SECTION 7: ECONOMIC ANALYSIS

This section discusses subprograms which provide you with economic analyses of your data. It is recommended that you obtain a copy of the CIMMYT Information Bulletin No. 27, "From Agronomic Data to Farmer Recommendations - An Economics Training Guide", for a thorough understanding of the analyses generated by ECON. The Bulletin applies directly to analyses provided by the ECON subprogram and may be obtained from a local CIMMYT office (free).

In addition to the ECON subprogram, there is a series of SEASON subprograms which provide seasonal and storage analysis of data. The Season subprograms are: SEAPLAN, SEACALC, SEASONAL, SEASTORE and SEATABLE. For a thorough understanding of how the analyses are obtained, you will want to obtain a copy of "Fundamentals of Price Analysis in Developing Countries' Food Systems: A Training Manual to Accompany the Microcomputer Software Program 'MSTAT'" by Stephan Goetz and Michael Weber, 1986. Contact the MSU Department of Agricultural Economics, 7 Agriculture Hall, MSU, East Lansing, Michigan 48824-1039 to get a copy of the Guide.

The following subprograms and guides are included in this section:

7.1 ECON	Computes marginal rates of returns, risk benefit tables and economic returns for agronomic practices.

7.2 Seasonal Analysis of Price Data.

7.2.1 SEAPLAN	Defines variables for a price file used in the Season series of subprograms.
7.2.2 SEACALC	Calculates values for some of the variables in a price file.
7.2.3 SEASONAL	Produces tables of regression results and various seasonal indices using data in a price file.
7.2.4 SEASTORE	Produces tables relevant to storage analysis using data in a price file.
7.2.5 SEATABLE	Produces tables of price changes using data in a price file.
7.3 LP	Linear programming to assist economists.

7.1 ECON

7.1 ECON

Purpose: Economic analysis of data from agronomic experiments. Calculates the net benefit, measures of variability or risk, and marginal returns for each treatment. ECON is also used to create the ECON parameters file.

Before starting, you need:

1. An active data file with yield and treatment variable values.
2. A parameters file which will be used to analyze the active data file. If you are going to create or update a parameters file, the following price information is needed:

   a. Field-level prices of products.
   b. Expected percent of harvest losses per product.
   c. Variable cash costs per treatment.
   d. Value of farmer-supplied materials per treatment.
   e. Value of farmer- and family-supplied labor per treatment.

Results: Depending on which options you select, data may be analyzed or information entered in an ECON parameters file.

Sample Files: The MAIZE data file is used in the example as well as the MAIZE-0 parameter file.

ECON is a program which performs economic analysis of data from agronomic experiments. ECON implements the methodology suggested in "From Agronomic Data to Farmer's Recommendations - An Economics Training Manual", CIMMYT Information Bulletin No. 27 by Richard Perrin et. al. It is designed for experiments which examine factors affecting crop yield, such as the use of fertilizer, herbicide, or insecticide, or method of tillage or planting.

The program calculates the economic net benefit (net return) of each treatment, based upon information supplied by the researcher on the value (price) of the crop output and the cash, in-kind, and family labor costs associated with each treatment. If data are available from several experiments incorporating the same treatments, ECON will also calculate minimum yield and yield variance for each treatment. In addition, the marginal rate of return associated with increasing input levels is calculated.

If there are no or minimal increased costs associated with the use of a new technology, the economically preferred treatment is the highest yielding treatment. However, analysis of yield variability is crucial to identify the risks likely to be faced by farmers who adopt the new technology. In some instances, yield stability may be more important than yield increases.

Usually the economic analysis will reinforce the results of the agronomic analysis. High economic returns will generally be associated with a statistically significant treatment response. However, there may be instances when a recommendation is justified on the basis of expected economic returns even though a treatment effect is not statistically significant. For instance, a new input or practice may lower the costs of achieving a given level of output. A recommendation might also be warranted when there is a statistically insignificant increase in the production of a particularly valuable commodity.

When evaluating a technology, overreliance on statistical significance levels is a mistake. Significance levels should be used as initial indicators of the influence of various sources of variance, but should not be treated as answers in themselves. The statistical relationship must be evaluated in terms of the underlying agronomic and economic factors faced by farmers.

An analysis from ECON is acquired through the following steps:

1. Create a data file that contains yield data (Product) and treatment data. You may use ASCII to transfer data from an existing file into an MSTAT-C data file.
2. Use the ECON program to:
   a. Enter price and cost data (monetary units) and other necessary parameters (such as yield unit of measure) into a new parameters file, or retrieve a previously created parameters file.
   b. Check and edit the parameters file.
   c. Run an analysis - calculate and display results:
      -net benefits
      -marginal returns
      -risk analysis
   d Sensitivity analysis (return to step a).

7.1 ECON

Creating a Data File to be used by ECON

ECON uses a data file that contains data for at least two types of variables:
1 Numeric variables with yields from each product in an experiment (e.g., metric tons per hectare of corn, straw or beans). To obtain the best estimate of yields which would be measured on the entire field where the experiment was located, values of the yield variable should be the average of within-site replications. Do not average across locations or across years, since the variability in such yields will tend to represent the variability faced by farmers adopting the new technology. The yield data for each year or for each location is entered in separate cases in the data file.

2 Variables whose values are a unique numeric treatment code; each code represents a treatment used in the experiment (e.g., phosphorus levels or nitrogen levels). Since you may be entering data from various experiments, there may be a different number of observations for each treatment. This is not a problem for ECON.

If your data are not yet in a computer file, you will have to create a new data file, and define numeric variables and enter data using SEDIT. The data file should contain at least two numeric variables, one for treatment code (unique for each treatment) and one or more for yield values from each product.

Transformation of Data to be used by ECON

ECON uses the following data for analysis:

1. One or more numeric variables of yield data (it will be averaged over replications within each unique treatment value).
2. A numeric treatment variable with unique values for each combination of experimental factors.

There are a variety of methods to get these variables, depending upon the nature of the experiment, whether yields are already averaged across within-site replications and whether the data file already contains a treatment variable.

The data may be computed or transformed with the aid of various MSTAT-C subprograms. The MEAN option can be used to generate average yield data. The CALC option may be used to create a treatment variable from the variables representing experimental factors and their levels.

The sample data file used in the following ECON example is MAIZE. The data, taken from CIMMYT Information Bulletin 27, are results of fertilizer trials in maize production. The trials tested four levels of nitrogen and three levels of phosphorus fertilizer in eight experiment locations. The following table shows the average yields obtained from the treatments listed (averaged within site).

Treatment Levels kgs. Nitrogen(N) or Phosphorus(P)	Average Yield (Product) tons/hct
1 000 N - 00 P	2.21
2 050 N - 00 P	3.14
3 100 N - 00 P	3.91
4 150 N - 00 P	4.01
5 000 N - 25 P	2.44
6 050 N - 25 P	3.88
7 100 N - 25 P	4.40
8 150 N - 25 P	4.84
9 000 N - 50 P	2.36
10 050 N - 50 P	4.05
11 100 N - 50 P	4.74
12 150 N - 50 P	5.16

The values for each treatment at each of the eight locations have been entered into the sample data file MAIZE. To view the file MAIZE, use SEDIT or consult Appendix E for a printed copy of the data file. All variables in the MAIZE data file are defined as numeric.

Variable no. 1 (TREATMENT CODE) contains the treatment code (this may differ from the treatment number)

Variable no. 2 (YIELD) contains the maize yield in units of Tons/Hectare. This is the product for this experiment.

Variable no.s 3 and 4 represent the levels of Nitrogen and Phosphorus applied to the experimental plot. Therefore, value 1 of variable no. 3 and value 1 of variable no. 4 indicate that the corresponding maize yield of that particular row is the response to 0 pounds of Nitrogen and 0 pounds of Phosphorus (Treatment 1).

Variables 3 and 4 of this sample data file can be used to create a new treatment variable. There are several ways to create the treatment variable;

the most cumbersome is direct keyboard entry of new values for each case. If the experimental design is complicated, this method may be the only viable alternative. However, this design is somewhat simpler and the CALC option can be used to create the treatment variable by multiplying the value of the Nitrogen variable by ten and adding the resultant value to that of the Phosphorus variable. This gives unique treatment codes 11, 12, 13 and so on up to 41, 42, 43 for Variable 1 TREATMENT. The treatment codes must be unique, but do not have to be sequential. The CALC formula in this example would be:

10 * V3 + V4

<u>PARAMETERS</u> - The Cost/Price Parameters File

The next step in the process is to collect and enter price and other data for the products and inputs used in the experiments. This information is entered into a special "parameters" file you can create using the ECON program; the file will then only be used in ECON and is identified by the .PAR extension of the file name.

The basic principle of partial budgeting is to assign field-level prices to those inputs and outputs which vary from one treatment to another. The prices should be those relevant to the farmers for whom the recommendation is intended. The required information should be derived by determining the field price per unit of output (Product) (e.g. $1000/ton) which may be either the market price minus harvest, storage, transportation and marketing costs or the opportunity price of the product (if the product is not sold). Next, calculate the value of all inputs which vary in magnitude from one treatment to another. The relevant cost categories are the cash costs, the value of farmer-supplied materials, and the opportunity cost of farmer and family labor:

Cash costs - What are the out-of-pocket costs associated with each treatment? In the MAIZE experiment, these are the costs of the various kinds and levels of fertilizer for each treatment. If the farmers typically hire labor to apply fertilizer, there might also be labor application costs which vary from treatment to treatment. Costs that will be incurred regardless of which treatment is used can be ignored as they will be constant in value.

Value of farmer-supplied materials (in-kind costs) - These are opportunity costs--that is, costs which are not cash, but nevertheless require the farmer to give up the opportunity to use the resource to do something else. An example might be a herbicide experiment which requires varying use per treatment of the farmer's oxen for tillage. Each day the oxen work on his own land means that the farmer gives up the opportunity to earn money from renting the oxen to a neighbor. For this example, this value will be zero.

Opportunity cost of family and farmer labor - You must identify family labor which varies from treatment to treatment, and therefore requires the family to give up the opportunity to do something else during that time. The 'something else' might be working on another crop, on a neighbor's farm, off-farm employment in town, or handicraft production in the home.

Two other parameters must be determined for economic analysis: the expected percent of harvest losses per treatment and the relevant target rate of return. This target rate is the estimated minimum rate of return which a farmer may consider necessary in order to make a change of technology worthwhile. Small farmers may not easily adopt a new technology unless it promises a very substantial increase in income or net returns. In general, the rate used in ECON analysis should be at least equal to the effective loan rate (direct cost of capital) plus a risk premium to provide an income safety margin over a period of time. CIMMYT suggests 40%.

MAIZE-0 parameters file

#	Code	Label	Cash Costs	Value Materials	Value Family Labor
1	11	0 N - 0 P	0	0	0
2	21	50 N - 0 P	400	0	50
3	31	100 N - 0 P	800	0	100
4	41	150 N - 0 P	1200	0	100
5	12	0 N - 25 P	250	0	50
6	22	50 N - 25 P	650	0	50
7	32	100 N - 25 P	1050	0	100
8	42	150 N - 25 P	1450	0	100
9	13	0 N - 50 P	500	0	50
10	23	50 N - 50 P	900	0	50
11	33	100 N - 50 P	1300	0	100
12	43	150 N - 50 P	1700	0	100

7.1 ECON

Product:

	MSTAT				%Harvest
#	Var	Label	Unit	Price	Loss
1	2	Maize Yield	T/HCT	10.00	10.00

MSTAT Variable indicating Treatment = 1
Target Rate of Return (in percent) = 40
Net Benefit/Cost Unit = $ / HCT

ECON Example

Before selecting ECON from the MSTAT-C Main Menu, load the sample data file MAIZE. In this example, you will first create a parameters file named MAIZTST. Then, to continue with the example, you will get the MAIZE-0 parameters file. MAIZE is the active data file.

```
ECON Options Menu
  1  Enter cost/price parameters
  2  Get cost/price parameters from disk
  3  Edit cost/price parameters
  4  Write cost/price parameters to disk
  5  Proceed with analysis
  Q  Quit
```

OPTION 1

To enter cost/price parameters, highlight option 1 and press <ENTER> or simply press 1. The following box appears on the screen:

```
Enter the number of products (1-50):     1
Enter the number of treatments  (1-50): 2
```

For this example, only one product and two treatments will be specified. There can be more than one product (yield) per experiment, for example, an inter-cropping experiment testing the total yield of corn and beans for different levels of fertilizer application would have data in two product variables. Enter the number of different yields being analyzed. Then enter the number of unique combinations of factors for the experiment (treatments).

Next, you will be prompted for the information for each treatment. Enter the requested information.

For the first product, enter the MSTAT variable number which contains the product or yield that will be analyzed. Enter the "label' for the product. This label is the name of the product for which you are testing the yield. In this example, the yield recorded is for maize. Continue to enter the information as requested.

```
Product : 1
Enter the MSTAT variable number (1- 4) for yield: 2
Enter the product label:                MAIZE
Enter the yield unit of measurement:    T/HCT
Enter the price of product per unit:    10
Enter the % harvest loss (0 - 100):     10
```

Express all benefit and cost data in terms of the same units. If one cost is calculated in $/HCT or pesos/ton, for example, they all should be in that unit. The value entered for target rate of return is used in the calculation for the marginal rate of return.

```
Enter the MSTAT variable number (1-4) indicating treatment: 1
Enter the target rate of return (%):                      40
Enter the net benefit/cost unit label:                 $/HCT
```

After the parameters have been entered, the user is requested for specific information for each treatment designated. The treatment label consists of text that is printed on the output to identify the treatments in more detail. For this example, the treatment label lists the rates of nitrogen and phosphorus applied. In previous screens, we identified MSTAT variable 1 as the variable indicating treatment. In this example, the value in variable 1 can be a code used to quickly indicate levels of nitrogen and phosphorus applied to each of 12 treatments (fertilizer combinations).

```
Treatment: 1
Enter the treatment label :                   00 N - 00 P
Enter the value of variable 1 indicating this treatment: 11
Enter the total cash costs:                         0.00
Enter the value of farmer-supplied materials:    0.00
Enter the value of family labor:                    0.00
```

7.1 ECON

```
Treatment: 2
Enter the Treatment label :            50 N - 00 P
Enter the value of variable 1 for this treatment: 21
Enter the total cash costs:                  400.00
Enter the value of farmer-supplied materials:  0.00
Enter the value of family labor:              50.00
```

Enter a meaningful text label to identify each treatment. Without such a label, the subsequent tables will be confusing. When you have entered all the treatment information, the program displays the Econ Options Menu.

```
ECON Options Menu
  1   Enter cost/price parameters
  2   Get cost/price parameters from disk
  3   Edit cost/price parameters
  4   Write cost/price parameters to disk
  5   Proceed with analysis
  Q   Quit
```

OPTION 2

To recall a previously created and saved parameter file, use option 2 - Get cost/price parameters from disk. You are prompted for a file name. It is not necessary to type in the .PAR extension. The <F1> key can be pressed to pull down a list of files in the default data directory or the file name can be typed in.

```
Enter the name of the PARAMETER file without
extension:
```

OPTION 3

Editing cost/price parameters is the next step prior to writing the values to disk. The screen lists 5 choices.

```
Edit Parameters Menu
  1   Print out parameters
  2   Edit a product
  3   Edit a treatment
  4   Edit miscellaneous parameters
  Q   Quit
```

The panels presented are identical to the ones used when initiating a parameter file.

```
Product : 1
Enter the MSTAT variable number (1- 4) for yield: 2
Enter the product label:                    MAIZE
Enter the yield unit of measurement:  T/HCT
Enter the price of product per unit:  10
Enter the % harvest loss (0 - 100):   10
```

```
Treatment: 1
Enter the treatment label :             00 N - 00 P
Enter the value of variable 1 indicating this treatment: 11
Enter the total cash costs:                        0.00
Enter the value of farmer-supplied materials:  0.00
Enter the value of family labor:                   0.00
 * = more than 2 s.d.'s from mean
```

```
Enter the MSTAT variable number indicating treatment: 1
Enter the target rate of return:                 40
Enter the net benefit/cost unit label:           $/HCT
```

Press Q when finished with the options shown and the program will continue with the introductory panel:

```
Econ Options Menu
  1   Enter cost/price parameters
  2   Get cost/price parameters from disk
  3   Edit cost/price parameters
  4   Write cost/price parameters to disk
  5   Proceed with analysis
  Q  Quit
```

Save the just entered parameters to a new file by choosing Option 4 to write cost/price parameters to disk before continuing. Enter only up to 8 characters for the name of the parameters file (MAIZTST). MSTAT-C will append the .PAR extension. You should print a reference copy of your parameters each time they are changed (Option 1 - Print out parameters of

Edit Parameters Menu). You may or may not wish to save the edited parameters to disk. In order that parameters files can be easily associated with the correct data files, we suggest that you use a logical, easy-to-remember system of naming both data and parameters files.

```
Enter the name of the PARAMETER file without
extension:
 MAIZTST
```

To continue with the example, the MAIZE-O.PAR file should be loaded in place of the MAIZTST.PAR file. This is a file that is distributed with MSTAT-C.

Select option 3 - Get cost/price parameters from disk and enter the name for the file in the field provided. If you press <F1>, a list of files with PAR extensions will be presented and can be subsequently loaded by highlighting and pressing <ENTER>.

```
Enter the name of the PARAMETER file without
extension:
 MAIZE-O
```

```
The following parameters were read from the selected
parameter file
The number of treatments is: 12.
The number of products is: 1.
If these are incorrect and you wish to return to the
beginning of ECON you may press <ESC> at this time.
If you don't press <ESC> the program will continue to
read data and will do the analysis with the parameters
from the file.
If you wish to continue with the program press <ENTER>.
```

Select option 5 (Proceed with analysis) from the ECON Options menu to analyze your data. The following message will be displayed:

```
The data file contains  96 cases.
Do you wish to use all cases ? Y/N
```

```
Print Analysis Menu
 1   Net Benefit Table
 2   Risk Table
 3   Marginal Return Table
 Q   Quit
```

Net Benefit Analysis

When you select option 1 from the analysis menu shown above, the Mean Net Benefit Table is printed. The following table is created for the MAIZE data file, using the MAIZE-0 parameters file.

MEAN NET BENEFIT TABLE FOR: MAIZE

	Treatment	Value of Output $/HCT	Cash Costs $/HCT	In-kind Costs $/HCT	Family Labor Costs $/HCT	Total Costs $/HCT	Net Benefit $/HCT
01	000N-00P	1991.25	0	0	0	00	1991.25
02	050N-00P	2826.00	400	0	50	450	2376.00
03	100N-00P	3519.00	800	0	100	900	2619.00
04	150N-00P	3612.38	1200	0	100	1300	2312.38
05	000N-25P	2199.38	250	0	50	300	1899.38
06	050N-25P	3492.00	650	0	50	700	2792.00
07	100N-25P	3955.50	1050	0	100	1150	2805.50
08	150N-25P	4351.50	1450	0	100	1550	2801.50
09	000N-50P	2126.25	500	0	50	550	1576.25
10	050N-50P	3648.38	900	0	50	950	2698.38
11	100N-50P	4263.75	1300	0	100	1400	2863.75
12	150N-50P	4646.25	1700	0	100	1800	2846.25

The table basically lists the treatment costs which you entered in the parameters file.

Treatment - Each Treatment is assigned a sequential number, followed by the treatment label (e.g., 000 N - 000 P) which you entered in the parameters file.

7.1 ECON

Value of Output - This is the mean yield of maize for each treatment times the entered price or the amount of income expected from this specific combination or treatment. If there is more than one product for the experiment (e.g., an inter-cropping experiment), the value of output for each product would be summed in this column.

Cash, In-kind (farmer-supplied material), Family Labor Costs - These column values were entered for each treatment.

Total Costs - This is the sum of the previous three cost columns.

Net Benefit - This is the result of subtracting the total costs from the value of output.

Note that the unit of measurement of both costs and benefits is the label ("$/HCT") entered in the "miscellaneous parameters" selection of the "Enter parameters" option.

Now, what does the table tell us? First, the net benefit column indicates the treatments with the highest average return over cost. In the absence of any other information, we might consider recommending to farmers the treatment which offers the highest average net benefit. In this case, treatment 11 (100 pounds nitrogen and 50 pounds phosphorus) gives the best return.

Note that if we had made a recommendation on the basis of highest average yield per treatment, we would have selected treatment 8 (150 pounds nitrogen and 25 pounds phosphorus). However, the additional costs of treatment 8 ($150) outweigh the additional return when we include economic factors in the analysis.

Marginal Return Analysis

When you highlight the marginal return table option, the Marginal Return Analysis is created. In this option, the incremental benefits and costs and the marginal rates of return of non-dominated treatments are calculated.

A 'non-dominated' treatment is one for which there is no alternative treatment with both a higher net benefit and a lower variable cost. The ECON program determines the non-dominated treatments by sorting the treatment means by net benefit value in descending order then (starting from the highest) eliminating treatments with variable costs equal to or higher than the previous treatment with the higher benefit.

After arranging undominated treatments in descending order of costs and benefits, ECON calculates the incremental benefits and costs of moving from each treatment to the next lower (more cost, more return) treatment. Then this marginal net benefit is divided by the marginal variable costs and (expressed as a percentage) displayed in the table as the marginal rate of return for the higher treatment when compared to the lower.

MARGINAL RETURN TABLE FOR: MAIZE
(Non-dominated treatments only)
Target rate of return: 40
Base:

Rank	Treatment		Net Benefit $/HCT	Total Variable Costs $/HCT	Marginal Net Benefit $/HCT	Marginal Var Cost $/HCT	Marginal Rate of Return %
1	11	100N-50P	2863.75	1400	58.25	250	23.3
2	7	100N-25P	2805.50	1150	13.50	450	3.0
3	6	050N-25P	2792.00	700	416.00	250	166.4*
4	2	050N-00P	2376.00	450	384.75	450	85.5*
5	1	000N-00P	1991.25	0	0.00	0	0.0

* These treatments meet or exceed target rate of return

Marginal net benefit of each treatment - This is calculated by subtracting from its own net benefit the net benefit of the next lower ranking treatment. For example, the marginal net benefit of treatment 2 is $2376 - $1991.25 = $384.75.

Marginal costs - This is calculated in the same manner as the marginal net benefit. The lowest ranking treatment has zero marginal net benefits and costs.

Marginal rate of return for each treatment - This is the result of dividing the marginal net benefit by its corresponding marginal cost (multiplied by 100 to express a percentage). For example, the marginal rate of return of treatment 2 is ($384.75 / $450) * 100 = 85.5%. This indicates that, compared with treatment 1, the incremental investment required from applying 50 pounds of nitrogen returns (on the average) $1.85 per dollar of variable cost. If this confuses you, think of it this way: we invest $450 in the additional fertilizer application, and get back production worth $450 + $384.75 = $834.75 more. Treatment 6 is calculated similarly, but uses treatment 2 as a base. Its rate of return for the additional investment in fertilizer is even higher: 166.4%. Treatments 7 and 11 both have much lower marginal rates of return. We have assumed that farmers will not want to make an investment unless the rate of return is at least 40%. Using that 'target rate of return', only treatments 6 and 2 have high enough returns to be selected as the recommended treatment.

Risk Analysis

	Treatment	N of Obs	Mean NB	Standard Dev	Index of Var*	Minimum NB	Ave Lowest 2
1	000N-00P	8	1991.25	1344.93	67.54	360	724.50
2	050N-00P	8	2376.00	1302.44	54.82	666	972.00
3	100N-00P	8	2619.00	1150.33	43.92	873	922.50
4	150N-00P	8	2312.38	1084.11	46.88	671	689.00
5	000N-25P	8	1899.38	1308.99	68.92	411	744.00
6	050N-25P	8	2792.00	889.93	31.87	1622	1712.00
7	100N-25P	8	2805.50	1137.11	40.53	1091	1374.50
8	150N-25P	8	2801.50	1465.39	52.31	970	1028.50
9	000N-50P	8	1576.25	1532.08	97.20	512	597.50
10	050N-50P	8	2698.38	865.78	32.09	1309	1727.50
11	100N-50P	8	2863.75	1072.16	37.44	1552	1570.00
12	150N-50P	8	2846.25	1114.26	39.15	1458	1476.00

* Index of variability = Std. Dev./Mean X 100

This table first shows the treatment numbers, labels and number of observations. The mean net benefit per treatment is then repeated from the net benefit table. The final four columns of the risk table give various measures of net benefit variability, which we will examine as an indicator of risk to the farmer from adoption of one or another treatment. We have pooled the within-site replications and consider the average yields of each treatment representative of the site-to-site, year-to-year and management variability in yields. The standard deviation of the pooled observations can be used to calculate the probability of the farmer securing a net benefit of at least some specified level. This is particularly important when deriving recommendations for subsistence farmers, for whom a "statistically possible" crop failure can mean disaster.

Assuming that yields (and hence net benefits) are normally distributed around the treatment mean, we can calculate that using treatment 1 the chances are approximately 83 out of 100 that net benefits would be greater than $650 (1991.25 - 1344.93). Using a table of probabilities, it can also be shown that chances are roughly 93 out of 100 that the net benefits from a plot with this treatment would have net benefits which at least pay for costs (are greater than zero). In general, we are looking for treatments which offer high net benefits but low standard deviations. An easy way to evaluate this trade off is to examine the index of variability. This allows direct comparisons of the variability of different treatments. We see that treatments 6 and 10 have the lowest variability in net benefit within treatment observations, while producing middle-range yields from the experiment. Treatment 11 has the third-lowest variability and also the highest net benefit of any treatment.

A further indication of variability is given by the final two columns of the table. The "Minimum NB" gives the net benefit of the worst observation within a particular treatment. All else equal, we would prefer a treatment with a "worst case" benefit higher than other treatments. We see from the table that treatment 6 has the highest "worst case" benefit, the lowest index of variability, and a reasonably good mean net benefit. Treatment 11, however, also has a high "worst case" benefit and the highest mean net benefit of any treatment. In case the lowest net benefit is a result of 'abnormal' conditions (a fluke), we should also examine the final column, which presents the average of the net benefits of the two lowest observations for each treatment. Treatment 10 has the highest figure, with treatments 6 and 11 following. All else equal, we would prefer treatments with high "Average lowest 2" figures, high minimum net benefit, low index of variability, and high mean net benefits. Treatment 6 has a particularly high marginal rate of return plus low variability. Taking into account net benefits, marginal returns and risk considerations, this treatment appears to be the best among those tested for the farmers in the recommendation domain. You may now print the table by responding "Y" to the prompt for printing a table.

7.2 ANALYSIS OF PRICE DATA

7.2 SEASONAL ANALYSIS OF PRICE DATA

The SEASON subprograms are used to conduct seasonal and storage analysis of nominal, inflated or deflated time series price data. One of the subprograms only defines variables for a price file. The other subprograms which use the price file, calculate calender, market and weighted market year averages as well as trend, moving average, seasonal, cyclical and grand seasonal indices of a historical price series. In addition, the historical gross storage margins for each market year are computed.

