	17146.31698	2857.71950	6.60	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
DrugDose	3	335.73526	111.91175	0.26	0.8551
Disease	2	18742.62386	9371.31193	21.66	<.0001
DrugDose*Disease	6	17146.31698	2857.71950	6.60	<.0001

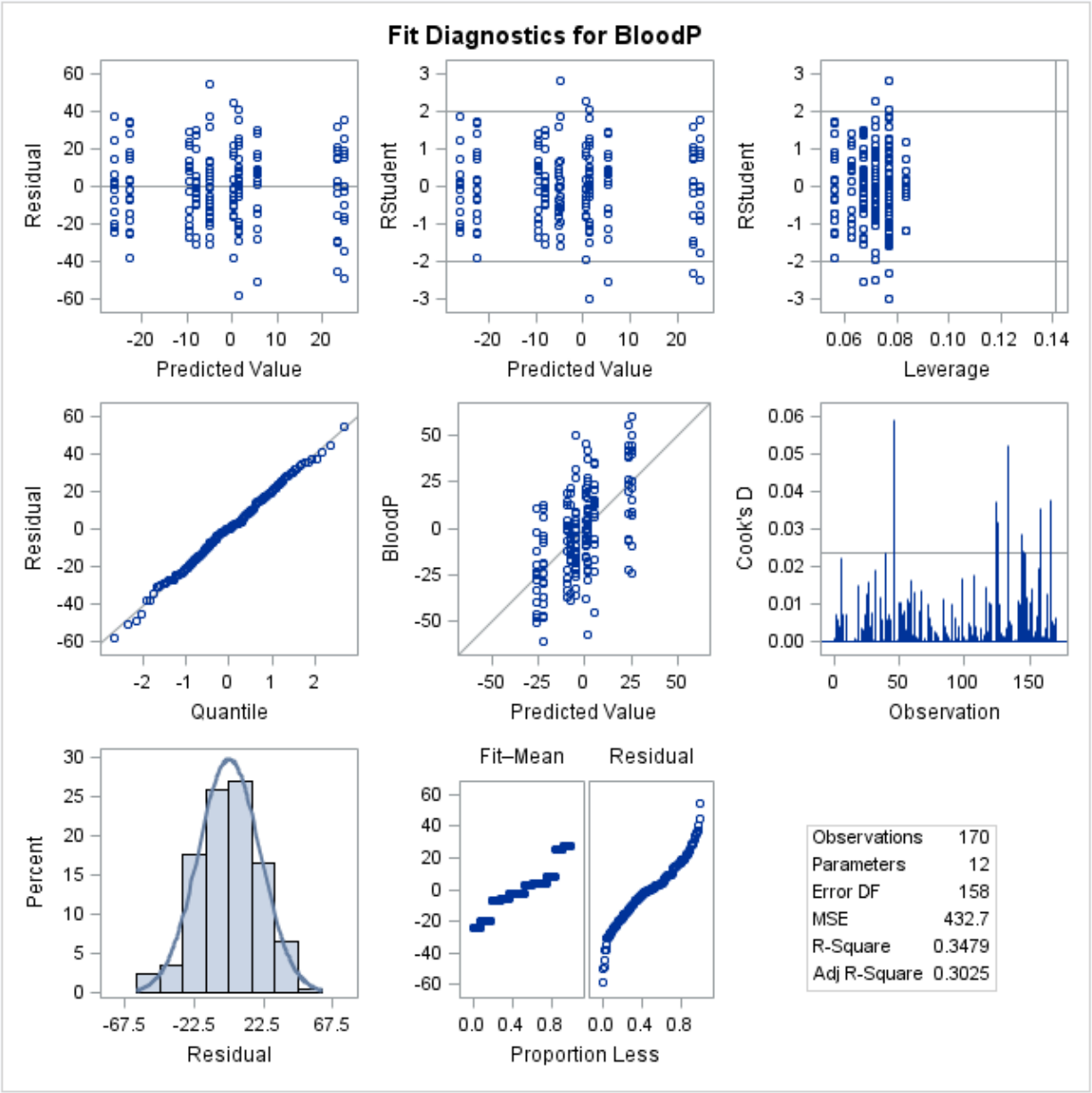
The Type I SS are model-order dependent; each effect is adjusted only for the preceding effects in the model. They are known as *sequential sums of squares*. They are useful in cases where the marginal (additional) effect for adding terms in a specific order is important. An example is where X, X*X, and X*X*X are in the model statement. Each term is only tested controlling for a lower order term. The TYPE I SS values are additive. They sum to the Model Sum of Squares for the overall model.

