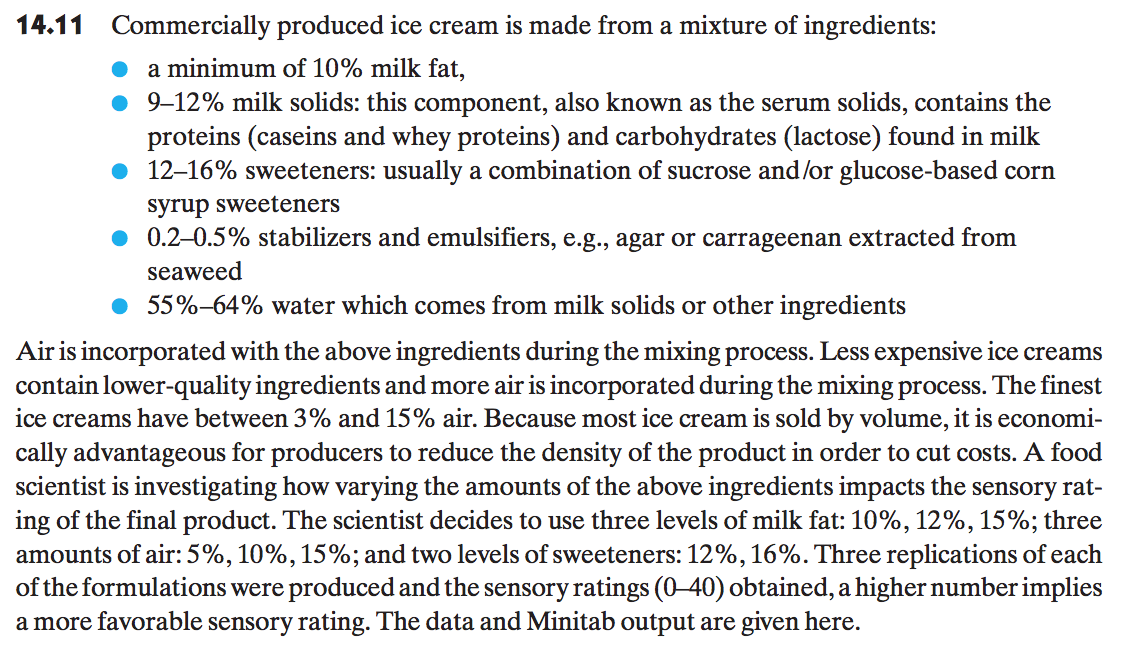
**CSUEB – STAT 6305 – Winter 2017 - Prof Yan Zhou**

**Homework 4 - Henry Lankin, Gui Larangeira**

February 09, 2017

**HW 4: 14.11, 14.12, 14.13, 14.20, 14.21**

14.11



1. **Identify the design and treatment structure for this study.**

The experiment is setup as a CRD factorial treatment design with the following factors: sweetener percentage, milk fat percentage and air percentage. The design uses three replications for each combination.

1. **Write a model for this study, identifying all terms in the model.**

Factorial design model:

Using a factorial treatment structure with:

– overall sensory rating mean

– the sensory rating from each sweetener-milk fat-air combination: 54 observations from 3 replications of the 18 treatment combinations

– the effect due to the percentage of sweetener: 2 treatment levels

– the effect due to the percentage of milk fat: 3 treatment levels

– the effect due to the percentage of air: 3 treatment levels

– the effect due to the two-way interaction between the levels of sweetener and milk fat factors: 6 interactions

– the effect due to the two-way interactions between the levels of sweetener and air factors: 6 interactions

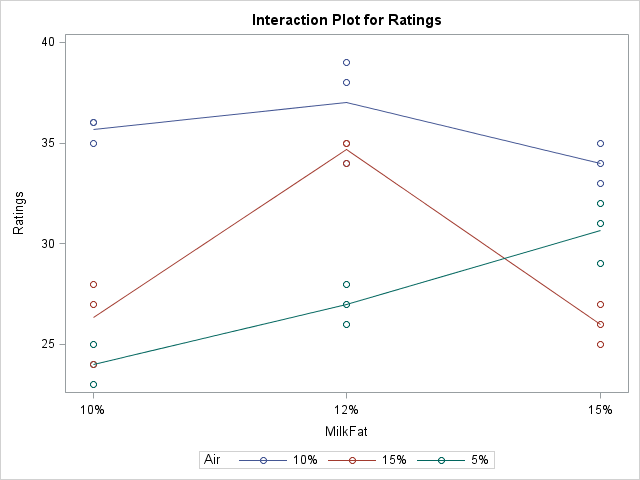
– the effect due to the two-way interactions between the levels of milk fat and air factors: 9 interactions

