
**BE640 - Intermediate Biostatistics
Computer Illustration**

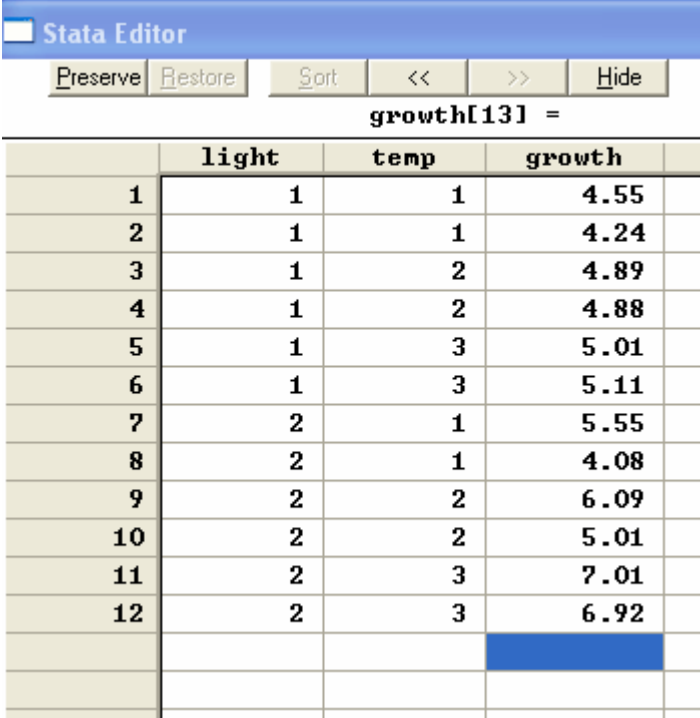
**Unit 7 – Analysis of Variance
Software: Stata version 7**

Two Way Analysis of Variance

Setting:

The data are n=12 observations from a two analysis of variance that investigates the separate and joint effects of light and water temperature on the growth of fish.

1. Using the DATA EDITOR, read in the data.



The screenshot shows the Stata Editor interface. At the top, there is a blue title bar with the text "Stata Editor". Below the title bar, there are several buttons: "Preserve", "Restore", "Sort", "<<", ">>", and "Hide". Below these buttons, the text "growth[13] =" is displayed. The main part of the screenshot is a data table with the following structure:

	light	temp	growth
1	1	1	4.55
2	1	1	4.24
3	1	2	4.89
4	1	2	4.88
5	1	3	5.01
6	1	3	5.11
7	2	1	5.55
8	2	1	4.08
9	2	2	6.09
10	2	2	5.01
11	2	3	7.01
12	2	3	6.92

2. Look at the data. Obtain side by side box and whisker plots of GROWTH

Be sure to sort your data first (as in this example)

I decided to look at growth two ways: (1) by light; and (2) by temperature

Also, notice that I have formatted the values of light and temperature for readability.

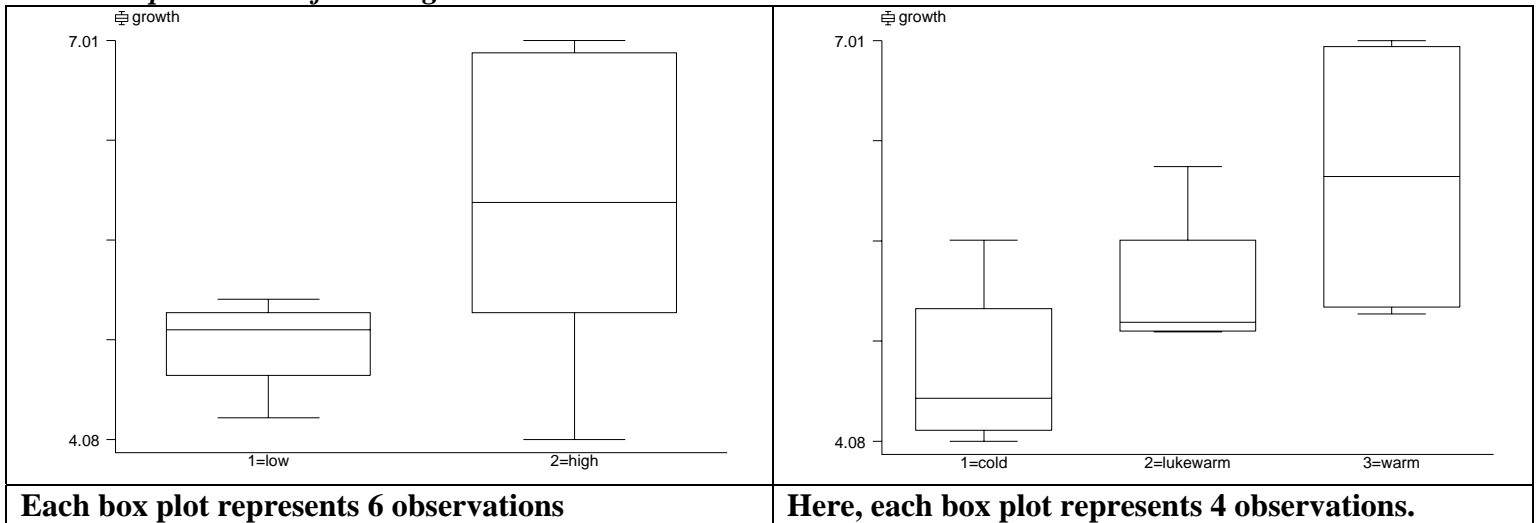
PROC BOXPLOT produces images that can be exported for pasting into a document later.

```
.* Dictionary of discrete variable value codes
. label define lightf 1 "1=low" 2 "2=high"
. label define tempf 1 "1=cold" 2 "2=lukewarm" 3 "3=warm"

. *Box and Whisker of Growth by Light
. sort light
. label values light lightf
. graph growth, box by(light)

. *Box and Whisker of Growth by Temperature
. sort temp
. label values temp tempf
. graph growth, box by(temp)
```

This will produce the following.