The Type III SS are commonly called *partial sums of squares*. The Type III sum of squares for a particular variable is the increase in the model sum of squares due to adding the variable to a model that already contains all the other variables in the model. Type III sums of squares, therefore, do not depend on the order in which the explanatory variables are specified in the model. The Type III SS values are not generally additive (except in a completely balanced design). The values do not necessarily sum to the Model SS.

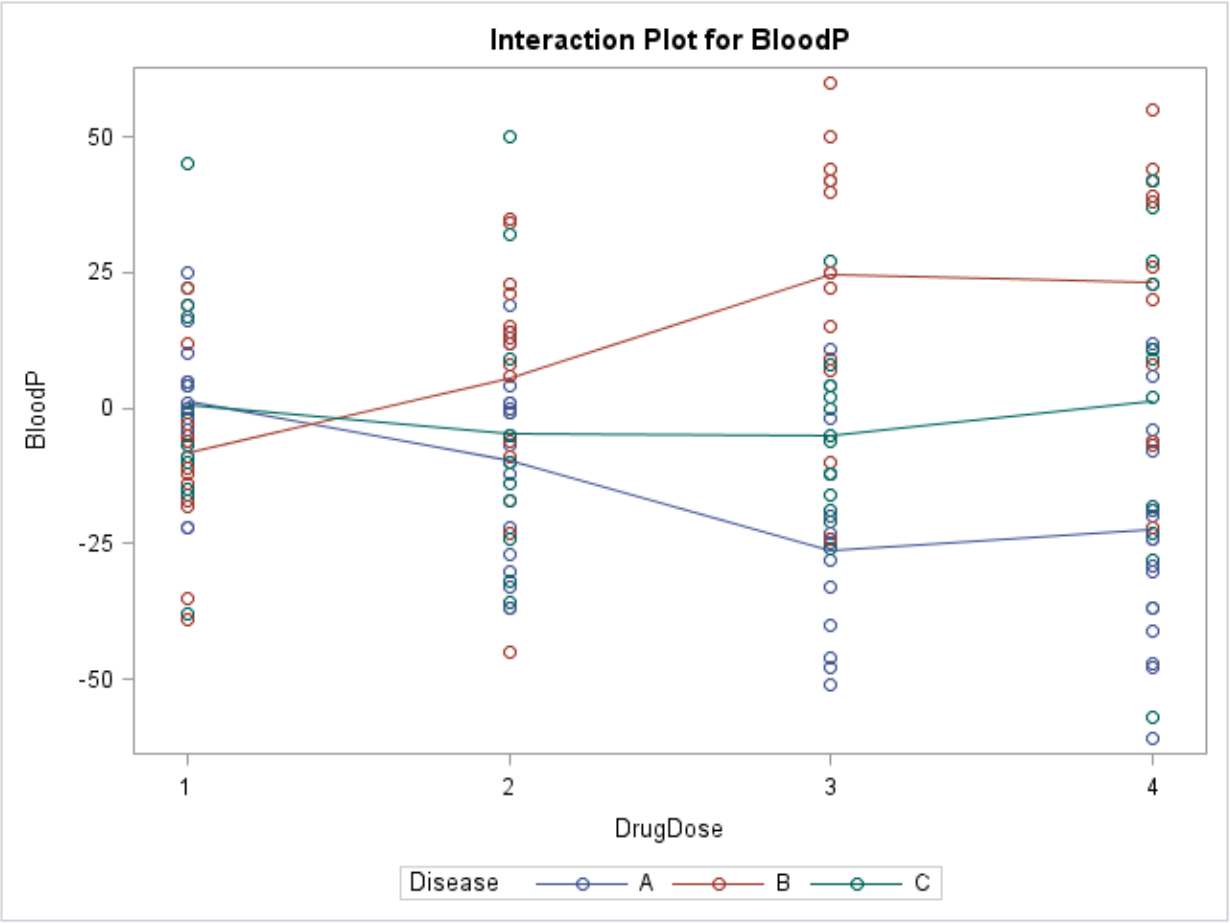
Look at tests using the Type III SS.