– the effect due to the three-way interactions between the levels of sweetener, milk fat and air factors: 18 interactions

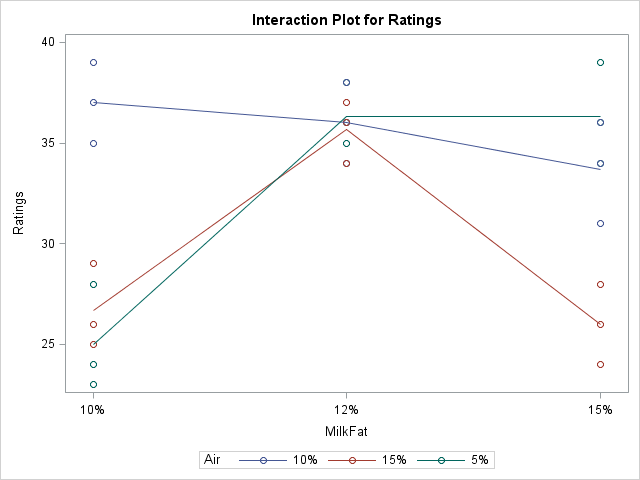
– random error associated with each sweetener-milk fat-air combination: 54 residual errors

1. **For each of the two levels of sweetener, draw profile plots of the effects of the percentage of air and milk fat on the sensory rating of ice cream.**

Profile plot for the interaction between milk fat and air percentage for **sweetener at**



Profile plot for the interaction between milk fat and air percentage for **sweetener at .**



1. **From the profile plots, does there appear to be a three-way interaction between the effects percentage of sweetener, air, and milk fat in ice cream on the mean sensory rating?**

Examining each profile plot – one for each level of sweetener – separately, it is clear that milkfat and air content interact as the lines cross each other and are not even close t parallel. Moreover, we observe comparing the two plots that the lines (especially the blue line, air = 5) has a different profile between different levels of sweetener. Thus, we can infer that air and sweetener factors probably interact in this experiment. However, the profiles of milkfat do not change as much between the two profile plots, which leads us to believe this interaction between milkfat and sweetener might not be as significant.

14.12



1. **Using the output given above, perform appropriate *F*-tests and draw conclusions from these tests concerning the effect of the percentage of sweetener, air, and milk fat on the sensory rating of ice cream. Use .05.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Type I SS** | **Mean Square** | **F Value** | **Pr > F** |
| **Sweetener** | 1 | 50.0740741 | 50.0740741 | 15.28 | 0.0004 |
| **MilkFat** | 2 | 261.3333333 | 130.6666667 | 39.86 | <.0001 |
| **Air** | 2 | 436.0000000 | 218.0000000 | 66.51 | <.0001 |
| **Sweetener\*MilkFat** | 2 | 11.2592593 | 5.6296296 | 1.72 | 0.1939 |
| **Sweetener\*Air** | 2 | 78.8148148 | 39.4074074 | 12.02 | 0.0001 |
| **MilkFat\*Air** | 4 | 355.6666667 | 88.9166667 | 27.13 | <.0001 |
| **Sweetener\*MilkFat\*Air** | 4 | 46.1851852 | 11.5462963 | 3.52 | 0.0159 |

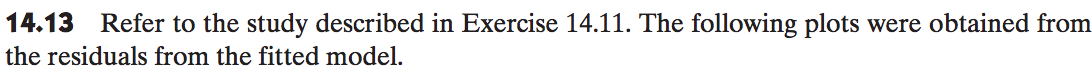
From the AOV table above, we have a -value of for the three-way interaction, implying that it has a significant effect on the mean sensory rating. We also have a

-value of 0.001 for both the Milk Fat-Air interaction and the Sweetener-Air interaction. Therefore, we would conclude that the three-way and two-way interactions must be included in the model to explain the variance of the data.

