**MAT 449**

**HW1**

**Inna WIlliams**

1. **Import the *Birthweight* data set into SAS using the INFILE statement or PROC IMPORT statement. Be sure to clearly define all variables using the PROC FORMAT and LABEL procedure. Save this file. Use the PROC PRINT procedure to print data that includes formats and submit. Also, submit your SAS code via email so I can check your work.**

Program Code:

**proc format;**

**value ETHNIC**

**1='White'**

**2='Black'**

**3='Hispanic'**

**;**

**value SMOKE**

**1='non-smoker'**

**2='light smoker'**

**3='heavy smoker'**

**;**

**value $BREASTFED**

**'Y'='Baby will be breastfed'**

**'N'='Baby will not be breastfed'**

**;**

**;**

**run;**

**options pagesize=80 linesize=80 pageno=1 nodate;**

**data birthweight;**

**infile '/folders/myfolders/code/sdm/data/hw1/birthweight.txt' truncover;**

**input Ethnic 1 Age 2-3 Smoke 4 preweight 5-10 delweight 11-16 breastfed $ 17 brthwght 18-21 brthlngth 22-26;**

**proc print data=birthweight LABELnoobs split='/' noobs;**

**TITLE "Hw 1 birthweigh";**

**label**

**Ethnic = ' Political/Party '**

**Age='Age of/mother'**

**Smoke='Smoking status/of mother'**

**preweight='Weight of mother/before pregnancy'**

**delweight='Weight of/mother at/delivery'**

**breastfed='Indicator if baby was/breastfed'**

**brthwght='Weight/of baby/at birth'**

**brthlngth='Length/of baby/at birth'**

**;**

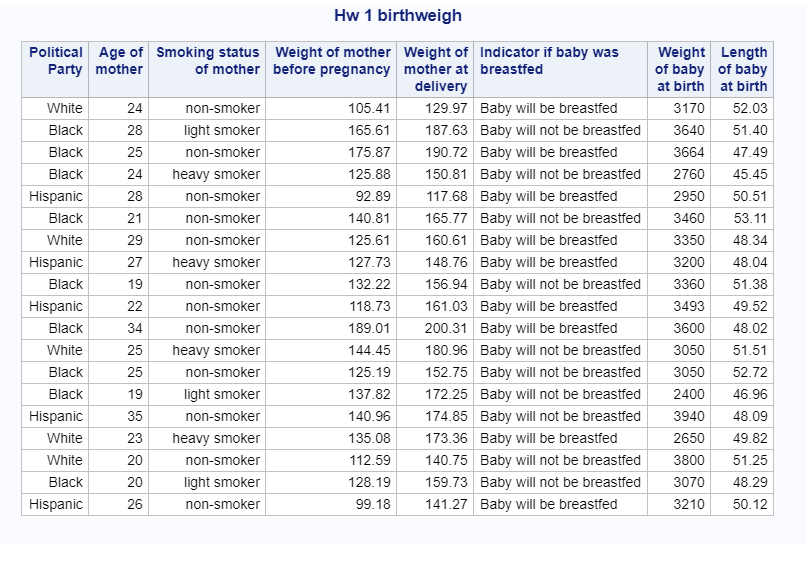