You should consider the test for the interaction first, because if there is an interaction, then by definition that means that the effect of each main effect is different at each level of the other main effect. The *p*-value for **`DrugDose*Disease`** is <0.0001. Presuming an alpha of 0.05, you reject the null hypothesis. You have sufficient evidence to conclude that there is an interaction between the two factors. As shown in the graph, the effect of the level of drug changes for different disease types.

You get an interaction plot, as well as a diagnostics plot.



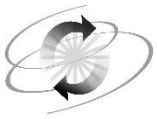
The Quantile-Quantile plot of the residuals indicates no great departure from normality. The Residual by Predicted Value plot shows no pattern in the residuals and the variability of the Residual seems fairly similar across all values of the Predicted Value.



This plot shows the interaction. Drug A shows a pattern of lower blood pressure at higher doses. Drug B shows the opposite pattern. Drug C displays a fairly constant effect across all doses.

Because of the interaction, you do not know the significance of the **`DrugDose`** effect at any particular level of **`Disease`**. (If we were to go further, the LSMEANS tests could be used to test the effect of **`DrugDose`** at each level of **`Disease`**.)

Given all of this information, it seems that one would want to aggressively treat blood pressure in people with disease A with high doses of the drug. For those with disease B (perhaps caused by a traumatic event), treating with the drug at all would be a mistake. For those with Disease C, there seems to be no effect on blood pressure.



Block Design Exercises

1. Analyzing Data in a Randomized Block Design

a. Test the hypothesis that the **Sales** means are equal. Include all of the variables in your MODEL statement. What can you conclude from your analysis? Was adding the blocking factor **Area** into the design and analysis detrimental to the test of **Ad**?

Perform an ANOVA on the **Ads1** data set to look for differences in sales across advertisement mode, controlling for **Area**, which is the blocking variable in the design. Check for normality of the residuals.

1) Open the data table **Ads1**.

2) Select **Tasks** ⇨ **ANOVA** ⇨ **Linear Models…**.

3) Under Data, assign **Sales** to the dependent variable task role and **Ad** and **Area** to the classification variable role.