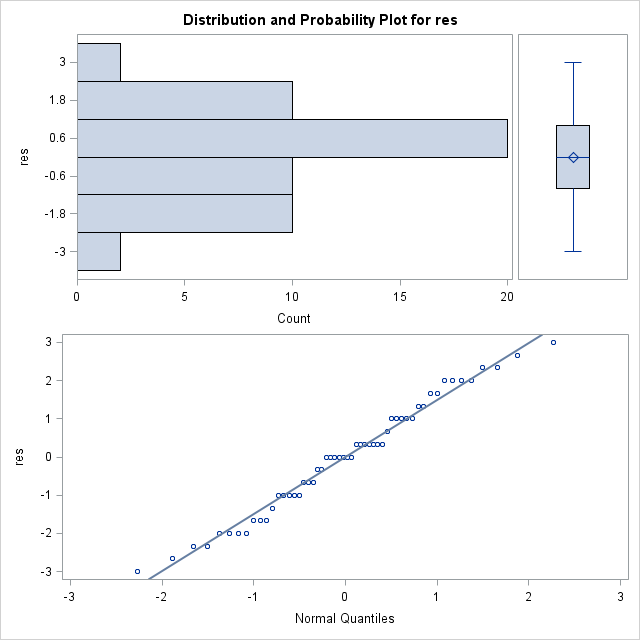
1. **Are the conclusions from the *F*-test consistent with your observations from the profile plots?**

Yes, the conclusion from the -test in part (a) agree with the conclusion made from examining the profile plots in question 14.11 that there is a three-way interaction between the three factors. Moreover, we had observed that while there was probably a two-way factor interaction between sweetener-air and milkfat-air, the one between milkfat-sweetener was not as strong, which we could confirm with non-significant p-value.

14.13



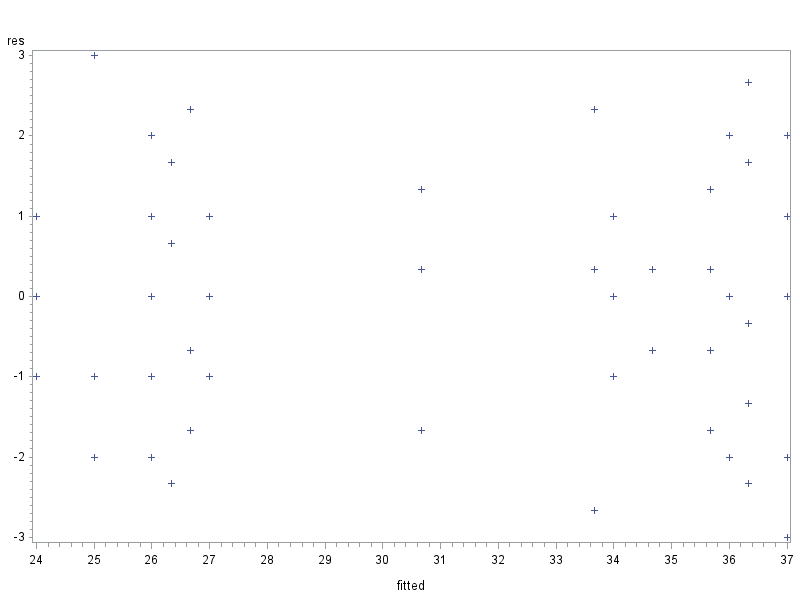
1. **Is there significant evidence that the residuals have a non-normal distribution?**



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tests for Normality** | | | | |
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | **0.97742** | **Pr < W** | **0.3978** |

We see from the Normal QQ-plot of the residuals that there is very little deviation from the line. Further, we have a -value for the normality test of . **Thus, we do not reject non-normality and conclude that the residuals follow a normal distribution.**

1. **Is there significant evidence that the residuals do not have constant variances?**



We see from the residuals vs. fitted values plot above that, although the fitted values are more concentrated on the extremes, there is no particularly strong vertical pattern in the residuals. To examine further, we run a Levene test on the data set where we consider each combination of Sweetener-Milk Fat-Air to be different levels of one factor.

