2-Way Factorial Design ANOVA
Example 10.9

Brand A is 1, B is 2, C is 3 and D is 4.
Club 1 is the driver is the 5-iron.

```
MTB > read 'mcclave/c10t11.dat' c1-c3
Entering data from file: mcclave/c10t11.dat
     32 rows read.

MTB > name c1 'brand' c2 'club' c3 'distance'
MTB > anova 'distance' = 'brand'|'club';
SUBC> means 'brand'|'club'.

Factor     Type Levels Values
brand     fixed      4     1     2     3     4
club      fixed      2     1     2

Analysis of Variance for distance

Source         DF          SS          MS        F      P
brand           3       800.7       266.9     7.79  0.001
club            1     32093.1     32093.1   936.75  0.000
brand*club      3       766.0       255.3     7.45  0.001
Error          24       822.2        34.3
Total          31     34482.0


    MEANS

 brand   N  distance
     1   8    199.86
     2   8    208.20
     3   8    205.14
     4   8    195.12

 club    N  distance
    1   16    233.75
    2   16    170.41

 brand club   N  distance
     1    1   4    228.43
     1    2   4    171.30
     2    1   4    233.73
     2    2   4    182.68
     3    1   4    243.10
     3    2   4    167.18
     4    1   4    229.75
     4    2   4    160.50
```

MINITAB does not compute SST for us. We have to do it ourselves.
SST = SSA + SSB + SSA(AB) = 800.7 + 32093.1 + 766.0 = 33659.8.
MST = 33659.8/7 = 4808.5 and $F_{treatment}$ = 4898.5/34.3 = 140.2

Since $F_{treatment}$ is 140.2 which is greater than $F_{.10, 7, 24}$ which is 1.98, there is enough evidence to indicate that the mean distances differ for at least 2 of the 8 brand and club combinations. A test to determine whether the brand of ball and the type of club interact to affect the distance should be done.

Since $F_{brand*club}$ = 7.45 with p-value = 0.001, there is enough evidence to indicate that the brand of golf ball and the type of club used interact to affect the distance traveled. Therefore, we need to compare all possible pairs of means. Since p = 8, there are 28 possible pairs to test. Since this would be a very boring and time consuming task to do by hand, I have analyzed the data using a one-way ANOVA with 8 treatments and used Bonferroni with α/c = .10/8 to compare the 28 possible pairs of means.

```
MTB > let c4 = 4*c2 +c1 -4
MTB > name c4 'treat'
MTB > print c1 c2 c4


 ROW  brand   club   treat

   1      1      1    1      Treatment 1 is brand A with a driver. (Ad)
   5      1      2    5      Treatment 5 is brand A with a 5-iron. (A5)
   9      2      1    2      Treatment 2 is brand B with a driver. (Bd)
  13      2      2    6      Treatment 6 is brand B with a 5-iron. (B5)
  17      3      1    3      Treatment 3 is brand C with a driver. (Cd)
  21      3      2    7      Treatment 7 is brand C with a 5-iron. (C5)
  25      4      1    4      Treatment 4 is brand D with a driver. (Dd)
  29      4      2    8      Treatment 8 is brand D with a 5-iron. (D5)

MTB > let k1 = .10/28
MTB > print k1

K1        0.00357143

MTB > oneway 'distance' 'treat';
SUBC> fisher k1.

ANALYSIS OF VARIANCE ON distance
SOURCE     DF        SS        MS        F        p
treat       7   33659.8    4808.5   140.35    0.000
ERROR      24     822.2      34.3
TOTAL      31   34482.0
                                    INDIVIDUAL 95 PCT CI'S FOR MEAN
                                    BASED ON POOLED STDEV
 LEVEL      N      MEAN     STDEV   ---------+---------+---------+-------
     1      4    228.43      6.12                         (-*-)
     2      4    233.73      4.95                           (-*-)
     3      4    243.10      3.25                              (-*-)
     4      4    229.75      7.56                          (-*-)
     5      4    171.30      6.67       (-*-)
     6      4    182.68      3.15           (-*-)
     7      4    167.18      9.28      (-*-)
     8      4    160.50      1.96   (--*-)
                                    ---------+---------+---------+-------
POOLED STDEV =     5.85                     180       210       240
```