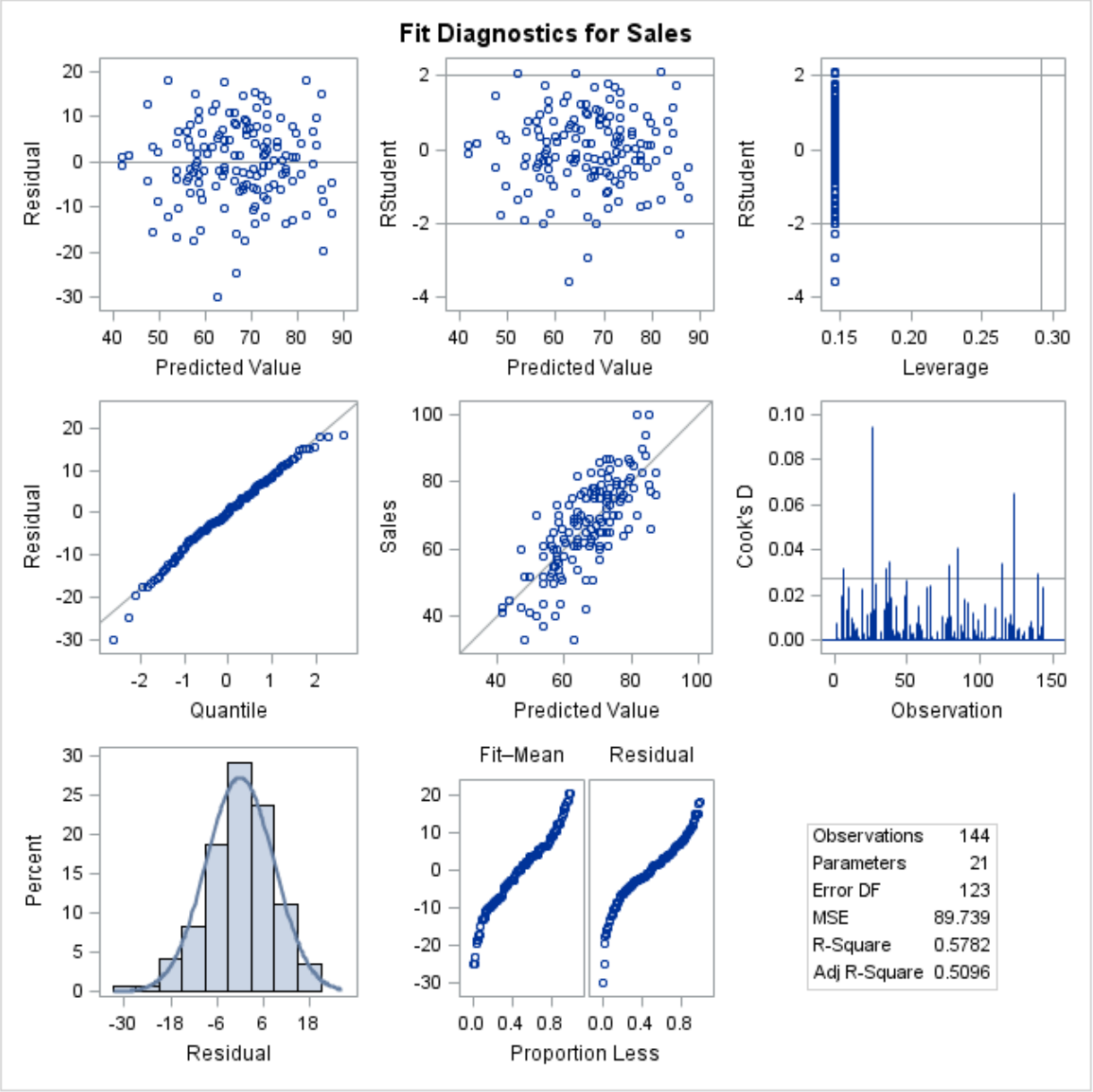
4) Under Model, click **Ad** and **Area** and then click Main.

Note that there is NO interaction with the Block design.

5) Under Model Options, uncheck **Type I** and **Show parameter estimates**.

6) Click Run.

Partial Output



The QQ Plot of Residuals indicates that the normality assumption for ANOVA is met.

Class Level Information		
Class	Levels	Values
Ad	4	display paper people radio
Area	18	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Number of Observations Read	144
Number of Observations Used	144

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	20	15131.38889	756.56944	8.43	<.0001
Error	123	11037.91667	89.73916		
Corrected Total	143	26169.30556			

R-Square	Coeff Var	Root MSE	Sales Mean
0.578211	14.17712	9.473076	66.81944

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Ad	3	5866.083333	1955.361111	21.79	<.0001
Area	17	9265.305556	545.017974	6.07	<.0001

*The p-value for **Ad** (<.0001) indicates that there was a difference in sales among the advertising campaign types, when controlling for **Area**.*

*The large (statistically significant) F-Value for **Area** gives evidence that area of the country was a useful factor to block on. It was definitely not detrimental.*

Post Hoc Pairwise Comparison

2. (optional exercise) Post Hoc Pairwise Comparisons – where is the difference?