**format Ethnic ETHNIC.**

**Smoke SMOKE.**

**breastfed $BREASTFED.**

**;**

**run;**

****

**2. You should create a SAS program which does the following:**

**(a) Calculate the initial diameter of each tree.**

**(b) Similarly calculate the diameter of each tree at the end of the study**

**(c) Calculate the difference (diff) between the initial diameter and the diameter**

**when the study ended.**

**(d) Run the program using PROC PRINT to make sure that it works and that**

**the data values are being read in correctly**.

**Program:**

**data treeda;**

**infile '/folders/myfolders/code/sdm/data/hw1/treeda.txt' DELIMITER =' ';**

**input number species $ circumference\_start code $ circumference\_end;**

**Run;**

**data treeda;**

**set treeda;**

**diameter\_start=circumference\_start/3.14;**

**diameter\_end=circumference\_end/3.14;**

**difference=diameter\_end-diameter\_start;**

**Run;**

**proc sort data=treeda;**

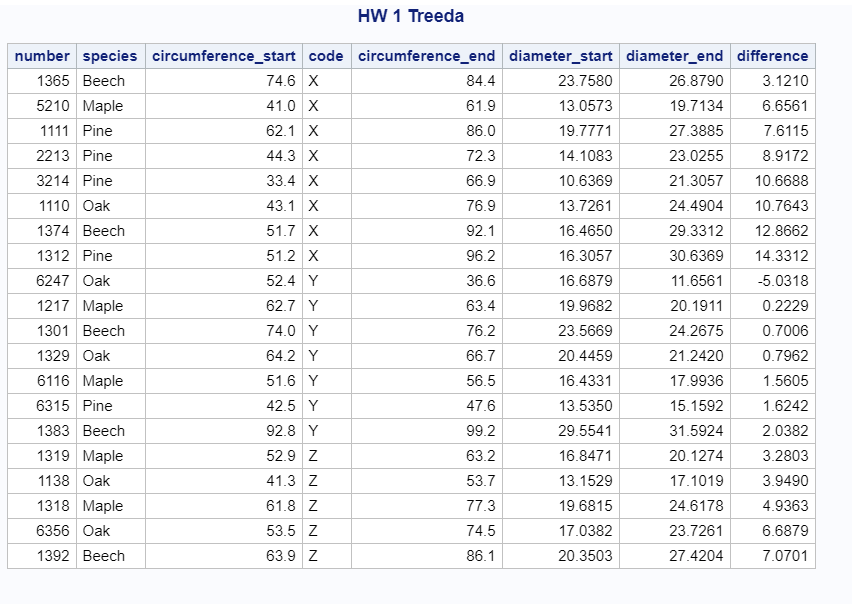
**By code difference;**

**run;**

**proc print data=treeda noobs;**

**title "HW 1 Treeda";**

**Run;**



(e) Use PROC ANOVA to determine whether there is a significant difference

between the means of the three fertilizers.

**proc anova data=treeda;**

**title 'Treeda Anova One way';**

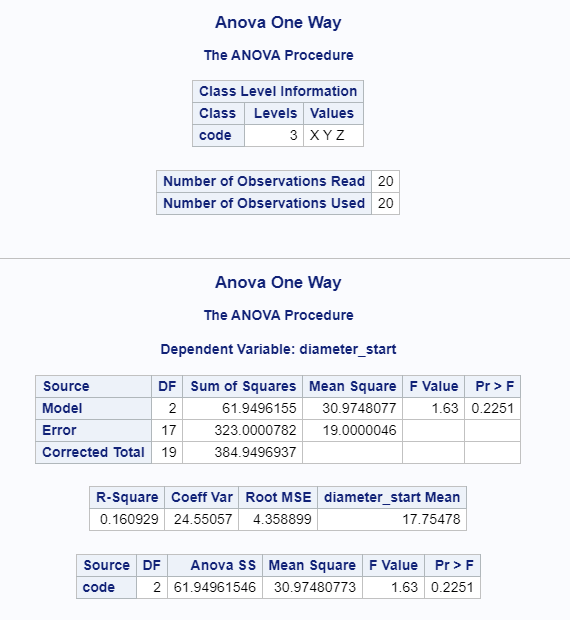
**class code ;**

**model diameter\_start diameter\_end difference=code ;**

**means code / snk;**

**run;**

Anova for diameter at the beginning



\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Ho-> The effects of fertilizers on means of the diameters at the beginning = zero

Ha-> The effects of fertilizers on means of the diameters at the beginning are not = zero

