BE640 - Intermediate Biostatistics
Computer Illustration

Unit 7 – Analysis of Variance
Software: SAS

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. Read in the data using cards.

```
data temp;
  input id growth light temp;
  cards;
1 4.55 1 1
2 4.24 1 1
3 4.89 1 2
4 4.88 1 2
5 5.01 1 3
6 5.11 1 3
7 5.55 2 1
8 4.08 2 1
9 6.09 2 2
10 5.01 2 2
11 7.01 2 3
12 6.92 2 3
;
run;

*2.  Dictionary of variable values for readability;
proc format;
  value lightf 1='1=low'
               2='2=high';
  value tempf 1='1=cold'
              2='2=lukewarm'
              3='3=warm';
 run;
 quit;
```

You should see something like:

```
NOTE: The data set WORK.TEMP has 12 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds
```

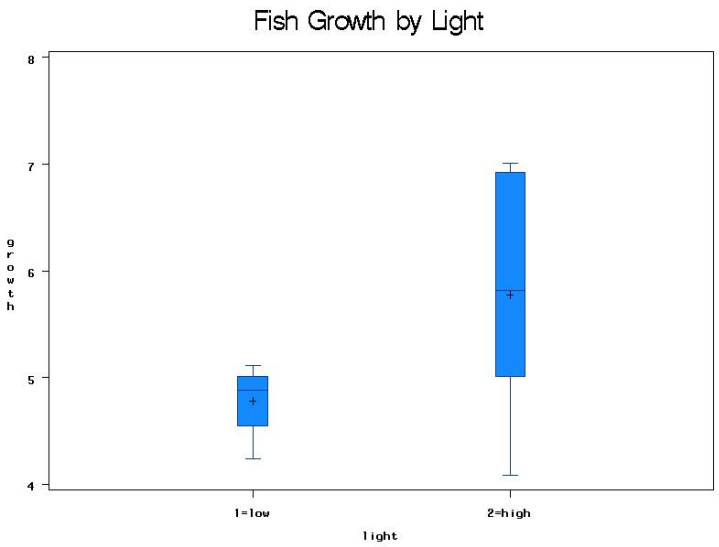
2. Look at the data. Obtain side by side box and whisker plots of GROWTH

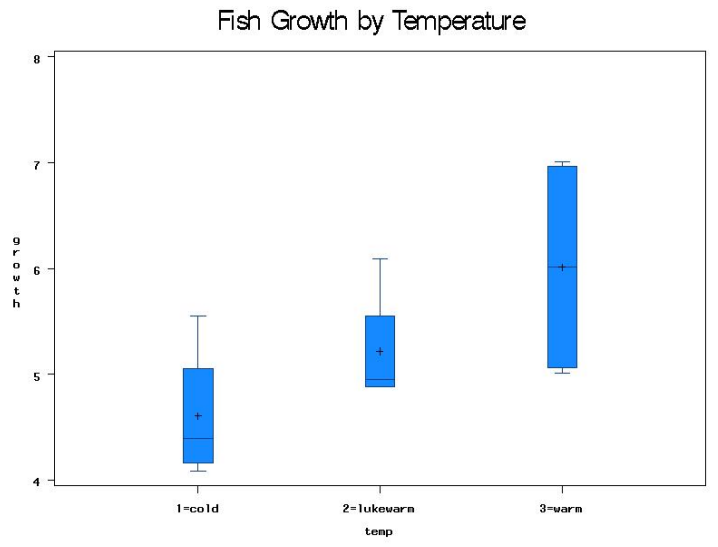
Here, I've used PROC BOXPLOT. Be sure to sort your data first (as in this example)
I decided to look at growth two ways: (1) by light; and (2) by temperature
Also, notice that I have formatted the values of light and temperature for readability.
PROC BOXPLOT produces images that can be exported for pasting into a document later.

```
*2.  Dictionary of variable values for readability;
proc format;
  value lightf 1='1=low'
               2='2=high';
  value tempf 1='1=cold'
              2='2=lukewarm'
              3='3=warm';
 run;
 quit;

*3.  Side by Side Box and Whisker Plots;
proc sort data=temp;
  by light;
run;
proc boxplot data=temp;
  plot growth*light
    /boxstyle=skeletal
       cboxes=cx153e7e
       cboxfill=cx1589ff;
   title 'Fish Growth by Light';
   format light lightf.;
run;
proc sort data=temp;
  by temp;
run;
proc boxplot data=temp;
  plot growth*temp
    /boxstyle=skeletal
       cboxes=cx153e7e
       cboxfill=cx1589ff;
   title 'Fish Growth by Temperature';
   format temp tempf.;
run;
```

This will produce the following.