Conduct pairwise comparisons with an experimentwise (use the Tukey adjustment) error rate of α=0.05.

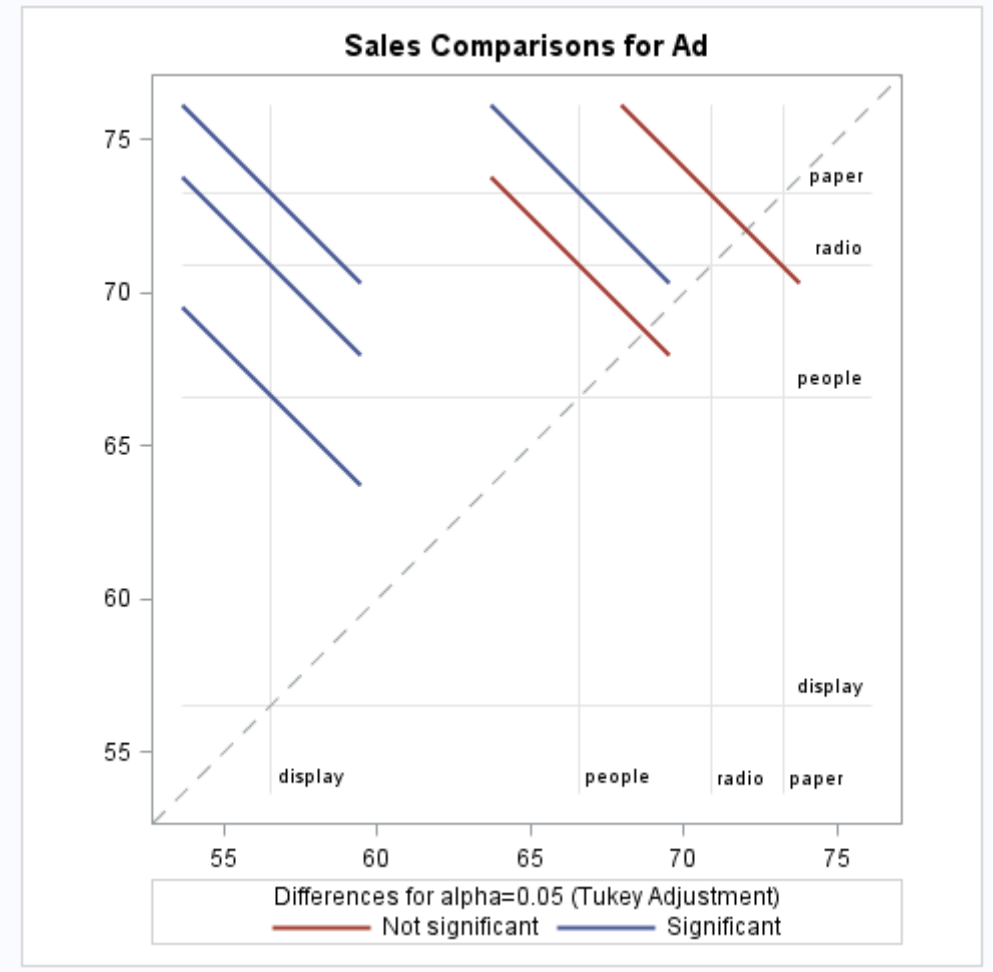
- Re-open the previous Linear Models task by right-clicking the icon for it and selecting **Modify…** from the drop-down menu.
- Select **Least Squares** under **Post Hoc Tests** and then click Add under the white space for `Effects to estimate`.
- A 0 line will appear and the `Options for means tests` area will be populated with several options. Click **False** next to **Ad** from the `Options for means tests` and then click that appears next to it. Change this value to **True** by scrolling to that value and clicking on it.
- In a similar fashion, under comparisons, change the `Show p-values for differences` option from None to **All pairwise differences**. The `Adjustment method for comparison` can be selected as **Tukey** or left at **Default** because Tukey is the default method used when **All pairwise differences** is chosen above.
- Click Add under the white space for `Effects to estimate` and effect 1 will appear. Choose **Ad** again as the `Class effects to use` variable. This time the choice for `Show p-values for differences` should be **Control using first level**. The default `Adjustment method` will be **Dunnett** and the default control group will be **display**.
- Click Run and do not replace the previous results.

a. Which types of advertising are significantly different?

