1. Two cholesterol lowering medications (statins) and a placebo were given to each of 10 volunteers with total cholesterol readings of 240 or higher. After 6 weeks, the following total cholesterol values were recorded:

**Statin A: 220** **190** **180** **185** **210** **170** **178** **200** **177** **189**

**Statin B: 160** **168** **178** **200** **172** **155** **159** **167** **185** **199**

**Placebo: 240** **220** **246** **244** **198** **238** **277** **255** **190** **188**

Create a SAS data set by reading these data. That is, have the program create the treatment variable (call it TREAT) with values of ‘A’, ‘B’, and ‘Placebo’. Next run a oneway ANOVA.

Code:

**data** cholesterol;

DO GROUP = 'STATIN\_A','STATIN\_B','PLACEBO ';

DO Patient = **1** TO **10**;

INPUT treat @;

OUTPUT;

END;

END;

datalines;

220 190 180 185 210 170 178 200 177 189

160 168 178 200 172 155 159 167 185 199

240 220 246 244 198 238 277 255 190 188

;

**run**;

**proc** **print** data=cholesterol;

**run**;

**PROC** **ANOVA** DATA=cholesterol;

CLASS GROUP ;

MODEL treat = GROUP;

MEANS GROUP/SNK alpha=**0.05**;

TITLE "Analysis of Cholesterol Treatment Data";

**RUN**;

Log file:

1 data cholesterol;

2 DO GROUP = 'STATIN\_A','STATIN\_B','PLACEBO ';

3 DO Patient = 1 TO 10;

4 INPUT treat @;

5 OUTPUT;

6 END;

7 END;

8 datalines;

NOTE: SAS went to a new line when INPUT statement reached past the end of a line.

NOTE: The data set WORK.CHOLESTEROL has 30 observations and 3 variables.

NOTE: DATA statement used (Total process time):

real time 2.00 seconds

cpu time 0.09 seconds

12 ;

13 run;

14

15 proc print data=cholesterol;

NOTE: Writing HTML Body file: sashtml.htm

16 run;

NOTE: There were 30 observations read from the data set WORK.CHOLESTEROL.

NOTE: PROCEDURE PRINT used (Total process time):

real time 3.84 seconds

cpu time 0.35 seconds

17

18

19 PROC ANOVA DATA=cholesterol;

20 CLASS GROUP ;

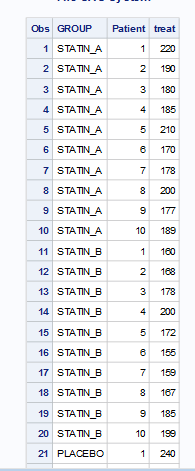
21 MODEL treat = GROUP;

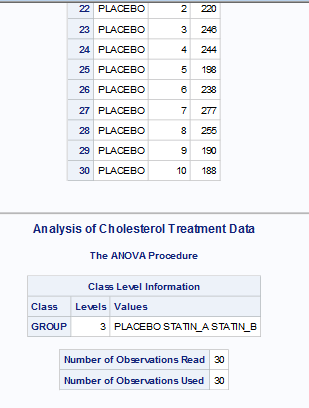
22 MEANS GROUP/SNK alpha=0.05;

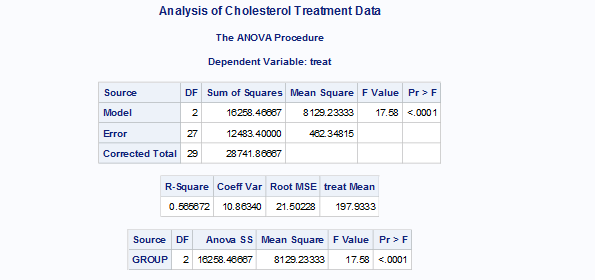
23 TITLE "Analysis of Cholesterol Treatment Data";

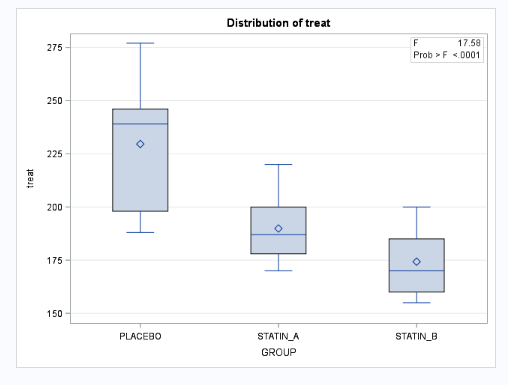
24 RUN;