F-value = 1,63 p-value=0.2251 > alpha=0.05

Fail to Reject Ho.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

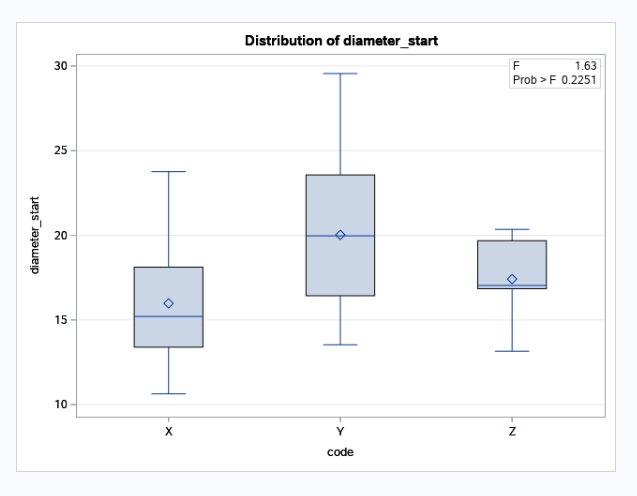
Conclusion about the effects of fertilizers on means of the diameters at the beginning :

At significance level alpha=0.05

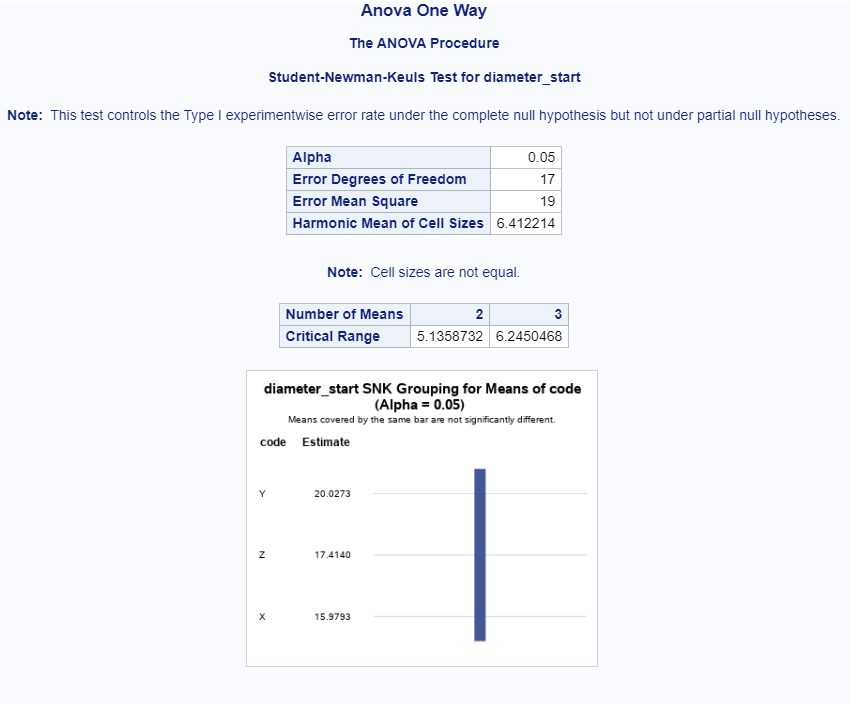
effects of fertilizers on means of the diameters at the beginning are not significantly

different from zero.(effect of fertilizers is not significant on a diameter at the beginning)

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*



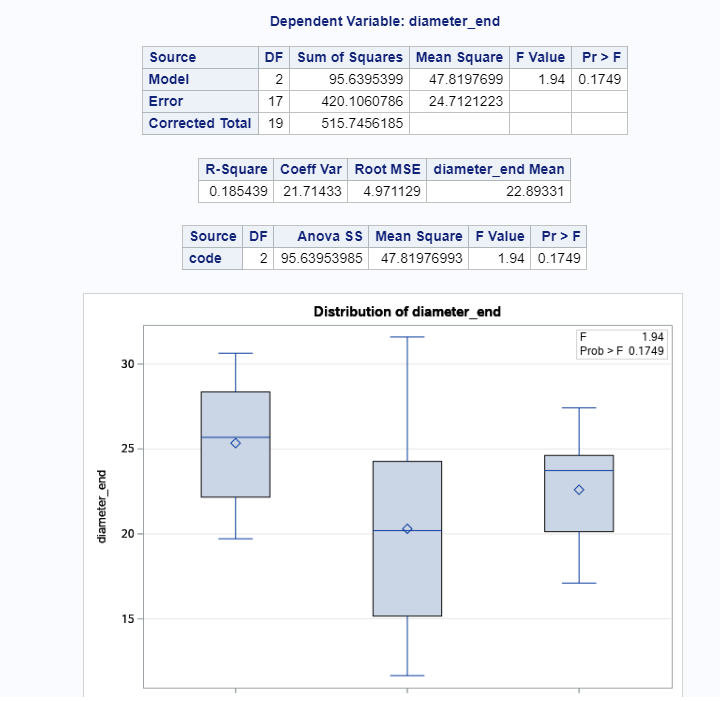
From the box plot we can see that the difference in means for 3 fertilizers are not significant.