Ad	Sales LSMEAN	LSMEAN Number
display	56.5555556	1
paper	73.2222222	2
people	66.6111111	3
radio	70.8888889	4

Least Squares Means for effect Ad
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: Sales

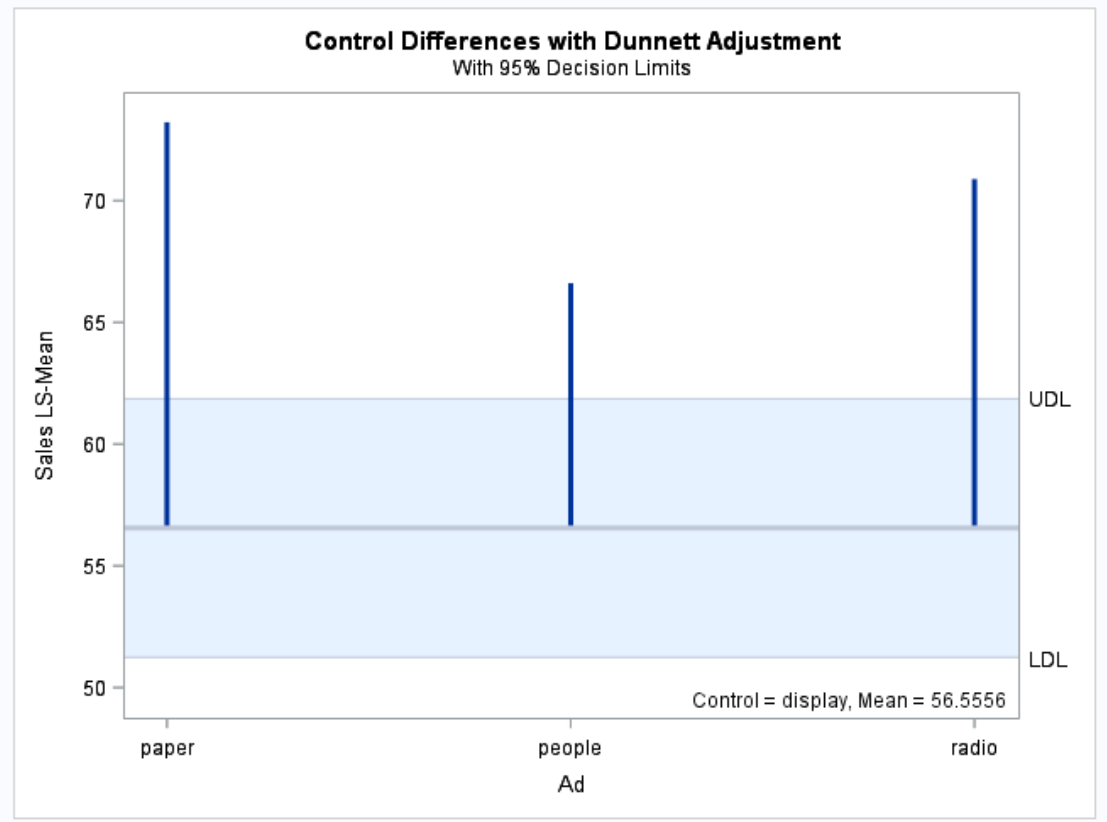
i/j	1	2	3	4
1		<.0001	<.0001	<.0001
2	<.0001		0.0190	0.7233
3	<.0001	0.0190		0.2268
4	<.0001	0.7233	0.2268	



The Tukey comparisons show significant differences between **`display`** and all other types of advertising and between paper and people (p=0.0190).

b. Use **`display`** as the control group and do a Dunnett comparison of all other advertising methods to see if those methods improved sales over just display ads in stores.

Ad	Sales LSMEAN	H0:LSMean=Control Pr > |t|
display	56.5555556	
paper	73.2222222	<.0001
people	66.6111111	<.0001
radio	70.8888889	<.0001



All other advertising campaigns resulted in significantly better average sales (statistically significant) than **`display`**.

Another Example - Two-Way ANOVA (if you want more practice)

Test the hypothesis that the means are equal, making sure to include an interaction term. What conclusions can you reach?

- Open the **Concrete** data set.
- Use the summary statistics task, choosing **Strength** for the analysis variables role and **Brand** and **Additive** for the classification variables role.
- Run the task.

Analysis Variable : Strength							
Brand	Additive	N Obs	Mean	Std Dev	Minimum	Maximum	N
Consolidated	reinforced	5	25.8000000	2.3727621	22.7000000	29.3000000	5
	standard	5	22.6000000	1.5313393	20.4000000	24.2000000	5
EZ Mix	reinforced	5	27.2600000	1.9603571	25.5000000	30.2000000	5
	standard	5	24.4000000	3.7729299	19.8000000	28.0000000	5
Graystone	reinforced	5	30.6600000	1.3390295	29.5000000	32.6000000	5
	standard	5	25.2800000	3.1451550	21.2000000	29.8000000	5

The means do seem to vary from group to group. However, it is not obvious if there is a consistent difference between additives across different levels of **Brand**. The differences between **reinforced** and **standard** range from about 3 for `Consolidated` and `EZ Mix` and 5 for `Graystone`.

To test the hypothesis that effect of **Additive** differs across brands, perform a 2-way ANOVA with an interaction.

- Select **Tasks** ⇨ **ANOVA** ⇨ **Linear Models…**.
- Under Data, assign **Strength** to the dependent variable task role and **Brand** and **Additive** to the classification variables role.
- Select **Model** at the left and highlight both **Brand** and **Additive** in the `Class and quantitative variables` pane (by clicking on one and then holding down the CTRL key on the keyboard and then clicking on the other variable).
- Click 

and see the terms, **Brand**, **Additive** and **Brand*Additive** displayed in the Effects pane.
- Under Model Options, deselect **Show parameter estimates** and **Type I**.
- Click 

.
- You will now see a window showing the SAS code created by the t Test task. This window is where you can directly edit the Code generated by SAS Task.
- Select the Show custom code insertion points and…
- Scroll down in the code to just under the words PLOTS(ONLY)=ALL.
- Click on the <insert custom code here> in the area just after PLOTS(ONLY)=ALL

- Type INTPLOT. This is the option to produce an interaction plot.
- Uncheck the Show custom code insertion points and close the window by clicking ☒ in the upper corner.
- Click Run.

Class Level Information		
Class	Levels	Values
Brand	3	Consolidated EZ Mix Graystone
Additive	2	reinforced standard

Number of Observations Read	30
Number of Observations Used	30

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	189.9080000	37.9816000	6.04	0.0009
Error	24	150.9520000	6.2896667		
Corrected Total	29	340.8600000			

R-Square	Coeff Var	Root MSE	Strength Mean
0.557144	9.645849	2.507921	26.00000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Brand	2	71.4980000	35.7490000	5.68	0.0095
Additive	1	109.0613333	109.0613333	17.34	0.0003
Brand*Additive	2	9.3486667	4.6743333	0.74	0.4862

There is no significant interaction between ***Additive*** *and* ***Brand****, even though the plot shows slightly different slopes among the two additives. At this point, you may choose to remove the interaction term from the model and, if still significant, conclude that there is a difference in additive types.*