How to save a graph in STATA. Execute the STATA instruction that produces the graph. From there,

FILE → SAVE GRAPH →

At the SAVE AS TYPE dialogue box drop down menu, choose "windows metafile" (*.wmf)

3. Get a graphical feel for interaction

I chose to look at the effect of temperature – separately at each level of light.

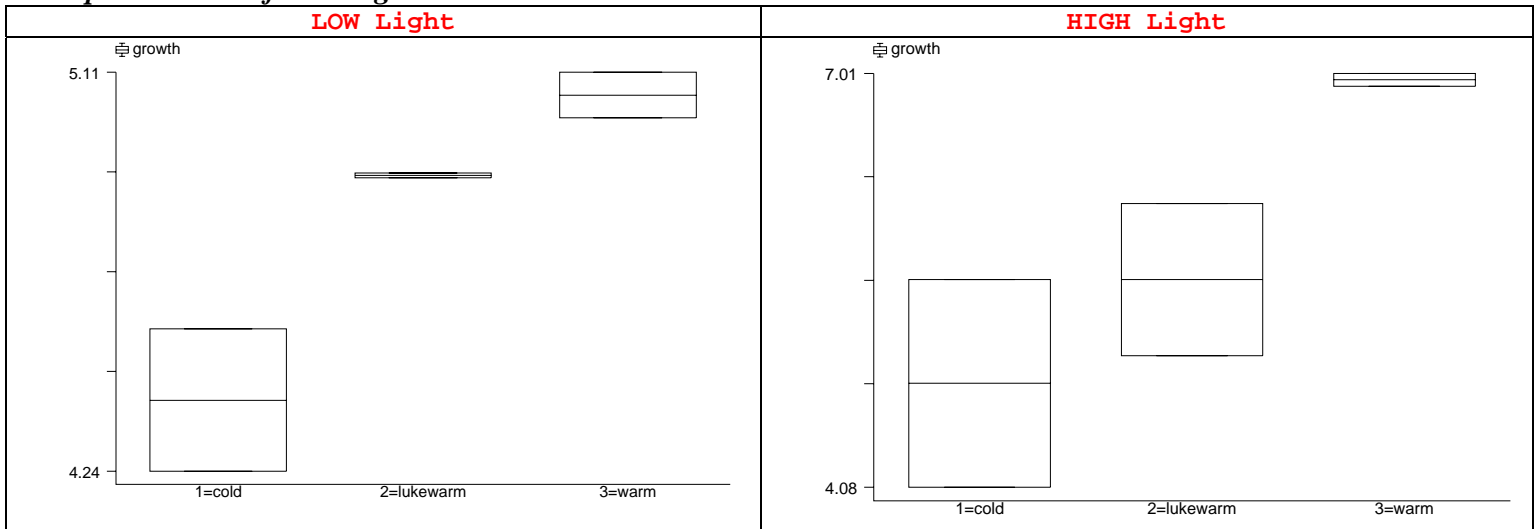
Equivalently, you could look at the effect of light – separately at each temperature condition.

```

. *Box and Whisker of Growth by Temperature - Separately for LOW and HIGH Light
. sort light temp
. by light: graph growth, box by(temp)

```

This produced the following.



The sample sizes are small so it is a little difficult to conclude much from these graphs; each box plot represents only n=2 observations. However, there does not appear to be much evidence that the effect of temperature on growth changes depending on the level of light.

Note - Be careful. The vertical axes are not quite the same.

3. Do a 2 way analysis of variance.

In the following code, I ask for the fit of an analysis of variance that has main effects for each of light and temperature plus an interaction of light and temperature.

```
. *Two Way Analysis of Variance
. anova growth light temp light*temp
```

This will produce the following. Note – I have omitted some output.

	Number of obs =	12	R-squared =	0.8271	
	Root MSE =	.535537	Adj R-squared =	0.6830	
Source	Partial SS	df	MS	F	Prob > F
Model	8.23196772	5	1.64639354	5.74	0.0276
light	2.98003383	1	2.98003383	10.39	0.0181
temp	3.9843175	2	1.99215875	6.95	0.0274
light*temp	1.26761639	2	.633808197	2.21	0.1909
Residual	1.72080044	6	.286800074		
Total	9.95276816	11	.904797106		

From the above output, we see that the analysis of variance table is the following

Source	Df	SSQ	MSQ	F	p-value
Due LIGHT	1	2.98	2.98	10.39	.018
Due TEMP	2	3.984	1.992	6.95	.027
Due Interaction	2	1.268	0.634	2.21	.191
Error	6	1.721	0.287		
Total (Corrected)	11	9.953	-		

The anova table confirms what we saw in the picture; these (albeit limited) data do not suggest an interaction of light and temperature on growth. We do see main effects of each of light and temperature; both associations are positive.