All the output components of this series of subprograms and the uses to which the output may be put are discussed in greater detail in the Economic Analysis Guide for the SEASON Subprograms. This Guide is available upon request from the MSTAT office (for a charge). Samples of the output from the subprograms are attached to the end of this section.

Defining a Price File

Before you use the SEASON subprograms, you will need to create a new data file, as follows:

File Name - As with any data file, you should use a file name that indicates what type of data will be stored in the file. For example, if you have Maize data beginning in 1972, you may want to name your file MAIZE-72.

File Width - The file width must be at least 127 characters. This will allow SEAPLAN to define the required 28 numeric variables, and allow you to define additional variables with the remaining 15 characters.

File Title - Enter a meaningful title; the file title is displayed in all of the tables created by the SEASON subprograms.

After defining the file name, width and title, you will use SEAPLAN to define the file variables. Then you must use SEDIT to enter data for variable no.s 1 (Quantity Marketed), 2 (Consumer Price Index) and 3 (Raw Commodity Price). You will then use SEACALC to calculate most of the remaining variable values.

If you are interested in decomposing a quantity series (as opposed to a price series) into its time series components, you can also use SEACALC to do this. You should enter your quantity series into variable no. 3 (instead of variable no. 1). Then you would enter 1's into cases for variable no.s 1 and 2; the subprogram requires that values be present in these variables.

After using SEACALC, you would print your results using SEASONAL which gives you all the seasonal analysis results. You could then also print your weights in a tabular format by choosing option 2 (Raw Price Data - Actual) from the SEATABLE Selection Menu.

Variables in a Price File

The sole function of the SEAPLAN subprogram is to define the variables you will need in a price file. The following variables are defined by SEAPLAN and used by the other SEASON subprograms:

Variable Number	Name/Description	No. of Cases*
1	Quantity Marketed	n
2	Consumer Price Index	n
3	Raw Commodity Prices	n
4	Raw Trend	n
5	Raw Moving Average	n-12
6	Raw Calendar Average	n/12
7	Raw Market Average	n/12
8	Raw Weighted Market Average	n/12
9	Raw Storage Price Changes	n
10	Deflated Commodity Prices	n
11	Deflated Trend	n
12	Deflated Moving Average	n-12
13	Deflated Calendar Average	n/12
14	Deflated Market Average	n/12
15	Deflated Weighted Market Average	n/12
16	Deflated Storage Price Changes	n
17	Inflated Commodity Price	n
18	Inflated Trend	n
19	Inflated Moving Average	n-12
20	Inflated Calendar Average	n/12
21	Inflated Market Average	n/12
22	Inflated Weighted Market Average	n/12
23	Inflated Storage Price Changes	n
24	Seasonal Index	n-12
25	Cyclical Index	n-12
26	Inflated CPI	n
27	Quantity as a Percent of Market Year Total	n
28	Time Count (used in regression & plotting)	n

* This last column shows how many cases are stored in each variable:

n - refers to the number of months.
n/12 -shows the number of years (for the yearly averages).
n-12 - refers to the moving averages and indices, since six months are lost at the beginning and the end of the series when these variables are calculated.

At the main menu, choose SEAMENU:

```
        1        Return to subprogram menu
        2        SEAPLAN
        3        SEACALC
        4        SEASONAL
        5        SEASTORE
        6        SEATABLE

WHICH OPTION WOULD YOU LIKE? (enter the number) 1
```

Tables Created by SEASONAL, SEATABLE and SEASTORE

The tables produced by the SEASON subprograms are currently designed to fit on one page using normal (uncompressed) print. This limits the number of digits which may be displayed for your price, and CPI and weight data. The display format used in most of the tables is "###.#", indicating that a maximum of three digits should be displayed to the left of the decimal point and that only one digit is displayed after the decimal point. When table values exceed this format, the value is preceded by a "%" and table data will not properly align under column headings. You may want to scale your raw data after you have entered it using the CALC subprogram; note that when you scale your data using CALC, you do not lose any decimal accuracy in the computations.

The storage tables will generally produce minus signs, which use up one of the "#" spaces. If your prices are in a xxx.x format and fall by more than 100 relative to the harvest month in any given post-harvest month (e.g., 500.0 - 600.0 = -100.0), the table data will not properly align under column headings. To avoid this problem you can in this case simply scale (i.e. divide) your price data by 10 using the CALC subprogram.

The tables printed by the SEASON subprograms are designed so that they can readily be adapted for publications, extension meetings (e.g., as overheads), etc. For this reason, the variable names appear as a part of the title of each table. The titles are composed of two parts: the first is the variable name (e.g., "Raw Commodity Price"), followed by a "for", which is in turn followed by the name of the file (e.g., "Michigan Navy Bean Data 1967-1982").

We recommend that you change the variable names for each of the three price tables (raw, inflated and deflated - variable no.s 3 through 23) so that the titles in the tables are tailored to your particular commodity and country. Use SEDIT-NEWTXT to handle file title and variable description changes: you may use a maximum of 64 characters for each title or variable description.

7.2.1 SEAPLAN

7.2.1 SEAPLAN

Purpose: Defines variables for a price file.

Before starting, you need:

1. A newly created data file with a file width of at least 127 characters.

Results: Variables are defined for your price file.

Sample Data File: SEATST. To define this data file, select FILES from the subprogram menu. When the FILES menu is displayed, enter *2<ENTER>* (2 - Create a new data file). Define the file name as SEATST and define the maximum number of characters for this file as 127.

When you select SEAPLAN from the subprogram menu, the subprogram first checks that the data file width is at least 127. After the subprogram defines the price file variables, the subprogram menu is displayed. SEAPLAN does not display any prompts.

After SEAPLAN has defined the file variables, notice that your file information has changed to the following:

```
File Width: 127  Used: 112 Unused:  15
Defined Variables:  28      Data Cases:   0
```

You may now use SEDIT to enter data for variable no.s 1, 2 and 3. If you want to change the variable name/descriptions, you must use SEDIT-NEWTXT. You may also define additional variables for the file.

SEAPLAN Example

When you select SEAPLAN from the subprogram menu, the subprogram header will be displayed. SEAPLAN defines the price file variables automatically and displays each variable number and description as the variable is defined. When the variables are defined, the subprogram menu is displayed.

7.2.2 SEACALC

Purpose: Calculates most of the remaining variables in a price file (variable no.s 4 through 28).

Before starting, you need:

1. An active data file, with variables defined using SEAPLAN. Variable no.s 1, 2 and 3 must contain data.
2. If only a portion of your file will be used in the calculations, you should know which cases you want used. Know month and year numbers for your analyses.

Results: Data are calculated for most of the remaining variables, according to the cases you specify (all or a range).

Sample Data File: NAVYB-67.

SEACALC calculates most of the remaining values that are entered in variable no.s 4 through 28. The subprogram first asks initial questions about data cases and then asks for the month of your first observation, the month which starts your market year and the year of your first observation. When you respond to these prompts, you may want to make a note of your responses since you will be asked the same questions again when you run the remaining three SEASON subprograms.

When the program has completed its calculations, the subprogram menu is displayed. All the variables calculated by SEACALC are now stored in your price file.

SEACALC Example

In the following example, SEACALC is used to calculate most of the remaining variables for the file NAVYB-67.

The SEACALC subprogram will analyze one major variable, (Price) with supporting information being provided by two additional variables, (Weight and CPI). These variables may have been previously manipulated in any fashion required, as long as there are the same number observations for each variable. The SEACALC program does not limit the number of observations that can be analyzed. The actual limit depends on the capacity of your disk drive.

7.2.2 SEACALC

```
WILL YOU USE ALL CASES OF DATA? (Y or N) N<ENTER>
```

In this example, only the seven most recent years in a 17-year long series will be analyzed: therefore, only a range of data cases will be used. Respond "yes" if you want to analyze your entire series.

```
WHAT IS THE FIRST CASE, LAST CASE? (e.g. 5,17)
121,204<ENTER>
```

```
WHAT MONTH IS YOUR FIRST OBSERVATION ? 1<ENTER>
```

Enter the number of a month. If, as in the case of the data used for this example, your first observation (case or month) corresponds to January, enter a "1". This information is important since it allows the program to calculate the calendar year averages for your data.

```
WHAT MONTH DOES YOUR MARKET YEAR START ? 9<ENTER>
```

Enter a month number. If, as in this example, the market year runs from September to August, enter a "9", which corresponds to the month of September. The answer to this prompt is important since it allows the program to calculate the market (and weighted market) year averages, as well as the historical storage tables.

```
WHAT YEAR IS THE FIRST OBSERVATION (YYYY) ? 1977<ENTER>
```

Enter the year in which your series begins (use the 4-digit year number). In this example, the series started in 1977.

```
Your data starts JAN, 1977 and represents a SEP, 1977
to AUG, 1978 market year
IS THE ABOVE CORRECT ? Y<ENTER>
```

If you respond "no", SEACALC will again ask you for this information. If you respond "yes", the program will begin to calculate some of the values for variables 4 through 28. This calculation will take some time, depending on the length of your data series. You will see the following on the screen as the program calculates:

```
Finding and Calculating the Last Year's Average CPI
 130  140  150  160  170  180  190  200
Calculating the Overall Regression for Trend Variables
 130  140  150  160  170  180  190  200
Calculating the Variables for All Observations
 130  140  150  160  170  180  190  200
Computing Observation  1 2 3 4 5 6 7 8 9 10 20 30 40 50
60 70 80 90 100 110 120 130 140 150 160 170 180 190 200
```

When the program has completed its calculations, the subprogram menu is displayed.

7.2.3 SEASONAL

7.2.3 SEASONAL

Purpose: Produces tables of regression results and various seasonal indices using data in a price file.

Before starting, you need/should know:

1. An active data file; this must be a price file with variables defined by SEAPLAN and certain variable values calculated by SEACALC.
2. If only a portion of your file will be used in the calculations, you should know which cases you want used. Know month and year numbers for your analyses.

Results: The requested SEASONAL tables are created for the specified time periods. Depending on which print/list options you select, these tables may be viewed on the screen, printed or stored in a disk file.

Sample Data File: NAVYB-67.

SEASONAL uses data from a price file to calculate values and construct price tables. You may create any or all of the following tables:

- Regression results.
- Various indices.
- Grand seasonal index.
- Seasonal index.

The subprogram asks initial questions about data cases, the starting month of observation, the starting month of your market and the year of your first observation. Then a menu of tables is displayed. To make the subprogram print only one table on each page of paper, select one table from this menu. After the requested table is printed, the menu of tables is again displayed.

The tables printed by this subprogram are discussed in the Economic Analysis Guide for the SEASON Subprograms.

SEASONAL Example

In the following example, all of the possible SEASONAL tables are created for NAVYB-67. Notice that the responses to prompts are the same as for the SEACALC example. To produce tables which reflect data for the same period of time, you should respond the same to the initial prompts used by four of the SEASON subprograms (SEAPLAN is the exception).

The tables created in this example are shown along with the other SEASON series tables at the end of this section.

```
WILL YOU USE ALL CASES OF DATA? (Y or N) N<ENTER>
```

```
WHAT IS THE FIRST CASE, LAST CASE? (e.g. 5,17)
121,204<ENTER>
```

```
WHAT MONTH IS YOUR FIRST OBSERVATION ? 1<ENTER>
```

Enter the number of the month which begins your observations. In this example, the first observation is in January.

```
WHAT MONTH DOES YOUR MARKET YEAR START ? 9<ENTER>
```

Enter the number of the month which starts your market year. In this example, the market year starts in September.

```
WHAT YEAR IS YOUR FIRST OBSERVATION (YYYY) ?
1977<ENTER>
```

Enter a year number. In this example, the series starts in 1977.

```
Your data starts JAN, 1977 and represents a SEP,1977 to
AUG, 1978 market year.

IS THE ABOVE CORRECT ? <ENTER>
```

If you respond "no", you are again prompted for the dates you want analyzed. When you respond "yes", the following menu is displayed.

7.2.3 SEASONAL

```
Seasonal Analysis Output Selection Menu
        1  Quit this subprogram
        2  Regression Results
        3  Various Indices
        4  Grand Seasonal Index
        5  Grand Seasonal Graph
        6  Seasonal Index

        (* to select all tables)

 YOUR CHOICE: *<ENTER>
```

Here you will enter the number of the desired table(s), or "*" if you want all tables. If you want tables 2 and 4, enter **2,4<ENTER>**. You may want to refer to the Guide or to the sample output in case you are unsure of which table to select.

After you have selected your tables the program will begin to carry out further computations to set up the tables. This will also take a while, depending on the length of your time series and the number of tables selected. As the subprogram computes and creates the requested tables, messages are displayed on the screen indicating that the program is working.

```
Preparing Tables


Preparing Regression
 130  140  150  160  170  180  190  200
Preparing Indices
 130  140  150  160  170  180  190  200
Preparing Grand Seasonal Index
 130  140  150  160  170  180  190  200
Preparing Grand Seasonal Bargraph
 130  140  150  160  170  180  190  200
Preparing Seasonal Analysis
 130  140  150  160  170  180  190  200
```

When the requested tables are created, the print/list menu of options is displayed. You may view the tables on the screen, print them or store them in a disk file. When you select an option that returns you to the subprogram, the Seasonal Analysis Output Selection Menu is again displayed. You may create another table or use option 1 to exit from the SEASONAL subprogram; in this case, the subprogram menu is displayed.

7.2.4 SEASTORE

7.2.4 SEASTORE

Purpose: Produces tables relevant to storage analysis using data in a price file.

Before starting, you need/should know:

1. An active data file; this must be a price file with variables defined by SEAPLAN and certain variable values calculated by SEACALC.
2. If only a portion of your file will be used in the calculations, you should know which cases you want used. Know month and year numbers for your analyses.

Results: The requested SEASTORE tables are created for the specified time periods. Depending on which print/list options you select, these tables may be viewed on the screen, printed or stored in a disk file.

Sample Data File: NAVYB-67.

SEASTORE uses data from a price file to calculate values and construct price tables. You may create any or all of the following tables:

- Changes in raw price from storage.
- Changes in inflated price from storage.
- Changes in deflated price from storage.
- Effective interest return from storage.

The subprogram asks initial questions about data cases, the starting month of observation, the starting month of your market and the year of your first observation. Then a menu of tables is displayed. To make the subprogram print only one table on each page of paper, select one table from this menu. After the requested table is printed, the menu of tables is again displayed.

The tables printed by this subprogram are discussed in the Economic Analysis Guide for the SEASON Subprograms.

SEASTORE Example

In the following example, all of the possible SEASTORE tables are created for NAVYB-67. Notice that the responses to prompts are the same as for the SEACALC example. To produce tables which reflect data for the same period of time, you should respond the same to opening prompts used by four of the SEASON subprograms (SEAPLAN is the exception).

The tables created in this example are shown at the end of this section along with the other SEASON series tables.

```
WILL YOU USE ALL CASES OF DATA? (Y or N) N<ENTER>
```

```
WHAT IS THE FIRST CASE, LAST CASE? (e.g. 5,17)
121,204<ENTER>
```

```
WHAT MONTH IS YOUR FIRST OBSERVATION ? 1<ENTER>
```

Enter the number of the month which begins your observations. In this example, the first observation is in January.

```
WHAT MONTH DOES YOUR MARKET YEAR START ? 9<ENTER>
```

Enter the number of the month which starts your market year. In this example, the market year starts in September.

```
WHAT YEAR IS YOUR FIRST OBSERVATION (YYYY) ?
1977<ENTER>
```

Enter a year number. In this example, the series starts in 1977.

```
Your data starts JAN, 1977 and represents a SEP,1977 to
AUG, 1978 market year.

IS THE ABOVE CORRECT ? <ENTER>
```

If you respond "no", you are again prompted for the dates you want analyzed. When you respond "yes", the following menu is displayed.

7.2.4 SEASTORE

```
Storage Analysis Output Selection Menu

        1   Quit this subprogram

        2   Changes in Raw Price from Storage
        3   Changes in Inflated Price from Storage
        4   Changes in Deflated Price from Storage
        5   Effective Interest Return from Storage

        (* to select all tables)

 YOUR CHOICE: *<ENTER>
```

You may, for example enter **2,5<ENTER>** for tables 2 and 5. Once you have selected the desired tables, the program begins to compute and construct the requested tables. As the program works, the following messages are displayed on the screen.

```
Preparing Tables

Raw Storage Price Change Table
 130 140 150 160 170 180 190 200 130 140 150 160 170 180 190
200
Inflated Storage Price Change Table
 130 140 150 160 170 180 190 200 130 140 150 160 170 180 190
200
Deflated Storage Price Change Table
 130 140 150 160 170 180 190 200 130 140 150 160 170 180 190
200
Interest Table
 130 140 150 160 170 180 190 200 130 140 150 160 170 180 190
200
```

When the requested tables are created, the print/list menu of options is displayed. You may view the tables on the screen, print them or store them in a disk file. When you select an option that returns you to the subprogram, the menu of tables is again displayed. You may create another table or use option 1 to exit from the SEASTORE subprogram; in this case, the subprogram menu is displayed.

7.2.5 SEATABLE

Purpose: Produces tables of price changes using data in a price file.

Before starting, you need:

1. An active data file; this must be a price file with variables defined by SEAPLAN and certain variable values calculated by SEACALC
2. If only a portion of your file will be used in the calculations, you should know which cases you want used. Know month and year numbers for your analyses.

Results: The requested SEASONAL tables are created for the specified time periods. Depending on which print/list options you select, these tables may be viewed on the screen, printed or stored in a disk file.

Sample Data File: NAVYB-67.

SEATABLE uses data from a price file to calculate values and construct price tables. You may create any or all of the following tables:

- Raw Price Data - actual, inflated and/or deflated price changes.
- Consumer Price Index - actual and/or inflated prices changes.
- Quantities Marketed - actual price changes.
- Weights as Percent of Quantities Marketed - actual price changes.

The subprogram asks initial questions about data cases, the starting month of observation, the starting month of your market and the year of your first observation. Then a menu of tables is displayed. To make the subprogram print only one table on each page of paper, select one table from this menu. After the requested table is printed, the menu of tables is again displayed.

The tables printed by this subprogram are discussed in the Economic Analysis Guide for the SEASON Subprograms.

7.2.5 SEATABLE

SEATABLE Example

In the following example, all of the possible SEATABLE tables are created for NAVYB-67. Notice that the responses to prompts are the same as for the SEACALC example. To produce tables which reflect data for the same period of time, you should respond the same way to initial prompts used by four of the SEASON subprograms (SEAPLAN is the exception).

The tables created in this example are shown along with the other SEASON series tables at the end of this section.

```
WILL YOU USE ALL CASES OF DATA? (Y or N) N<ENTER>
```

```
WHAT IS THE FIRST CASE, LAST CASE? (e.g. 5,17)
121,204<ENTER>
```

```
WHAT MONTH IS YOUR FIRST OBSERVATION ? 1<ENTER>
```

Enter the number of the month which begins your observations. In this example, the first observation is in January.

```
WHAT MONTH DOES YOUR MARKET YEAR START ? 9<ENTER>
```

Enter the number of the month which starts your market year. In this example, the market year starts in September.

```
WHAT YEAR IS YOUR FIRST OBSERVATION (YYYY) ?
1977<ENTER>
```

Enter a year number. In this example, the series starts in 1977.

```
Your data starts JAN, 1977 and represents a SEP,1977 to
AUG, 1978 market year.

IS THE ABOVE CORRECT ? <ENTER>
```

If you respond "no", you are again prompted for the dates you want analyzed. When you respond "yes", the following menu is displayed.

```
Tabular Analysis Output Selection Menu
                               Actual  Inflated  Deflated
Quit this subprogram              1
Monthly Tables
   Raw Price Data                 2        3         4
   Consumer Price Index           5        6
   Quantities Marketed            7
   Weights as Percent
   of Quantities Marketed         8
   (* to select all tables)

YOUR CHOICE: *<ENTER>
```

You may, for example, enter **5,7<ENTER>** for tables 5 and 7. Once you have selected the desired tables, the program begins to compute and construct the requested tables; it displays the following messages on the screen:

```
Preparing Tables


Raw Price Tables
 130 140 150 160 170 180 190 200 130 140 150 160 170 180 190
200
Inflated Price Tables
 130 140 150 160 170 180 190 200 130 140 150 160 170 180 190
200
Deflated Price Tables
 130 140 150 160 170 180 190 200 130 140 150 160 170 180 190
200
Consumer Price Index Tables
 130 140 150 160 170 180 190 200 130 140 150 160 170 180 190
200
Inflated Consumer Price Index Tables
 130 140 150 160 170 180 190 200 130 140 150 160 170 180 190
200
Quantity Marketed Tables
 130 140 150 160 170 180 190 200 130 140 150 160 170 180 190
200
Percent Weights Table
 130  140  150  160  170  180  190  200
```

7.2.5 SEATABLE

When the tables are created, the print/list menu of options is displayed. You may view the tables on the screen, print them or store them in a disk file. When you select an option that returns you to the subprogram, the menu of tables is again displayed. You may create another table or select option 1 to exit from the subprogram; in this case, the subprogram menu is displayed.

TABLES CREATED BY THE SEASON SUBPROGRAMS

Table 1: Regression Results -- option 2 from menu

REGRESSION RESULTS FOR MICH NAVY BEAN DATA 1967 - 1982

NO OF OBS = 84	STD ERR OF EST = 6.431
MEAN OF Y = 19.518	STD DEV OF Y = 6.658
A = 16.272	T = 2.639
B = 0.076	STD ERR OF B = 0.029

7.2.6 SEASON TABLES

Table 2: Various Indices -- option 3

MICHIGAN NAVY BEAN DATA 1967 - 1982

OBSERV. NUMBER	ACTUAL DATA	TREND	MOVING AVERAGE	SEASONAL INDEX	CYCLICAL INDEX
1	13.50	16.35			
2	13.63	16.42			
3	15.20	16.50			
4	17.25	16.58			
5	17.12	16.65			
6	17.10	16.73			
7	15.25	16.81	16.14	94.49	96.03
8	12.75	16.88	16.58	76.89	98.21
9	12.37	16.96	17.02	72.70	100.33
10	17.75	17.04	17.31	102.54	101.61
11	19.50	17.11	17.49	111.48	102.22
12	20.00	17.19	17.61	113.59	102.44
13	18.00	17.26	17.67	101.85	102.37
14	19.75	17.34	17.75	111.26	102.36
15	19.50	17.42	17.84	109.32	102.41
16	20.00	17.49	17.70	113.01	101.16
17	18.75	17.57	17.27	108.58	98.28
18	18.25	17.65	16.72	109.15	94.75
19	15.67	17.72	16.21	96.68	91.45
20	14.20	17.80	15.68	90.55	88.10
21	13.00	17.88	15.12	85.96	84.61
22	13.75	17.95	14.64	93.95	81.52
23	13.20	18.03	14.25	92.61	79.06
24	13.13	18.11	14.07	93.33	77.71
25	12.60	18.18	14.24	88.46	78.34
26	12.50	18.26	14.67	85.19	80.36
27	13.38	18.33	15.06	88.84	82.15
28	14.38	18.41	15.40	93.39	83.64
29	15.20	18.49	15.74	96.56	85.15
30	17.38	18.56	16.11	107.86	86.80
31	20.75	18.64	16.51	125.68	88.57
32	19.40	18.72	17.00	114.11	90.84
33	17.13	18.79	17.52	97.76	93.24
34	17.70	18.87	17.87	99.08	94.68

(Table 2 cont'd)

35	17.50	18.95	18.05	96.96	95.27
36	17.75	19.02	18.13	97.88	95.34
37	17.50	19.10	18.13	96.52	94.93
38	19.38	19.17	18.24	106.25	95.12
39	19.00	19.25	18.59	102.22	96.55
40	17.00	19.33	19.10	89.01	98.82
41	17.00	19.40	19.68	86.39	101.41
42	17.63	19.48	20.20	87.28	103.70
43	20.40	19.56	20.69	98.59	105.81
44	22.37	19.63	21.23	105.38	108.13
45	22.50	19.71	21.92	102.64	111.22
46	24.60	19.79	22.89	107.48	115.68
47	24.50	19.86	24.31	100.77	122.41
48	23.30	19.94	26.12	89.21	131.00
49	23.75	20.01	27.70	85.73	138.41
50	26.00	20.09	28.34	91.76	141.04
51	29.00	20.17	28.41	102.08	140.86
52	30.20	20.24	28.78	104.92	142.18
53	38.00	20.32	29.36	129.42	144.50
54	40.00	20.40	29.98	133.44	146.96
55	36.00	20.47	30.47	118.14	148.85
56	22.00	20.55	30.78	71.49	149.76
57	24.60	20.63	30.90	79.61	149.81
58	31.50	20.70	30.68	102.66	148.21
59	31.50	20.78	29.93	105.26	144.02
60	31.00	20.85	28.55	108.58	136.90
61	28.00	20.93	26.97	103.83	128.84
62	29.00	21.01	25.99	111.59	123.71
63	29.00	21.08	25.30	114.62	120.00
64	25.00	21.16	23.98	104.27	113.31
65	25.00	21.24	22.21	112.55	104.60
66	20.00	21.31	20.42	97.96	95.79
67	18.00	21.39	18.75	96.12	79.97
69	13.60	21.54	15.55	87.48	72.17
70	10.75	21.62	14.23	75.55	65.82
71	9.90	21.69	13.27	74.63	61.15
72	9.50	21.77	12.57	75.57	57.74

7.2.6 SEASON TABLES

(Table 2 cont'd)

73	9.50	21.85	12.30	77.25	56.29
74	9.50	21.92	12.65	75.07	57.72
75	9.62	22.00	13.60	70.74	61.82
76	12.75	22.08	14.75	86.43	66.82
77	14.13	22.15	15.93	88.68	71.92
78	14.20	22.23	17.10	83.02	76.94
79	17.25	22.31			
80	25.80	22.38			
81	27.00	22.46			
82	25.00	22.53			
83	24.00	22.61			
84	23.50	22.69			

Table 3: Grand Seasonal Index -- option 4

GRAND SEASONAL INDEX FOR MICH NAVY BEAN DATA 1967-1982

MONTH	AVERAGE SEASONAL INDEX	STANDARD ERROR	GRAND SEASONAL INDEX	CORRECTED STANDARD ERROR	GSI + CSE	GSI + CSE
JAN	92.27	10.26	94.83	10.84	105.67	83.99
FEB	96.85	15.16	99.54	16.01	115.55	83.52
MAR	97.97	15.91	100.68	16.81	117.49	83.88
APR	98.50	10.46	101.23	11.05	112.28	90.19
MAY	103.70	16.37	106.57	17.29	123.86	89.28
JUN	103.12	18.22	105.98	19.25	125.22	86.73
JUL	104.93	13.43	107.84	14.19	122.02	93.65
AUG	92.42	16.34	94.98	17.26	112.24	77.73
SEP	87.69	11.12	90.12	11.74	101.86	78.38
OCT	96.88	11.37	99.56	12.01	111.57	87.55
NOV	96.95	12.74	99.64	13.46	113.10	86.18
DEC	96.36	13.72	99.03	14.49	113.52	84.54

7.2.6 SEASON TABLES

Table 4: Grand Seasonal Bar Graph -- option 5

```
% 110.6+       GRAPH OF THE GRAND SEASONAL INDEX
       |       FOR MICHIGAN NAVY BEAN DATA 1967 - 1982
o      |
f 106.5+                        XXX        XXX
       |                        XXX  XXX   XXX
A      |                        XXX  XXX   XXX
n 102.4+                        XXX  XXX   XXX
n      |                   XXX  XXX  XXX   XXX
u 100% ============XXX===XXX===XXX===XXX===XXX=========================
a  98.3+     XXX  XXX  XXX  XXX  XXX  XXX            XXX  XXX  XXX
l      |     XXX  XXX  XXX  XXX  XXX  XXX            XXX  XXX  XXX
       |     XXX  XXX  XXX  XXX  XXX  XXX            XXX  XXX  XXX
A  94.2+ XXX XXX  XXX  XXX  XXX  XXX  XXX  XXX       XXX  XXX  XXX
v      | XXX XXX  XXX  XXX  XXX  XXX  XXX  XXX       XXX  XXX  XXX
e      | XXX XXX  XXX  XXX  XXX  XXX  XXX  XXX       XXX  XXX  XXX
r  90.1+ XXX XXX  XXX  XXX  XXX  XXX  XXX  XXX  XXX  XXX  XXX  XXX
a GSI. | 94.8 99.5 100.7 101.2 106.6 106.0 107.8 95.0 90.1 99.6 99.6 99.0
g CSE. | 10.8 16.0 16.8 11.0 17.3 19.2 14.2 17.3 11.7 12.0 13.5 14.5
e 86.0+--+-----+-----+-----+-----+-----+-----+----+----+----+-----+-----+---
  Months J   F     M     A     M     J     J    A    S    O     N     D
```

GSI. IS THE GRAND SEASONAL INDEX

CSE. REPRESENTS THE CORRECTED STANDARD ERROR

Notes for Table 4

THE 100% LINE IN THE BARCHART REPRESENTS THE AVERAGE OF 19.488 CURRENCY UNITS OVER THE 7 YEAR PERIOD OF ANALYSIS OF COMMODITY PRICE DATA.