SNK test also shows no difference in means of the diameters at the beginning

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Anova for diameter at the end



\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Ho-> The effects of fertilizers on means of the diameters at the end = zero

Ha-> The effects of fertilizers on means of the diameters at the end are not = zero

F-value = 1,94 p-value=0.1749 > alpha=0.05

Fail to Reject Ho.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

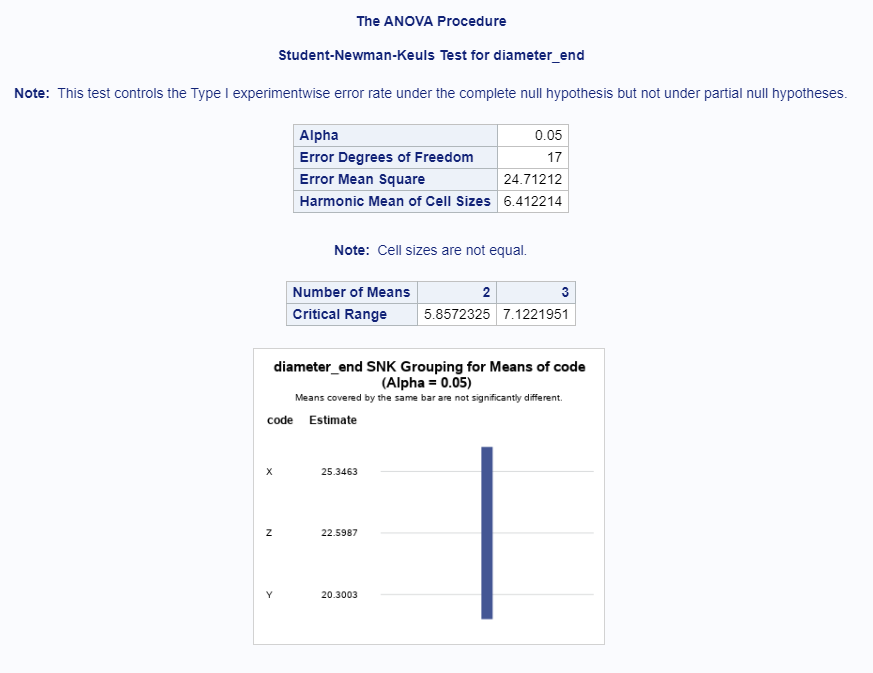
Conclusion about the effects of fertilizers on means of the diameters at the end :

At significance level alpha=0.05

effects of fertilizers on means of the diameters at the end are not significantly

different from zero.(effect of fertilizers is not significant on a diameter at the end)

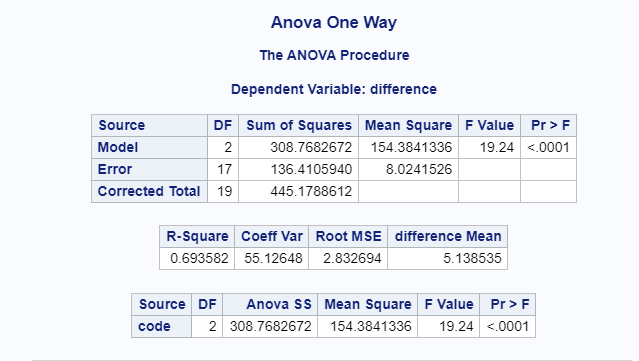
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*



SNK test also shows no difference in means of the diameters at the beginning

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Anova for difference in diameters



\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Ho-> The effects of fertilizers on means of the difference of the diameters = zero

Ha-> The effects of fertilizers on means of the difference of the diameters are not = zero

**F-value = 19,24 p-value=0.0001 < alpha=0.05**

**Reject Ho.**

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Conclusion about the effects of fertilizers on means of the difference of the diameters :

**At significance level alpha=0.05**

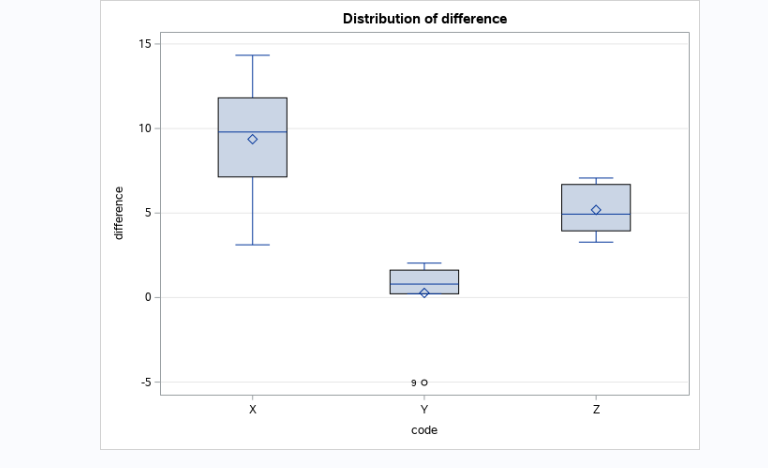
effects of fertilizers on means of the difference of the diameters are significantly

different from zero

.(**effect of fertilizers are very significant on the difference in diameters**)

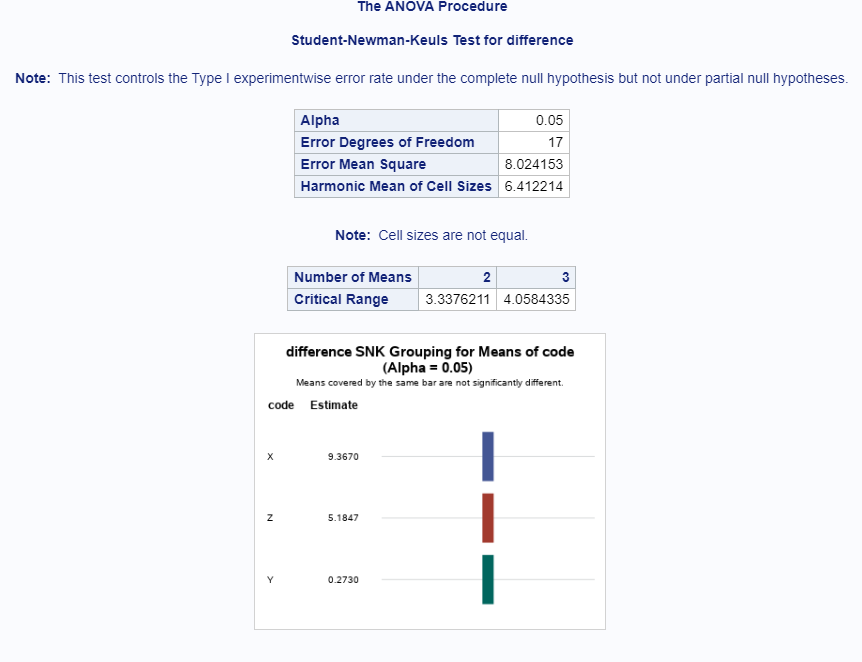
R-Square is 0.693, which means that the model accounts for 69% of data

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*



Box plot also shows the difference in means of the difference of the diameters for

3 fertilizers

SNK test also shows that each of 3 fertilizes belong to the 3 distinct groups.