Fisher's pairwise comparisons

Family error rate = 0.0596
Individual error rate = 0.00357

Critical value = 3.230

Intervals for (column level mean) - (row level mean)

	1 Ad	2 Bd	3 Cd	4 Dd	5 A5	6 B5	7 C5
2 Bd	-18.67 8.07						
3 Cd	-28.04 -1.31	-22.74 3.99					
4 Dd	-14.69 12.04	-9.39 17.34	-0.02 26.72				
5 A5	43.76 70.49	49.06 75.79	58.43 85.17	45.08 71.82			
6 B5	32.38 59.12	37.68 64.42	47.06 73.79	33.71 60.44	-24.74 1.99		
7 C5	47.88 74.62	53.18 79.92	62.56 89.29	49.21 75.94	-9.24 17.49	2.13 28.87	
8 D5	54.56 81.29	59.86 86.59	69.23 95.97	55.88 82.62	-2.57 24.17	8.81 35.54	-6.69 20.04

Each pair in this box compares a brand of ball hit by a driver to one hit by a 5-iron.

There is enough evidence to indicate that the mean distance traveled for each brand of ball using a driver is greater than that for every brand of ball using a 5-iron. For the driver, there is enough evidence to indicate that the mean distance traveled is greater for brand C than for brand A. No other pairs of means differ when a driver is used. For the 5-iron, there is enough evidence to indicate that the mean distance traveled is greater for brand B than for brands C and D.

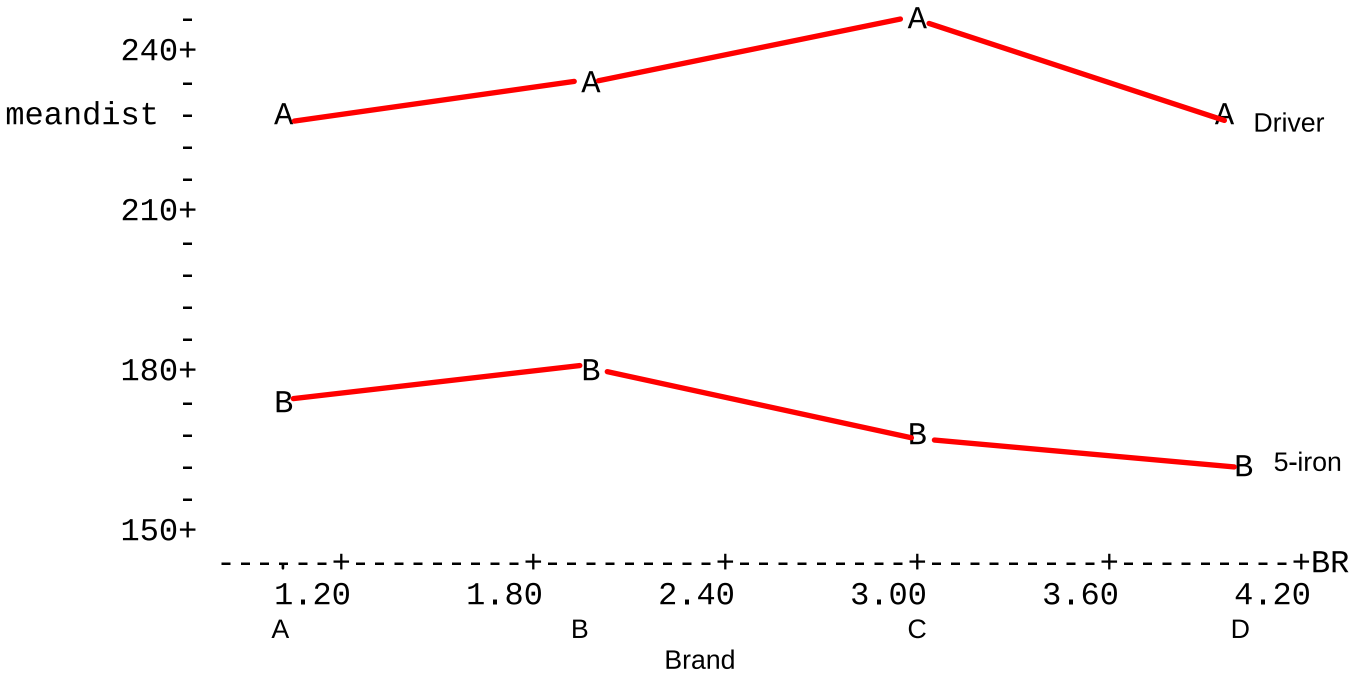
Next we will have MINITAB plot the interaction graphs.

Enter the data for the sample means.

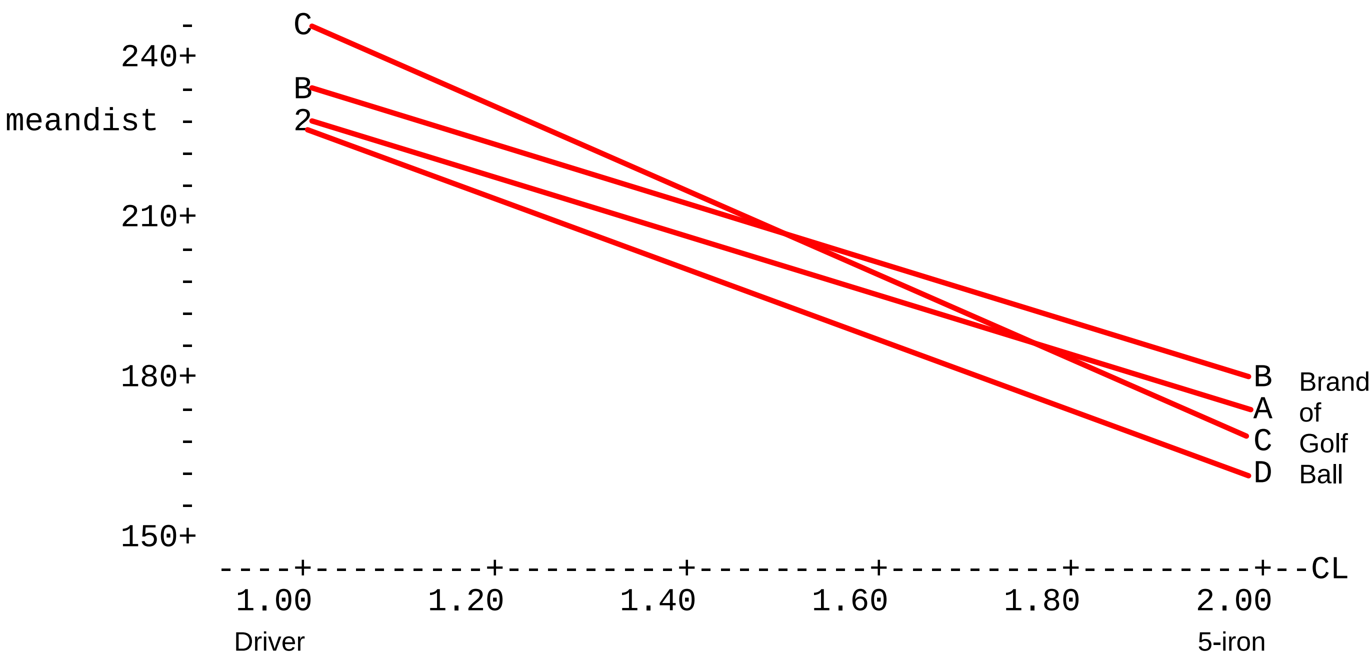
```
MTB > read c21-c23
DATA> 1 1 228.43
DATA> 1 2 233.73
DATA> 1 3 243.10
DATA> 1 4 229.75
DATA> 2 1 171.30
DATA> 2 2 182.68
DATA> 2 3 167.18
DATA> 2 4 160.50
DATA> end
      8 rows read.
```

Plot the interaction graphs for this experiment.

```
MTB > name c21 'CL' c22 'BR' c23 'meandist'
MTB > lplot c23 c22 c21
```




```
MTB > lplot c23 c21 c22
```



2-Way Factorial Design ANOVA
Example 10.10

Brand E is 1, F is 2, G is 3 and H is 4.
Club 1 is the 5-iron and 2 is the driver.

```
MTB > read 'mcclave/9thed/golffac2.dat' c1-c3
Entering data from file: mcclave/9thed/golffac2.dat
     32 rows read.
MTB > name c1 'club' c2 'brand' c3 'distance'
MTB > anova 'distance' = 'club' | 'brand';
SUBC> means 'club' | 'brand'.

Factor     Type Levels Values
club      fixed      2     1     2
brand     fixed      4     1     2     3     4


Analysis of Variance for distance

Source         DF         SS         MS       F      P
club            1    46443.9    46443.9 1887.94  0.000
brand           3     3410.3     1136.8   46.21  0.000
club*brand      3      105.2       35.1    1.42  0.260
Error          24      590.4       24.6
Total          31    50549.8



    MEANS

 club    N  distance
    1   16    172.82
    2   16    249.01

 brand   N  distance
     1   8    202.44
     2   8    218.44
     3   8    223.59
     4   8    199.20

 club brand   N  distance
    1     1   4    164.38
    1     2   4    177.73
    1     3   4    188.00
    1     4   4    161.18
    2     1   4    240.50
    2     2   4    259.15
    2     3   4    259.17
    2     4   4    237.23
```

MINITAB does not compute SST for us. We have to do it ourselves.
SST = SSA + SSB + SSA(AB) = 46443.9 + 3410.3 + 105.2 = 49959.4.
MST = 49959.4/7 = 7137.1 and $F_{treatment}$ = 7137.1/24.6 = 290.1

Since $F_{treatment}$ is 290.1 which is greater than $F_{.10, 7, 24}$ which is 1.98, there is enough evidence to indicate that the mean distances differ for at least 2 of the 8 brand and club combinations. A test to determine whether the brand of ball and the type of club interact to affect the distance should be done.

Since $F_{brand*club}$ = 1.42 with p-value = 0.260, there is not enough evidence to indicate that the brand of golf ball and the type of club used interact to affect the distance traveled. Hence, tests for main effects should be done.

Since F_{club} = 1887.94 with p-value = 0.000, there is enough evidence to indicate that the mean distances traveled differ for the 2 types of golf clubs. Since there are only two golf clubs an examination of the data is enough to determine which mean is greater. Since $\bar{x}_{5-iron}$ = 172.82 and $\bar{x}_{driver}$ = 249.01, we can conclude that the mean distance traveled when hit by a driver is greater than the mean distance traveled when hit by a 5-iron.

Since F_{brand} = 46.21 with p-value = 0.000, there is enough evidence to indicate that the mean distances traveled differ for at least 2 of the 4 brands of golf balls used. Bonferroni confidence intervals should be constructed to determine which of the 6 pairs of means for brands differ.

$$\bar{x}_i - \bar{x}_j \pm t_{\frac{\alpha}{2c}} \sqrt{\frac{2MSE}{ra}}$$

In this formula, r = number of observations per treatment and a = the number of golf clubs used. Therefore, ra is the number of observations per ball. In this example, r = 4, a = 2 and ra = 8.

To find the value of $t_{\frac{\alpha}{2c}}$, use the Bonferroni table with α = .10, c = 6 and df = 24. Therefore, $t_{\frac{\alpha}{2c}}$ = 2.5736.

Since MSE = 24.6, $t_{\frac{\alpha}{2c}} \sqrt{\frac{2MSE}{ra}} = 2.5736\sqrt{\frac{2(24.6)}{8}} = 6.38$.

The Bonferroni Confidence Intervals

$$-16.00 - 6.38 < \mu_E - \mu_F < -16.00 + 6.38$$
$$-22.38 < \mu_E - \mu_F < -9.62$$

$$-21.15 - 6.38 < \mu_E - \mu_G < -21.15 + 6.38$$
$$-27.53 < \mu_E - \mu_G < -14.77$$

$$3.24 - 6.38 < \mu_E - \mu_H < 3.24 + 6.38$$
$$-3.14 < \mu_E - \mu_H < 9.62$$

$$-5.15 - 6.38 < \mu_F - \mu_G < -5.15 + 6.38$$
$$-11.53 < \mu_F - \mu_G < 1.23$$

$$19.24 - 6.38 < \mu_F - \mu_H < 19.24 + 6.38$$
$$12.86 < \mu_F - \mu_H < 25.62$$

$$24.39 - 6.38 < \mu_G - \mu_H < 24.39 + 6.38$$
$$18.01 < \mu_G - \mu_H < 30.77$$

What conclusions can be drawn about the mean distances traveled by the four brands of golf balls?

There is enough evidence to indicate that the mean distance traveled by brand E golf balls is greater than that for brands F and G and that the mean distance traveled by brand H golf balls is less than that for brands F and G. There are no other pairs of means which differ significantly.

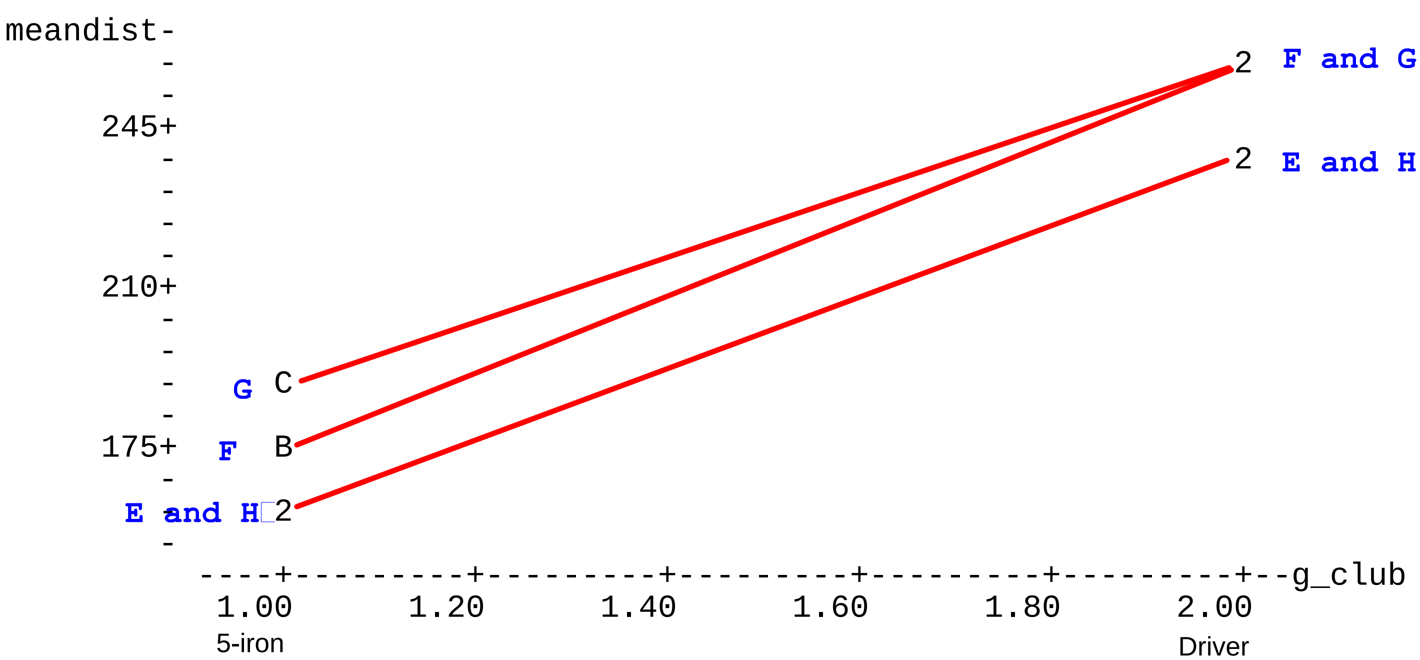
The diagram for these results is given.

```
              H    E    F    G
             ________
                      __________
```

Next we will have MINITAB plot the interaction graphs.

Enter the data for the sample means.

```
MTB > read c11- c14
DATA>  1     1   4    164.38
DATA>  1     2   4    177.73
DATA>  1     3   4    188.00
DATA>  1     4   4    161.18
DATA>  2     1   4    240.50
DATA>  2     2   4    259.15
DATA>  2     3   4    259.17
DATA>  2     4   4    237.23
DATA> And
      8 rows read.
MTB > name c11 'g_club' c12 'b_brand' c14 'meandist'
MTB > lplot c14 c11 c12
```



```
MTB > lplot c14 c12 c11
```