THE MONTH INDEX VALUE INDICATES BY HOW MANY PERCENTAGE POINTS EACH MONTH'SVALUE LIES ABOVE OR BELOW THE ANNUAL AVERAGE.

THE STANDARD ERROR OF 10.8 FOR JANUARY INDICATES THAT THE JANUARY VALUE WILL LIE WITHIN PLUS OR MINUS 10.8 PERCENTAGE POINTS OF ITS MEAN IN 7 OUT OF 10 YEARS

Table 5: Seasonal Index -- option 6

SEASONAL INDEX FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1977							94.5	76.9	72.7	102.5	111.5	113.6
1978	101.8	111.3	109.3	113.0	108.6	109.2	96.7	90.6	86.0	94.0	92.6	93.3
1979	88.5	85.2	88.8	93.4	96.6	107.9	125.7	114.1	97.8	99.1	97.0	97.9
1980	96.5	106.3	102.2	89.0	86.4	87.3	98.6	105.4	102.6	107.5	100.8	89.2
1981	85.7	91.8	102.1	104.9	129.4	133.4	118.1	71.5	79.6	102.7	105.3	108.6
1982	103.8	111.6	114.6	104.3	112.5	98.0	96.0	96.1	87.5	75.6	74.6	75.6
1983	77.2	75.1	70.7	86.4	88.7	83.0						
Mean	92.3	96.9	98.0	98.5	103.7	103.1	104.9	92.4	87.7	96.9	97.0	96.4
Trend	-2.5	-3.3	-3.3	-2.4	-0.2	-3.3	1.3	0.9	1.7	-2.9	-4.1	-4.4
t	-1.0	-0.9	-0.8	-1.0	-0.1	-0.7	0.4	0.2	0.6	-1.1	-1.5	-1.5

7.2.6 SEASON TABLES

Table 6: Changes in Raw Price from Storage -- option 2

RAW STORAGE PRICE CHANGES FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG
1977	0.0	5.4	7.1	7.6	5.6	7.4	7.1	7.6	6.4	5.9	3.3	1.8
1978	0.0	0.8	0.2	0.1	-0.4	-0.5	0.4	1.4	2.2	4.4	7.8	6.4
1979	0.0	0.6	0.4	0.6	0.4	2.3	1.9	-0.1	-0.1	0.5	3.3	5.2
1980	0.0	2.1	2.0	0.8	1.3	3.5	6.5	7.7	15.5	17.5	13.5	-0.5
1981	0.0	6.9	6.9	6.4	3.4	4.4	4.4	0.4	0.4	-4.6	-6.6	-8.1
1982	0.0	-2.9	-3.7	-4.1	-4.1	-4.1	-4.0	-0.9	0.5	0.6	3.6	12.2
1983	0.0	-2.0	-3.0	-3.5								

Statistics for above Table

Mean	0.0	1.6	1.4	1.1	1.0	2.2	2.7	2.7	4.1	4.0	4.1	2.8
S.D.	0.0	3.6	4.3	4.5	3.3	4.0	4.2	3.9	6.1	7.5	6.6	6.9
Low	0.0	-2.9	-3.7	-4.1	-4.1	-4.1	-4.0	-0.9	-0.1	-4.6	-6.6	-8.1
High	0.0	6.9	7.1	7.6	5.6	7.4	7.1	7.7	15.5	17.5	13.5	12.2
Rises	0.0	5.0	5.0	5.0	4.0	4.0	5.0	4.0	5.0	5.0	5.0	4.0
Falls	0.0	2.0	2.0	2.0	2.0	2.0	1.0	2.0	1.0	1.0	1.0	2.0

Statistics for Last 5 Years of Table 6

Mean	0.0	0.9	0.5	0.1	0.1	1.1	1.8	1.7	3.7	3.7	4.3	3.0
S.D.	0.0	4.0	4.9	5.1	3.9	4.7	4.8	4.5	6.8	8.4	7.4	7.7
Low	0.0	-2.9	-3.7	-4.1	-4.1	-4.1	-4.0	-0.9	-0.1	-4.6	-6.6	-8.1
High	0.0	6.9	6.9	6.4	3.4	4.4	6.5	7.7	15.5	17.5	13.5	12.2
Rises	0.0	4.0	4.0	4.0	3.0	3.0	4.0	3.0	4.0	4.0	4.0	3.0
Falls	0.0	2.0	2.0	2.0	2.0	2.0	1.0	2.0	1.0	1.0	1.0	2.0

Table 7: Changes in Inflated Price from Storage -- option 3

(Past data values inflated using 1983 as the base CPI of 100.[a]

INFLATED STORAGE PRICE CHANGES FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG
1977	0.0	8.6	11.3	12.0	8.6	11.2	10.6	11.1	8.9	7.8	3.7	1.4
1978	0.0	1.0	0.0	-0.2	-1.1	-1.5	-0.4	0.8	1.7	4.5	8.8	6.7
1979	0.0	0.6	0.1	0.2	-0.5	1.6	0.8	-2.0	-2.2	-1.6	1.7	3.9
1980	0.0	2.2	1.9	0.2	0.5	2.8	6.0	7.1	15.5	17.3	12.5	-2.9
1981	0.0	7.3	7.2	6.6	3.3	4.3	4.3	-0.0	-0.3	-5.7	-7.9	-9.5
1982	0.0	-2.9	-3.8	-4.1	-4.2	-4.2	-4.1	-1.0	0.4	0.4	3.4	11.8
1983	0.0	-2.0	-3.1	-3.6								

Statistics for above Table

	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG
Mean	0.0	2.1	2.0	1.6	1.1	2.4	2.9	2.7	4.0	3.8	3.7	1.9
S.D.	0.0	4.4	5.5	5.8	4.4	5.3	5.2	5.2	6.8	8.1	7.0	7.5
Low	0.0	-2.9	-3.8	-4.1	-4.2	-4.2	-4.1	-2.0	-2.2	-5.7	-7.9	-9.5
High	0.0	8.6	11.3	12.0	8.6	11.2	10.6	11.1	15.5	17.3	12.5	11.8
Rises	0.0	5.0	5.0	4.0	3.0	4.0	4.0	3.0	4.0	4.0	5.0	4.0
Falls	0.0	2.0	2.0	3.0	3.0	2.0	2.0	3.0	2.0	2.0	1.0	2.0

Statistics for Last 5 Years of Table 7

	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG
Mean	0.0	1.0	0.4	-0.2	-0.4	0.6	1.3	1.0	3.0	3.0	3.7	2.0
S.D.	0.0	5.0	6.3	6.7	5.3	6.3	6.1	6.2	7.7	9.2	7.8	8.3
Low	0.0	-2.9	-3.8	-4.1	-4.2	-4.2	-4.1	-2.0	-2.2	-5.7	-7.9	-9.5
High	0.0	7.3	7.2	6.6	3.3	4.3	6.0	7.1	15.5	17.3	12.5	11.8
Rises	0.0	4.0	4.0	3.0	2.0	3.0	3.0	2.0	3.0	3.0	4.0	3.0
Falls	0.0	2.0	2.0	3.0	3.0	2.0	2.0	3.0	2.0	2.0	1.0	2.0

Table 8: Changes in Deflated Price from Storage -- option 4

DEFLATED STORAGE PRICE CHANGES FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG
1977	0.0	2.9	3.8	4.0	2.9	3.8	3.6	3.7	3.0	2.6	1.2	0.5
1978	0.0	0.3	0.0	-0.1	-0.4	-0.5	-0.1	0.3	0.6	1.5	3.0	2.3
1979	0.0	0.2	0.0	0.1	-0.2	0.5	0.3	-0.7	-0.7	-0.5	0.6	1.3
1980	0.0	0.7	0.6	0.1	0.2	0.9	2.0	2.4	5.2	5.8	4.2	-1.0
1981	0.0	2.4	2.4	2.2	1.1	1.4	1.4	-0.0	-0.1	-1.9	-2.6	-3.2
1982	0.0	-1.0	-1.3	-1.4	-1.4	-1.4	-1.4	-0.3	0.1	0.1	1.1	4.0
1983	0.0	-0.7	-1.0	-1.2								

Statistics for above Table

	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG
Mean	0.0	0.7	0.7	0.5	0.4	0.8	1.0	0.9	1.3	1.3	1.2	0.6
S.D.	0.0	1.5	1.8	1.9	1.5	1.8	1.7	1.7	2.3	2.7	2.3	2.5
Low	0.0	-1.0	-1.3	-1.4	-1.4	-1.4	-1.4	-0.7	-0.7	-1.9	-2.6	-3.2
High	0.0	2.9	3.8	4.0	2.9	3.8	3.6	3.7	5.2	5.8	4.2	4.0
Rises	0.0	5.0	5.0	4.0	3.0	4.0	4.0	3.0	4.0	4.0	5.0	4.0
Falls	0.0	2.0	2.0	3.0	3.0	2.0	2.0	3.0	2.0	2.0	1.0	2.0

Statistics for Last 5 Years of Table 8

	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG
Mean	0.0	0.3	0.1	-0.1	-0.1	0.2	0.4	0.3	1.0	1.0	1.2	0.7
S.D.	0.0	1.7	2.1	2.2	1.8	2.1	2.0	2.1	2.6	3.1	2.6	2.8
Low	0.0	-1.0	-1.3	-1.4	-1.4	-1.4	-1.4	-0.7	-0.7	-1.9	-2.6	-3.2
High	0.0	2.4	2.4	2.2	1.1	1.4	2.0	2.4	5.2	5.8	4.2	4.0
Rises	0.0	4.0	4.0	3.0	2.0	3.0	3.0	2.0	3.0	3.0	4.0	3.0
Falls	0.0	2.0	2.0	3.0	3.0	2.0	2.0	3.0	2.0	2.0	1.0	2.0

Table 9: Effective Interest Return from Storage -- option 5

INTEREST RATE OF RETURN FROM STORAGE FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG
1977	0.0	43.5	57.6	61.7	45.5	59.7	57.6	61.7	51.6	47.5	26.7	14.8
1978	0.0	5.8	1.5	1.0	-3.1	-3.8	2.9	10.6	16.9	33.7	59.6	49.2
1979	0.0	3.3	2.2	3.6	2.2	13.1	10.9	-0.8	-0.8	2.9	19.1	30.6
1980	0.0	9.3	8.9	3.6	5.6	15.6	28.9	34.2	68.9	77.8	60.0	-2.2
1981	0.0	28.0	28.0	26.0	13.8	17.9	17.9	1.6	1.6	-18.7	-26.8	-32.9
1982	0.0	-21.0	-27.2	-30.1	-30.1	-30.1	-29.3	-6.3	3.9	4.4	26.8	89.7
1983	0.0	-7.4	-11.1	-13.0								

Statistics for above Table

Mean	0.0	8.8	8.6	7.5	5.6	12.0	14.8	16.9	23.7	24.6	27.6	24.9
S.D.	0.0	21.5	27.5	29.4	24.6	29.5	28.8	26.2	29.5	35.2	32.0	42.4
Low	0.0	-21.0	-27.2	-30.1	-30.1	-30.1	-29.3	-6.3	-0.8	-18.7	-26.8	-32.9
High	0.0	43.5	57.6	61.7	45.5	59.7	57.6	61.7	68.9	77.8	60.0	89.7
Rises	0.0	5.0	5.0	5.0	4.0	4.0	5.0	4.0	5.0	5.0	5.0	4.0
Falls	0.0	2.0	2.0	2.0	2.0	2.0	1.0	2.0	1.0	1.0	1.0	2.0

Statistics for Last 5 Years of Table 9

Mean	0.0	3.0	0.4	-1.5	-2.3	2.5	6.3	7.9	18.1	20.0	27.7	26.9
S.D.	0.0	24.5	31.7	33.9	29.2	35.0	33.8	31.3	33.7	39.7	35.7	47.5
Low	0.0	-21.0	-27.2	-30.1	-30.1	-30.1	-29.3	-6.3	-0.8	-18.7	-26.8	-32.9
High	0.0	28.0	28.0	26.0	13.8	17.9	28.9	34.2	68.9	77.8	60.0	89.7
Rises	0.0	4.0	4.0	4.0	3.0	3.0	4.0	3.0	4.0	4.0	4.0	3.0
Falls	0.0	2.0	2.0	2.0	2.0	2.0	1.0	2.0	1.0	1.0	1.0	2.0

7.2.6 SEASON TABLES

Table 10: Raw Price Data (Monthly Tables) -- option 2 - actual

RAW COMMODITY PRICE FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1977	13.5	13.6	15.2	17.3	17.1	17.1	15.3	12.8	12.4	17.8	19.5	20.0
1978	18.0	19.8	19.5	20.0	18.8	18.3	15.7	14.2	13.0	13.8	13.2	13.1
1979	12.6	12.5	13.4	14.4	15.2	17.4	20.8	19.4	17.1	17.7	17.5	17.8
1980	17.5	19.4	19.0	17.0	17.0	17.6	20.4	22.4	22.5	24.6	24.5	23.3
1981	23.8	26.0	29.0	30.2	38.0	40.0	36.0	22.0	24.6	31.5	31.5	31.0
1982	28.0	29.0	29.0	25.0	25.0	20.0	18.0	16.5	13.6	10.8	9.9	9.5
1983	9.5	9.5	9.6	12.8	14.1	14.2	17.3	25.8	27.0	25.0	24.0	23.5

Year	Calendar Year Average	Market Year Average	Weighted Market Year Average
1976		15.225 *	15.685 *
1977	15.952	17.812	17.932
1978	16.433	14.889	14.387
1979	16.306	18.363	18.079
1980	20.432	28.321	25.482
1981	30.296	25.758	28.478
1982	19.521	13.042	12.628
1983	17.688	24.875 *	25.173 *

* Denotes a value computed from less than 12 months data.

Table 11: Raw Price Data - Inflated -- option 3

INFLATED COMMODITY PRICE FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1977	23.0	23.0	25.5	28.7	28.3	28.1	24.9	20.8	20.1	28.7	31.4	32.1
1978	28.7	31.3	30.7	31.2	28.9	27.9	23.8	21.4	19.5	20.4	19.5	19.3
1979	18.4	18.0	19.1	20.3	21.2	23.9	28.3	26.2	22.9	23.4	23.0	23.0
1980	22.4	24.5	23.6	20.9	20.7	21.2	24.6	26.8	26.7	28.9	28.5	26.9
1981	27.2	29.5	32.6	33.8	42.2	44.0	39.2	23.7	26.3	33.6	33.5	32.9
1982	29.6	30.5	30.6	26.2	26.0	20.5	18.4	16.8	13.8	10.9	10.1	9.7
1983	9.7	9.7	9.8	12.9	14.2	14.2	17.2	25.6	26.7	24.7	23.6	23.1

Year	Calendar Year Average	Market Year Average	Weighted Market Year Average
1976		25.263 *	26.067 *
1977	26.194	28.005	28.497
1978	25.211	21.172	20.707
1979	22.307	23.086	23.220
1980	24.646	31.933	29.448
1981	33.198	27.072	30.185
1982	20.263	13.146	12.746
1983	17.611	24.522 *	24.828 *

* Denotes a value computed from less than 12 months data.

Past data values inflated using 1983 as the base CPI of 100.

7.2.6 SEASON TABLES

Table 12: Raw Price Data - Deflated -- option 4

DEFLATED COMMODITY PRICE FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1977	7.7	7.7	8.5	9.6	9.5	9.4	8.4	7.0	6.7	9.6	10.5	10.7
1978	9.6	10.5	10.3	10.4	9.7	9.3	8.0	7.2	6.5	6.8	6.5	6.5
1979	6.2	6.0	6.4	6.8	7.1	8.0	9.5	8.8	7.7	7.9	7.7	7.7
1980	7.5	8.2	7.9	7.0	6.9	7.1	8.2	9.0	8.9	9.7	9.6	9.0
1981	9.1	9.9	10.9	11.3	14.1	14.7	13.1	8.0	8.8	11.3	11.2	11.0
1982	9.9	10.2	10.2	8.8	8.7	6.9	6.2	5.6	4.6	3.7	3.4	3.2
1983	3.2	3.2	3.3	4.3	4.8	4.8	5.8	8.6	8.9	8.3	7.9	7.7

Year	Calendar Year Average	Market Year Average	Weighted Market Year Average
1976		8.466 *	8.735 *
1977	8.778	9.385	9.550
1978	8.448	7.095	6.939
1979	7.475	7.736	7.781
1980	8.259	10.701	9.868
1981	11.125	9.072	10.115
1982	6.790	4.405	4.271
1983	5.902	8.217 *	8.320 *

* Denotes a value computed from less than 12 months data.

Table 13: Consumer Price Index - actual -- option 5

CONSUMER PRICE INDEX FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1977	175.3	177.1	178.2	179.6	180.6	181.8	182.6	183.3	184.0	184.5	185.4	186.1
1978	187.2	188.4	189.8	191.5	193.3	195.3	196.7	197.8	199.3	200.9	202.0	202.9
1979	204.7	207.1	209.1	211.5	214.1	216.6	218.9	221.1	223.4	225.4	227.5	229.9
1980	233.2	236.4	239.8	242.5	244.9	247.6	247.8	249.4	251.7	253.9	256.2	258.4
1981	260.5	263.2	265.1	266.8	269.0	271.3	274.4	276.5	279.3	279.9	280.7	281.5
1982	282.5	283.4	283.1	284.3	287.1	290.6	292.2	292.8	293.3	294.1	293.6	292.4
1983	293.1	293.2	293.4	295.5	297.1	298.1	299.3	300.3	301.8	302.6	303.1	303.5

Year	Calendar Year Average
1977	181.542
1978	195.425
1979	217.442
1980	246.817
1981	272.350
1982	289.117
1983	298.417

Table 14: Consumer Price Index - inflated -- option 6

INFLATED CPI FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1977	58.7	59.3	59.7	60.2	60.5	60.9	61.2	61.4	61.7	61.8	62.1	62.4
1978	62.7	63.1	63.6	64.2	64.8	65.4	65.9	66.3	66.8	67.3	67.7	68.0
1979	68.6	69.4	70.1	70.9	71.7	72.6	73.4	74.1	74.9	75.5	76.2	77.0
1980	78.1	79.2	80.4	81.3	82.1	83.0	83.0	83.6	84.3	85.1	85.9	86.6
1981	87.3	88.2	88.8	89.4	90.1	90.9	92.0	92.7	93.6	93.8	94.1	94.3
1982	94.7	95.0	94.9	95.3	96.2	97.4	97.9	98.1	98.3	98.6	98.4	98.0
1983	98.2	98.3	98.3	99.0	99.6	99.9	100.3	100.6	101.1	101.4	101.6	101.7

Year	Calendar Year Average
1977	60.835
1978	65.487
1979	72.865
1980	82.709
1981	91.265
1982	96.884
1983	100.000

Past data values inflated using 1983 as the base CPI of 100.

Table 15: Quantities Marketed -- option 7

QUANTITY MARKETED FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1977	43.0	34.7	30.6	65.2	36.5	49.8	27.4	11.2	62.0	102.4	108.3	48.0
1978	35.9	39.6	24.7	39.1	20.9	32.5	17.0	23.7	114.5	81.8	44.9	51.3
1979	44.7	32.3	32.5	38.6	39.7	48.3	32.4	23.9	170.8	82.8	50.2	36.5
1980	42.3	55.0	36.6	14.3	13.1	30.2	47.1	20.5	177.2	218.3	74.3	56.7
1981	67.4	42.2	45.1	22.0	27.8	24.3	7.6	8.7	109.1	171.7	111.6	64.5
1982	45.7	31.3	23.5	13.5	40.9	24.8	8.7	9.3	63.9	179.7	49.1	39.2
1983	43.9	34.1	21.0	34.4	71.1	49.6	33.1	37.8	68.7	102.5	49.4	29.8

Year	Calendar Year Average	Market Year Average
1976		37.292 *
1977	51.579	46.166
1978	43.825	48.737
1979	52.722	49.946
1980	65.462	64.304
1981	58.504	54.555
1982	44.141	54.740
1983	47.946	62.608 *

* Denotes a value computed from less than 12 months data.

7.2.6 SEASON TABLES

Table 16: Weights as Percent of Quanties Marketed -- option 8

QUANTITY AS A PERCENT OF YEAR TOTAL FOR MICHIGAN NAVY BEAN DATA 1967 - 1982

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1977	14.4	11.6	10.3	21.8	12.2	16.7	9.2	3.8	11.2	18.5	19.5	8.7
1978	6.5	7.2	4.5	7.1	3.8	5.9	3.1	4.3	19.6	14.0	7.7	8.8
1979	7.7	5.5	5.6	6.6	6.8	8.3	5.5	4.1	28.5	13.8	8.4	6.1
1980	7.1	9.2	6.1	2.4	2.2	5.0	7.9	3.4	23.0	28.3	9.6	7.4
1981	8.7	5.5	5.8	2.8	3.6	3.1	1.0	1.1	16.7	26.2	17.0	9.9
1982	7.0	4.8	3.6	2.1	6.2	3.8	1.3	1.4	9.7	27.4	7.5	6.0
1983	6.7	5.2	3.2	5.2	10.8	7.6	5.0	5.7	27.4	40.9	19.7	11.9

APPENDIX A

DATA ENTRY FORM

DATA ENTRY FORM NAME

CROPS						FEEDCROPS OR OTHER CROPS				
TABLE_1	CROP 1	CROP 2	CROP 3	CROP 4	CROP 5	CROP 6	CROP 7	CROP 8	CROP 9	CROP 10
1.ENTERPR										
Descrip										
2.EXP.PR										
3.EXP.Yi										
4.UNITS 0										
5.AVG PRI	CROP 1	CROP 2	CROP 3	CROP 4	CROP 5	CROP 6	CROP 7	CROP 8	CROP 9	CROP 10
Year 1										
Year 2										
Year 3										
Year 4										
Year 5										
6.AVG YIE	CROP 1	CROP 2	CROP 3	CROP 4	CROP 5	CROP 6	CROP 7	CROP 8	CROP 9	CROP 10
Year 1										
Year 2										
Year 3										

Year 4										
Year 5										
7.VAR.EX										

8.Labor Hours

	CROP 1	CROP 2	CROP 3	CROP 4	CROP 5	CROP 6	CROP 7	CROP 8	CROP 9	CROP 10
Planting										
JAN-FEB										
MAR-APR										
MAY-JUN										
JUL-AUG										
SEP-OCT										
NOV-DEC										
TOTAL	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ
Harvest										
JAN-FEB										
MAR-APR										
MAY-JUN										
JUL-AUG										
SEP-OCT										
NOV-DEC										
TOTAL	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ	EQ

LIVESTOCK ENTRY

9. TYPE(NA ________

10.SALE WG ________

11. PRICE($ ________

12. # Farr ________

13. VAR.EX ________

14.AVG PR ________

YEAR 1 ________

YEAR 2 ________

YEAR 3 ________

YEAR 4 ________

YEAR 5 ________

15.AVG WG

YEAR 1 ________

YEAR 2 ________

YEAR 3 ________

YEAR 4 ________

YEAR 5 ________

16. LABOR HOURS/BREEDING FEMALE

JAN-FEB ________

MAR-APR ________

MAY-JUN ________

JUL-AUG ________

SEP-OCT ________

NOV-DEC ________

TOTAL ________EQ

17. FEED REQUIREMENTS

CORN BU ________

PASTURE ACRE ________

HAY TONS ________

FAMILY LABOR AVAILABLE FOR FARM

TYPE	JAN-FEB	MAR-APR	MAY-JUN	JUL-AUG	SEP-OCT	NOV-DEC	TOTAL
18. FARM H	______	______	______	______	______	______	EQ______

FAMILY OFF-FARM EMPLOYMENT

TYPE	JAN-FEB	MAR-APR	MAY-JUN	JUL-AUG	SEP-OCT	NOV-DEC	TOTAL	WAGE	TAX RATE
19. OFF-FA	______	______	______	______	______	______	______	EQ______	______

BORROWING PERCENT

20.INTERE ______

LAND	BA ACRES	$/ACRE	TOB.LBS	$/LB.
21.RENTED	______	______	______	______
22.OWNED	______		______	

END OF STRUCTURAL DATA

INPUT FOR SCENARIO DEVELOPMENT

ADDITIONAL INFORMATION FOR TAB

ITEM	VALUE	
23.LIVESTOCK?	Y=	______
24.TOBACCO?	Y=1,	______
25. CASH O		______
26.MAXIMU		______
27.MAXIMUTOB.LB		______
28.MAXIMU		______
29.FARM W		______
30.MAX.N		______

COMMODITY DEMANDS AND QUOTAS

COMMODITY UNIT	CROP 1	CROP 2	CROP 3	CROP 4	CROP 5	CROP 6	CROP 7	CROP 8	CROP 9	CROP 10
31.DEMAND	0	0	0	0	0	0	0	0	0	0
	ACRES									
32.QUOTA	0	0	0	0	0	0	0	0	0	0
(3,=),(	3	3		3	3	3	3	3	3	3
	PRICE/YIELD VARIATIONS									
33.PRICE	0	0	0	0	0	0	0	0	0	0
34.YIELD	0	0	0	0	0	0	0	0	0	0

END OF DATA ENTRY

APPENDIX B

Data Interpretation Form: Run explaination

Objective function solution:

ACTIVITIES IN THE SOLUTION

TYPE	ACTIVITY VALUE	
PRODUCTION FOR MARKET		
NO. NAME ACRES		
1. CROP 1	______	______
2. CROP 2	______	______
3. CROP 3	______	______
4. CROP 4	______	______
5. CROP 5	______	______
6. CROP 6	______	______
7. CROP 7	______	______
8. CROP 8	______	______
9. CROP 9	______	______
10. CROP 10	______	______
11. #ANIMAL	______	______
	ACRES	
12. R. LAND	______	
	LBS	
13. R. TOB LB	______	
HIRED FARM LABOR	(HOURS)	

SHADOW PRICES AND SLACKS

RESOURCE	PRICE	EXCESS
1. LAND	______	______
2.CAPITAL	______	______
3. TOB.LBS	______	______
4. RENT.LAND	______	______
5. RENT.TOB	______	______
6. FARM LAB. J-F	______	______
7. FARM LAB. M-A	______	______
8. FARM LAB. M-J	______	______
9. FARM LAB. J-A	______	______
10. FARM LAB. S-O	______	______
11. FARM LAB. N-D	______	______
12. TOTAL HIRE	______	______
13. OFF-FARM J-F	______	______
14. OFF-FARM M-A	______	______
15. OFF-FARM M-J	______	______
16. OFF-FARM J-A	______	______
17. OFF-FARM S-O	______	______
18. OFF-FARM N-D	______	______

Objective function solution:

ACTIVITIES IN THE SOLUTION

TYPE	ACTIVITY VALUE
14. HIRED LABOR J-F	________
15. HIRED LABOR M-A	________
16. HIRED LABOR M-J	________
17. HIRED LABOR J-A	________
18. HIRED LABOR S-O	________
19. HIRED LABOR N-D	________

OPER. CAPITAL BORROWING

20. BORROWED MONEY ________

ACRES FOR CONSUMPTION

	NAME	ACRES
21. CROP 1	______	______
22. CROP 2	______	______
23. CROP 3	______	______
24. CROP 4	______	______
25. CROP 5	______	______
26. CROP 6	______	______
27. CROP 7	______	______
28. CROP 8	______	______
29. CROP 9	______	______
30. CROP 10	______	______
31. #ANIMAL	______	______

SHADOW PRICES AND SLACKS

RESOURCE	PRICE	EXCESS
19. TOT.OFF-FARM	______	______

CONSUMPTION DEMAND

	NAME	PRICE	SLACK
20. CROP 1	______	______	______
21. CROP 2	______	______	______
22. CROP 3	______	______	______
23. CROP 4	______	______	______
24. CROP 5	______	______	______
25. CROP 6	______	______	______
26. CROP 7	______	______	______
27. CROP 8	______	______	______
28. CROP 9	______	______	______
29. CROP 10	______	______	______
30. #ANIMAL	______	______	______

GOVERNMENT QUOTA

31. CROP 1	______	______	______
32. CROP 2	______	______	______
33. CROP 3	______	______	______
34. CROP 4	______	______	______
35. CROP 5	______	______	______
36. CROP 6	______	______	______

OFF-FARM LABOR (HOURS)

32. OFF-FARM LAB J-F ______
33. OFF-FARM LAB M-A ______
34. OFF-FARM LAB M-J ______
35. OFF-FARM LAB J-A ______
36. OFF-FARM LAB S-O ______
37. OFF-FARM LAB N-D ______

NEGATIVE DEVIATIONS

38. YR 1 ______
39. YR 2 ______
40. YR 3 ______
41. YR 4 ______
42. YR 5 ______

43. YR 2

GOVERNMENT QUOTA

37. CROP 7 ______ ______ ______
38. CROP 8 ______ ______ ______
39. CROP 9 ______ ______ ______
40. CROP 10 ______ ______ ______
41. ANIMAL ______ ______ ______

NEGATIVE DEVIATIONS

42. YR 1 ______ ______ ______

______ ______ ______

44. YR 3 ______ ______ ______
45. YR 4 ______ ______ ______
46. YR 5 ______ ______ ______
47. MOTAD ______ ______ ______

DATA ENTRY FORM EXPLANATION

The data entry form below will be described in terms of item numbers 1 thru 32.

1. Enterprise - name of possible crop enterprises (activities) such as Tobacco. Description - other identifying characteristics such as Flue Cured (Note: If livestock is included as an activity, you must include feed crops - field corn, pasture and hay)

2. Exp. Price - average price or expected price to be received for each crop commodity.

3. Exp. Yield - average yield or expected yield for each crop.

4. Units of Measure - type of crop output units such as "bushels".

5. Avg price - past price information for the past five years.

6. Avg Yield - past crop yield information for the past five years.

7. Var.Expenses - Dollar operating expenses per acre of each crop.

8. Labor Hours - Bi-monthly labor requirements for planting and harvesting per acre of crop.

9. TYPE(NAME) - livestock enterprise type such as COW-CALF.

10. SALE WGT CWT (HEAD) - expected sale weight per head in hundred weights.

11. PRICE ($/CWT) - expected price per hundred weight.

12. # Farrowed/female - breeding rate per year per breeding female.

13. VAR.EXPENSE - dollar operating production cost per breeding female.

14. AVG PRICE - price received for past five years.

15. AVG WGT - average sale weight of animals for past five years.

16. LABOR HOURS/BREEDING FEMALE - bi-monthly labor hours needed per breeding female.

17. FEED REQUIREMENTS - feed requirements per breeding female in terms of corn bushels, acres of pasture and tons of hay.

18. FAMILY LABOR AVAILABLE FOR FARM - bi-monthly hours of labor available from the farm family that can be used for farm work.

19. FAMILY OFF-FARM EMPLOYMENT - bi-monthly hours of family labor employed off the farm.
    WAGE - hourly wage rate
    TAX RATE - income tax rate such as .25 meaning that 25 cents of every dollar goes to pay income tax and other deductions from the pay check.

20. INTEREST RATE - the cost of borrowing operating capital in terms of a percent such as 10, meaning 10 percent.

21. RENTED - acres presently rented, cost per acre, tobacco pounds rented (if any) and tobacco rental charge per pound of allotment (if applicable).

22. OWNED - tillable acres owned and pounds of tobacco allotment owned (if applicable).

This is the end of the information needed for the structural development of the farm model. The following information is needed for scenario development or manipulation of some of the constraints. This of course does not mean that the structural information can not be altered if desired, however the following section will probably be the one in which most alternative runs will be developed for a given farm.

23. LIVESTOCK? Y=1, N=0 - if you intend for the model to include the livestock enterprise as a possible activity enter "1". This is an important item since the response determines if there will be any internal transfer of crops to be consumed by the livestock. Based on the feed requirements, the model will force production of the required farm produced feeds to support a given activity of livestock size. Unfortunately, the model does not allow for purchasing of corn, pasture or hay. The cost of any other type of feeds should be included in the

VAR.EXPENSES section above (item 13). Also, if livestock is to be included, you must have corn, pasture and hay included as the last 3 enterprise activities under item 1.

24. TOBACCO? Y=1,N=0 - if you include tobacco as a possible enterprise you must answer with a "1" in this space. If you include tobacco you must also place the tobacco enterprise information as the first crop in item 1 above. If you are not including either livestock or tobacco in the model, you may place the 10 possible crop enterprises in any order that you please.

25. CASH ON HAND - indicate the amount of capital that the farm operator has on hand at the beginning of the year that can be used to purchase production inputs.

26. MAXIMUM RENTED LAND - indicate the maximum rented land the farmer will lease for a given year.

27. MAXIMUM RENTED TOB.LBS - indicate the maximum tobacco poundage available to be rented in a given year, if applicable.

28. MAXIMUM HIRED LABOR - indicate the maximum amount of hired available to the farmer in terms of total hours per year.

29. FARM WAGE - indicate the average farm wage that the operator can expect to pay hired labor (not family labor).

30. MAXIMUM NEGATIVE INC. - indicate the maximum negative deviation from the mean that a farmer is willing to risk for a given year. For instance a figure of 10,000 would indicate that the farmer will not accept a combination of enterprises that may result in him receiving 10,000 dollars less that what can be expected in a typical year. The typical year is determined by the expected price and yields indicated in items 2 and 3 for crops and items 10 and 11 for the livestock enterprise. The possible deviations from those means are determined using the past historical price and yield information supplied in items 5 and 6 for crops and items 14 and 15 for the livestock enterprise.

31. DEMAND (UNITS) - for each crop indicate any household consumption in terms of the specific units of the commodity (ie. 5 bushels of wheat to be consumed getting support prices.)

(3,=),(2,=) - the default is set at "3" meaning that the quota requires one to grow at least the amount prescribed. The other alternative is a "2" which indicated that the farm must grow exactly a certain amount, and no more and no less. This option is especially useful if one is trying to force a solution out of contention or forcing in a prescribed amount of acreage of crop due to rotational requirements.

33. PRICE (+,-PERCENT) - this allows for sensitivity analysis in terms of price variations. For each of the possible enterprises one can see how the optimal solution of activities change if any price went up by 10% (indicate 10) or down by 10% (-10).

34. YIELD (+,-PERCENT) - this allows for sensitivity analysis in terms of yield variations.

FORM EXPLANATION

Objective Function (representing gross returns minus variable operating expenses) goes in solution blank on the form and refers to the value of the objective function under the optimal solution for the present scenario.

Activities in Solution. The activities are designated in terms of numbers from 1 to 42. What gives these activities meaning in terms of precesses should be written in on the interpretation form. For instance, if Tobacco is the first crop, then the "NAME" Tobacco should be written in next to Crop 1 and the level of activity next to it. In the example of Farmlp presented in the manual, activity 5 is "POTATOES". Activity 5 has a value of 24.096 representing 24 acres of potatoes. The activities reflected in columns 38 to 42 corresponds to the given year that contributes to negative deviations from the objective function solution. The sum of activities 38 to 42 should equal the "MOTAD' constraint level indicated when developing the scenario.

Shadow Prices and Slacks (Excess) which are placed on the Interpretation form in items 1 through 47 under "PRICES" and "EXCESS". "PRICES" refers to the opportunity cost of the resource which is how much one more unit of the resource would contribute to the objective function.

These are all data that is used for the interpretation form. The other data are used for sensitivity analysis. A brief explanation is given for the example problem (TRY) described in appendix C.

UNDERSTANDING LINEAR PROGRAMMING

Assumptions for the model

a. Additivity of resources and activities -- The sum of the resources used by different activities must equal the total quantity of resources used. For example, the objective function cannot add net income from a land activity and pounds of milk from a dairy activity. If the object is to maximize dollars then the objective function must include only dollar coefficients. The structure of the model can take care of the interaction among activities of the resources.

b. Linearity of the objective function -- For example, the price (and therefore net income) of corn or soybeans cannot be a function of quantity sold. The user can structure the model by adding activities to take care of the problem.

c. Non-negativity of the decision variables -- Cannot have negative acres or labor.

d. Divisibility of activities and resources -- Implies continuity of resources and output and have fractional uses (for example, 3.3 milk cows or .3 of a tractor). For problems requiring integrated resources or outputs, integer programming can be used.

e. Finiteness of the activities and resource restriction -- There is no optimal solution unless something limits the model.

f. Proportionality of activity levels to resources -- Proportionality assumption implies linear relationships between activities and resources. For example, if we use one acre of land, 5 hours of labor, 200 pounds of fertilizer and .25 bushels of seed to produce 60 bushels of wheat, then 120 bushels of wheat can be produced with twice the resources. The assumption implies constant resource productivity and constant returns to scale. The problems can be handled by structuring the model with different activity sets.

g. Single-valued expectations -- all values are known with certainty. The model is deterministic -- The problem can be handled with stochastic models.

EXAMPLE : TRY

The basic linear programming structure for the example model TRY is:

a. Objective function

Max y = 150 X1 + 140 X2

where y = gross margin (excluding fixed costs)

X1 = corn activity (acres produced)

X2 = soybean activity (acres produced)

b. Subject to restrictions

X1 + X2 <= b1 = acres of land

.45X1 + .6X2 <= b2 = hours harvest labor

110X1 + 60X2<= b3 = dollars operating capital

c. and X1 >=0 and X2 >=0

The linear programming tableau would then be as follows:

Row	Column Activities resource 1	2	Type Constraint	Constraint Level
1	1.0000	1.0000		400
2	0.4500	0.6000		225
3	110.0000	60.0000		35000
4	1.0000	0.0000	=	0
5	0.0000	1.0000	=	0
Object	150	140		

The MSTAT LP program displays this same tableau with a slightly different format but the meanings are the same:

7.3 LINEAR PROGRAMMING

Function : Linear Programming Package
Problem Size : 5 Rows, 2 Columns

Initial Matrix

	Column	
Row	1	2
1	1.0000	1.0000
2	0.4500	0.6000
3	110.0000	60.0000
4	1.0000	0.0000
5	0.0000	1.0000

Row		Restriction	Right Hand Side
1	|		400.0000
2	|		225.0000
3	|		35000.0000
4	|	=	0.0000
5	|	=	0.0000

Column	|	Objective Value
1	|	150.0000
2	|	140.0000

Below are the outputs similar to screens 6 through 9 illustrated earlier. The optimal solution of 58200 represents gross returns minus variable operating expenses. The solution in terms of activities are to grow 220 acres of corn (activity 1) and 180 acres of soybeans (activity 2). All of the land is utilized (0 excess for constraint 1) and all of the capital is utilized (0 excess for constraint 3). However, 18 hours of labor are left not utilized.

The price of 128 on constraint 1 (land) indicates that the addition of one more acre of land could increase the objective function by $128 or in other words he could afford to pay $127 for one more acre of land and still increase his income by $1.

The upper and lower bound on the objective function coefficients help determine how sensitive the given solution is. For example, as long as the objective function on corn stays within bound and 256.67, the solution would not change. But if the returns dropped to below 140, he would cut back on corn production.

OPTIMAL SOLUTION
(4 iterations)

OBJECTIVE FUNCTION = 58200.0000

Activities in solution :

Column	Activity
1	220.0000
2	180.0000

Shadow Prices and Slacks :

Row	Price	Excess
1	128.0000	0.0000
2	0.0000	18.0000
3	0.2000	0.0000
4	0.0000	220.0000
5	0.0000	180.0000

Cost of Forcing in Nonoptimal Activities :

-- There are no non-optimal activities

Price Range Over Which Optimal Solution Holds :

Column	Lower Bound		Upper Bound	
	Activity	Price	Activity	Price
1	5	140.0000	3	256.6667
2	3	81.8182	5	150.0000

8.1 STAT

Purpose: Computes number of observations, maximum and minimum values, sum, mean, standard deviation, skewness and kurtosis for the requested variables.

Before starting, you need:

1. An active data file with data cases.

Results: A table of statistics is created for all requested numeric variables - number of observations, minimum and maximum values, sum, mean, variance, standard deviation, standard error, skewness with T-value test and probability value and kurotsis with T-value test and probability value. Either print the resultant table or store the table in a disk file after viewing it.

Sample Data File: HURON1.

STAT computes summary statistics for individual variables. You are prompted for a list of variables.

STAT computes the statistics for each variable and arranges the statistics in a table. Then the print/list menu of options is displayed. Since the statistics are not stored with the active data file, you will want to print the table and/or store the table in a disk file.

The skewness value indicates the degree to which a distribution of values approximates a normal curve. A value of zero is obtained with a completely bell-shaped curve. A positive value indicates that the tested values cluster to the right of the mean. You will need to consult a t-value table for more specific information on the amount of skewness.

Kurtosis is a measure of the relative peakedness or flatness of the curve. A normal distribution will have a kurtosis value of zero. If the kurtosis is positive, then the distribution is more peaked (narrow) than would be true for a normal distribution, while a negative value means that it is flatter. You will need to consult a statistics text for more specific information on the amount of kurtosis.

The user must enter a data file to access the STAT program. This can be done in the initial menu FILES program. Then highlight STAT and press

<ENTER>. Following the first screen, a listing of the variables in the active data file is presented on the screen. Press the spacebar to highlight each variable to be used (calculation of basic statistics to be presented in tabular format). Press <ENTER> to finish the selection process.

```
Do you want to include skewness and kurtosis : Y/N
```

```
The data file contains 120 cases
Do you wish to use all cases? Y/N
```

Because some of the MSTAT programs add cases to data files , the user is prompted to indicate the range of cases to use in the calculations. If the active data file contains only the cases with the necessary information, then the user should press <ENTER> when the above prompt is presented.

Next the subprogram displays the print/list menu of options. You may view the table on the screen, print it or store it in a disk file.

The following table is created from this example.

Data file :HURON1
Title : BEAN EXPERIMENT

Function: STAT
Data case no. 1 to 120

Variable Number	N of Cases	Minimum	Maximum	Sum
YIELD IN GRAMS PER PLOT				
3	120	784.000	3500.000	244761.000

Variable Number	Mean	Variance	Standard Deviation	Standard Error
3	2039.675	316830.843	562.877	51.383

Var No	Skewness	T-val	Prob	Kurtosis	T-val	Prob
3	-0.0415	-0.1878	0.4257	-0.3602	-0.8217	0.206

8.2 MEAN

Purpose: Calculates means and does scaling (multiplying by a constant).

Before starting, you need:

1. An active data file with data cases.

Results: The means of requested variables are computed for each level of a group variable. These data are added to the end of the active data file.`

Sample Data File: COMPACT.

MEAN computes the averages for each requested variable over any group variable. You may use a scaling factor for any of the variables used in computing means. The subprogram ignores any text variables.

One method of comparing research data is to study the average of each research unit (variety, fertilizer treatment, irrigation application, etc.). In the COMPACT data file example, there is a variable that indicates a variety (variable no. 9). This variable has group levels of 1 through 15 (or 15 distinct varieties). The information recorded in the COMPACT data file is pod length, yield, seed per pod and seed weight. It is possible to calculate the average of each of these variables for each variety tested. The MEAN program computes the averages of each variable for each group level; in this example, computation of the means for 15 levels is done and 15 cases are added to the end of the active data file. At any point in the program, the user can press <ESC> to return to the Initial MSTAT menu.

The first screen is:

```
Press <F1> for a list of variables

Enter the number of the GROUP variable (1-9) : 9
Enter the lowest and highest values in the GROUP
variable
     Lowest: 1            Highest : 15
```

After entering the values for the group variable, a list of variables in the active data file is presented on the screen. Move the cursor to a variable that contains gathered data (dependent values) and press the spacebar to highlight that item. Continue to place the cursor at different data variables and press the spacebar to highlight them. Do not place a highlight on the grouping variable or other grouping (independent) variables.

```
Choose up to 8 variables (Press ESC to quit)
  01    (NUMERIC)      REPLICATION
  02    (NUMERIC)      COMPACTION
  03    (NUMERIC)      ENTRY
  04    (NUMERIC)      10-POD LENGTH
  05    (NUMERIC)      YIELD
  06    (NUMERIC)      SEED PER POD
  07    (NUMERIC)      1000-SEED WEIGHT
  08    (NUMERIC)      PODS PER PLANT
  09    (NUMERIC)      variety
```

For this example, move the cursor down the screen and press the spacebar to highlight variables 4, 5, 6 and 7. Press <ENTER> to continue with the program after the independent variables have been selected.

```
The data file contains 120 cases
Do you wish to use all cases? Y/N
```

Because some of the MSTAT programs add cases to data files , the user is prompted to indicate the range of cases to use in the calculations. If the active data file contains only the cases with the necessary information, then the user should press <ENTER> when the above prompt is presented.

Press <ENTER> to respond <Y>. In this data file, the information is stored in the first 120 cases of the file. If there are more than 120 cases in the COMPACT active data file, type <N>, press <ENTER>, then enter the first and last case range for this analysis (1, 120).

It is sometimes desirable to change the values of all data collected on a specific dependent variable by multiplying or dividing all the entered data by a constant value. If all the recorded values in one variable are less than .01, it may be easier to see their relationship to each other if they are all multiplied (scaled) by 100. This is scaling. To divide all the values by a constant, multiply (or scale) by a value less than 1.0. For example, to divide by 100, use a scale value of 0.01. Press <RETURN> to enter the scale value and to continue the program.

```
Do you want to use scaling : Y/N
```

```
Enter the scaling factor that you wish for each
variable
If you do not want a variable to be scaled input a 1.0
Variable : 5 YIELD
Scale : 0.01
```

The means are calculated and are added to the end of the active data file. In this example, there are 15 levels in the group variable, so 15 data cases are added to the end of COMPACT. The means for each independent variable are placed in the same column as the raw data.

The following are the first six cases of the 15 cases of data added to the end of COMPACT. The group variable for the example was variable no. 9 (variety) which has values ranging from 1 through 15. When you list the COMPACT data file, there are four replications (variable no. 1) and all 15 varieties are in each replication for each of two compaction levels. Therefore, each mean value under variable no. 4 is obtained from eight values; each mean value under variable no. 5 is obtained from eight values, etc.

Function : MEAN
Data case no. 121-135

CASE NO	4	5	6	7	9
121	76.5	2.5	6.0	163.0	1
122	87.6	2.4	6.1	184.4	2
123	84.6	2.2	6.0	186.9	3
124	93.4	2.6	6.1	194.6	4
125	85.1	2.1	5.9	188.6	5
126	90.5	2.8	5.8	205.9	6

8.3 GROUPIT

8.3 GROUPIT

Purpose: Transforms a quantitative variable into a group variable by dividing data into numbered ranges.

Before starting, you need:

1. An active data file with data cases.
2. A printout of the data file or a STAT analysis so that you can specify how to divide up data values.

Results: Depending on how you respond to GROUPIT prompts, group levels (ranges) are either created as new variables or overwrite existing variables. A table indicating the group levels (ranges), number of data points in each range and frequency of occurrence is also created.

Sample Data File: HURON1.

GROUPIT is used to transform a quantitative variable (collected data) into a group variable (assigned levels). To do this, the various quantities of a variable are divided into ranges; each range is assigned a number, beginning with "1". There can ba a maximum of 100 distinct ranges in this program. For example, you have a variable that indicates maize yields with values ranging from 40 tons/hectare to 700 tons/hectare. You would like to label all yields of 0-200 tons/hectare as "group 1", all yields of 201-350 tons/hectare as "group 2", and so forth. By this method you create what is called a group variable (or grouping variable). The group variable may then be used in statistical analyses as an independent variable. Up to a maximum of 100 groups can be formed.

When creating a grouping variable, the information can be stored in an existing variable or in a newly created variable.

In this example, variable no. 3 is used to create a group variable (variable no. 4). Variable no. 3 has values ranging from 784 to 3500.

At any point in the program, the user can press <ESC> to return to the Initial MSTAT menu.

```
Press <F1> for a list of variables
Enter the number of the QUANTITATIVE variable   (1-4): 3
```

The quantitative variable is the variable that is to be separated into ranges. The values that are used as the low and high ends of each range are entered by the user from a general stat table of the data or from other knowledge of the data file contents. In this example, the data point values had a minimum value of 784 and a maximum value of 3500. The majority of the data values were in the 1000-2500 range with an overall range of approximately 2700 units. It was decided to separate the yield into 500 unit increments with all values under 1000 in the first grouping and all values greater than 2500 in the last grouping. So the high limits were set to 1000, 1500, 2000, 2500 and all yield values above 2500. As the data file contains no valuable information in variable 4, it was decided to put the group level designate into that variable and in the process, overwrite the information currently stored in that position.

```
You may define a new variable to contain the groups
or you may use an existing variable

Enter the number of the variable you want transformed
into a GROUP variable
You may overwrite an existing variable
You may not use your QUANTITATIVE variable number

Press <F1> for a list of variables
Enter the number of the GROUP variable (1-4) : 4
```

```
Enter the number of groups that you will form
The last group's high limit is set at 999999
Enter the number of groups (1-100) : 5

Enter the high limits for each group in ascending order
The last group's high limit is set at 999999
Enter the high limit for group ( 1) in the GROUP
variable
        Low : 0.00           High : 1000
Enter the high limit for group ( 2) in the GROUP
variable
        Low : 1000           High : 1500
```

8.3 GROUPIT

Continue to enter the upper limits for each group. The user will not be prompted for the fifth group in this example as the program automatically assigns the value 999999 as the high limit of the last range.

3 - **2000**
4 - **2500**
5 - 999999

The group levels are divided into the following ranges:

Group One	0 - 1000
Group 2	1001 - 1500
Group 3	1501 - 2000
Group 4	2001 - 2500
Group 5	2501 - 999999 (GROUPIT creates this group automatically.)

Following the entry of the last highest level for the 4th group, the user will be prompted for the range of cases to use in the grouping process.

```
Get Case Range
The data file contains 120 cases.
Do you wish to use all cases? Y/N
```

For this example, all 120 cases are grouped.

The print/list menu of options is displayed next. In addition to creating new values for a variable, the subprogram creates a table showing the group levels and the frequency of them. You may view this table on the screen, print it or store it in a disk file. For this example, it creates the following table:

```
Data file   HURON1
Title: BEAN EXPERIMENT

Function: GROUPIT
Data case no.  1  to  120
```

Var. 3 High limit	Var. 4 Group	Freq
1000.000	1	4
1500.000	2	19
2000.000	3	31
2500.000	4	38
999999.000	5	28

8.4 FREQ

Purpose: Creates one- and two-way frequency tables.

Before starting you need/should know:

1. An active data file with data cases.
2. You may want a STAT analysis of your data before creating a frequency table.

Results: Depending on how you respond to the prompts, one-way frequency tables which may include histograms and two-way frequency tables of groups or quantitative variables will be created. These tables may be displayed on the screen, edited, printed or stored in a disk file.

Sample Data File: HURON1.

FREQ divides variables into frequency classes so that the distance between group levels remains the same. For each variable, you specify the lowest and highest values in the data file and give the "divisor", or value needed to divide the data into groups. This procedure also is automated if you want to see a preliminary grouping. The STAT analysis of each variable will help determine the proper frequency classification.

FREQ asks some preliminary questions about the data you want to use in tables and then gives you the option to create a one-way or a two-way table.

The examples on the following pages demonstrate how you may use these options. If you choose to create a one-way table, you can also get a histogram of the frequency of values occurring within each group level. With the histogram you can visualize the actual distribution of your data as well as find separate segregation or modal patterns. If you choose to create a two-way table, FREQ displays the frequency of occurrences for combinations of the group levels of two variables.