Each box plot represents 6 observations	**Here, each box plot represents 4 observations.**

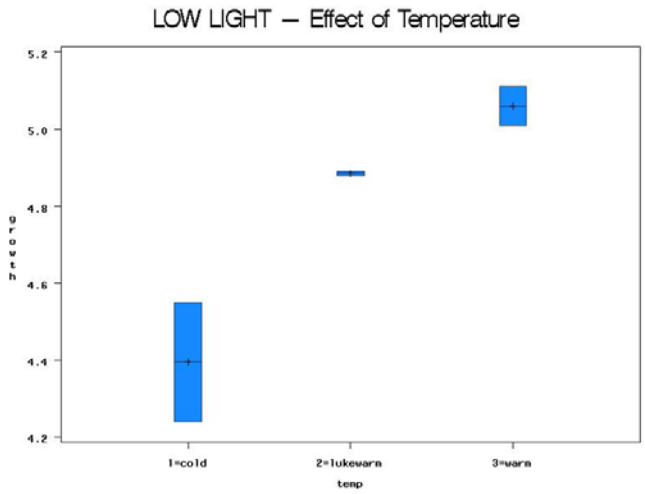
How to save a graph: (1) Activate the GRAPH output window. From there,
(2)FILE → EXPORT AS IMAGE. (3) At the "SAVE AS TYPE" dialog box – choose jpeg

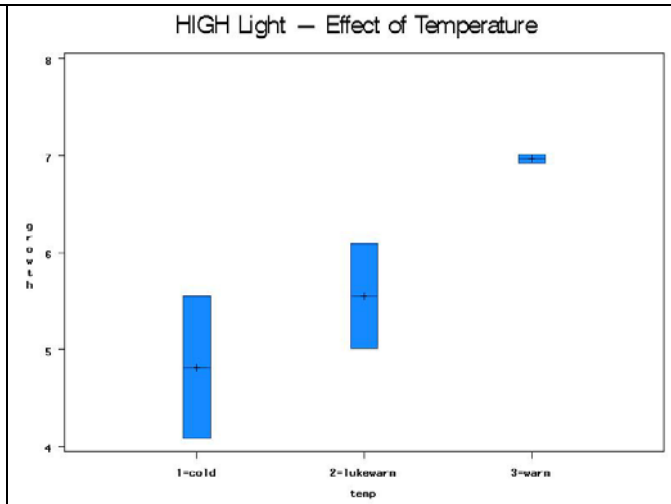
3. Get a graphical feel for interaction

I asked SAS for box and whisker plots of the effects of temperature, separately at each level of light. Equivalently, you might look at box and whisker plots of the effects of light at each temperature.

```
proc sort data=temp;
  by light temp;
run;
proc boxplot data=temp;
  where light=1;
  plot growth*temp
    /boxstyle=skeletal
       cboxes=cx153e7e
       cboxfill=cx1589ff;
   title 'LOW LIGHT - Effect of Temperature';
   format temp tempf.;
run;
proc boxplot data=temp;
  where light=2;
  plot growth*temp
    /boxstyle=skeletal
       cboxes=cx153e7e
       cboxfill=cx1589ff;
   title 'HIGH Light - Effect of Temperature';
   format temp tempf.;
run;
```

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. The MEANS option requests the output of the sample means, separately for each group.

```
proc anova data=temp;
  class light temp;
  model growth = light temp light*temp;
  means light temp;
  title 'Analysis of Fish Growth, by Light and Temperature';
  format light lightf. temp tempf.;
run;
quit;
```

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

```
Analysis of Fish Growth, by Light and Temperature
The ANOVA Procedure
Dependent Variable: growth

                                    Sum of
Source                      DF         Squares     Mean Square    F Value    Pr > F

Model                        5      8.23196667      1.64639333       5.74    0.0276

Error                        6      1.72080000      0.28680000

Corrected Total             11      9.95276667


R-Square     Coeff Var      Root MSE    growth Mean
0.827103      10.14595      0.535537       5.278333


Source                      DF        Anova SS     Mean Square    F Value    Pr > F
light                        1      2.98003333      2.98003333      10.39    0.0181
temp                         2      3.98431667      1.99215833       6.95    0.0274
light*temp                   2      1.26761667      0.63380833       2.21    0.1909

The ANOVA Procedure
Level of           ------------growth-----------
light        N             Mean          Std Dev

1=low        6       4.78000000       0.32508460
2=high       6       5.77666667       1.13528264


Level of             ------------growth-----------
temp           N             Mean          Std Dev

1=cold         4       4.60500000       0.65952003
2=lukewarm     4       5.21750000       0.58465802
3=warm         4       6.01250000       1.10122281
```

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..