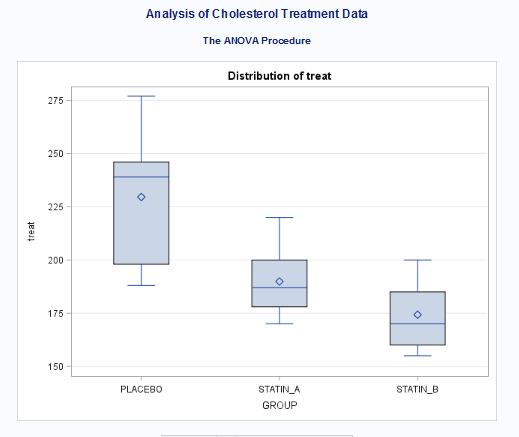
Output:

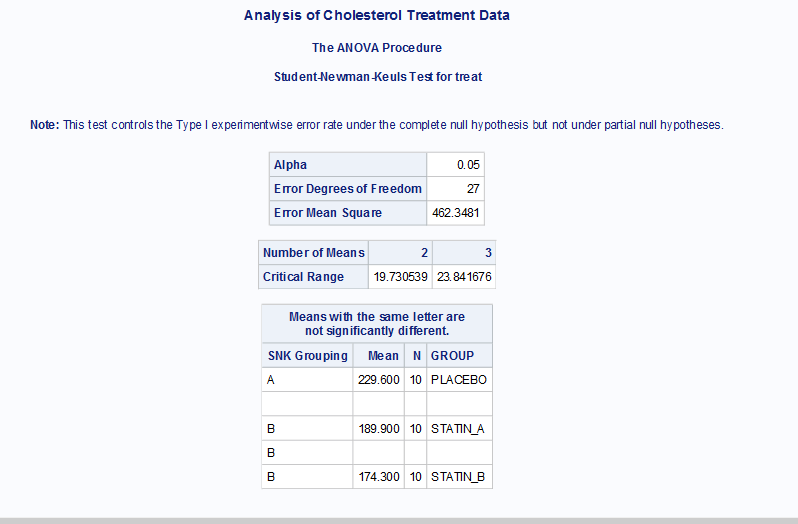












1. Four different methods of preparing for a college entrance exam were compared. They are labeled ‘A’, ‘B’, ‘C’, and ‘D’, The following exam scores were obtained for each of the four programs:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Program |  |  |  |  |  |  |  |  |
| A | 560 | 520 | 530 | 525 | 575 | 527 | 580 | 620 |
| B | 565 | 522 | 520 | 530 | 510 | 522 | 600 | 590 |
| C | 512 | 518 | 555 | 502 | 510 | 520 | 516 |  |
| D | 505 | 508 | 512 | 520 | 543 | 523 | 517 |  |

Compare the four preparation methods. Use a multiple-comparison method of your choice to make pairwise comparisons. Create two contrasts: One to compare method A and B to C and D, the other to compare method D to the other three.

Code:

**DATA** entrance\_exam;

INPUT PROGRAM $ SCORE;

DATALINES;

A 560

A 520

A 530

A 525

A 575

A 527

A 580

A 620

B 565

B 522

B 520

B 530

B 510

B 522

B 600

B 590

C 512

C 518

C 555

C 502

C 510

C 520

C 516

D 505

D 508

D 512

D 520

D 543

D 523

D 517

;

**run**;

**proc** **print** data=entrance\_exam;

title 'Entrance Exam';

**run**;

Title;

**PROC** **GLM** DATA=entrance\_exam;

TITLE "Analysis of College Entrance Exam";

CLASS PROGRAM;

MODEL SCORE = PROGRAM;

MEANS PROGRAM/SNK alpha=**0.05**;

CONTRAST 'A B VS C D' PROGRAM -**1** -**1** **1** **1**;

CONTRAST 'A B C VS D' PROGRAM -**1** -**1** -**1** **3**;

**RUN**;

Title;

Log file:

25 DATA entrance\_exam;

26 INPUT PROGRAM $ SCORE;

27 DATALINES;

NOTE: The data set WORK.ENTRANCE\_EXAM has 30 observations and 2 variables.