Groups, Columns and Rows

When creating a one-way table without automatic grouping, you are prompted for variable number, lowest and highest value as well as the grouping width. When creating a two-way table, you are prompted for the variable numbers for row and column as well as high and low values and grouping width. FREQ prints each row as a line of text on your paper. Columns are printed across the width of the paper (they are printed horizontally): therefore paper width limits you to the following:

8.5 inch wide paper with normal print: 80/11= 7 groups maximum.
13 inch wide paper with normal print: 132/11 = 12 groups maximum.

Each group or column requires 10 character spaces. If you use condensed print, the maximums are larger.

Using a Divisor

Use a divisor to put variable values into groups. For example, if a variable has values ranging from 1 to 120, you may want a divisor of 10 so that you have groups of 1-10, 11-20, etc. If a variable only has values 1 to 10, then you will probably use a divisor of 1. In the HURON1 data file, the values in variable 3 range from 740 to 3700. These low and high limits were rounded to 500-4000 then grouped in units of 500 each. If the user has already established groups in the GROUPIT program, the divisor will be 1.

FREQ Examples

The HURON1 sample data file is used for the following examples. Variable no.s 1 and 3 will be used in the examples.

Example 1: A One-Way Table.

For this example, use the first 120 data cases.

```
Case Range 1-180
First selected case    1
Last selected case     120
```

```
Would you like a one-way table or a two-way table? O/T
Would you like to have automatic grouping? Y/N
```

A variable selection panel allows the user to highlight the variables that are to be tabulated if automatic grouping is selected by pressing <ENTER> to accept the default of Y. If <N> is entered, the user will be prompted for the number of variables to be analyzed.

```
How many variables do you want to analyse ? 1
```

8.4 FREQ

If automatic grouping is chosen the next four prompts will not appear on the screen. They are for information if the user wants to control the group division. For this example, the values are inserted as shown.

```
What is the variable number?                    3
What is the lowest value in the data file?   500
What is the highest value in the data file?  4000
What is the distance between groups (divisor)?  500
```

If either automatic or manual grouping is selected, the next prompt concerns displaying of the histogram associated with the data.

```
Would you like to have a histogram displayed? Y\N
```

When FREQ is finished computing, the print/list menu of options is displayed on the screen. You may view the table on your screen, edit, print or save it on a diskette.

The following table and a similar graph are generated from the example if automatic grouping is selected.

```
Data file    HURON1
Title: BEAN EXPERIMENT

Function: FREQ
Data case no.   1  to  120

      ONE-WAY FREQUENCY TABLE

  Variable Number:          3
    Low Value:            784.00
    High Value:          3503.00
    Divisor:              340.00
```

Group Number	Begin	End	Frequency
1	784	1123	10
2	1124	1463	9
3	1464	1803	19
4	1804	2143	32

5	2144	2483	22
6	2484	2823	19
7	2824	3163	8
8	3164	3503	1

HISTOGRAM

```
Frequency

32                                   **
31                                   **
30                                   **
29                                   **
28                                   **
27                                   **
26                                   **
25                                   **
24                                   **
23                                   **
22                                   **      **
21                                   **      **
20                                   **      **
19                           **      **      **      **
18                           **      **      **      **
17                           **      **      **      **
16                           **      **      **      **
15                           **      **      **      **
14                           **      **      **      **
13                           **      **      **      **
12                           **      **      **      **
11                           **      **      **      **
10           **              **      **      **      **
9            **      **      **      **      **      **
8            **      **      **      **      **      **
7            **      **      **      **      **      **      **
6            **      **      **      **      **      **      **
5            **      **      **      **      **      **      **
4            **      **      **      **      **      **      **
3            **      **      **      **      **      **      **
2            **      **      **      **      **      **      **
1            **      **      **      **      **      **      **      **
```

Begin:	784	1124	1464	1804	2144	2484	2824	3164
End:	1123	1463	1803	2143	2483	2823	3163	3503
Frequency:	10	9	19	32	22	19	8	1

Example 2: A Two-Way Table.

For a two-way frequency table with automatic grouping; using variables 1 and 3, the parameter screen would d be as shown below:

```
Would you like a one-way table or a two-way table? O/T
Would you like to have automatic grouping? Y/N
```

If automatic grouping is choosen, the next screen will present the variables in the data file. Cursor through the list and highlight those variables to be used as row variables. Each of the selected variables will be paired with each of the column variables which will be selected on the next screen. From the second screen, cursor down and select the variables to be used as the column variables (to be paired with each of the row variables previously selected.).

For a two-way frequency table **without** automatic grouping; using variables 1 and 3, the parameter screen would have the additional prompts shown below:

```
You will now be asked to enter some row variables and
some column variables.  Each row variable will be
paired with each column variable to produce your two
way frequency tables.

How many row variables would you like to use?      1
How many column variables would you like to use?   1
```

```
What is the ROW variable number?                   1
What is the lowest value?                          1
What is the highest value?                         30
How wide should the groupings be (divisor)         1
```

```
What is the COLUMN variable number?                3
What is the lowest value?                          500
What is the highest value?                         4000
How wide should the groupings be (divisor)         500
```

8.4 FREQ

Enter the range of values for both the row and the column variable as well as the desired divisors. For example, the column variable no. 3 is a quantitative variable with values ranging from 784 through 3700; a divisor of 500 is used to group the quantitative values.

When the two-way frequency table is printed, it will look like the sample on the following page.

Data file HURON1

Title: BEAN EXPERIMENT

Function: FREQ

Data case no. 1 to 120

TWO-WAY FREQUENCY TABLE

Var 3 Grouping= 500

Var 1 Value	500 999	1000 1499	1500 1999	2000 2499	2500 2999	3000 3499	3500 3999	:	Total
1	0	0	1	2	1	0	0	:	4
2	0	1	1	2	0	0	0	:	4
3	0	2	0	2	0	0	0	:	4
4	0	0	2	2	0	0	0	:	4
5	0	1	0	1	2	0	0	:	4
6	0	0	1	2	1	0	0	:	4
7	0	0	2	1	1	0	0	:	4
8	0	0	2	0	2	0	0	:	4
9	0	1	1	1	1	0	0	:	4
10	0	0	0	1	3	0	0	:	4
11	0	1	1	0	2	0	0	:	4
12	0	0	2	2	0	0	0	:	4
13	3	1	0	0	0	0	0	:	4
14	0	1	3	0	0	0	0	:	4
15	0	3	0	1	0	0	0	:	4
16	0	0	2	1	1	0	0	:	4
17	0	1	2	1	0	0	0	:	4
18	0	0	2	0	2	0	0	:	4
19	0	0	0	2	2	0	0	:	4
20	0	0	1	2	0	0	1	:	4
21	0	2	1	0	0	0	0	:	4
22	0	2	2	0	0	0	0	:	4
23	0	0	0	3	1	0	0	:	4
24	0	0	1	2	1	0	0	:	4
25	0	1	3	0	0	0	0	:	4

26	1	0	2	1	0	0	0	:	4
27	0	1	2	1	0	0	0	:	4
28	0	0	0	3	1	0	0	:	4
29	0	0	0	2	1	1	0	:	4
30	0	0	0	0	1	3	0	:	4
Total	4	19	31	38	23	4	1	:	120

8.5 PLOT

8.5 PLOT

Purpose: Creates an x,y plot with regression analyses.

Before starting, you need/should know:

1. An active data file with data cases.
2 GRAPHICS.COM must have been loaded before running MSTAT-C.

Results: Creates an x,y graph between two variables with regression statistics. Depending on which print/list options you choose, the graph may be viewed, printed or stored in a disk file.

Sample Data File: HURON1.

PLOT creates the diagram with x as the horizontal axis and y as the vertical axis. The regression statistics are printed in table format before the graph is printed. These are the same as those printed by the CORR subprogram, except that the information is formatted in a slightly different way.

The PLOT program will use the graphics mode of the printer and the output will be sideways to the other output. The regression line will be printed as well as the data points. The range of the X-axis and the Y-axis is determined by the program. PLOT asks for the variables that will be used for the x and y axes then the range of cases to include on the graph..

PLOT Example

The following is an example of creating an x,y plot with regression. The data file contains 120 cases and the graph will be of the first 60 cases. At the Case Range prompt, enter 1 as the lower case range value and 60 as the upper case range value.

Variable no. 1 of the HURON1 data file is used for the x-axis, and variable no. 3 is used for the y-axis. Either type in these values or press <F1> to recall the variable list bar menu and highlight the variable desired.

```
PLOT: Variable Numbers
Select variable numbers to use as the Coordinates (1-4)
     (Press  F1  for a list of variables)
  X-Coordinate:   1    Y-Coordinate:     3
```

When viewing the plot after it has been created, the user can first see the plotted data then call up a second screen with the plotted data and the regression line. To see the regression line, respond <Y> to the following prompt.

```
Would you like to view the linear regression line on your
graph?
```

Either view, print or save the output as prompted on the screen.

This is the output generated for this example:

Regression Line: y = a + bX

Title of X Variable	:	Entry
Title of Y Variable	:	Yield in Grams per Plot
Number of Data Points	(K):	60
Mean of X Variable	(X-bar):	15.500
Mean of Y Variable	(Y-bar):	1955.917
Variance of X Variable	:	76.186
Variance of Y Variable	:	336606.722
Coefficient of Correlation	(r):	0.107
Regression Line Intercept	(a):	1845.739
Regression Line Slope	(b):	7.108
Standard Error of Slope	(s):	8.678
t Test Value	(t):	0.819
Probability	(P):	1.000

8.6 CURVES

Purpose: To print one or more curves on the same figure. Each curve represents the values of one variable, with successive data cases on different lines.

Before starting, you need:

1. An active data file.

Results: A plot of the variables chosen is produced, using ASCII symbols or characters to represent the data points plotted. More than one variable can be plotted against the same axes. The Y-Axis is the variable case number; the X-Axis increases from 0 to 100. The output will be in compressed print mode.

Sample Data File: COMPACT

CURVES can be used to visibly compare one variable with another. This program is most effective when means of the variables are used.

Each plotted value within the variable must fall within the range of 0 and 100. This transformation of the data is done within the program. The transformed values are not written to the file so the shifting of the data is redone each time the CURVES program is accessed. The data are first transformed to values between 0 and 100 by subtracting a constant (zero point or subtrahend), and then dividing the remainder with a scale factor.

For each curve you will have to give the following information:

Variable number - Enter the variable number that contains the values to be plotted. Only numeric data can be plotted by this program.

Symbol - any character on the keyboard can be used to indicate the specific points on the plot. One method is to use the variable number.

Constant to be subtracted from each data value (=SUBTRAHEND). Because the final value to be plotted must fall in the range of 0 and 100, the data may have to be altered by subtracting a positive or negative constant integer from all data values of a specific variable. Each variable plotted can have a different constant as the subtrahend.

Scale factor (divisor) - To further alter the data so that it can be plotted between 0 and 100, the data value can be multiplied by a constant. This constant can be greater than 1 to decrease the data value or can be less than 1 to increase the value. The scale factor operation will be performed on the data after the subtrahend has been subtracted from the actual data.

If the result after division (step 4 above) is between 0 and 100, the symbol is printed in the correct column. If the result is negative, the symbol is printed to the left of the zero (left) margin; if it is more than 100, the symbol is printed to the right of the right margin.

If more than one curve happens to fall on the exact same spot, only the symbol for the last one will show up in the figure.

CURVES Example

For this example, means were calculated for the first 60 cases of the data file COMPACT. The resultant 15 means were stored in cases 121 to 135 of the data file.

Case	4	5	8
121	77	267	27
122	85	255	22
123	84	227	21
124	92	275	23
125	84	214	19
126	91	265	21
127	84	248	23
128	84	216	22
129	83	328	29
130	81	222	19
131	100	250	21
132	91	289	25
133	100	330	27
134	85	255	23
135	86	265	30

8.6 CURVES

The variables in the COMPACT data file that will be plotted are:

04	(NUMERIC)	10-POD LENGTH
05	(NUMERIC)	YIELD
08	(NUMERIC)	PODS PER PLANT

For the data given, variables 4 and 8 are within the 0 to 100 range required for the program. In order to place the values for variable no. 4 further from the right-hand margin of the plot, a constant value of 20 was subtracted from each of the recorded data values. No changes were made on variable no. 8 values.

Variable no. 5 had to be transformed The range of values for variable no. 5 is 214 to 330. All of the data values were divided by 10 and the resultant range was 21 to 33. This is the same range as the values in variable no. 8 so a comparison of Yield and Pods per Plant on the mean of each of the 15 varieties tested would be more clearly envisioned.

Before starting the program CURVES, from the Initial MSTAT menu, highlight FILES and load the data file COMPACT. Return to the Initial MSTAT menu and highlight MEAN. Calculate means for variables 4, 5 and 8 using the first 60 cases with variable no. 9 as the grouping variable (levels of 1 to 15.) When the Initial MSTAT menu is presented, highlight CURVES. The first screen requests the number of curves that will be plotted on the graph at the same time. For this example, variables 4, 5 and 8 will be plotted for a total of three curves.

```
Enter the number of curves you want to plot on one graph:    3
```

Once the number of curves has been established, each curve must be defined as to variable number, plotted symbol, subtrahend and divisor. If <F1> is pressed when the Variable number block is highlighted, a variable list bar menu will be superimposed on the screen and the user can cursor to the variable desired and press <ENTER> to select that variable.

```
Press <F1> for a list of variables
For each curve enter variable number, 'ZERO-POINT'
(=SUBTRAHEND), SCALE (=DIVISOR) and SYMBOL (one
character)
Variable number to plot  : 4
       SYMBOL to use     : 4
ZERO-POINT (SUBTRAHEND)  : 20
       SCALE (DIVISOR)   : 1
```

As each curve is defined, the user is prompted for the information for the next curve. For this example, fill in the information as indicated.

```
Variable number to plot : 5
        SYMBOL to use   : 5
ZERO-POINT (SUBTRAHEND) : 0
        SCALE (DIVISOR) : 10
```

```
Variable number to plot : 8
        SYMBOL to use   : 8
ZERO-POINT (SUBTRAHEND) : 0
        SCALE (DIVISOR) : 1
```

Once the curves are defined, the range of cases to include must be determined. For this example, the means of the variables are stored in the data file in cases 121 to 135 for a total of 15 values. Do not use all the cases in the data file rather press <N> then enter the case numbers 121 and 135 as the lower and upper limits of cases to be included.

8.6 CURVES

```
The data file contains  135 cases.
Do you wish to use all cases? Y/N
```

```
First selected case     121
Last selected case      135
```

Because the printed plot is in compressed print (17 characters per inch) the screen view will only show up to the 60 mark. The printed output generated by this subprogram is presented below. For cases numbers 121 and 128, the values generated were the same for variables 5 and 8 so the last symbol (8) was printed on the plot.

```
Data file   COMPACT
Title: SOIL COMPACTION

Function: CURVES
Data case no.  121  to  135

Var.No.    0-Point     Divisor      Symbol  Text
 4            20         1            4     10-POD LENGTH
 5             0        10            5     YIELD
 8             0         1            8    PODS PER PLANT
```

```
Case No. 0        10        20        30        40        50        60        70        80        90        100
------------------------------------------------------------------------------------------------------------------------
  121  |                           8                             4                                                |
  122  |                        8  5                                         4                                    |
  123  |                      85                                           4                                      |
  124  |                        8    5                                               4                            |
  125  |                     85                                            4                                      |
  126  |                      8    5                                                 4                            |
  127  |                        8                                        4                                        |
  128  |                        8                                        4                                        |
  129  |                             8  5                               4                                         |
  130  |                   8  5                                     4                                             |
  131  |                      8   5                                                      4                        |
  132  |                          8    5                                           4                              |
  133  |                           8        5                                            4                        |
  134  |                      85                                         4                                        |
  135  |                             5  8                                  4                                      |
------------------------------------------------------------------------------------------------------------------------
```

8.7 T-TEST

Purpose: Given two samples, this program will perform an F-test at alpha levels 0.10, 0.05 or 0.01 to see if the variances are equal. Based on the result of the F-test, an appropriate t-test will be performed to see if the means are equal.

Before starting, you need:

1. An active data file with a separate variable for each set of data.

Results: For each sample, mean, variance and standard deviation will be calculated. F-test for the hypothesis "Variance 1 = Variance 2" will be performed with a result of either significant F (reject the hypothesis) or non-significant F (accept the hypothesis). T-Test for the hypothesis "Mean 1 = Mean 2" will be performed with a result of either significant T (reject the hypothesis) or non-significant T (accept the hypothesis).

Sample Data File: TTEST for non-paired, PAIRED for paired sets.

The T-test is used when two means are to be compared to see if they are statistically the same. For this data file, the two samples to compare are in variables 3 and 4, cases 1 to 10. This example is at the 0.05 alpha value. If other alpha values are to be used, press the space bar to toggle between the three choices (0.01, 0.05 and 0.10). The data is as follows:

Case No.	Variable Number 3	4
1	42.60	489.71
2	50.54	489.22
3	57.59	501.14
4	45.38	509.47
5	44.82	497.10
6	48.44	516.63
7	43.68	498.67
8	60.68	501.62
9	50.32	493.45
10	57.19	508.67

8.7 T-TEST

The first prompt in the T-Test program is:

```
SAMPLE 1:
       What is the variable number?      3
       What is the first case number?    1
       What is the last case number?    10

SAMPLE 2:
       What is the variable number?      4
       What is the first case number?    1
       What is the last case number?    10
```

Choose the significance level for testing the hypotheses "Variance 1 = Variance 2" and "Mean 1 = Mean 2" by either accepting the default value of 0.05 or pressing the space bar to toggle between the choices of 0.05, 0.10 and 0.01. As these data samples are not paired, press N in response to the appropriate prompt. (To see the output generated by a paired sample, run through the T-TEST program with the data file PAIRED.)

```
Press the space bar to select an alpha value: 0.05
```

```
Are these paired observations ? Y/N
```

The output is as follows:

<u>SAMPLE ONE</u>		<u>SAMPLE TWO</u>	
Variable 3: Set 2a		Variable 4: Set 2b	
Cases 1 through 10		Cases 1 through 10	
Mean:	50.124	Mean:	500.568
Variance:	40.991	Variance:	79.698
Standard Deviation:	6.402	Standard Deviation:	8.927

<u>F-TEST FOR THE HYPOTHESIS "VARIANCE 1 = VARIANCE 2"</u>

F Value	1.9443
Numerator degrees of freedom:	9
Denominator degrees of freedom:	9
Probability:	0.3363

Result: Non-Significant F - Accept the Hypothesis

<u>T-TEST FOR THE HYPOTHESIS "MEAN 1 = MEAN 2"</u>

Pooled s squared:	60.3444
Variance of the difference between the means:	12.0689
Standard Deviation of the difference:	3.4740
t Value:	-129.6604
Degrees of freedom:	18
Probability of t:	0.0000

Result: Significant t - Reject the Hypothesis
Confidence limits for the difference of the means
(for alpha = 0.05) = 450.444
plus or minus 7.299 (443.145 through 457.743)

8.8 TABLES

Purpose: Computes means of data over several grouping variables (up to 10) for up to 5 data variables as well as calculate the standard deviations, standard error and coefficients of variation.

Before starting, you need:

1. An active data file with separate variables for all factors and replications as well as data.

Results: Depending on how you respond to prompts, the analysis may be viewed on the screen, printed or stored in a disk file. In addition, the means may be added to the end of the active data file.

Sample Data File: COMPACT.

MEAN is the only program in MSTAT-C that can list out the data side-by-side but only for one factor. All of the analysis of variance programs calculate on one data variable at a time. This program, TABLES, presents means side-by-side and for more than one factor. There are limitations. Up to ten factors can be designated and up to five data variables can be selected.

In this example, the sample data file COMPACT is used. This is an experiment that studies the effect of soil compaction levels on dry bean yield components. It is a randomized complete block split-plot design with four replications and two factors. The main plots consisted of the two compaction treatments, and the subplots consisted of the 15 dry bean varieties being tested. Since all possible treatment combinations are included in this experiment, 4 * 2 * 15 = 120 plots or cases are observed in the experiment.

Variable no.s 2 and 3 are the factors in COMPACT. Note that the levels in variable no. 3 are not numbered sequentially but can be used.

```
The data file contains 120 cases
Do you wish to use all cases? Y/N
```

Once the user has determined the number of cases to use, the grouping variables are selected. The TABLES program can handle up to 10 individual grouping variables. The COMPACT data file has 9 variables with variables 2 and 3 designated as the treatment or grouping variables for this example.

```
Choose up to 9 variables (Press ESC to quit)
 01    (NUMERIC)     REPLICATION
 02    (NUMERIC)     COMPACTION
 03    (NUMERIC)     ENTRY
 04    (NUMERIC)     10-POD LENGTH
 05    (NUMERIC)     YIELD
 06    (NUMERIC)     SEED PER POD
 07    (NUMERIC)     1000-SEED WEIGHT
 08    (NUMERIC)     PODS PER PLANT
 09    (NUMERIC)     variety
```

Highlight the variables to use (2 and 3) by moving the cursor to the variable desired then pressing the spacebar. The variable will be highlighted. Continue to highlight variables for the analysis. When all the grouping variables have been selected, press <ENTER> to enter the list. Other group variables can be used to create a variety of means tables. Be sure that at least one group or factor variable is not selected or the final table in the output will list all cases in the file with a count of one.

Next, select the data variables by the same method. Be sure that the grouping variables are not selected a second time. If a grouping variable is selected along with the data variables, an error message will appear when the list is entered stating that the group variables and the data variables overlap. Both the group and the data variables will have to be selected again. There can only be five data variables selected at a time.

```
Choose up to (5) variables (Press ESC to quit)
    01   (NUMERIC)        REPLICATION
    02   (NUMERIC)        COMPACTION
    03   (NUMERIC)        ENTRY
    04   (NUMERIC)        10-POD LENGTH
    05   (NUMERIC)        YIELD
    06   (NUMERIC)        SEED PER POD
    07   (NUMERIC)        1000-SEED WEIGHT
    08   (NUMERIC)        PODS PER PLANT
    09   (NUMERIC)        variety
```

8.8 TABLES

Highlight the data variables to use (4 , 5, 6 and 7) by moving the cursor to the variable desired then pressing the spacebar. The variable will be highlighted. Continue to highlight variables for the analysis. When all the data variables have been selected, press <ENTER> to enter the list.

Means will be generated for the various combinations of factors and levels. These "tables" of means can be appended to the end of the data file in the order that they appear on the output. If you select to have the standard deviations, standard error and coefficients of variation calculated, these additional values will not be stored in the data file.

For this example, only means will be calculated for the data values by responding <Y> to the prompt for computing the means and <N> to the remaining prompts presented.

```
Would you like to compute the means? Y/N

Would you like to compute standard deviations ? Y/N

Would you like to compute standard errors ? Y/N

Would you like to compute coefficients of variation ?
Y/N
```

Before the calculations are started, the user may choose to have the means added to the end of the data file.

Now the subprogram begins calculating and constructing tables. When TABLES has completed its calculations, the print/list menu of options is displayed. You may view the values on the screen, edit them, print them or store them in a disk file.

The output generated for this sample follows:

Mean tables for these group variables:
Variable 2: COMPACTION
Variable 3: ENTRY

And this data variable:
Variable 4: 10 POD LENGTH
Variable 5: YIELD
Variable 6: SEED PER POD
Variable 7: 1000-SEED WEIGHT

Group Variables					Data Variables			
			mean	mean	mean	mean		
2	3	:	4	5	6	7	:	Count
*	*	:	87.71	255.14	5.86	188.83	:	120

Group Variables					Data Variables			
			mean	mean	mean	mean		
2	3	:	4	5	6	7	:	Count
*	6	:	76.50	248.13	6.00	163.88	:	8
*	7	:	87.63	242.00	6.13	184.38	:	8
*	8	:	84.63	222.50	6.00	186.88	:	8
*	9	:	93.38	259.38	6.13	194.63	:	8
*	10	:	85.13	210.00	5.88	188.63	:	8
*	11	:	90.50	281.13	15.75	205.88	:	8
*	12	:	83.38	261.00	6.13	170.13	:	8
*	13	:	83.75	246.75	4.75	183.75	:	8
*	17	:	82.25	330.13	6.13	178.63	:	8
*	28	:	83.63	219.63	6.00	189.25	:	8
*	39	:	100.13	257.88	6.13	203.50	:	8
*	50	:	92.00	268.63	6.00	199.50	:	8
*	61	:	99.13	287.13	6.38	202.25	:	8
*	72	:	86.63	241.75	5.75	186.63	:	8
*	83	:	85.00	250.50	4.75	194.50	:	8

8.8 TABLES

Group Variables			Data Variables				
			mean	mean	mean	mean		
2	3	:	4	5	6	7	:	Count
1	*	:	87.43	260.85	5.93	182.93	:	60
2	*	:	87.98	249.43	5.78	194.72	:	60

Group Variables			Data Variables					
			mean	mean	mean	mean		
2	3	:	4	5	6	7	:	Count
1	6	:	77.25	267.00	6.25	160.00	:	4
1	7	:	85.50	255.50	6.25	181.25	:	4
1	8	:	84.50	227.75	6.00	179.25	:	4
1	9	:	92.25	275.75	6.25	187.50	:	4
1	10	:	84.50	214.50	5.50	190.50	:	4
1	11	:	91.50	265.50	6.00	201.00	:	4
1	12	:	84.25	248.50	6.25	158.75	:	4
1	13	:	84.25	216.50	4.75	181.75	:	4
1	17	:	83.00	328.00	6.25	177.50	:	4
1	28	:	81.00	222.25	6.00	186.00	:	4
1	39	:	100.00	250.75	6.25	194.00	:	4
1	50	:	91.00	289.50	6.00	197.00	:	4
1	61	:	100.50	330.50	6.25	195.50	:	4
1	72	:	85.50	255.25	6.25	167.25	:	4
1	83	:	86.50	265.50	4.75	186.75	:	4
2	6	:	75.75	229.25	5.75	167.75	:	4
2	7	:	89.75	228.50	6.00	187.50	:	4
2	8	:	84.75	217.25	6.00	194.50	:	4
2	9	:	94.50	243.00	6.00	201.75	:	4
2	10	:	85.75	205.50	6.25	186.75	:	4
2	11	:	89.50	298.00	5.50	210.75	:	4
2	12	:	86.50	273.50	6.00	181.50	:	4
2	13	:	83.25	277.00	4.75	185.75	:	4
2	17	:	81.50	332.25	6.00	179.75	:	4
2	28	:	86.25	217.00	6.00	192.50	:	4
2	39	:	100.25	265.00	6.00	213.00	:	4
2	50	:	93.00	247.75	6.00	202.00	:	4
2	61	:	97.75	243.75	6.50	209.00	:	4
2	72	:	87.75	228.25	5.25	206.00	:	4
2	83	:	83.50	235.50	4.75	202.25	:	4

8.9 MISVALEST

Purpose: Performs estimation of missing values in a completely randomized block or split plot type design with up to six factors in the design.

Before starting, you need:

1. The active data file must have separate variables for all factors and replications as well as data.
2. The data file must be sorted on all factor variables with the slowest varying factor as the first key, the next slowest varying factor as the second key, etc. Factor variables which are not to be used as indicators of missing values must be the outermost (first listed) keys of the sort. Appropriate cases of the file are then used.
3. Missing factor levels are not allowed (no missing values or cases within the design parameters even if the data is missing).
4. Factor level of 0 (zero) is not allowed.
5. Factor levels must be in consecutive integer order.

Results: Depending on how you respond to prompts, the analysis may be viewed on the screen, edited, printed or stored in a disk file.

Sample Data File: COMPACT. Before starting the example in this unit, enter SEDIT and delete the values in cases 20, 65, and 86 for Variable 4.

Missing values are the bane of research. They can be the result of lost samples, poor soil conditions, insect damage, breakdown of equipment, etc. With the MSTAT analysis of variance programs, the problem of missing data values can be handled in the simpler one factor designs (ANOVA-2) or by analyzing the data as a nonorthogonal design. The ANOVA-2 program will print a list of estimates for missing values then use those values in the analysis of variance calculations.

The MSTAT subprogram FACTOR will not analyze a data file with missing data and accordingly prompts the user to supply the value before calculations are performed. The MISVALEST subprogram was written to estimate a missing value to be entered at the FACTOR prompt when the design contains more than one factor in addition to replicate.

The missing values are estimated as if the experiment contains one factor (treatment) with one block (replication). Therefore, if the experiment is designed as a split-plot, the portion of the data file containing the missing value for one level of the split should be used with this program. If a data value is missing in more than one level of a split, run the program for each level.

MISVALEST Example

From the Initial MSTAT menu, go to the FILES program and open a file. For this program, open the file COMPACT. Quit the FILES program and highlight the SEDIT program and edit the data so that the values in Variable 4 for cases 20, 64 and 86 are missing. Exit SEDIT and highlight the MISVALEST program, press <ENTER>.

After the informational screen has been presented, the user will type in a list of variables that have been identified as factor variables. For this file, variables 1, 2 and 9 are used. Variable 1 is replications (1-4), Variable 2 is soil condition (non-compacted or compacted) and Variable 9 is plant variety (15 varieties total). Because the soil condition varies the slowest, it is listed first followed by the variable for variety then the variable for replicate.

```
Enter up to 6 FACTOR variables (1-9)
   List : 2, 9, 1
```

The factor variables are entered in the order of the sort keys - outermost key listed first (factor level which varies the slowest).

```
Enter the number of the factor variable representing
blocks ( 1 - 9): 1
```

Next enter the variable number for the factor that is the primary block in the experimental design (factor that varies the fastest). For the data file COMPACT, the field was blocked as replicates (Variable 1).

At the beginning of this unit, it was requested that values be deleted from Variable 4. Select this variable by highlighting it then pressing <ENTER>. If a variable is selected that does not have missing values, the program will return the user to the Initial MSTAT menu.

```
Choose up to 6 variables (Press ESC to quit)

 01    (numeric)    REPLICATION
 02    (numeric)    COMPACTION
 03    (numeric)    ENTRY
 04    (numeric)    10-POD LENGTH
 05    (numeric)    YIELD
 06    (numeric)    SEED PER POD
 07    (numeric)    1000-SEED WEIGHT
 08    (numeric)    PODS PER PLANT
 09    (numeric)    variety
```

```
Enter the variables which contain the MISSING VALUES to
be calculated
```

Once Variable 4 has been selected, accept the default value of <Y> by pressing <ENTER> to use all 120 cases of the data file COMPACT. If there are more than 120 cases in the data file, respond <N> and use only the first 120 cases.

```
The data file contains 120 cases
Do you wish to use all cases? Y/N
```

If the calculated value is to be placed in the data file at the proper cell, respond <Y> to the next prompt. If the data value is inserted into the file then there will be no means for MSTAT-C to recognize the value as being any different from any other value. If the file is then used to perform a FACTOR analysis, the degrees of freedom will not be decreased in the analysis of variance table. Respond <N> if you will be running the FACTOR program and desire that the degrees of freedom be adjusted for missing values.

```
Do you want the calculated values placed in your data file : Y/N
```

Often, the values of one variable are related to the values in another variable. By indicating the related or correlated variables, a more accurate estimation can be calculated. The variable with missing values must again be highlighted even if there are no related or correlated variables.

8.9 MISVALEST

```
Choose up to (6) variables (Press ESC to quit)

 01    (numeric)    REPLICATION
 02    (numeric)    COMPACTION
 03    (numeric)    ENTRY
 04    (numeric)    10-POD LENGTH
 05    (numeric)    YIELD
 06    (numeric)    SEED PER POD
 07    (numeric)    1000-SEED WEIGHT
 08    (numeric)    PODS PER PLANT
 09    (numeric)    variety
```

```
Enter the list of related or correlated variable
numbers to be read for this analysis.  This list must
contain Variable 4 which has the missing value.
```

Once the variable or variables have been highlighted, press <ENTER> to continue. For this example, only highlight Variable 4.

As the missing values are calculated, the program "fine tunes" the value by repetitive iterations. The user selects the maximum number of passes for the program to attempt before presenting the calculated estimate. The range of values that can be entered for the number of iterations is 1-20.

The estimated value can be calculated with accuracy up to 6 places to the right of the decimal. With large values, the calculation for this degree of accuracy would be too complex so it is suggested that an accuracy value of 1 or 2 be entered. The range of values accepted is -6 to 6.

```
Enter the maximum number of passes before iteration is
terminated : 5

Enter the accuracy desired for estimate : -2
```

The output can be viewed, edited, printed or saved to a file. Using the output, the user can use the FACTOR program to analyze the experiment and type in the estimated value when prompted.

Variable number 4
10-POD LENGTH

Convergence obtained during pass 3

Case No.	Factor Level Var. No. 2	9	1	Estimated value
20	1	5	4	84.858
65	2	2	1	85.075
86	2	7	2	85.856

9.1 ANOVA-1

Purpose: Calculates analysis of variance for a one-way comparison of means.

Before starting, you need:

1. An active data file with a group variable.
2. A STAT analysis of the data file.

Results: Depending on how you respond to prompts, the analysis may be viewed on the screen, printed or stored in a disk file. In addition, the means may be added to the end of your file.

Sample Data File: CUCUMBER.

ANOVA-1 computes a one-way analysis of variance. The group variable over which the analysis is computed may have up to 1000 different group levels; this number actually depends on how much memory you have in your computer. The group levels do not have to be consecutive, as long as the gaps between levels are not too large. The number of cases per group level does not have to be consistent, since ANOVA-1 determines the number of observations for each group and calculates the analysis accordingly.

In addition to the ANOVA table, ANOVA-1 computes and prints the number of observations, sum, average, standard deviation (SD) and standard error (SE) for each group level.

Group variable or grouping variable is a term used to denote a defined variable in MSTAT that separates out observations into units of similar characteristics. The most common group variable in agronomic research is the treatment. For more information on group or grouping variables, refer to Section 2 of this manual.

ANOVA-1 Example

From the Initial MSTAT menu, go to the FILES program and open a file. For this program, open the file CUCUMBER. Quit the FILES program and highlight the ANOVA-1 program, press <ENTER>.

In this example, the data file CUCUMBER is used. The data represent a completely randomized design with five replications. The experiment is based on comparing population yields of pickling cucumbers. Each case, or plot, consisted of 30 plants. The number of fruits per 30 plants per plot were counted for each case; variable no. 2 is this fruit number. Variable no. 1 is population; there are 11 populations in this study.

```
Enter the number of the GROUP variable ( 1 - 2)   1

Enter the lowest and highest value in the GROUP
variable
     Lowest: 1              Highest:  11
   (Press <F1> for a list of variables)
```

When the GROUP variable block is highlighted, the user can enter a value in the highlighted area then press <F1>. A list of the variables in the data file will appear on the screen. Move the cursor to the variable number desired then press <ENTER> to select that variable.

```
The data file contains 150 cases.
Do you wish to use all cases ? Y/N  N
```

Because some of the MSTAT programs add cases to data files so that the data file then contains more that the raw data, the user is prompted to indicate the range of cases to use in the calculations. If the active data file contains only the cases with the necessary information, then the user should press <RETURN> when the above prompt is presented. For this analysis, there are 5 samples within 11 different groups for a total of 55 raw data cases. Enter the values 1 and 55 for for first selected and last selected case.

```
Choose up to 1 variables (Press ESC to quit)
01 (NUMERIC) POPULATION
02 (NUMERIC) FRUIT NO
```

```
Enter the list of the DEPENDENT variable number
You may not choose the GROUP variable number
```

When the menu of the list of variables in the data file is on the screen, the user can move the cursor with the arrow keys to the variable number or numbers to be analyzed. Pressing the spacebar will highlight the choice or choices. Press <ENTER> to continue the program. In this example, variable no. 2 will be the dependent variable. If the independent or group variable number is highlighted and the spacebar pressed, a message will appear on the screen and the dependent variables will have to be highlighted again. Press <ENTER> to continue the program.

```
Do you want to store means at the end of your data file (Y/N)
N
```

If you respond "yes", then cases are added to the end of the data file (one case for each group level).

```
Variable: FRUIT No.
Do you want to perform single DF orthogonal comparisons ?
(Y/N)
```

If you do not understand the statistical procedure referenced as single DF orthogonal comparisons (contrasts), respond with "no". If you respond "yes", the next screen will appear to prompt for the coefficients for the number of treatments in the GROUP variable. The information on the screen will scroll to accommodate all the treatments necessary.

```
Enter the values for the following:
(Press <ESC> to abandon, <F10> to finish)

treatment number          coefficient
           1                       -5
           2                       -4
           3                       -3
           4                       -2
           5                       -1
           6                        0
           7                        1
           8                        2
           9                        3
          10                        4
          11                        5

(The treatment coefficients must sum to zero)
```

9.1 ANOVA-1

After entering the coefficients for the 11 treatments in the CUCUMBER data file, press <F10> to end the entry of coefficients.

```
Do you want to perform more orthogonal comparisons:(Y/N)  N
```

As there is only one independent variable in the data file, press <N> to continue with the program.

Next the print/list menu of options is displayed. You may view the analysis on the screen, print it or store it in a disk file.

Data file CUCUMBER

Function: ANOVA-1

Data case no. 1 to 55

One way ANOVA grouped over variable 1 (POPULATION)
with values from 1 to 11

Variable 2 (FRUIT NO)

A N A L Y S I S O F V A R I A N C E T A B L E

	df	Sum of Squares	M.S.	F-val	Prob.
Between	10	967.345	96.735	34.888	0.0000
Within	44	122.000	2.773		
Total	54	1089.345			

Coefficient of Variation= 8.21%

Var. 1	V A R I A B L E No.	No. 2 Sum	Ave	SD	SE
1	5.00	104.000	20.800	2.17	0.74
2	5.00	59.000	11.800	1.30	0.74
3	5.00	113.000	22.600	2.51	0.74
4	5.00	99.000	19.800	1.64	0.74
5	5.00	77.000	15.400	1.14	0.74
6	5.00	127.000	25.400	1.14	0.74
7	5.00	88.000	17.600	1.14	0.74
8	5.00	126.000	25.200	1.79	0.74
9	5.00	111.000	22.200	1.30	0.74
10	5.00	125.000	25.000	2.00	0.74
11	5.00	87.000	17.400	1.52	0.74
Total	55.00	1116.000	20.291	4.49	0.61
Within				1.67	

Bartlett's Test

Chi-square = 5.889

Number of Degrees of Freedom = 10

Approximate Significance = 0.825

9.1 ANOVA-1

ORTHOGONAL CONTRASTS

Treat.	Coeff.	Treat.	Coeff.
1	-5.00	7	1.00
2	-4.00	8	2.00
3	-3.00	9	3.00
4	-2.00	10	4.00
5	-1.00	11	5.00
6	0.00		

Sum Of Squares:	102.989
Effect:	0.433
Error:	0.071
F value:	37.144
Prob:	0.000

9.2 ANOVA-2

Purpose: Computes a two-way analysis of variance (with or without missing variables).

Before starting, you need:

1. An active data file with at least two group variables.
2. A STAT analysis of the data file.

Results: Depending on how you respond to prompts, the analysis may be viewed or edited on the screen, printed or stored in a disk file. In addition, the means over the second group variable (usually treatment levels) may be added to the end of the active data file.

Sample Data File: HURON1

ANOVA-2 is used to compute a two-way analysis of variance for experimental designs in which only one observation exists for each combination of levels in two group variables. An example of this type of experiment is a randomized complete block design with blocks (replications) as one group variable and treatments (i.e., varieties) as the second group variable. If more than one observation exists for each combination of treatments, then one of the other programs which provide an analysis of variance should be used instead. If an experiment is not a complete design and has unequal numbers of observations per combination, then use the program NONORTHO or use MEAN to compute the average and then use ANOVA-2.

Group variable or grouping variable is a term used to denote a defined variable in MSTAT that separates out observations into units of similar characteristics. The most common group variable in agronomic research is the treatment applied to a test site. Treatments may be in the form of irrigation, fertilizer, specific variety of crop, or any other independent action applied to a test site. If the study in progress concerns varietal yield then a variable named variety would be created and the various varieties to be planted would be numbered in 1-2-3 order until each specific variety had a unique value assigned. These would be the group levels of the group variable 'variety'. A second group variable can also be defined as replication and be assigned values of 1, 2, 3, .. etc within each level in the main group variable

9.2 ANOVA-2

Covariance analysis is available for a one factor experiment in the FACTOR program. If your experiment is normally analyzed in the ANOVA-2 program, you can obtain the same results using FACTOR with the RCBD 1 Factor design.

ANOVA-2 estimates the values of missing data (plots) using an iteration procedure. The ANOVA is adjusted by one degree of freedom being subtracted from the residual error degrees of freedom for each missing value. However, the actual F value and probability of significance are not properly adjusted and are incorrect if there are many missing values.

ANOVA-2 Example

From the Initial MSTAT menu, go to the FILES program and open a file. For this program, open the file HURON1. Quit the FILES program and highlight the ANOVA-2 program, press <ENTER>.

In this example, the sample data file HURON1 is used. The data are from a randomized complete block design used for a dry bean yield trial. There are four complete replications or blocks, with 30 varieties being tested. Variable no. 3 is being analyzed. In this example, use only data cases 1 to 120 (4 x 30)

```
Get Case Range
   The data file contains 120 cases.
   Do you wish to use all cases? Y/N
```

Because some of the MSTAT programs add cases to data files, the user is prompted to indicate the range of cases to use in the calculations. If the active data file contains only the cases with the necessary information (120), then the user should press <ENTER> when the above prompt is presented.

```
First group variable number:    2
Lowest level:                   1
Highest level:                  4

Second group variable number:   1
Lowest level:                   1
Highest level:                  30
```

In this example, variable no. 2 (replication) is the first group variable, with levels ranging from 1 through 4.

Variable no. 1 (entry) is the second group variable, with levels ranging from 1 through 30. After you specify the second group variable, the menu of the list of variables in the data file is on the screen. The user can move the cursor with the arrow keys to the variable number or numbers to be analyzed. Pressing the spacebar will highlight the choice or choices. Press <ENTER> to continue the program. In this example, only variable no. 3 will be the dependent variable. You may not choose the group (independent) variable numbers. If that happens, a message will appear on the screen and the dependent variables will have to be highlighted again. Press <ENTER> to continue the program.

```
Choose up to 2 Variables (Press ESC to quit)

01    (NUMERIC)    ENTRY
02    (NUMERIC)    REPLICATION
03    (NUMERIC)    YIELD IN GRAMS PER PLOT
04    (NUMERIC)    YIELD IN GROUP LEVELS
```

Besides the analysis of variance table, it is possible to have the means of the first group variable (in HURON1, this is REPLICATIONS) printed on the output as well as the means for the second group variable (in HURON1, this is ENTRY or variety).

```
Output options

Do you want to see means over the
first group variable? Y/N
Do you want to see means over the
second group variable? Y/N
Do you want to save the means over the
second group variable at the end of your
MSTAT data file?  Y/N
```

When you respond <Y> to the question of saving the means over the second group variable, the means are stored at the end of your data file. In this example, 30 cases would be added to the end of the data file.

```
Do you want to perform single DF orthogonal comparisons ? Y/N
```

If you do not understand the statistical procedure referenced as single DF orthogonal comparisons (contrasts), respond with <N>. If you respond <Y>, these additional messages are shown:

```
Enter the values for the following:
treatment number          coefficient
             1
             2
             3
             4
```

The number of prompts for treatment numbers will be determined by the highest level of the second group variable. To see output generated by this routine, look at the output for ANOVA-1.

When the calculations are complete, the print/list menu of options is displayed. You may view the analysis on the screen, print it or store it in a disk file.
The analysis generated in this example is shown below.

Data file HURON1
Title: BEAN EXPERIMENT

Function: ANOVA-2
Data case no. 1 to 120

Two-way analysis of Variance over
variable 2 (REPLICATION)with values from 1 to 4 and over
variable 1 (ENTRY) with values from 1 to 30

Variable 3 : YIELD IN GRAMS PER PLOT

ANALYSIS OF VARIANCE TABLE

	Degrees of Freedom	Sum of Squares	Mean Square	F-value	Prob
REPLICATON	3	1264900.69	421633.564	2.45	0.0686
ENTRY	29	21490397.08	741048.175	4.31	0.0000
Error	87	14947572.56	171811.179		
Non-additivity	1	46935.21	46935.214	0.27	
Residual	86	14900637.34	173263.225		
Total	119	37702870.33			

Grand Mean= 2039.675 Grand Sum= 244761.000 Total Count= 120

Coefficient of Variation= 20.32%

Means for variable 3 (YIELD IN GRAMS PER PLOT)
for each level of value of variable 2 (REPLICATION)

Var 2 Value	Var 3 Mean
1	1974.733
2	1937.100
3	2041.600
4	2205.267

9.2 ANOVA-2

**Means for variable 3 (YIELD IN GRAMS PER PLOT)
for each level of variable 1 (ENTRY)**

Var 1 Value	Var 3 Mean	Var 1 Value	Var 3 Mean	Var 1 Value	Var 3 Mean
----	-----	----	-----	----	-----
1	2163.75	11	2169.25	21	1530.00
2	1838.75	12	2028.75	22	1370.25
3	1836.25	13	912.50	23	2470.75
4	1956.00	14	1584.50	24	2124.50
5	2218.00	15	1305.25	25	1929.50
6	2271.00	16	2204.25	26	1788.00
7	2132.50	17	1788.75	27	1729.75
8	2254.00	18	2214.00	28	2389.50
9	2039.50	19	2494.75	29	2500.25
10	2520.75	20	2378.75	30	3046.50

9.3 ANOVALAT

Purpose: Computes analysis of variance on two-dimensional one-restrictional lattice designed experiments and obtain adjusted treatment means under the circumstance of an efficient lattice.

Before starting, you need:

1. An active data file with the data in a lattice design (no missing data).
2. A STAT analysis of the data file is suggested.
3. The variables in the data file **must** be in the following order:
   Variable 1 - Replication, Variable 2 - Block, Variable 3 - Treatment.
   Use ASCII to rearrange the variables if they are not in this order.

Results: An analysis is computed for the specified variables. The analysis may be viewed on the screen, printed or stored in a print file. In addition, you may have the means added to the end of the active data file.

Sample Data File: LATTEST.

ANOVALAT is designed to analyze square and rectangular lattice designs: these experimental designs may have more than three replications. An MSTAT data file for any lattice design may be analyzed. All factors must be consecutive numeric codes. For example, if 16 varieties are entered, the treatment codes should range from 1 through 16. If any gaps exist in the codes, the missing codes will be treated as missing values.

ANOVALAT asks for the number of arrangements, number of replications and the block size. After specifying a variable list, the program presents the option of storing the means over treatments in the active data file (the data would be added to the end of your file). When ANOVALAT completes its calculations, the print/list menu of options is displayed.

ANOVALAT Example

The sample data file used in this example is LATTEST; this is a square lattice design with three arrangements, three replications and four blocks. Variable no. 9 (yield in grams/plot) will be analyzed.

From the Initial MSTAT menu, go to the FILES program and open a file. For this program, open the file LATTEST. Quit the FILES program and highlight the ANOVALAT program then press <ENTER>.

9.3 ANOVALAT

A message screen will state the reuirement that variables 1, 2 and 3 must be replication, block and treatment in that order. If your data file was created by VARPLAN, the variables will be in the correct order.

If the variables in your data file are in a different order, use the ASCII program to rearrange the variables to correspond to this order.

```
The Replication, Block and Treatment variables must be
present in the following order:
   Variables:  1=Replication, 2=Block, 3=Treatment.

If the variables in your file are in a different order, use
ASCII to rearrange the variables to correspond to this order.
Press <ESC> at this time if variables need to be rearranged.

Note: if the data file was created by VARPLAN, the variables
are in the correct order.

Press <ESC> to abort or any other key to continue
```

The number of cases of the MSTAT data file to be used is calculated from three values: the number of arrangements, the number of replications and the block size. The structure of the data file is very important to the correct operation of the analysis.

1) The number of arrangements is the number of replications (treatment tables) in the design (non-replicated).
2) The number of replications is the total number of replications in the design. This value should be the maximum value in the replication variable.
3) The block size is the number of rows in each treatment table in the design. Using *n* as the block size, we have:
   for a square design, the number of treatments = *n* x *n*.
      (for example: 4 x 4, 5 x 5, 9 x 9)
   for a rectangular design, the number of treatments = *n* x (*n*+1)
      (for example: 4 x 5, 5 x 6)

```
Enter the lattice design used in experiment :   1
   1) Square lattice
   2) Rectangular lattice
```

For this data file (LATTEST) there are three arrangements, three replications and the block size size is 4. As can be seen from the following screen, the ANOVALAT program is capable of handling up to 10 arrangements, up to 50 replications and a block size of up to 40.

```
Enter the number of arrangements used in design (1 - 10) : 3
Enter the number of replications used in design (1 - 50) : 3
Enter the block size used in design ( 1- 40)             : 4
```

For this data file, each replication has a different arrangement of the plots within that replication. If the three replications were duplicated for a total of six replications, there would still be only three arrangements.

In this example, the cases containing the data to be analyzed starts with case 1. Because some of the MSTAT data files contain data from more than one experiment (such as other locations, other years), the user is prompted to indicate the first case number to use in the calculations. If the active data file contains only the cases with the necessary information, then the user should type in 1 when the above prompt is presented. Based on the number of cases that would have to be used in the analysis to meet the design stated above (3 arrangements, 3 replications and 4 blocks), the program presents the range of cases in the data file that could be used as the first case.

```
Enter the number of the first case (1 - 65) :   1
```

```
Do you want to print the unadjusted treatment means : Y/N
```

The user has the option to print only the adjusted means for the treatments or to print both the unadjusted and the adjusted means if both are calculated. For this example, print the unadjusted means by pressing <ENTER> when the cursor is flashing on the Y. Both the unadjusted means and the adjusted means will be printed. If no adjustments are made, unadjusted means will be printed whether requested or not.

```
Do you want to sort the treatment means for printing : Y/N
```

The adjusted means can be presented in sorted order from highest value to lowest value (decending) or from lowest value to highest value (ascending). Depending on the purpose of the experiment, the user will sometimes use an ascending sort; sometimes a decending sort order. For this example, the decending sorted order will be chosen as the quantity being analyzed is yield capabilities of specific varieties and the highest yielding variety is being sought.

```
Do you want an ascending sort : Y/N
```

```
The treatment means will be sorted in decending order
```

```
Do you want to store the means over treatments : Y/N
```

If further analysis of the mean values will be performed (LSD, Duncan's etc) respond <Y> and the means will be appended to the end of the data file. The option to store the data in any variable or to create a new variable to store the means will be presented if this option is chosen.

The variable list is presented along with the instructions for selecting a variable or a group of variables to be analyzed by the ANOVALAT program. Do not highlight the treatment, replication or block variables. Only highlight the raw data collected during the experiment or the calculated data generated within the experiment. For this example, move the cursor to variable no. 9 and press <ENTER>.

When ANOVALAT completes its calculations, the print/list menu of options is displayed.

Data file LATTEST
Title: NEW LATTICE DESIGN GENERATED BY VARPLAN

Function: ANOVALAT
Data case no. 1 to 48

Variable 9
YIELD GMS/PLOT

ANALYSIS OF VARIANCE TABLE
For Square Lattice Design

	Degrees of Freedom	Sum of Squares	Mean Square	F-value	Prob
Replications	2	1047586.167	523793.083		
Treatments					
-Unadjusted	15	666541.917	44436.128	0.56	0.881
Blocks within					
Reps (adj.)	9	322182.333	35798.037		
Error					
-RCD Design	30	2376751.833	79225.061		
-Intrablock	21	2054569.500	97836.643		
Total	47	4090879.917			

Efficiency of Lattice : Less efficient than Randomized Complete Block design. No adjustments made to treatment means

Grand Sum = 90578.00 Grand Mean= 1887.0417 Total Count= 48

Coefficient of Variation= 14.9159 percent

Least Significant Differences
P = 0.05 LSD = 469.3527
P = 0.01 LSD = 632.0009

9.3 ANOVALAT

TABLE OF UNADJUSTED MEANS

Treatment Number	Treatment Mean
1	1610.000
2	1995.000
3	1916.667
4	1823.667
5	1801.667
6	1862.667
7	1874.667
8	1809.333
9	2043.667
10	1897.667
11	1767.667
12	2079.667
13	1835.000
14	2009.667
15	1841.000
16	2024.667

No Adjustments were made to treatment means

TABLE OF UNADJUSTED MEANS
Sorted on Variable 9 in Descending Order

Treatment Number	Treatment Mean
12	2079.667
9	2043.667
16	2024.667
14	2009.667
2	1995.000
3	1916.667
10	1897.667
7	1874.667
6	1862.667
15	1841.000
13	1835.000
4	1823.667
8	1809.333
5	1801.667
11	1767.667
1	1610.000

No Adjustments were made to treatment means

9.4 FACTOR

Purpose: Computes an analysis of variance for split-plot and factorial experiments.

Before starting, you need:

1. An active data file with separate variables for all factors and replications with one by one incrementaion in the factor levels.
2. A STAT analysis of the data file.

Results: Depending on how you respond to prompts, the analysis may be viewed on the screen, edited, printed or stored in a disk file. In addition, the means may be added to the end of the active data file.

Sample Data File: COMPACT.

FACTOR computes a factorial analysis of variance for two to six factors. The subprogram can also analyze all types of split-plot designs with up to four splits. Confounded designs or designs with unequal observations per treatment combination can not be completely analyzed by FACTOR, but the sums of squares needed to complete an analysis are computed.

FACTOR will not compute missing values. If there are missing values, estimate them using ANOVA-2, NONORTHO or MISVALEST prior to using FACTOR. When FACTOR detects a missing value, a message is displayed along with a prompt for the estimated value; you may then insert the missing value. For each missing value, the ANOVA final error degrees of freedom are reduced by one. Since there may be more than one error term in a factorial analysis, the final error term is always the last one presented in the table. If split-plot designs have missing values, neither NONORTHO or ANOVA-2 will give an estimated value taking into effect the main plot or split-plot bias accounted for in the estimation iteration. The standard deviation values following the analysis of variance table will all be calculated as if the design is not split due to the extreme complexity encountered with splits.

FACTOR will also do covariance analysis. This procedure is applicable wherever a treatment sum of squares is to be partitioned into main effects and interaction components. For example, F-tests of the three null hypothesis are made with the adjusted error mean square in the denominator. This procedure could be used in studies with varied plant populations using yield as the covariate.

If the data file contains duplicate treatment combinations in a replication, the subprogram aborts (ie: if a file has two cases with factor A level as 2 by factor B level as 3).

FACTOR initially displays a design menu from which can be chosen your design. The last menu option will allow you to enter K-values for a design not presented on this screen (explanation for this procedure follows the FACTOR example). A text description of each design is displayed at the top of the screen as the cursor is moved. When a design is selected, a sample format is displayed before FACTOR begins calculations. When calculations are complete, the print/list menu of options is displayed.

FACTOR Example

In this example, the sample data file COMPACT is used. This is an experiment that studies the effect of soil compaction levels on dry bean yield components. It is a randomized complete block split-plot design with four replications and two factors. The main plots consisted of the two compaction treatments, and the subplots consisted of the 15 dry bean varieties being tested. Since all possible treatment combinations are included in this experiment, 4 * 2 * 15 = 120 plots or cases are observed in the experiment.

Variable no.s 2 and 9 are the factors in COMPACT. Variable no. 9 was actually created from variable no. 3 using the TABTRANS subprogram; the codes in variable no. 3 were not numbered sequentially and, therefore, could not be used.

```
Would you like to do covariance analysis? Y/N
```

If you are not familiar with statistical procedure of covariance analysis, respond <N>. If you choose to do covariance analysis, the variable to use as the covariate will be requested after the listing of variables for the factors. The table of means that is calculated will contain the mean value of the covariate as well as the unadjusted and the adjusted means of the variable being analyzed. The FACTOR Design Menu follows.

```
 1. CRD 2 Factor (a)             19. RCBD 2 Factor Comb (a)
 2. CRD 2 Factor (b)             20. RCBD 2 Factor Comb (b)
 3. CRD 3 Factor (a)             21. RCBD 2 Factor Comb (c)
 4. CRD 3 Factor (b)             22. RCBD 2 Factor Comb (d)
 5. CRD 3 Factor (c)             23. RCBD 2 Factor Comb (e)
 6. CRD 4 Factor                 24. RCBD 2 Factor Comb (f)
 7. RCBD 1 Factor                25. RCBD 2 Factor Comb (g)
 8. RCBD 2 Factor (a)            26. RCBD 2 Factor Comb (h)
 9. RCBD 2 Factor (b)            27. RCBD 3 Factor Comb (a)
10. RCBD 3 Factor (a)            28. RCBD 3 Factor Comb (b)
11. RCBD 3 Factor (b)            29. RCBD 4 Factor Comb (a)
12. RCBD 3 Factor (c)            30. RCBD 4 Factor Comb (b)
13. RCBD 3 Factor (d)            31. RCBD 2 Factor Strip
14. RCBD 4 Factor                32. RCBD 3 Factor Strip
15. RCBD 1 Factor Comb (a)       33. RCBD 4 Factor Strip
16. RCBD 1 Factor Comb (b)       34. RCBD 4 Factor Strip
17. RCBD 1 Factor Comb (c)       35. Other Type of Design
18. RCBD 1 Factor Comb (d)
```

```
Factor: ANOVA Table for this model

K Value Source   Degrees of Freedom           Is this what
──────────────────────────────────────────── you had in
  1  Replication     r-1                      mind? Y/N
  2  Factor A        a-1
 -3  Error           (r-1)(a-1)
  4  Factor B        b-1
  6  AB              (a-1)(b-1)
 -7  Error           a(r-1)(b-1)
```

The design that is chosen determines the order that the variables are entered. In the above design, the first K Value of 1 is associated with the replicate or block of the design. The K Value of 2 is associated with the Main Effect or Factor A. The K Value of 4 is associated with Factor B or the Sub Effect of the experiment. In those situations where data has been combined over year and/or location, the ANOVA Table for the model will assist the user in determining the order to enter the factors.

If the user enters <N>, the FACTOR design menu is presented and the user can select a different design. Following the order of the model chosen, the user is prompted for the factors and their levels as follows:

9.4 FACTOR

```
FACTOR: Enter data for the first variable (Replication)
Enter the desired Variable Number:          1
Enter the lowest level for this Variable:   1
Enter the highest level for this Variable:  4
```

```
FACTOR: Enter data for the second variable (Factor A)
Enter the desired Variable Number:          2
Enter the lowest level for this Variable:   1
Enter the highest level for this Variable:  2
```

```
FACTOR: Enter data for the third variable (Factor B)
   Enter the desired Variable Number:          9
   Enter the lowest level for this Variable:   1
   Enter the highest level for this Variable:  15
```

A list of variables can be pulled down by pressing <F1> after filling in a value in the variable number box. The correct variable can then be highlighted and will be used by the program. In this example, the first entry (1,1,4) is variable no. 1 (replication) which has values 1 through 4 for four replications. The second entry is variable no. 2 (compaction) which has values 1 and 2. The third entry is variable no. 9 (variety) which has values ranging from 1 through 15. Note that the first variable entered is for replications.

As a checking mechanism, the following table is presented:

```
FACTOR: Selected Variables

  Number of Factors:    3
                                         Lowest  Highest
Variable Description      Anova Use      Level   Level
1     REPLICATION       Replication        1      4
2     COMPACTION        Factor A           1      2
3     variety           Factor B           1      15

      Is this correct?  Y/N
```

By pressing <ENTER>, the information shown is accepted as correct. If the incorrect variables are listed, press <N> and reenter the data for the three variables again. Once the correct values are shown and accepted, the program prompts for the range of cases then the variables to be analyzed. Since all factors are being used in to analysis, the total number of cases is 4 * 2 * 15 or a total of 120 cases. Note that the number of variables to be chosen is three less than the number of variables in the data file. The user is not to select any of the three variables that have been used as the independent grouping variables (in this case, variables 1, 2, and 9).

```
The data file contains 120 cases
Do you wish to use all cases? Y/N
```

```
Choose up to (6) variables (Press ESC to quit)

  01    (NUMERIC)        REPLICATION
  02    (NUMERIC)        COMPACTION
  03    (NUMERIC)        ENTRY
  04    (NUMERIC)        10-POD LENGTH
  05    (NUMERIC)        YIELD
  06    (NUMERIC)        SEED PER POD
  07    (NUMERIC)        1000-SEED WEIGHT
  08    (NUMERIC)        PODS PER PLANT
  09    (NUMERIC)        variety
```

Highlight the variables to use by moving the cursor to the variable desired then pressing the spacebar. The variable will be highlighted. Continue to highlight variables for the analysis. When all the dependent variables have been selected, press <ENTER> to enter the list. For this example, highlight variables 4, 6, 7 and 8.

```
Do you want all means stored at the end of your file?
Y/N
```

If the default of <Y> is chosen by pressing <ENTER>, the means are added to the end of the active data file. If you have just finished entering K-values, the above prompt is the next prompt displayed by the subprogram.

Now the subprogram begins calculating and constructing tables. When FACTOR has completed its calculations, the print/list menu of options is displayed. You may view the analysis on the screen, print them or store them in a disk file.

The factorial analysis generated for variable 8 follows:

Data file COMPACT
Title: SOIL COMPACTION

Function: FACTOR
Experiment Model Number 9:
Randomized Complete Block Design for Factor A, with
Factor B a Split plot on A
Data Case no.1 to 120

Factorial ANOVA for the factors:
Replication (Var 1:REPLICATION) with values from 1 to 4
Factor A (Var 2: COMPACTION) with values from 1 to 2
Factor B (Var 9: variety) with values from 1 to 15

Variable 8: PODS PER PLANT

Grand Mean= 23.400 Grand Sum= 2808.000 Total Count= 120

TABLE OF MEANS

1	2	9	8	Total
		THE MEANS OF FOUR REPLICATIONS		
1	*	*	22.400	672.000
2	*	*	22.567	677.000
3	*	*	22.600	678.000
4	*	*	26.033	781.000
		THE MEANS OF TWO COMPACTION LEVELS		
*	1	*	23.867	1432.000
*	2	*	22.933	1376.000
		THE MEANS OF FIFTEEN VARIETIES		
*	*	1	25.750	206.000
*	*	2	21.500	172.000
*	*	3	20.375	163.000
*	*	4	21.500	172.000
*	*	5	18.375	147.000
*	*	6	23.750	190.000
*	*	7	24.625	197.000
*	*	8	27.875	223.000
*	*	9	29.125	233.000
*	*	10	19.625	157.000
*	*	11	21.000	168.000
*	*	12	22.750	182.000
*	*	13	22.750	182.000
*	*	14	23.125	185.000
*	*	15	28.875	231.000

			MEANS FOR EACH VARIETY AT EACH COMP. LEVEL	
*	1	1	27.000	108.000
*	1	2	22.750	91.000
*	1	3	21.500	86.000
*	1	4	23.250	93.000
*	1	5	19.000	76.000
*	1	6	21.750	87.000
*	1	7	23.750	95.000
*	1	8	22.250	89.000
*	1	9	29.000	116.000
*	1	10	19.750	79.000
*	1	11	21.750	87.000
*	1	12	25.000	100.000
*	1	13	27.000	108.000
*	1	14	23.750	95.000
*	1	15	30.500	122.000
*	2	1	24.500	98.000
*	2	2	20.250	81.000
*	2	3	19.250	77.000
*	2	4	19.750	79.000
*	2	5	17.750	71.000
*	2	6	25.750	103.000
*	2	7	25.500	102.000
*	2	8	33.500	134.000
*	2	9	29.250	117.000
*	2	10	19.500	78.000
*	2	11	20.250	81.000
*	2	12	20.500	82.000
*	2	13	18.500	74.000
*	2	14	22.500	90.000
*	2	15	27.250	109.000

ANALYSIS OF VARIANCE TABLE

K Value	Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Prob
1	Replication	3	278.067	92.689	11.9513	.0356
2	Factor A	1	26.133	26.133	3.3696	.163
-3	Error	3	23.267	7.756		
4	Factor B	14	1219.800	87.129	3.9697	.000
6	AB	14	541.867	38.705	1.7634	.0581
-7	Error	84	1843.667	21.948		
	Total	119	3932.800			

Coefficient of Variation: 20.02%

$s_{\bar{y}}$ for means group 1: .5084 Number of obs: 30

$s_{\bar{y}}$ for means group 2: .3595 Number of obs: 60

$s_{\bar{y}}$ for means group 4: 1.6564 Number of obs: 8

$s_{\bar{y}}$ for means group 6: 2.3425 Number of obs: 4

If you have chosen to do the covariance analysis, the printed output will include the following statement: (Variable A is the variable that is entered to be analyzed, Variable B is the covariate variable)

```
The numbers in the "Unadjusted A" column are the means
based on the actual data in the file.  The "Adjusted
A" column contains the means adjusted with the
following formula:  Adj Yj = Unadj Yj - b(Xj-X) where
Adj Yj is the adjusted Variable A mean, Unadj Yj is the
unadjusted Variable A mean, Xj is the Variable B mean
for that treatment combination, X is the Variable B
grand mean and b is Exy/Exx (the slope of the data).
```

An additional table will be printed containing the unadjusted sums of products (Var. BxVar. B, Var. BxVar. A, Var. AxVar. A). The analysis of covariance table will be inserted for the analysis of variance table with a source line for the covariate degrees of freedom, sum of squares, mean square and F value.

9.4 FACTOR

There are four new designs on the menu that are too large to present on the screen in the FACTOR program. They are entered here in their entirety.

27. RCBD 3 Factor factorial combined over locations and years, new location each year.

K Value	Source	Degrees of Freedom
1	Y	y-1
3	L(Y)	y(l-1)
7	R(LY)	ly(r-1)
8	A	(a-1)
9	YA	(y-1)(a-1)
11	LA(Y)	y(l-1)(a-1)
16	B	(b-1)
17	YB	(y-1)(b-1)
19	LB(Y)	y(l-1)(b-1)
24	AB	(a-1)(b-1)
25	YAB	(y-1)(a-1)(b-1)
27	LAB(Y)	y(l-1)(a-1)(b-1)
32	C	(c-1)
33	YC	(c-1)(y-1)
35	LC(Y)	(c-1)y(l-1)
40	AC	(c-1)(a-1)
41	YAC	(c-1)(y-1)(a-1)
43	LAC(Y)	(c-1)y(l-1)(a-1)
48	BC	(c-1)(b-1)
49	YBC	(c-1)(y-1)(b-1)
51	LBC(Y)	(c-1)y(l-1)(b-1)
56	ABC	(c-1)(a-1)(b-1)
57	YABC	(c-1)(y-1)(a-1)(b-1)
59	LABC(Y)	(c-1)y(l-1)(a-1)(b-1)
**	Error	by subtraction

28. RCBD 3 Factor factorial combined over locations and years, same location but randomized each year

K Value	Source	Degrees of Freedom
1	Y	y-1
2	L	l-1
3	YL	(y-1)(l-1)
7	R(LY)	ly(r-1)
8	A	(a-1)
9	YA	(y-1)(a-1)
10	LA	(l-1)(a-1)
11	YLA	(y-1)(l-1)(a-1)
16	B	(b-1)
17	YB	(y-1)(b-1)
18	LB	(l-1)(b-1)
19	YLB	(y-1)(l-1)(b-1)
24	AB	(a-1)(b-1)
25	YAB	(y-1)(a-1)(b-1)
26	LAB	(l-1)(a-1)(b-1)
27	YLAB	(y-1)(l-1)(a-1)(b-1)
32	C	(c-1)
33	YC	(c-1)(y-1)
34	LC	(c-1)(l-1)
35	YLC	(c-1)(y-1)(l-1)
40	AC	(c-1)(a-1)
41	YAC	(c-1)(y-1)(a-1)
42	LAC	(c-1)(l-1)(a-1)
43	YLAC	(c-1)(y-1)(l-1)(a-1)
48	BC	(c-1)(b-1)
49	YBC	(c-1)(y-1)(b-1)
50	LBC	(c-1)(l-1)(b-1)
51	YLBC	(c-1)(y-1)(l-1)(b-1)
56	ABC	(c-1)(a-1)(b-1)
57	YABC	(c-1)(y-1)(a-1)(b-1)
58	LABC	(c-1)(l-1)(a-1)(b-1)
59	YLABC	(c-1)(y-1)(l-1)(a-1)(b-1)
**	Error	by subtraction

9.4 FACTOR

29. RCBD 4 Factor factorial combined over locations and years, new location each year

K Value	Source	Degrees of Freedom
1	Y	y-1
3	L(Y)	y(l-1)
7	R(LY)	ly(r-1)
8	A	(a-1)
9	YA	(y-1)(a-1)
11	LA(Y)	y(l-1)(a-1)
16	B	(b-1)
17	YB	(y-1)(b-1)
19	LB(Y)	y(l-1)(b-1)
24	AB	(a-1)(b-1)
25	YAB	(y-1)(a-1)(b-1)
27	LAB(Y)	y(l-1)(a-1)(b-1)
32	C	(c-1)
33	YC	(c-1)(y-1)
35	LC(Y)	(c-1)y(l-1)
40	AC	(c-1)(a-1)
41	YAC	(c-1)(y-1)(a-1)
43	LAC(Y)	(c-1)y(l-1)(a-1)
48	BC	(c-1)(b-1)
49	YBC	(c-1)(y-1)(b-1)
51	LBC(Y)	(c-1)y(l-1)(b-1)
56	ABC	(c-1)(a-1)(b-1)
57	YABC	(c-1)(y-1)(a-1)(b-1)
59	LABC(Y)	(c-1)y(l-1)(a-1)(b-1)
64	D	(d-1)
65	YD	(d-1)(y-1)
67	LD(Y)	(d-1)y(l-1)
72	AD	(d-1)(a-1)
73	YAD	(d-1)(y-1)(a-1)
75	LAD(Y)	(d-1)y(l-1)(a-1)
80	BD	(d-1)(b-1)
81	YBD	(d-1)(y-1)(b-1)
83	LBD(Y)	(d-1)y(l-1)(b-1)
88	ABD	(d-1)(a-1)(b-1)
89	YABD	(d-1)(y-1)(a-1)(b-1)
91	LABD(Y)	(d-1)y(l-1)(a-1)(b-1)
96	CD	(d-1)(c-1)
97	YCD	(d-1)(c-1)(y-1)
99	LCD(Y)	(d-1)(c-1)y(l-1)

104	ACD	(d-1)(c-1)(a-1)
105	YACD	(d-1)(c-1)(y-1)(a-1)
107	LACD(Y)	(d-1)(c-1)y(l-1)(a-1)
112	BCD	(d-1)(c-1)(b-1)
113	YBCD	(d-1)(c-1)(y-1)(b-1)
115	LBCD(Y)	(d-1)(c-1)y(l-1)(b-1)
120	ABCD	(d-1)(c-1)(a-1)(b-1)
121	YABCD	(d-1)(c-1)(y-1)(a-1)(b-1)
123	LABCD(Y)	(d-1)(c-1)y(l-1)(a-1)(b-1)
**	Error	by subtraction

30. RCBD 4 Factor factorial combined over locations and years, same location but randomized each year

K Value	Source	Degrees of Freedom
1	Y	y-1
2	L	l-1
3	YL	(y-1)(l-1)
7	R(LY)	ly(r-1)
8	A	(a-1)
9	YA	(y-1)(a-1)
10	LA	(l-1)(a-1)
11	YLA	(y-1)(l-1)(a-1)
16	B	(b-1)
17	YB	(y-1)(b-1)
18	LB	(l-1)(b-1)
19	YLB	(y-1)(l-1)(b-1)
24	AB	(a-1)(b-1)
25	YAB	(y-1)(a-1)(b-1)
26	LAB	(l-1)(a-1)(b-1)
27	YLAB	(y-1)(l-1)(a-1)(b-1)
32	C	(c-1)
33	YC	(c-1)(y-1)
34	LC	(c-1)(l-1)
35	YLC	(c-1)(y-1)(l-1)
40	AC	(c-1)(a-1)
41	YAC	(c-1)(y-1)(a-1)
42	LAC	(c-1)(l-1)(a-1)
43	YLAC	(c-1)(y-1)(l-1)(a-1)
48	BC	(c-1)(b-1)
49	YBC	(c-1)(y-1)(b-1)

9.4 FACTOR

50	LBC	(c-1)(l-1)(b-1)
51	YLBC	(c-1)(y-1)(l-1)(b-1)
56	ABC	(c-1)(a-1)(b-1)
57	YABC	(c-1)(y-1)(a-1)(b-1)
58	LABC	(c-1)(l-1)(a-1)(b-1)
59	YLABC	(c-1)(y-1)(l-1)(a-1)(b-1)
64	D	(d-1)
65	YD	(d-1)(y-1)
66	LD	(d-1)(l-1)
67	YLD	(d-1)(y-1)(l-1)
72	AD	(d-1)(a-1)
73	YAD	(d-1)(y-1)(a-1)
74	LAD	(d-1)(l-1)(a-1)
75	YLAD	(d-1)(y-1)(l-1)(a-1)
80	BD	(d-1)(b-1)
81	YBD	(d-1)(y-1)(b-1)
82	LBD	(d-1)(l-1)(b-1)
83	YLBD	(d-1)(y-1)(l-1)(b-1)
88	ABD	(d-1)(a-1)(b-1)
89	YABD	(d-1)(y-1)(a-1)(b-1)
90	LABD	(d-1)(l-1)(a-1)(b-1)
91	YLABD	(d-1)(y-1)(l-1)(a-1)(b-1)
96	CD	(d-1)(c-1)
97	YCD	(d-1)(c-1)(y-1)
98	LCD	(d-1)(c-1)(l-1)
99	YLCD	(d-1)(c-1)(y-1)(l-1)
104	ACD	(d-1)(c-1)(a-1)
105	YACD	(d-1)(c-1)(y-1)(a-1)
106	LACD	(d-1)(c-1)(l-1)(a-1)
107	YLACD	(d-1)(c-1)(y-1)(l-1)(a-1)
112	BCD	(d-1)(c-1)(b-1)
113	YBCD	(d-1)(c-1)(y-1)(b-1)
114	LBCD	(d-1)(c-1)(l-1)(b-1)
115	YLBCD	(d-1)(c-1)(y-1)(l-1)(b-1)
120	ABCD	(d-1)(c-1)(a-1)(b-1)
121	YABCD	(d-1)(c-1)(y-1)(a-1)(b-1)
122	LABCD	(d-1)(c-1)(l-1)(a-1)(b-1)
123	YLABCD	(d-1)(c-1)(y-1)(l-1)(a-1)(b-1)
**	Error	by subtraction

Entering K-Values

Since most experimental designs are listed in the FACTOR experimental design menu, you will not usually need to be concerned with entering K-values. However, in the unlikely event that you not find your design in the menu, select option 35 (Other Type of Design) from the menu and enter K-values. K-values designate which sources of variation should be allocated as factor main effects or interactions, which should be pooled and which should be allocated as error terms. For example, COMPACT consists of a replication variable and two factors (A and B). The K-values should be allocated as K=1 (rep), 2(A), 4(B), 6(A x B).

1. The K-value for an interaction is the sum of the K-values for each involved factor main effect (e.g. AB=3, BC=6, ABC=7, etc.).
2. K-values should be entered in ascending order:

Effect=	A	B	AB	C	AC	BC	ABC	D	AD	BD
K-value	1	2	3	4	5	6	7	8	9	10

Effect=	ABD	CD	ACD	E	F	ABCDEF
K-value	11	12	13	16	32	63

3. If a K-value is omitted for a main effect, the degrees of freedom and the sums of squares for the omitted factor will be included in the next interaction. For example, if B is nested in A, as occurs in multi-environmental studies, the B and AB sums of squares and degrees of freedom may be pooled by specifying K- values of 1 (A) and 3(A x B) for the analysis. Normally, a nested DF is based on the sum of the Rep DF and E x R interaction DF. The combination of these two sources of variation is automatically combined by not listing the Rep K-value.

4. If an interaction is going to be used for the error of the sources of variation above it in the table (as in a split-plot design), a minus (-) sign must be used with the K-value.

5. The last K-value in a table, invariably an interaction, should not be entered. The resulting interaction will automatically pool all nonassigned sources of error. For example, in factorial designs with 3, 4, 5 and 6 factors, the last K-value of each should not be entered (7, 15, 31, or 63, respectively).

6 To include all positive K-values for a given design automatically, enter the range of K-values from the lowest (x) to the highest (y), as x,y format. For example, all K-values for a five factor experiment can be entered as 1,30<ENTER> at the K value prompt.

If a special pooling needs to be performed for the ANOVA to match the research design, and you are not sure of the proper K-value, then have all of the effects analyzed separately using the 1,y format described under #6. The x values for the possible factorial designs (including reps) analyzable by FACTOR are as follows:

factor level-	3	4	5	6
K-value-	06	14	30	62

The required pooling of the degrees of freedom and the sums of squares can then be done by hand. In the following example, the K-values for COMPACT are entered. This is actually unnecessary since the exact experimental design is available in a menu. However, it will help explain how you enter K-values. COMPACT is a split-plot design with two factors in addition to replications. The actual sources of variation and degrees of freedom of a traditional ANOVA table based on this design are as follows:

Source	DF	K-value
total	119	—
reps	3	1
factor A	1	2
error a	3	-3
factor B	14	4
A x B	4	6
error b	84	-7

The final error term (error b), which is not entered, will consist of the rep x B and rep x A x B interaction. Error a is used to test the significance of the A factor main effect, and the error b is used to test the significance of the B factor and the A x B interaction.

After you respond to initial prompts for variable numbers and factor levels, the subprogram will display a menu of experimental designs. If you do not find your design in the menu, you will select option 31 to enter K-values. The following prompts will then be displayed:

```
Factor: Number of Factors
   Please enter the number of Factors (including
      Replications) in the Analysis:    3
```

```
FACTOR: Enter data for the first variable (Factor A)
   Enter the desired Variable Number:          1
   Enter the lowest level for this Variable:   1
   Enter the highest level for this Variable:  4
```

```
FACTOR: Enter data for the second variable (Factor B)
   Enter the desired Variable Number:          2
   Enter the lowest level for this Variable:   1
   Enter the highest level for this Variable:  2
```

```
FACTOR: Enter data for the third variable (Factor C)
   Enter the desired Variable Number:          9
   Enter the lowest level for this Variable:   1
   Enter the highest level for this Variable:  15
```

```
FACTOR: Enter Your K Values (Enter 0 to End)


   K Value Source  Degrees of freedom
```

As you enter each K value, the program will fill in the information for source and degrees of freedom. To enter a range of K values, enter the lowest and highest numbers, separated by a comma, for that range. Enter the value 0 (zero) to end this process.

9.4 FACTOR

```
Factor: ANOVA Table for this model

K Value Source   Degrees of Freedom
-------------------------------------
    1      Factor A     ar-1
    2      Factor B     b-1
   -3      Error      (a-1)(b-1)
    4      Factor C     c-1
    6      B C        (b-1)(c-1)
   -7      Error     b(a-1)(c-1)

Factor: Selected Variables
   Number of Factors:   3
                                         Lowest  Highest
  Variable Description      Anova Use   Level   Level
1     REPLICATION           Factor A    1       4
2     COMPACTION            Factor B    1       2
3     variety               Factor C    1       15

     Is this correct? Y/N
```

The prompts then continue as in the preceding FACTOR example.

9.5 HIERARCH

Purpose: Computes a hierarchical analysis of variance and creates a table with numbers and averages over each group and subgroup variable.

Before starting, you need:

1. An active data file with group variables.
2. A STAT analysis of the data file.

Results: A hierarchical analysis of variance is computed for the selected variables. Depending on which print/list options are selected, the analysis may be viewed on the screen, printed or stored in a disk file.

Sample Data File: HIERTEST.

HIERARCH computes a hierarchical analysis of multiple classifications with samples within samples. The calculations for the sums of squares follow the same pattern described in earlier programs. This type of hierarchical classification can be calculated on from two to six groupings of samples within samples. The most complex design would be similar to the following example:

States	Group A
Counties within States	Group B
Fields within Counties	Group C
Acres within Fields	Group D
Plants within Fields	Group E
Leaves within Plants	Group F

The program asks for the number of hierarchical group variables to analyze then the group variable number, and the highest and lowest levels within each group; begin with the innermost group (in the above example, start with Group F). After entering a list of the dependent variables to be analyzed, HIERARCH begins its calculations. When its calculations are complete, the print/list menu of options is displayed.

HIERARCH Example

The sample data file used in this example is HIERTEST. The data for this file are from a horticultural grafting project involving cherry tree clones.

9.5 HIERARCH

The data are for three clones; two trees were sampled in each of the three clones; 20 subsamples were collected on each tree (20 within 2 within 3). The numerous characters studied are included in the file: harvest pull force, fruit diameter, fruit length, fruit firmness, percent soluble solids, pit length, pit width, fruit dry weight and pit dry weight.

From the Initial MSTAT menu, go to the FILES program and open a file. For this program, open the file HIERTEST. Quit the FILES program and highlight the HIERARCH program, press <ENTER>.

Two hierarchical grouping variables will be analyzed. In the example, the group variables are variable no. 4 (tree) and variable no. 3 (clone). The dependent variable is variable no. 6 (pull force).

```
Number Hierarchical Grouping Variables :   2
```

```
For each grouping give the variable number,
and lowest and highest values
```

Be sure to start with the innermost grouping. By entering a variable number in the box then pressing <F1>, a list of variables in the file can be pulled down and the correct variable chosen.

```
Group 1
Variable:  4     Lowest: 1     Highest:   2
```

```
Group 2
Variable:  3     Lowest: 1     Highest:   3
```

The innermost variable is variable no. 4 which has group levels of 1 and 2; then variable no. 3 which has values ranging from 1 through 3.

```
Chose up to 14 variables (Press ESC) to quit

01   (NUMERIC)    Line Number
02   (NUMERIC)    Year
03   (NUMERIC)    Clone
04   (NUMERIC)    Tree
05   (NUMERIC)    Sample
06   (NUMERIC)    Pull Force
07   (NUMERIC)    Fruit Diameter
08   (NUMERIC)    Fruit Length
09   (NUMERIC)    Fruit Firmness
10   (NUMERIC)    Solsol
11   (NUMERIC)    Pit Length
12   (NUMERIC)    Pit Width
13   (NUMERIC)    Fruit Dry Weight
14   (NUMERIC)    Pit Dry Weight
```

Move the highlight bar with the arrow keys. Press the spacebar when the highlighted bar is on the variables to be included in the analysis. For this example, only variable 6 will be used.

```
(NOTE: MEANS MAY BE STORED ON THE END OF YOUR DATA
FILE!)
```

```
Get Case Range
   The data file contains 120 cases.
   Do you wish to use all cases?  Y/N
```

Because some of the MSTAT programs add cases to data files, the user is prompted to indicate the range of cases to use in the calculations. If the active data file contains only the cases with the necessary information, then the user should press <RETURN> when the above prompt is presented. If you respond <N>, you are prompted for a range of cases.

When the program completes its calculations, the print/list menu of options is displayed. View the analysis on the screen, print it or store it in a disk file. The analysis generated in this example is shown below.

9.5 HIERARCH

In the following table, note that the table is built in sequence, beginning with the innermost variable. Using the table, the source of variation can be described as including the following:

1. **Variation due to differences between the clones.**
2. **Variation due to differences between the two trees for each clone.**
3. **Variation due to differences among the 20 samples for each tree.**

```
Variable  6

Hierarchical table with number of observations and mean
in each group from the innermost:
Var  4  From  1  To  2    TREE
Var  3  From  1  To  3    CLONE
```

Variable 4 Value	Number	Mean	Variable 3 Value	Number	Mean
1	20	559.0			
2	20	479.5	1	40	519.25
1	20	523.0			
2	20	563.0	2	40	543.00
1	20	532.6			
2	20	600.5	3	40	566.53

Source of Variation	DF	SS	MS	F	P%
Between var 3	2	44699	22349.43	0.53	0.0000
Betw var 4 within var 3	3	125375	41791.51	2.75	0.1345
Within var 4	114	1730741	15181.94		

Note: The P%-values are correct only if you have the same number in each grouping.

9.6 NONORTHO

Purpose: Computes the frequency, sum and/or means for one or more variables in a two-way table. For non-orthogonal tables, corrected averages in both directions can be computed, along with an analysis of variance, giving corrected mean squares for both groupings, interactions and within cells.

Before starting, you need:

1. An active data file that contains group and quantitative variables.
2. A STAT analysis of the data file.

Results: Depending on how you respond to the prompts, you may have two-way tables of frequency, sum, means and analysis of variance. The analysis may be viewed on the screen, printed or stored in a disk file.

Sample Data File: CORN.

The group variables analyzed by NONORTHO may be locations and treatments. The quantitative variable may be the mean yields of the treatments in the same experiment grown in each region. The analysis would become non-orthogonal if there were missing yield results for some years and treatments. In these cases, the NONORTHO program must be used for the analysis.

If the frequency of observations in a given treatment combination is 0 or 1, as occurs when the data are the yield for varieties in a number of trials, then ANOVA-2 and NONORTHO will generate the same sums of squares for the main effect sources of variation and estimate the same missing values. However, NONORTHO computes an accurate F value and probability percentage with missing values.

The program prompts for the first and second group variables along with their respective range of group levels. The user is prompted for the list of dependent variables. NONORTHO then gives the option of including or excluding the following analyses:

- Table of frequencies
- Table of sums
- Table of means
- Analysis of variance
- Table of means over the first group variable
- Table of means over the second group variable

9.6 NONORTHO

NONORTHO Example

The data file used in this example is CORN. The experimental design is a 3 x 4 factorial in a randomized complete block using four replications and eight environments. Variable no. 1 indicates the three phosphorous levels of 0, 25 and 50 Kg of P205 per hectare; it also indicates the four nitrogen urea treatments which consisted of 0, 50, 100 and 150 Kg N per hectare. The CORN data file was created from the MAIZE data file; CORN was created with missing values (MAIZE has 96 cases and CORN has 91 cases).

In the following example, variable no. 3 (location code) is the first group variable and variable no. 4 (treatment code) is the second group variable. The dependent variable is variable no. 2 (maize yield). If a value is entered into the variable number box and <F1> is pressed, a list of variables in the file can be pulled down and the correct variable chosen.

```
Enter  the  number  of  the  first  GROUP  variable  (1-  4):  3
Enter the lowest and highest values in the first GROUP variable
      Lowest : 1            Highest : 8
```

```
Enter  the  number  of  the  second  GROUP  variable  (1-  4):  4
Enter the lowest and highest values in the second GROUP variable
      Lowest : 1            Highest : 12
```

In this example, variable no. 3 is the first group variable, with group levels ranging from 1 through 8. The second group variable is variable no. 4, with group levels ranging from 1 through 12.

```
Enter the list of the DEPENDENT variable numbers
You may not choose either GROUP variable number
```

In this example, only variable no. 2 (maize yield) is a dependent variable. Move the highlighted bar to Variable 2 and press the spacebar. If other variables are to be analyzed, place the highlighted bar at the variable desired and press the spacebar. Press <ENTER> when all dependent variables have been highlighted.

```
Choose up to 2 variables (Press ESC to quit)
01      (NUMERIC)      Treatment code 1st three digits
02      (NUMERIC)      Maize yield in t/ha of 14% mois
03      (NUMERIC)      Location code
04      (NUMERIC)      Treatment code
```

For this example, use all 91 cases in the CORN data file.

```
Which of the following do you want to print:
TABLE OF FREQUENCIES  Y/N:
TABLE OF SUMS         Y/N:
TABLE OF MEANS        Y/N:
ANALYSIS OF VARIANCE  Y/N:
```

If an analysis of variance table is created (user responds <Y> to the last prompt above), the program places the following prompts on the screen:

```
Table of Means Over First Group Variable     Y/N
Table of Means Over Second Group Variable    Y/N
```

```
Do you want to save means over second group variable
Y/N
```

The means will be saved at the end of the data file in the variable analyzed.

When NONORTHO completes its calculations, the print/list menu of options is displayed. View the analysis on the screen, print it or store it in a disk file.

9.6 NONORTHO

The following tables are generated by the NONORTHO program:

```
Two-way tables grouped over the variables  3  and  4
 3 LOCATION CODE
with values from  1  to  8
 4 TREATMENT CODE
with values from  1  to  12

Variable  2
MAIZE YIELD IN T/HA OF 14% MOISTURE GRAIN
```

TABLE OF FREQUENCY

Var 3 Levels	Var 4: 1	2	3	4	5	6	7
1	1	1	1	1	0	1	1
2	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1
4	0	1	1	1	1	1	1
5	1	1	0	1	1	1	1
6	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1
Total	7	8	7	8	7	8	8

TABLE OF FREQUENCY Continued

Var 3 Levels	Var 4 8	9	10	11	12	Total
1	1	1	1	1	1	11
2	1	1	1	1	1	12
3	1	0	1	1	1	11
4	1	1	1	1	1	11
5	1	1	1	1	1	11
6	1	1	1	0	1	11
7	1	1	1	1	1	12
8	1	1	1	1	1	12
Total	8	7	8	7	8	91

TABLE OF SUMS

Var 3 Levels	Var 4 1	2	3	4	5	6	7
1	1.21	1.53	1.37	1.24	0.00	5.66	2.08
2	4.74	4.91	5.10	1.92	3.21	3.51	5.23
3	4.15	4.44	5.12	3.82	4.54	4.88	1.97
4	0.00	2.04	1.44	2.94	5.00	3.89	4.29
5	2.42	2.36	0.00	2.60	3.79	4.13	4.80
6	1.61	1.81	1.18	5.41	5.22	2.51	5.14
7	1.53	1.67	1.41	4.86	3.92	3.44	3.63
8	0.40	0.79	1.67	2.33	2.58	4.41	4.14
Total	16.06	19.55	17.29	25.12	28.26	32.43	31.28

TABLE OF SUMS Continued.

Var 3 Levels	Var 4 8	9	10	11	12	Sum
1	2.49	3.28	2.23	2.80	4.35	28.24
2	4.97	5.38	4.92	5.14	5.28	54.31
3	5.10	0.00	4.87	5.28	6.62	50.79
4	5.38	4.54	2.19	2.93	3.62	38.26
5	3.61	6.36	4.48	7.17	6.58	48.30
6	3.12	3.32	4.34	0.00	4.92	38.58
7	6.26	5.38	5.32	6.83	6.27	50.52
8	4.23	3.75	3.76	4.72	3.66	36.44
Total	35.16	32.01	32.11	34.87	41.30	345.44

TABLE OF MEANS

Var 3 Levels	Var 4 1	2	3	4	5	6	7
1	1.21	1.53	1.37	1.24		5.66	2.08
2	4.74	4.91	5.10	1.92	3.21	3.51	5.23
3	4.15	4.44	5.12	3.82	4.54	4.08	1.97
4		2.04	1.44	2.94	5.00	3.89	4.29
5	2.42	2.36		2.60	3.79	4.13	4.80
6	1.61	1.81	1.18	5.41	5.22	2.51	5.14
7	1.53	1.67	1.41	4.86	3.92	3.44	3.63
8	0.40	0.79	1.67	2.33	2.58	4.41	4.14
Mean	2.29	2.44	2.47	3.14	4.04	4.05	3.91

TABLE OF MEANS Continued

Var 3 Levels	Var 4 8	9	10	11	12	Sum
1	2.49	3.28	2.23	2.80	4.35	2.57
2	4.97	5.38	4.92	5.14	5.28	4.53
3	5.10		4.87	5.28	6.62	4.62
4	5.38	4.54	2.19	2.93	3.62	3.48
5	3.61	6.36	4.48	7.17	6.58	4.39
6	3.12	3.32	4.34		4.92	3.51
7	6.26	5.38	5.32	6.83	6.27	4.21
8	4.23	3.75	3.76	4.72	3.66	3.04
Mean	4.39	4.57	4.01	4.98	5.16	3.80

ANOVA TABLE

	DF	SS	MS	F(W)	P%	F(I)	P%
TOTAL	90	225.028					
Var 3 uncor	7	45.315	6.474				
Var 4 uncor	11	79.457	7.223				
Var 3 cor	7	45.624	6.517			4.70	000
Var 4 cor	11	79.765	7.251			5.22	000
Interaction	72	99.947	1.388				

9.6 NONORTHO

MEANS FOR VAR 2 FOR EACH LEVEL OF VAR 3

Level Number	Frequency	Uncorrected Mean	Corrected Mean
1	11	2.567	2.582
2	12	4.526	4.533
3	11	4.617	4.708
4	11	3.478	3.344
5	11	4.391	4.285
6	11	3.507	3.621
7	12	4.210	4.217
8	12	3.037	3.044

MEANS FOR VAR 2 FOR EACH LEVEL OF VAR 4

Level Number	Frequency	Uncorrected Mean	Corrected Mean
1	7	2.294	2.235
2	8	2.444	2.448
3	7	2.470	2.545
4	8	3.140	3.145
5	7	4.037	3.869
6	8	4.054	4.058
7	8	3.910	3.915
8	8	4.395	4.400
9	7	4.573	4.708
10	8	4.014	4.018
11	7	4.981	4.962
12	8	5.162	5.167

The first table is the frequency table. The column and row totals represent the total number of observations for each level of the two group variables.

If there is more than one observation per cell (at least in some cells), the total DF and SS will be divided into two parts: Between cells and Within cells. In this example, where there is only one observation per cell, the Between cells DF and SS equal the total DF and SS. The Between cells

DF and SS are further divided into three parts: location, treatment and interaction. Note that the sum of the corrected SS for these three parts does not equal the total SS. The uncorrected SS for each of the two grouping variables is the same as would be obtained by two separate ANOVA-1 analyses.

If more than one observation exists for any cell, the total sums of squares is divided into Within and Between cell error. The CORN example does not fit this situation; the analysis presents only the analysis of variation for the uncorrected and corrected group variable effects. The degrees of freedom for each are calculated as (N-1), where N is the number of levels in each group. However, notice that the interaction DF is not the product of the group variables DF. Instead, the interaction DF is computed as the product of the two group variable DF values minus a degree of freedom for each missing value (e.g., 77-5). The interaction mean square is then used for the error term.

If multiple observations are present for at least one cell, the Within cell and Between cell analysis of variation is computed based upon a one-way classification in which the 8 x 12 treatment combinations of CORN are treated as 96 treatments, and the hypothetical multiple observations per treatment (i.e, replications) would be treated as observations Within cells. The Within cell source of variation is computed by the equation: Total SS - Between SS.

9.7 RANGE

Purpose: Calculates the separation of the means and ranks the means in ascending order.

Before starting, you need:

1. If data are from a data file, the first and last case number of the means must be known.
2. An analysis of variance for your data with values for error mean square, degrees of freedom, number of means to be separated and number of observations in each mean.

Results: The selected mean separation test is performed on data from a data file or entered from the keyboard. Depending on which print/list options you select, the analysis may be viewed on the screen, printed or stored in a disk file.

Sample Data File: CUCUMBER.

The RANGE subprogram gives you the option of using any of the following tests to analyze the separation of the means:

- LSD (Least Significant Difference)
- Duncan's Multiple Range Test
- Tukey's Test
- Student-Newman-Keul's Test

Whichever test you select, the following information is necessary to complete the means separation: alpha level (a choice of 0.10, 0.05, 0.01), the error mean square, the degrees of freedom for the error mean square value, the number of observations used to calculate a mean value, and the number of means to be separated.

The data may be either read from an MSTAT data file or entered directly through the keyboard. If your data have been added to the end of a data file, you will need to know the first case number of the data and which variable number the data are stored under. It is suggested that PRLIST be used to obtain a printed copy of your data file including the cases where the means are stored.

When RANGE has completed its calculations, the print/list menu of options is displayed.

In the analysis generated by the subprogram, the original order of the means is displayed alongside of the ranked order (highest to lowest value). The significance of difference is indicated by alphabetic letters. If two values are followed by the same alphabetic letter, they are not significantly different for the separation test chosen.

The sample data file CUCUMBER is used in this example. The data are taken from cases 56 through 66 of the file. These means were obtained from analyzing variable no. 2 (fruit number) against variable no. 1 (population). PRLIST was used to list the following means which are under variable no. 2.

CASE NO	1	2
56	1	21
57	2	12
58	3	23
59	4	20
60	5	15
61	6	25
62	7	18
63	8	25
64	9	22
65	10	25
66	11	17

If the data for the means have not been stored in a data file, use the cursor keys to place the highlighted bar on Parameters then press <ENTER>. If the means have been stored in an MSTAT data file, you must activate the data file in the FILES program prior to entering the RANGE program.

```
RANGE
Enter input parameters
Parameters   Range   Quit
```

9.7 RANGE

```
File to perform Range Tests on:
B:CUCUMBER

Mean Separation Test:LSD
Source of Means:     Disk    Number of means:       11
First Case (if disk): 56     Alpha Level to use:    0.05
Variable No for Means: 2     Error Mean Square:     2.77
Observations per Mean: 5     Degrees of Freedom:    44
```

For this example, enter or choose the values shown on the above screen. By pressing the spacebar when the cursor is in the highlighted area for Mean Separation Test, the contents of the box will change from LSD, Duncan's, S-N-K or Tukey's. When you have chosen a test, press <ENTER> to enter that choice.

The Source of Means highlighted area will change from Disk to Keyboard by pressing the spacebar when the cursor is in that highlighted block. If Keyboard is highlighted, the program will present an input screen for each mean value to be entered after all the other parameters have been entered. Exit the parameter option and highlight RANGE. After highlighting Range and pressing <ENTER>, the user will be prompted for the number of means indicated in the parameter screen. For this example, press <ENTER> when Disk is shown.

The Alpha Levels available in MSTAT are 0.01, 0.05 and 0.10 and those choices can be changed by pressing the spacebar when the cursor is in that highlighted block. Select 0.05 for this example.

The values for error mean square, Degrees of Freedom, the number of observations used to calculate a mean value, and the number of means to be separated can be obtained from an analysis of variance table either from an MSTAT file or other statistical packages.

When all of the highlighted blocks have been filled with accurate information, press <F10> and move the cursor to the RANGE block. Press <ENTER> and the means separation analysis will be performed.

The following ranking order and significant difference table was created in this example:

Data File: B:CUCUMBER
Title: C. R. Design
Case Range: 56-66
Variable 2: Fruit NO
Function: RANGE
Error Mean Square = 2.770
Error Degrees of Freedom = 44
No. of observations to calculate a mean = 5

Original Order			Ranked Order		
Mean 1=	20.80	BC	Mean 6=	25.40	A
Mean 2=	11.80	F	Mean 8=	25.20	A
Mean 3=	22.60	BC	Mean 10=	25.00	A
Mean 4=	19.80	C	Mean 3=	22.60	B
Mean 5=	15.40	E	Mean 9=	22.20	B
Mean 6=	25.40	A	Mean 1=	20.80	BC
Mean 7=	17.60	D	Mean 4=	19.80	C
Mean 8=	25.20	A	Mean 7=	17.60	D
Mean 9=	22.20	B	Mean 11=	17.40	DE
Mean 10=	25.00	A	Mean 5=	15.40	E
Mean 11=	17.40	DE	Mean 2=	11.80	F

9.8 NEIGHBOR

Purpose: Nearest neighbor analysis is used when (1) there is a linear trend in variation, (2) the experiment will not fit a lattice design or (3) the treatments do not have the same number of replications.

Before starting, you need:

1. An active data file with separate variables for plot number, variety number, and yield.

Results: Depending on how you respond to prompts, the analysis may be viewed on the screen, edited, printed or stored in a disk file. In addition, the means may be added to the end of the active data file.

Sample Data File: NBORTST.

NEIGHBOR uses an iterative process (as described by Wilkinson in 1983) to estimate the true yield of several varieties which were randomly arranged in a long, narrow field. It accomplishes this by comparing each plot with its neighbors on each side, and updating its yield estimate for the variety of that plot accordingly. It can iterate this process as many times as desired, producing better estimates at the expense of some computer time. When it is finished, it displays a chart of the variety means before the analysis, a chart of estimated variety means (what the varieties would have yielded if the soil had been exactly homogeneous), and a map of the field indicating the soil trends. This map can be used when planning future experiments on the same field.

```
The data file contains 300 cases
Do you wish to use all cases? Y/N
```

The NBORTST data file contains 300 cases. For this example, use all 300 cases.

```
Enter the variable number
which contains your varieties:    2
Enter the variable number
which contains your yields:       4
Enter the variable number
which contains your plot numbers: 1
```

The NEIGHBOR program includes two methods of treating marginal plots (those on the end of rows or series). Marginal plots do not have neighbors on one side and therefore have to be treated differently. The statistically cleanest solution is to exclude the marginal plots from the yield estimates of varieties on those plots (Wilkinson, 1983). However, depending on the actual situation, the loss of information may outweigh the clean solution. Therefore, this program offers you two options, as discussed in Schwarzbach, 1984. Press the space bar to cycle between the options and press <ENTER> to select one.

Options available are:

```
Ignore marginal plots
Compare marginal plots with two plots on one side only
```

Once the marginal plot treatment has been selected, (in this case, highlight Compare marginal plots ...) the user must enter the plot numbers that correspond to the breakpoints of the experiment. These are the values of *n* for which no physical relationship exists between plot *n* and plot *n+1*. These will most likely occur at the end of a row of plots, where the next plot number starts up at the opposite end of the field. In addition to the breakpoints entered here, the plots in the file for which the yield variable is missing will also be treated as breakpoints. In the analysis, these plots will become marginal plots and will be dealt with according to the response to the previous question.

In the NBORTST data file, the plots are numbered in a series of tiers of 75 plots each (101-175, 201-275, 301-375 and 401-475). Since there is a break already existing in the plot number sequence, it is not necessary to enter breakpoint values for this data file. As an example, the correct plot numbers have been entered.

```
Breakpoint Number 1: 175
Breakpoint Number 2: 275
Breakpoint Number 3: 375
Breakpoint Number 4:
Breakpoint Number 5:

(Press F10 when finished)
```

After entering the plot numbers for the breakpoints the user enters the number of iterations that will be performed. Using more iterations will result in greater accuracy, at the expense of extra computer time. Experimentation has found that 5 iterations generally produce sufficient accuracy.

```
How many iterations would you like (1-20)? 10
```

```
Would you like to add the means to the end of your datafile? Y/N
```

If the means are stored at the end of the data file for this experiment, they will be the estimated variety means (what the varieties would have yielded if the soil had been exactly homogeneous) as printed in the table of *Yields and Ranking after NEIGHBOR analysis*.

The results of a successful run are included here.
Preliminary Ranking, before NEIGHBOR analysis

Variety	Ave. Yield	Rel. Yield	Reps	Rank
1	10.363	0.893	20	14
2	10.267	0.885	20	15
3	10.997	0.948	20	13
4	11.140	0.960	20	12
5	11.173	0.963	20	11
6	11.217	0.967	20	10
7	11.323	0.976	20	9
8	11.493	0.991	20	8
9	11.700	1.009	20	7
10	12.377	1.067	20	4
11	11.933	1.029	20	6
12	12.517	1.079	20	3
13	12.023	1.036	20	5
14	12.603	1.089	20	2
15	12.873	1.110	20	1

Grand Mean: 11.60

<u>Components of Variance:</u>

Total	2.595
Varieties	0.747
Yield Trends	1.848
Plot Error	0.000

Yields and Ranking after NEIGHBOR analysis

Variety	Ave. Yield	Rel. Yield	Reps	Rank
1	10.207	0.880	20	15
2	10.400	0.897	20	14
3	10.589	0.913	20	13
4	10.801	0.931	20	12
5	11.009	0.949	20	11
6	11.200	0.966	20	10
7	11.395	0.982	20	9
8	11.596	1.000	20	8
9	11.799	1.017	20	7
10	11.999	1.034	20	6
11	12.199	1.052	20	5
12	12.400	1.069	20	4
13	12.600	1.086	20	3
14	12.799	1.103	20	2
15	13.007	1.121	20	1

Grand Mean: 11.600

(Note: The Ave. Yield values from the table above will be stored as means at the end of the data file if that option was chosen within the program.)

To locate a percent deviation value for a specific plot on the following map, look down the column that is headed by the first plot number in the tier or group of plots then add the values running down the right-hand side of the page to that number until the desired plot number is reached. For example, to find the value for plot 218, move down the column headed 201 to the cell across from 17.

9.8 NEIGHBOR

Map of Yield Trends:
(percent deviation from trial mean)

	Initial Plot Number			
BASE	101	201	301	401
+0	-23	-21	-21	-23
+1	-19	-20	-19	-18
+2	-18	-22	-20	-18
+3	-17	-19	-17	-19
+4	-18	-17	-18	-17
+5	-17	-19	-18	-20
+6	-15	-15	-16	-17
+7	-16	-15	-17	-16
+8	-17	-17	-16	-15
+9	-17	-15	-15	-17
+10	-14	-14	-16	-17
+11	-14	-16	-14	-17
+12	-13	-13	-13	-16
+13	-14	-14	-14	-12
+14	-11	-14	-14	-13
+15	-14	-12	-11	-12
+16	-13	-10	-13	-12
+17	-10	-11	-11	-12
+18	-11	-10	-11	-11
+19	-9	-9	-9	-9
+20	-10	-8	-8	-9
+21	-9	-8	-9	-9
+22	-7	-8	-7	-9
+23	-8	-9	-7	-7
+24	-7	-8	-6	-6
+25	-7	-6	-6	-6
+26	-6	-6	-6	-6
+27	-5	-5	-6	-6
+28	-5	-5	-5	-5
+29	-4	-4	-5	-4
+30	-4	-3	-3	-4
+31	-3	-3	-3	-3
+32	-2	-3	-3	-2
+33	-2	-2	-2	-2
+34	-1	-1	-1	-1
+35	-1	-1	-1	-1
+36	0	0	0	0

+37	0	0	0	0
+38	0	0	0	0
+39	1	0	1	1
+40	1	1	1	1
+41	2	2	2	2
+42	2	2	2	3
+43	3	3	3	3
+44	4	3	4	3
+45	4	5	5	4
+46	5	4	4	5
+47	5	5	6	5
+48	6	5	6	6
+49	7	6	6	6
+50	7	6	6	6
+51	7	8	9	9
+52	9	7	9	8
+53	8	8	9	8
+54	8	9	10	9
+55	10	10	9	9
+56	9	11	10	11
+57	11	12	11	12
+58	11	12	12	12
+59	13	11	11	11
+60	14	14	14	11
+61	14	13	15	13
+62	12	16	14	13
+63	14	15	14	13
+64	17	17	15	17
+65	17	16	17	15
+66	16	15	16	17
+67	16	17	18	19
+68	16	17	17	17
+69	19	16	19	19
+70	19	18	21	20
+71	18	19	18	19
+72	22	17	22	21
+73	22	23	18	19
+74	22	18	22	24