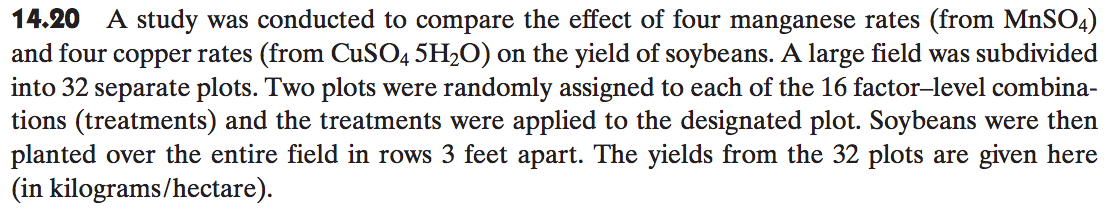
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Levene's Test for Homogeneity of ratings Variance ANOVA of Squared Deviations from Group Means** | | | | | |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **treatment** | 17 | 121.3 | 7.1329 | 1.35 | 0.2165 |
| **Error** | 36 | 189.6 | 5.2654 |  |  |

We have from the table above a -value of 0.2165, implying that we fail to reject the null hypothesis of equal variances.

1. **How could we assess whether the residuals are independently distributed?**

We could review the conditions under which the experiment was performed, to make sure the treatments were truly randomly assigned. Besides what we already did, which is examining the fitted vs. residuals for patterns, we could also plot the response data vs the residuals and make sure that has no pattern.

14.20



1. **Identify the design for this experiment.**

The experiment follows a CRD factorial treatment design with the following factors: manganese () and copper rates (). The design uses two replications for each combination.

1. **Write an appropriate statistical model for this experiment.**

**Factorial design model:**

– overall soybean yield mean

– the yield of soybeans from each - combination: 32 observations from 3 replications of the 16 treatment combinations

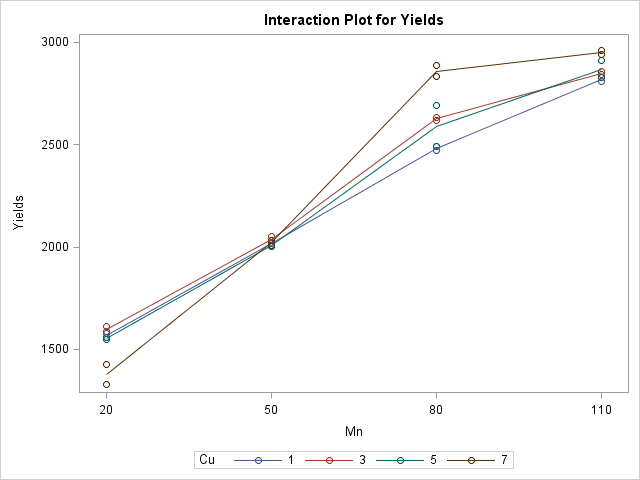
– the effect due to the manganese rate: 4 treatment levels

– the effect due to the copper rate: 4 treatment levels

– the effect due to the two-way interaction between the levels of manganese and copper factors: 16 interactions

– random error associated with each - combination: 32 residual errors

1. **Construct a profile plot and describe what this plot says about the effect of Mn and Cu on soybean yield.**



From the profile plot above, we see that the profiles are not parallel to one another and, in fact, intersect. Thus, we conclude there is likely a two-way interaction between the and , and the interaction is non-orderly.

14.21



1. **Using the computer printout given here, test for an interaction between the effect of and on soybean yield. Use .05.**

| **Source** | **DF** | **Type I SS** | **Mean Square** | **F Value** | **Pr > F** |
| --- | --- | --- | --- | --- | --- |
| **Mn** | 3 | 8935108.09 | 2978369.365 | 1486.70 | <.0001 |
| **Cu** | 3 | 28199.344 | 9399.781 | 4.69 | 0.0155 |
| **Mn\*Cu** | 9 | 204399.281 | 22711.031 | 11.34 | <.0001 |

From the AOV table above, we have a -value of less than 0.0001 for the two-way interaction \* Thus, we conclude that the interaction has a significant effect on soybean yield.

1. **What level of appears to produce the highest yield?**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameter** | **Estimate** |  | **Standard Error** | **t Value** | **Pr > |t|** |
| **Intercept** | 2859.500000 | B | 31.64920023 | 90.35 | <.0001 |
| **Mn 110** | 91.000000 | B | 44.75872820 | 2.03 | 0.0590 |
| **Mn 20** | -1482.000000 | B | 44.75872820 | -33.11 | <.0001 |
| **Mn 50** | -839.000000 | B | 44.75872820 | -18.74 | <.0001 |
| **Mn 80** | 0.000000 | B | . | . | . |

By examining the profile plot, the maximum yield is achieved at 110. Further, this is confirmed by the largest parameter estimate above of at this manganese level.