NOTE: DATA statement used (Total process time):

real time 0.25 seconds

cpu time 0.01 seconds

58 ;

59 run;

60

61 proc print data=entrance\_exam;

62 title 'Entrance Exam';

63 run;

NOTE: There were 30 observations read from the data set WORK.ENTRANCE\_EXAM.

NOTE: PROCEDURE PRINT used (Total process time):

real time 0.66 seconds

cpu time 0.04 seconds

64 Title;

65

66 PROC GLM DATA=entrance\_exam;

67 TITLE "Analysis of College Entrance Exam";

68 CLASS PROGRAM;

69 MODEL SCORE = PROGRAM;

70 means program/SNK alpha=0.05;

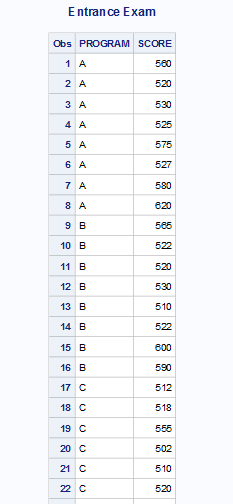
71 CONTRAST 'A B VS C D' PROGRAM -1 -1 1 1;

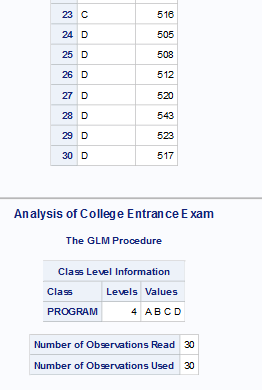
72 CONTRAST 'A B C VS D' PROGRAM -1 -1 -1 3;

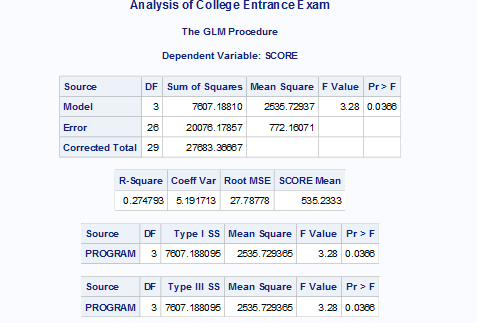
73 RUN;

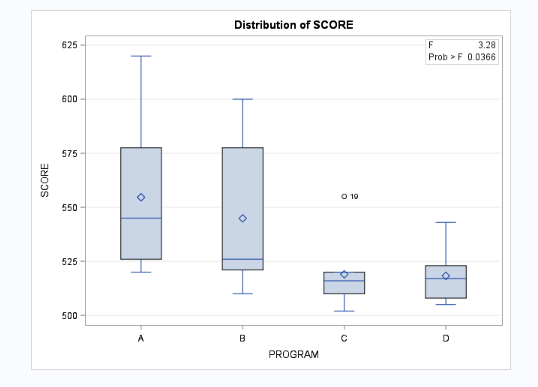
74 Title;

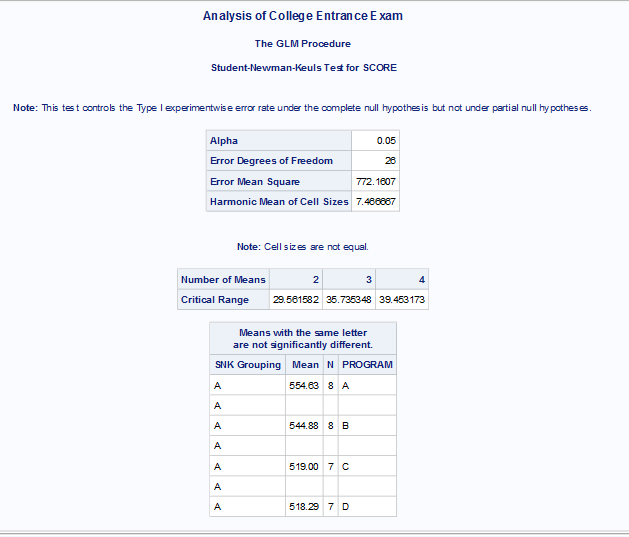
Output:

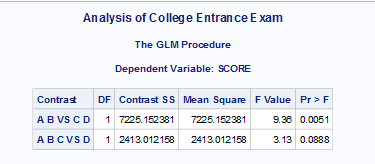












From the above output we can infer that Contrast 1 is significant as its p-value is 0.0051 which is less than alpha 0.05 and Contrast 2 is not significant as its p-value is 0.0888.

1. Two groups of patients, those with a genetic deficiency and the other “normal” group, are given either a drug to cure depression or a placebo. A standardized depression survey is given to all the subjects with the following results:

|  |  |  |
| --- | --- | --- |
|  | Anti-Depression Drug | Placebo |
| Genetic Deficiency | 9 | 9 |
| 11 | 6 |
| 10 | 6 |
| 10 | 7 |
| Normal | 5 | 12 |
| 4 | 11 |
| 7 | 10 |
| 7 | 11 |

Perform a two-way analysis of variance to test for drug and genetic differences.

Code:

**DATA** dep\_study;

INPUT GROUP $ **19.** DRUG $ **7.** DEPRESSION;

DATALINES;

Genetic\_Deficiency DRUG 9

Genetic\_Deficiency PLACEBO 9

Genetic\_Deficiency DRUG 11

Genetic\_Deficiency PLACEBO 6

Genetic\_Deficiency DRUG 10

Genetic\_Deficiency PLACEBO 6

Genetic\_Deficiency DRUG 10

Genetic\_Deficiency PLACEBO 7

Normal DRUG 5

Normal PLACEBO 12

Normal DRUG 4

Normal PLACEBO 11

Normal DRUG 7

Normal PLACEBO 10

Normal DRUG 7

Normal PLACEBO 11

;

**RUN**;

**proc** **print** data=dep\_study;

title 'Depression Dataset';

**run**;

Title;

**PROC** **ANOVA** DATA=dep\_study;

TITLE "Analysis of depression Data";

CLASS GROUP DRUG;

MODEL DEPRESSION = GROUP DRUG;

MEANS GROUP/SNK alpha=**0.05**;

**RUN**;

Title;

Log file:

75 DATA dep\_study;

76 INPUT GROUP $ 19. DRUG $ 7. DEPRESSION;

77 DATALINES;

NOTE: The data set WORK.DEP\_STUDY has 16 observations and 3 variables.

NOTE: DATA statement used (Total process time):

real time 0.01 seconds

cpu time 0.00 seconds

94 ;

95 RUN;

96 proc print data=dep\_study;

97 title 'Depression Dataset';

98 run;

NOTE: There were 16 observations read from the data set WORK.DEP\_STUDY.

NOTE: PROCEDURE PRINT used (Total process time):

real time 0.19 seconds

cpu time 0.01 seconds

99 Title;

100

101 PROC ANOVA DATA=dep\_study;

102 TITLE "Analysis of depression Data";

103 CLASS GROUP DRUG;

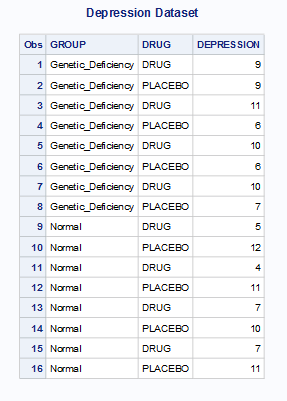
104 MODEL DEPRESSION = GROUP DRUG;

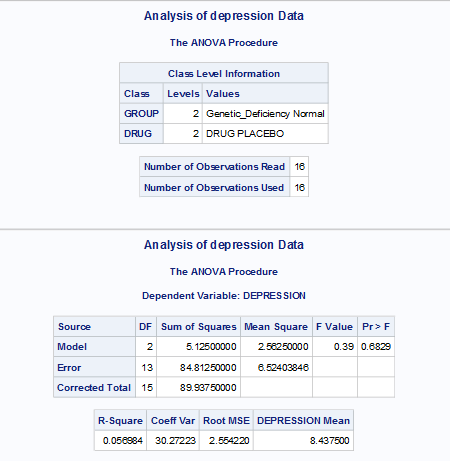
105 MEANS GROUP/SNK alpha=0.05;