1. **What level of appears to produce the highest yield?**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameter** | **Estimate** |  | **Standard Error** | **t Value** | **Pr > |t|** |
| **Intercept** | 2859.500000 | B | 31.64920023 | 90.35 | <.0001 |
| **Cu 1** | -379.500000 | B | 44.75872820 | -8.48 | <.0001 |
| **Cu 3** | -233.500000 | B | 44.75872820 | -5.22 | <.0001 |
| **Cu 5** | -269.500000 | B | 44.75872820 | -6.02 | <.0001 |
| **Cu 7** | 0.000000 | B | . | . | . |

By examining the profile plot, the maximum yield is achieved at . Further, this is confirmed by the parameter estimate above of at this copper level.

1. **What combination of - appears to produce the highest yield?**

**So this is interesting.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Parameter** | **Estimate** |  | **Standard Error** | **t Value** | **Pr > |t|** |
| **Intercept** | 2950.500000 | B | 31.64920023 | 93.23 | <.0001 |
| **Mn\*Cu 20 1** | 321.000000 | B | 63.29840045 | 5.07 | 0.0001 |
| **Mn\*Cu 20 3** | 322.500000 | B | 63.29840045 | 5.09 | 0.0001 |
| **Mn\*Cu 20 5** | 257.000000 | B | 63.29840045 | 4.06 | 0.0009 |
| **Mn\*Cu 20 7** | 0.000000 | B | . | . | . |
| **Mn\*Cu 50 1** | 128.000000 | B | 63.29840045 | 2.02 | 0.0602 |
| **Mn\*Cu 50 3** | 115.000000 | B | 63.29840045 | 1.82 | 0.0880 |
| **Mn\*Cu 50 5** | 66.500000 | B | 63.29840045 | 1.05 | 0.3091 |
| **Mn\*Cu 50 7** | 0.000000 | B | . | . | . |
| **Mn\*Cu 80 1** | -249.000000 | B | 63.29840045 | -3.93 | 0.0012 |
| **Mn\*Cu 80 3** | -133.500000 | B | 63.29840045 | -2.11 | 0.0510 |
| **Mn\*Cu 80 5** | -189.000000 | B | 63.29840045 | -2.99 | 0.0087 |
| **Mn\*Cu 80 7** | 0.000000 | B | . | . | . |
| **Mn\*Cu 110 1** | 0.000000 | B | . | . | . |
| **Mn\*Cu 110 3** | 0.000000 | B | . | . | . |
| **Mn\*Cu 110 5** | 0.000000 | B | . | . | . |
| **Mn\*Cu 110 7** | 0.000000 | B | . | . | . |

By examining the profile plot, the maximum yield is achieved at . But, this is actually not confirmed by the maximum parameter estimate above of at.

SAS code:

\* 14.11

\* input data;

**data** icecream;

input Ratings Sweetner$ MilkFat$ Air$;

cards;

23 12% 10% 5%

24 12% 10% 5%

25 12% 10% 5%

36 12% 10% 10%

35 12% 10% 10%

36 12% 10% 10%

28 12% 10% 15%

24 12% 10% 15%

27 12% 10% 15%

27 12% 12% 5%

28 12% 12% 5%

26 12% 12% 5%

34 12% 12% 10%

38 12% 12% 10%

39 12% 12% 10%

35 12% 12% 15%

35 12% 12% 15%

34 12% 12% 15%

31 12% 15% 5%

32 12% 15% 5%

29 12% 15% 5%

33 12% 15% 10%

34 12% 15% 10%

35 12% 15% 10%

26 12% 15% 15%

27 12% 15% 15%

25 12% 15% 15%

24 16% 10% 5%

23 16% 10% 5%

28 16% 10% 5%

37 16% 10% 10%

39 16% 10% 10%

35 16% 10% 10%

26 16% 10% 15%

29 16% 10% 15%

25 16% 10% 15%

38 16% 12% 5%

36 16% 12% 5%

35 16% 12% 5%

34 16% 12% 10%

38 16% 12% 10%

36 16% 12% 10%

36 16% 12% 15%

37 16% 12% 15%

34 16% 12% 15%

34 16% 15% 5%

36 16% 15% 5%

39 16% 15% 5%

34 16% 15% 10%

36 16% 15% 10%

31 16% 15% 10%

28 16% 15% 15%

26 16% 15% 15%

24 16% 15% 15%

;

**run**;

**proc** **print** data=icecream;

**run**;

\* run factorial design anova test;

**proc** **glm** data=icecream;

class Sweetner MilkFat Air;

model ratings = Sweetner MilkFat Air Sweetner\*MilkFat Sweetner\*Air MilkFat\*Air Sweetner\*MilkFat\*Air;

output out=residuals r=res p=fitted;

**run**;

**quit**;

\* plot residuals qq-plot;

**proc** **univariate** normal plot data = residuals;

var res;

**run**;

**quit**;

\* plot residuals vs. fitted values to check for equal variances;

**proc** **gplot** data=residuals;

plot res\*fitted;

symbol i=none v=star;

**run**;

**quit**;

\* time series graph to look at independence of residuals;

**proc** **gplot** data=residuals;

symbol i=line v=star;

plot res\*ratings=**1**;

**run**;

**quit**;

\* re-enter data set so as to have one factor: the three factor combination;

**data** icecream\_oneway;

length treatment $**12**;

input ratings treatment$;

cards;

23 12%10%5%

24 12%10%5%

25 12%10%5%

36 12%10%10%

35 12%10%10%

36 12%10%10%

28 12%10%15%

24 12%10%15%

27 12%10%15%

27 12%12%5%

28 12%12%5%

26 12%12%5%

34 12%12%10%

38 12%12%10%

39 12%12%10%

35 12%12%15%

35 12%12%15%

34 12%12%15%

31 12%15%5%

32 12%15%5%

29 12%15%5%

33 12%15%10%

34 12%15%10%

35 12%15%10%

26 12%15%15%

27 12%15%15%

25 12%15%15%

24 16%10%5%

23 16%10%5%

28 16%10%5%

37 16%10%10%

39 16%10%10%

35 16%10%10%

26 16%10%15%

29 16%10%15%

25 16%10%15%

38 16%12%5%

36 16%12%5%

35 16%12%5%

34 16%12%10%

38 16%12%10%

36 16%12%10%

36 16%12%15%

37 16%12%15%

34 16%12%15%

34 16%15%5%

36 16%15%5%

39 16%15%5%

34 16%15%10%

36 16%15%10%

31 16%15%10%

28 16%15%15%

26 16%15%15%

24 16%15%15%

;

**run**;

**proc** **print** data=icecream\_oneway;

**run**;

\* run one-anova procedure with levene test to test for equal variances;

**proc** **glm** data=icecream\_oneway;

class treatment;

model ratings = treatment;

means treatment / hovtest=levene;

**run**;

**quit**;

\* create new data set for only values associated with sweetner=12%;

**data** icecream\_sweetner12; set icecream;

if Sweetner = '12%';

**run**;

**proc** **print** data=icecream\_sweetner12;

**run**;

\* run two facotr factorial design anova test for sweetner=12% to look at interaction plot;

**proc** **glm** data=icecream\_sweetner12;

class MilkFat Air;

model ratings = MilkFat Air  MilkFat\*Air;

**run**;

**quit**;

\* create new data set for only values associated with sweetner=16%;

**data** icecream\_sweetner16; set icecream;

if Sweetner = '16%';

**run**;

**proc** **print** data=icecream\_sweetner16;

**run**;

\* run two factor factorial design anova test for sweetner=16% to look at interaction plot;

**proc** **glm** data=icecream\_sweetner16;

class MilkFat Air;

model ratings = MilkFat Air  MilkFat\*Air;

**run**;

**quit**;

\* 14.20

\* input data;

**data** soybeans;

input Yields Mn Cu;

cards;

1558 20 1

1578 20 1

1590 20 3

1610 20 3

1558 20 5

1550 20 5

1328 20 7

1427 20 7

2003 50 1

2033 50 1

2020 50 3

2051 50 3

2003 50 5

2010 50 5

2010 50 7

2031 50 7

2490 80 1

2470 80 1

2620 80 3

2632 80 3

2490 80 5

2690 80 5

2887 80 7

2832 80 7

2830 110 1

2810 110 1

2860 110 3

2841 110 3

2830 110 5

2910 110 5

2960 110 7

2941 110 7

;

**run**;

**proc** **print** data=soybeans;

**run**;

\* run factorial treatment ANOVA test;

**proc** **glm** data = soybeans;

class Mn Cu;

model Yields = Mn Cu Mn\*Cu / solution;

**run**;

**quit**;