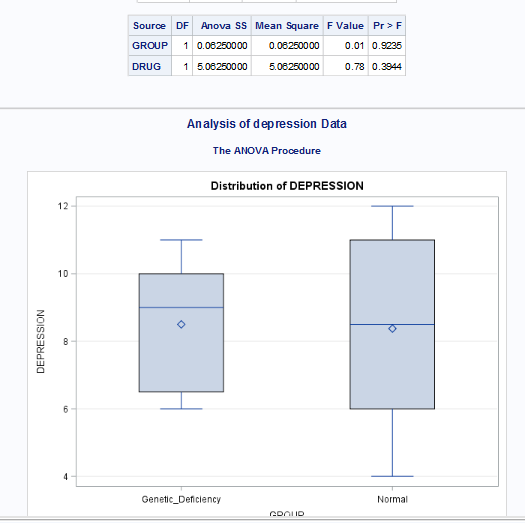
106 RUN;

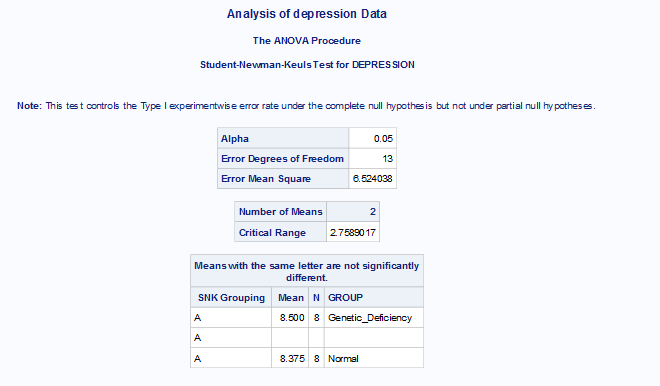
107 Title;

Output:









1. You are given data on the speed with which rats can negotiate a maze. The rats are grouped into three age groups and two genetic strains. Conduct a two-way ANOVA with AGE and STRAIN as the independent variables and SPEED as the dependent variable. The collected data are shown here.

|  |  |  |
| --- | --- | --- |
| Age | Strain | Speed |
| 3 | A | 12 |
| 3 | A | 14 |
| 3 | A | 9 |
| 3 | A | 17 |
| 3 | A | 10 |
| 3 | A | 11 |
| 3 | A | 9 |
| 3 | A | 10 |
| 3 | B | 24 |
| 3 | B | 17 |
| 3 | B | 22 |
| 3 | B | 16 |
| 3 | B | 18 |
| 6 | A | 22 |
| 6 | A | 20 |
| 6 | A | 12 |
| 6 | A | 12 |
| 6 | A | 17 |
| 6 | A | 14 |
| 6 | A | 17 |
| 6 | B | 23 |
| 6 | B | 26 |
| 6 | B | 34 |
| 6 | B | 20 |
| 9 | A | 14 |
| 9 | A | 14 |
| 9 | A | 10 |
| 9 | A | 15 |
| 9 | A | 17 |
| 9 | A | 12 |
| 9 | A | 19 |
| 9 | B | 27 |
| 9 | B | 29 |
| 9 | B | 27 |
| 9 | B | 23 |

Code:

**Data** rats;

Input Age Strain $ Speed;

datalines;

3 A 12

3 A 14

3 A 9

3 A 17

3 A 10

3 A 11

3 A 9

3 A 10

3 B 24

3 B 17

3 B 22

3 B 16

3 B 18

6 A 22

6 A 20

6 A 12

6 A 12

6 A 17

6 A 14

6 A 17

6 B 23

6 B 26

6 B 34

6 B 20

9 A 14

9 A 14

9 A 10

9 A 15

9 A 17

9 A 12

9 A 19

9 B 27

9 B 29

9 B 27

9 B 23

;

**run**;

**PROC** **ANOVA** DATA=rats;

TITLE "Analysis of Rats Data";

CLASS Age Strain;

MODEL Speed = Strain Age;

MEANS Age Strain/SNK alpha=**0.05**;

**RUN**;

Title;

Log file:

108 Data rats;

109 Input Age Strain $ Speed;

110 datalines;

NOTE: The data set WORK.RATS has 35 observations and 3 variables.

NOTE: DATA statement used (Total process time):

real time 0.01 seconds

cpu time 0.01 seconds

146 ;

147 run;

148

149

150 PROC ANOVA DATA=rats;

151 TITLE "Analysis of Rats Data";

152 CLASS Age Strain;

153 MODEL Speed = Strain Age;

154 MEANS Age Strain/SNK alpha=0.05;

155 RUN;

WARNING: PROC ANOVA has determined that the number of observations in each cell is not equal.

PROC GLM may be more appropriate.

156

157 Title;

Output:

