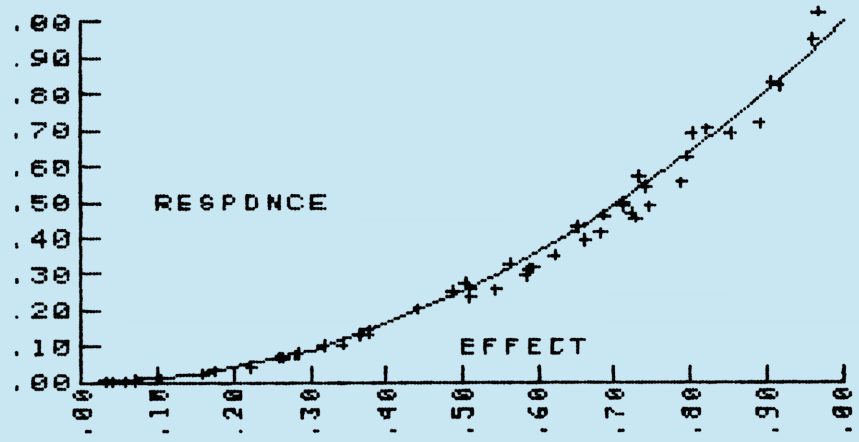

R E G 1 7 5

SHAMMAS SOFTWARE SERVICES
1533F NOHEY GROVE DRIVE
RICHMOND, VIRGINIA 23229

DISCLAIMER OF ALL WARRANTY AND LIABILITY

Shammas Software Services makes no warranties, either expressed or implied, with respect to this manual or with respect to the software described in this manual, its quality, performance, merchantability or fitness for any particular purpose. REG175 is sold "As Is". WHile every effort is made to make REG175 versatile, in no event will Shammas Software Services and/or the program author be liable for direct, indirect, incidental or consequential damages resulting from any defect in the software.

Table of Contents

FORWARD

The REG175 is a regression analysis program package for the HP75C handheld computer. The software that comes on a cassette tape contains programs that peform single, multiple and polynomial regression curve fitting. These programs are accompanied by data entry and manipulation programs that prepare the logged information.

To run the REG175 package the HP-IL cassette drive (82161A) is required. The following additional peripherals are recommended:

- HP-IL printer (82161A) or
- Video interface.

A user can very possibly involve more than one cassette drive or any other mass storage device.

The REG175 package is composed of the following programs:

1) DB11A.
2) DB12A.
3) REG1A.
4) REG21A
5) REG22.
6) TRSNF.
7) REG31A.

Shammas Software Services will modify and/or customize programs in the REG175 package. Custom regression programs can be also obtained to meet a specific need. The latter will be easier to use and faster to operate. Contact Shammas Software Services for the fees involved in such alterations.

OPERATION GUIDELINES

The first operational guideline is setting the DELAY command that informs the HP75C for how long message will be displayed. Using the HP75C liquid crystal display one would use a 'DELAY 2' or 'DELAY 3', while with a video interface a shorter DELAY can be used.

The second operation guideline is setting the HP-IL loop. To accomplish that key in,

ASSIGN IO

then assign the following device codes to the present HP-IL peripherals:

Device	Device Code
Printer	PT
Video display	TV
First mass storage device	C1

When examples are given the user's input will be underlined to distinguish it from displayed results or messages.

We would like to remind the user that all inputs must end with pressing the [RTN] key inorder for the copmuter to accept the input. This reminder is thought of as sufficient and will not be repeated.

INTRODUCTION

It is important for the user to share some of the views regarding the types of regression with which this package was planned and written.

We can say that regression analysis is used to explain the variations between two or more variables by assuming that a proposed equation (model), derived analytically or empirically, is adequate (at least for the user's purposes) for such correlation. The proposed model should state a relationship between the dependent variable and one or more independent variable. Constants and possibly suitable mathematical transformations will also be involved in the equations, and it is these constants that regression seeks to evaluate and give a measured effect as well as provide the user with an equation that can be applied in predicting values of unavailable data.

It is because of the presence of mathematical functions (transformations) that regression analysis is not always able to work in a simple straight-forward fashion. Indeed, it is how we obtain the regression constants (henceforth called regression coefficients) that determines the category of the model we are using. There are two global categories:
(1) General Linear Regression: This is characterized by the fact that all the variables involved will NOT be transformed at all. They are processed "as they are". This category is divided into two classes:
(1.1) Linear Regression: This is "the" simplest form where the dependent variable, call it Y, is affected by one independent variable, call it X, in the following manner,

$$Y = a + bX \qquad (1)$$

where a= intercept, b= slope.
"a" and "b" are the regression coefficients. The reader will most probably recognize this class, because of its popular implementation in pocket calculators.

(1.2) Multiple Linear Regression: Similar to the above case, we have here an extension in the number of independent variables. Examples are,

$$Y = a + b\,X1 + c\,X2 \qquad (2)$$

$$Y = a + b\,X1 + c\,X2 + d\,X3 \qquad (3)$$

where X1, X2 and X3 are the independent variables. The regression coefficients are "a", "b", "c" and "d".

In this category the regression coefficients are obtained by carrying out the proper summations and then solving the simulatenous linear equations yielding the desired results.

(2) General Nonlinear Regression: Due to the use of functions in describing the relationship between the variables, we can no

longer proceed as we did in the first category. The steps taken to overcome this problem will cause two classes to emerge. They are:

(2.1) Intrinsically Linear : Suitable transformations on our variables can succeed in linearizing the original model and make the situation look like a category 1 regression. Examples are:

Power curve: $Y = a X^b$ (4)

Transformations using logarithms on both sides of the equation gives,

$\log(Y) = \log(a) + b \log(X)$ (5)

Empirical curve: $Y = a X1^(b) X2^(c)$ (6)

Again a logarithmic transformation will give,

$\log(Y) = \log(a) + b \log(X1) + c \log(x2)$ (7)

Transforming equations (4) and (6) gave us equations (5) and (7) respectively, that possess linearity and thus are similar to equations (1) and (2). Thus the solution will consist of:

(1) Transformation of the variables.
(2) Summing up the proper terms.
(3) Solving the simultaneous equations.
(4) Possible "inverse transformation" of the regression coefficients obtained.

In equation (4) we note that because of the transformations some or all the regression coefficients obtained, will also be transformed and one may need to carry out the inverse transformation to obtain the values used in the "original model". In equation (5) and (7) the coefficient "a" will be obtained as log(a).

(2.2) Intrinsically nonlinear: Due to the nature of some models, the steps taken in the first class of this category may not work (i.e. no suitable transformation can be used to linearize the model). Iterative techniques are involved. The scope of this class falls outside what REG175 offers.

The two categories presented above are extreme cases. A "hybrid" combination of category 1 (class 2) and category 2 (class 1) exists. It is the case when:

(1) Variable transformation may not be applied to all the variables involved as in,

$Y = a + b X1 + c \log(X2) + d (1/X3)$ (8)

where Y and X1 remained untransformed.

(2) Variables may co-exist with their transformations side by side in the same model. This means having more "terms" in the model than the total number of variables. An example is the class of powerful polynomials,

$Y = a + b X + c X^2 + d X^3$ (9)

the above is basically a model that relates Y and X, where the

square and cube of X (its transformations) form additional terms in the model. Another example is,

$$Y = a + b X1 + c X1^2 + c \log(X2) + d X2 \qquad (10)$$

where more than one independent variable show up. It is a four term model that relates the dependent variable, Y, with two independent variables.

The goodness of fit is measured by the coefficient of determination (R^2). An adjusted value for the latter is also given for multiple and polynomial regression to be able to compare different models with a varying number of terms.

DATA MANAGEMENT

The first part of the REG175 package are two data entry and manipulation programs. They will enable the user to enter, delete, append and alter data .

The regression programs in REG175 process data from the HP75C data files. Each of the latter is merely a "table" of data with columns and rows. Each column represents a variable and each row represents an observation or data point.

Two programs are provided to perform data entry and manipulations. The first is DB11A and performs the following operations:

1) Create new data files.
2) Add variables/observations.
3) Delete variables/observations.
4) Change a data point.
5) Store/Load data to/from HP75C's data files or mass storage files.
6) View/Print data.

The second program, DB12A, performs the following tasks:

1) Merge two data files, space allowing.
2) Copy from one file to another with the ability to select data points that meet certain criteria.

DB11A

The first program in the REG175 package is the DB11A. It is loaded by the following command:

```
COPY "DB11A:C1" TO "DB11A"
```

To start running the program key in,

```
RUN "DB11A"
```

The following menu will appear, accompanied by an entry prompt:

```
MAIN MENU
0) Quit
1) New File
2) Add Point
3) Add Var.
4) Delete Point
5) Delete Var.
6) Change
7) View/Print
8) Store Data
9) Load Data

Enter choice by number?
```

Options (1) and (9) are the only ones that can be chosen when the program starts for obvious reasons. A "wrong" selection will show the following message and then take you back to the menu:

```
No data in memory
```

Here is what happens for each choice:
0) Quit : You quit the program and the following message appears:

```
End Program
```

1) New File : This tells the program that the user wishes to create a new data table within the limits assigned to the size of the table. The program will proceed by asking you to define your table specifications. You will see first,

MESSAGE : Title?

INPUT : Enter a suitable title for the data.

MESSAGE : No. of data?

INPUT : Enter the number of observations.

MESSAGE : No. of vars.?

INPUT : Enter number of variables (columns in the table).

MESSAGE : Input 6 letters max. names

The above message tells the user that it is time to name (i.e. label) each column with a name not to exceed 6 letters and no intermediate spaces are allowed.

A loop will start to input the names of each variable.

MESSAGE : Name for var. # (i)

INPUT : Enter the name for variable (i).

When the loop ends a message appears to ask if the data entered was correct, if not the user will start with a new title.

MESSAGE : All information correct?

INPUT : Enter 'Y' for yes and 'N' for no.

The program will now start prompting for the data entry, using the variable names assigned earlier and the observation number.

MESSAGE : (variable name) (i) ?

INPUT : Enter the (i)th observation for the corresponding variable.

When the above is carried out the main menu will appear again. The user should understand that at this stage the data entered are stored in program variables and not in any data file. Option number 8 will accomplish that task.

2) Add Point : This option enables the addition of more observations as long as the maximum allowed (default is 100 observations) is not exceeded. One obsevation (i.e. row) can be added at a time. The prompts used are the same for choice number one.

MESSAGE : (variable name) (i)?

INPUT : Enter the (i)th observations for the appropriate variable.

At the end of the row entry the main menu will appear.

3) Add Var. : While the second option enabled the addition of one row, this option enables the user to add a new column to the table. If the number of variables (column) is at the maximum specified (default is 10) an error message will appear to inform that the addition cannot be carried out. On the positive side,

the program will ask to name the new variable.

MESSAGE : Var. name ?

INPUT : Enter variable name (6 letters maximum).

A loop will start to enter the values for the new variable in as many observations there already are for the other variables (to update the table).

MESSAGE : (new variable name) (i)

INPUT : Enter the (i)th observation for the new variable.

When the loop ends you will return to the main menu.

4) Delete Point : This option will allow the user to delete any row in the data table. The row is selected by its sequence number.

MESSAGE : Delete point?

INPUT : Enter observation number.

The program will return to the main menu.

5) Delete Var. : This option will serve to delete a column in the table. The variable is selected by number.

MESSAGE : Delete var. #

INPUT : Enter the variable to be deleted by its "column" position.

MESSAGE : (variable name)? Y/N

The above prompt is showing the variable name and requesting to confirm with a 'Y' for the go-ahead.

INPUT : Enter 'Y' for yes or 'N' for no.

The program will return to the main menu.

6) Change : To change a specific location in the table. The location is determined by its current row and column position. By "current" we mean its up-to-date position and NOT NECESSARILY its original entry position, because deletion can alter the original position.

MESSAGE : Change var(I,J) ?

INPUT : Enter location of table element by its row position and its column position [e.g. 2,4 for row # 2 and column # 4].

MESSAGE : New X(i,j)= [value of old X(i,j)]

INPUT : Key in new value that will overwrite the old one, or simply press [RTN] to retain to old value.

MESSAGE : More changes ?

INPUT : Enter 'Y' for more changes, otherwise enter 'N'.

The program will return to the main menu when no further changes are desired.

7) View/Print : This option will allow the user to view or print the entire table. Viewing the data will depend on the DELAY command that determines the speed of viewing messages on the HP75C display.

MESSAGE : View Data ?

INPUT : Enter 'Y' to view data, else enter 'N' to print the data.

The data will be viewed/printed as follow:

```
Title : <your title>
No. vars= <number of columns>
No. of data= <number of rows>
. . . . .
<variable name> (i) = ......
. . . . .
```

At the end of the process the program will return to the main menu.

8) Store Data : This option gives a choice to store the current active table on a mass storage meduim or in an HP75C data file (in its memory).

MESSAGE : Filename ?

INPUT : Enter a filename:
1) <filename> to save in HP75C memory.
2) <filename:device code> to save on mass storage medui

It is worthy to note that the program will create a data file in the HP75C' memory anyway. Then your input will be examined to see if you included a device code. If the latter is true then the created data file will be copied to mass storage. The last step will be purging the "intermediate" file in the machine's memory.

The program will go back to the main menu.

9) Load Data : To load data from data files in the HP75C memory or from mass storage.

MESSAGE : Filename ?

INPUT : Enter filename:
1) <filename> to recall from HP75C memory.
2) <filename:device code> to recall from mass storage..

MESSAGE : File <filename> is loaded.

The program will go back to the main menu.

ERROR HANDLING

Program DB11A allows the user to interrupt the program and resume at the main menu. This is useful when one detects that inappropriate information, though acceptable by the program, has been entered. In this case it is a nuisance to continue entering further data, knowing that a fresh restart is needed. Examples may be entering a wrong variable name or title. The remedy is to press the [ATTN] key and then press the [CTRL] and [C] keys simultaneously. Assigned to [CTRL][C] is the command CONT 30, telling the program to resume at line 30 where the menu display begins. This measure saves a lot of additional error trapping precautions that lengthen the program at the expense of the available memory.

During program operation error messages may appear. Some of them are originated by the program itself. Consult appendix B for the the latter.

Example 1

Here is an example for using DB11A. There are two demonstration data files stored on the cassette you received. They are DEM01 and DEM02.
We will begin by loading DEM01 and viewing it. Then we will create a new file, DEM03. The latter will be merged with DEM01 in example 2. Here is how things would go:
(1) Load program DB11A.

```
COPY "DB11A:C1" TO "DB11A"
```

(2) Run the program,

```
RUN "DB11A"
```

(3) The main menu will appear as shown below. Select option 9 to load file DEM01 from mass storage.

```
0) Quit
1) New File
2) Add Point
3) Add Var.
4) Delete Point
5) Delete VAr.
6) Change
7) View/Print
8) Store Data
9) Load Data

Enter choice by number 9
```

(4) The program will ask for the filename. Our input should include a colon and a mass storage device code (DEM01:C1 in our case).

```
Filename ? DEM01:C1
File DEM01:C1 is loaded
```

(5) The main menu will appear again. This time we will select option number 7 to view or print the data.

```
Enter choice by number 7

View Data ? Y

Title : DEM01
No. vars= 2
No. of data= 4

TEMPF (1)= 50
TEMPC (1)= 10

TEMPF (2)= 72
```

```
TEMPC (2)= 25


TEMPF (3)= 88
TEMPC (3)= 30

TEMPF (4)= 93
TEMPC (4)= 35
```

(6) The main menu will appear again. Let us now select option number 1 to enter a new data file.

```
Enter choice by number 1
```

We want to enter the following tables:

TEMP (F)	TEMP (C)
105	40
212	100
218	105

```
Title : DEM03
```

(7) The program will start to ask for information concerning the data.

```
Title ? DEM03
No. of data ? 3
No. of vars ? 2
Input 6 letters max. names
Name for var. # 1  TEMPF
Name for var. # 2  TEMPC
All information correct ? Y
```

The program will now allow the data entry. The prompts use the names just supplied.

```
TEMPF (1) ? 105
TEMPC (1) ? 40
TEMPF (2) ? 212
TEMPC (2) ? 100
TEMPF (3) ? 218
TEMPC (3) ? 115  <----  MISTAKE !
```

(8) The main menu will appear. We want to view the data only to discover that a mistake has been made in TEMPC (3).

```
Enter choice by number 7

View Data ? Y
```

```
Title : DEM03
No. vars=2
No. of data= 3


TEMPF (1)= 105
TEMPC (1)= 40

TEMPF (2)= 212
TEMPC (2)= 100

TEMPF (3)= 218
TEMPC (3)= 115
```

(9) The main menu will appear again. Select option number 6 to correct the spotted mistake. The element sought is the second column (J=2) and the third row (I=3).

```
Enter choice by number 6

Change var(I,J) ? 3,2
New X(3,2)= 115
            ^
            |_____ cursor is here
```

Type 105 which will overwrite the 115 and then hit the [RTN] key. The following prompt will ask if more changes are desired. Our answer is a No.

```
More changes ? N
```

The main menu will appear again.

(10) Now is the time to save the information in a data file that will reside in the memory only. This means that we will select option number 8.

```
Enter choice by number 8
```

We will use the filename DEM03. No device code is included this time since we will not save it in mass storage.

```
Filename ? DEM03
```

The main menu will appear again. Select option number zero to exit the program. The following message will show up.

```
End program
```

(11) Purge DB11A to allow space for DB12A, unless you want to experiment with DB11A some more.

```
PURGE "DB11A"
```

This program performs only two operations, namely merging two files (space allowing) and copying a file while screening (i.e. rejecting) data points (rows) that fail to meet the criteria set by the user.

To load the program perform,

```
COPY "DB12A:C1" TO "DB12A"
```

Type in the following to start running the program,

```
RUN "DB12A"
```

The menu of the program will appear as,

```
0) Quit
1) Screen-copy
2) Merge

Enter choice by number
```

Here are the details for choices (1) and (2).

(1) Screen-copy : Suppose that a user wants to copy a data file from one existing file to another while dropping out every row that has a negative value in the first column. Here is where this part of DB12A is handy. As a matter of fact the user has a choice of;
a) Which column will be the center of attention, which the program calls the key variable.
b) The type of comparison. This can be the set (=,<,>,>=,<=,#,<>).
c) The value of a constant (any real or integer number) that will be involved in the comparison.

The program will start by asking the names of the source file (the existing file that actually contains the data AND is in the HP75C memory) and the target file (a new file that will be created and will contain the screened data).

MESSAGE : Source file ?

INPUT : name for source file.

MESSAGE : Target file ?

INPUT : name for target file.

The program will read the data file and display some file information,

Title : <title>
<# vars> vars and <# data observations> data points

MESSAGE : Enter key var no.

INPUT : Enter the number of the key variable.

MESSAGE : Comparison test ?

INPUT : Enter comparison test. If an erroneous entry is made a "Bad input" message will appear, then the above message will reappear.

MESSAGE : Test value

INPUT : Enter the numeric constant used to screen the data.

Should all the data fail the screening test a message "No data passed screening" will appear.

Upon a successful operation the program returns to the main menu.

(2) Merge : This option simply merges two files that are in the HP75C memory. The messages begin to ask the names of the candidate files,

MESSAGE : Enter first filename

INPUT : Key in the first filename .

MESSAGE : Enter second filename

INPUT : Key in the second filename.

Merging files is possible when:
a) Each data file starts with the "DATA1" file type identifier. If not, a "type error" message will appear to indicate which file is irregular.
b) The total number of data points does not exceed the maximum allowable (default is 100).
c) The number of variables (column) in both files is the same. If not, a "No. of vars mismatch" message will show up.

The merging will be such that the first file is enlarged. Only the values of the table in the second file are passed, no variable names or title is copied. When the merging takes place and is completed the main menu will appear again.

ERROR HANDLING

Program DB11A allows the user to interrupt the program and resume

at the main menu. This is useful when one detects that inappropriate information, though acceptable by the program, has been entered. In this case it is a nuisance to continue entering further data, knowing that a fresh restart is needed. Examples may be entering a wrong variable name. The remedy is to press the [ATTN] key and then press the [CTRL] and [C] keys simultaneously. Assigned to [CTRL][C] is the command CONT 35, telling the program to resume at line 35 where the menu display begins. This measure saves a lot of additional error trapping precautions that lengthen the program at the expense of the available memory.

During program operation error messages may appear. Some of them are originated by the program itself. Consult appendix B for the the latter.

Example 2

For this example we will merge data files DEM01 and DEM03 that are residing in the HP75C memory. Their presence can be checked by a CAT ALL command. Other programs should preferably be purged to allow adequate space for the data files.

(1) Load program DB12A.

```
COPY "DB21A:C1" TO "DB21A"
```

(2) Run the program,

```
RUN "DB21A"
```

The main menu will appear as shown below. Select option number two to merge the two files in question.

```
0) Quit
1) Screen-copy
2) Merge

Enter choice by number 2
```

(3) The program will ask for the filenames.

```
Enter first filename  DEM01
Enter second filename  DEM03
```

When the merging is completed the main menu will appear. Select option number zero to exit the program.

(4) Perform the following to save the enhanced version of DEM01 as DEM04 on mass storage and purge the current files in memory.

```
COPY "DEM01" TO "DEM04:C1"
PURGE "DB21A"
PURGE "DEM03"
PURGE "DEM01"
```

REGRESSION PROGRAMS

Having introduced the data handling programs we come now to the regression programs. We would like to talk about a few common features that the following three programs have.

The regression programs are all menu driven, with a "smart menu" feature. This means that the program monitors the on-going activities and decides to show the appropriate menu options only and hide, temporarily, the other "prematurely" requested options. The latter will gradually appear as activities progress. For example, when a regression program first starts the first expected thing to be done is to supply information about data files. At this stage the user is in no position to request the execution of regression calculations. Thus the option for regression calculations is temporarily hidden.

Another common feature between the regression programs is the number of digits displayed. The programs will request the number of significant digits. The programs will display results rounded using a pseudo-scientific mode. This will handle numbers in all ranges.

For each program a tabular data file (one or more) will be read. The user has the flexibility to select the variables (table columns) that will be considered (i.e. participate in the regression). This feature allows the grouping of data collected simultaneously while maintaining the ability to look at sub-tables for the data. This will save the user from having to delete variables.

The regression programs have been written to handle a large number of data. Due to limitations in the size of data files in the memory of the HP75C, chaining data files is required for large data. This also means that the memory will be DEDICATED to the regression. The programs will delete "foreign" program files. BE SURE TO SAVE ALL FILES THAT HAVE NOTHING TO DO WITH REGRESSION, because they will be erased ! This situation is overridden when there is one data file to be processed AND IT HAPPENS TO BE IN THE HP75C MEMORY ALREADY!

REG1A

This program deals with linear and linearized regression. It is used to relate one dependent variable to an independent variable. The two variables can be selected from others if the data source table has more than two variables.

The program will perform the following:

1) Calculate the averages and standard deviations of the regression variables (or their transformations).
2) Calculate the slope, intercept, coefficient of determination (R^2), standard error for the regression coefficients and the critical student-t values for the variable coefficients.
3) Produce an ANOVA table.
4) Calculate projections with confidence intervals, regression coefficients confidence intervals and student-t tests for regression coefficients.

To load the program key in,

```
COPY "REG1A:C1" TO "REG1A"
```

Before running the program you must save all unrelated files (key assignement, data, programs, appointtments). Another aspect worthy of mention is the ability to use user definable functions to linearize the model. Lines 7300 and 7301 are used for that purposes. Edit these lines before you start running the program.

To run the program press the [RUN] key and you will see the first prompt,

MESSAGE : Enter no. of digit display

INPUT : Key in the number of digits used for rounding the results.

MESSAGE : All data in HP75 memory ?

The above message is asking if there is one data file to process which happens to be already in memory, probably from a previous activity. If this is the case answer be a 'Y'.

INPUT : Y for yes, N for no.

The main menu in its entirety is,

```
MAIN MENU
 0) Quit
 1) Data Files.
 2) Variable Selection
 3) Regression
 4) More Stat
```

5) Additional Data

Enter choice by number

However not all the above options will always appear depending on which stage of the regression one is in. This is to inform the user of the available options at the status quo. Here is a complete description of what the options offer:

1) Data Files: This allows to load at least one data file stored on a mass storage meduim. One can also tell the program that a number of files are to be processed, which may or may not be retrieved from the same tape or disc.

The first prompt is,

MESSAGE : Enter number of files

INPUT : Enter the number of datafiles to be recalled from mass storage and then processed.

If more than one data file is to be loaded the following meassage will appear,

MESSAGE : Need to swap storage media?

INPUT : Enter a Y if you want the program to stop every time a new data file is loaded to enable placing the appropriate mass storage medium in the mass storage device. If all the data files are on the same medium answer with an N.

The program will ask the data file name(s).

MESSAGE : Enter filename # (i)

INPUT : Enter name for file number (i). A device code (':dc') is allowed, when more than one mass storage device is used.

Upon entering all the data filenames the main menu will appear, showing options 0,1 and 2 only.

2) Variable Selection: This is a very important step prior to calculating the regression coefficients. Here we must tell the program which variable is the dependent one and which is the independent one, keeping in mind that these two can be selected (pulled out) from a data table that contains more than two. The other variables are ignored.

The programs starts to display some file information TAKEN FROM THE FIRST DATA FILE READ! They are:

Title : <title>
Variable # 1 is <name>

```
Variable # 2 is <name>
. . . . . .
```

MESSAGE : Enter # of dependent var.

INPUT : Enter the variable number that will be the dependent variable. Use the information displayed above.

MESSAGE : Enter # of independent var.

INPUT : Enter the variable number that will be the independent variable. Use the information displayed above.

The program will now display a transformation menu to allow the user to linearize his model via mathemtical transformations. The menu is,

```
TRANSFORMATION MENU
0) None
1) (aX+b)^c+d
2) LOG(aX+b)+c
3) User defined
```

The above menu will be followed by the following prompt,

MESSAGE : Trnsf # for depnd. var. (Y)

INPUT : Enter the menu option chosen by number.

If option (1) is chosen,

MESSAGE : Enter a,b,c & d

INPUT : Enter the values for 'a','b','c' and 'd'.

If option (2) is chosen,

MESSAGE : Enter a,b & c

INPUT : Enter values for 'a','b' and 'c'.

If option (3) is chosen, line 7300 should have the user definable function in the form,

```
7300 IF C=1 THEN FNF=function of X(A)
```

The program now turn to the transformations of the independent variable.

MESSAGE : Trnsf # for independ. var.

INPUT : Enter the menu option chosen by number.

If option (1) is chosen,

MESSAGE : Enter a,b,c, & d

INPUT : Enter values for 'a','b','c' and 'd'.

If option (2) is chosen,

MESSAGE : Enter a,b & c

INPUT : Enter values for 'a','b' and 'c'.

If option (3) is chosen, line 7301 should have the user definable function in the form,

7301 IF C=2 THEN FNF=function of X(A)

Upon entering all the above information the program will ask the user to confirm the given information.

MESSAGE : All the information correct?

INPUT : Enter Y to confirm, N to re-enter the previous information.

The main menu will appear showing options zero to three.

3) Regression : This option will execute the regression calculations. A "Change storage medium & hit [RTN] " message will appear if the user had previously requested the program to stop every time a new data file was to be loaded from mass storage.

The following messages and results will appear:

Wait Please ...

X is <independent variable name and its transformation>
Y is <dependent variable name and its transformation>
Y = A + B*X
Number of points = <# of data points processed>
X BAR= <average X>
Y BAR= <average Y>
SDEV X= <stnd. deviation for X>
SDEV Y= <stnd. deviation for Y>
R^2 = <coefficient of determination>
A = <A, intercept>
B = <B, slope>
S.E. (B) = <stnd. error for B>

```
S.E. (A) = <stnd. error for A>
t(A) = <critical student-t value for A>
t(B) = <critical student-t value for B>

Hit [RTN] to continue

  A  N  O  V  A
-----------------
Regression
df = 1
SS = <sum of squares due to regression>

Residual
df = <degrees of freedom>
SS = <residual sum of squares>

Total, corrected
df = <degrees of freedom>
SS = <total, corrected, sum of squares>

S^2 = mean sum of squares
F = <Snedocor F statistic>


Hit [RTN] to continue
```

When the [RTN] key is pressed, the main menu appears showing options zero to four.

(4) More Stat: In addition to performing the basic regression, one can use REG11A to evaluate projection with optional confidence intervals, regression coefficient confidence intervals and student-t tests for thr regression coefficients.

The following menu appears,

```
ADDITIONAL STAT MENU
0) Quit
1) Project Y^
2) Calc Y^ & C.I.
3) C.I. for A & B
4) t-test for A
5) t-test for B

Enter choice by number
```

If option (0) is chosen the main menu will appear.

If option (1) is chosen,

```
MESSAGE : Enter <independent variable name>
```

INPUT : Enter 'X'.

MESSAGE : <dependent variable>^ = <value>

MESSAGE : Want to calc. another Y^ ?

INPUT : Enter Y to confirm, N to return to menu.

If option (2) is chosen,

MESSAGE : Enter % probability

INPUT : % probability for confidence interval.

MESSAGE : Enter <independent variable name>

INPUT : Enter 'X'.

MESSAGE : <independent variable>= <value> & <probab.> % Prob.

MESSAGE : <dependent variable>^ = <value>

MESSAGE : <dependent variable>^ Upper Limit = <value>

MESSAGE : <dependent variable>^ Lower Limit = <value>

MESSAGE : Want another Y^ & C.I. ?

INPUT : Enter Y to confirm, N to return to the menu.

If option (3) is chosen,

MESSAGE : Enter % probability

INPUT : Enter probability for confidence intervals.

The following results will appear:

@ <probab.> % Prob.

A Upper Limit= <value>

A = <value>

A Lower Limit= <value>

B Upper Limit= <value>

B = <value>

B Lower Limit= <value>

Hit [RTN] to continue

The menu will appear when the [RTN] key is pressed.

If option (4) is chosen,

MESSAGE : Enter % probability

INPUT : Enter % probability to calculate student-t.

The following results will appear:

MESSAGE : From regression
A= <value>
Enter test value

INPUT : Enter value for comparison.

MESSAGE : t-calc. = <value>
Student-t = <value>
@ <prob> % Probability & <degree freedom> df
H0 : <test value> = <A> REJECTED/ACCEPTED
Want to test another value?

INPUT : Enter Y for another test (at same % prob.), N to return to menu.

If option (5) is chosen,

MESSAGE : Enter % probability

INPUT : Enter % probability to calculate student-t.

MESSAGE : From regression
B = <value>
Enter test value

INPUT : Enter value for comparison.

MESSAGE : t calc. = <value>
Student-t = <value>
@ <prob> % Probability & <degree freedom> df
H0 : <test value> = <B> REJECTED/ACCEPTED
Want to test another value?

INPUT : Enter Y for anothr test, N to return to menu.

(5) Additional Data : In case a user is dealing with a very large number of data (i.e. a lot of data files too) and is looking for the most suitable model that will explain the data,

he may want to use a portion of his entire data first to quickly scan for the best model. Once a satisfactory candidate is found the user may want to go back and add more data to the currently selected model. This is where the option in question comes in handy. The messages and inputs are as in option (1). The only difference is that while processing data from this option, the statistical summations will not be reset to zero as is the case normally.

Error Handling : During program operation error messages may appear. Some of them are originated by the program itself. Consult appendix B for the the latter.

Example 3

In this example we will examine the data in file DEM04. The data in the latter has two variables, TempF and TempC. We will conduct a log-log fit as well as a linear fit to test the two models:

Temp F = a (Temp C)^b
Temp F = a + b*Temp C

The first model will require a logarithmic transformation for both variables. By contrast the second model is already linear.

(1) Load the program.

COPY "REG1A:C1" TO "REG1A"

(2) To run the program key in,

RUN "REG1A"

(3) The program will inquire about the number of digits to be displayed throughout the results' display. Using four digits, we will enter four (4).

Enter no. of digits display 4

The program also wants to know whether the data to be examined is already residing in the HP75C memory. In our case we had deliberately purged DEM01 from memory and saved it as DEM04.

All data in HP75 memory ? N

(4) The main menu will appear showing only options zero and one. We will select option number one to specify the filename.

MAIN MENU

0) Quit
1) Data Files

Enter choice by number 1
Enter number of files 1
Enter filename # 1 DEM04:C1

The file DEM04 will be loaded and the main menu will appear.

(5) The main menu will display options zero to two this time. We will select option number two to specify the independent and dependent variables as well as their transformations (when needed).

```
MAIN MENU

 0) Quit
 1) Data Files
 2) Variable Selection

Enter choice by number 2
```

(6) The program will display information that will define the variables by number (as X1, X2, etc., regardless to the name given to them).

```
Title : DEMO1
Variable # 1 is TEMPF
Variable # 2 is TEMPC
```

The above means that X1 is TEMPF and X2 is TEMPC. We want TEMPF (or X1) to be the dependent variable and TEMPC (or X2) to be the independent variable.

```
Enter # of dependent var.  1
Enter # of independent var.  2
```

What follows is the transformation menu.

```
TRANSFORMATION MENU
0) None
1) (aX+b)^c+d
2) LOG(aX+b)+c
3) User defined
```

For a logarithmic transformation with no shifting or scaling, we need option number two with the values for a,b and c as 1,0 and 0 respectively, for both variables.

```
Trnsf # for depnd. var. (Y)  2
Enter a,b & c  1,0,0
Trnsf # for indepnd. var. (X)  2
Enter a,b & c  1,0,0
All information correct ? Y
```

(7) The main menu will appear showing options zero to three. We will choose option number three and watch the results.

```
MAIN MENU

 0) Quit
 1) Data Files
 2) Variable Selection
 3) Regression
```

```
Enter choice by number 3


Wait Please ...

X is LOG(TEMPC)
Y is LOG(TEMPF)
Y = A+ B*X

Number of points = 7
X BAR= 3.6323
Y BAR= 4.6562
SDEV X= .81567
SDEV Y= .54254
R^2 = .97564
A = 2.2698
B = .65699
S.E.(A) = .046428
S.E.(B) = .17225
t(A) = 13.178
t(B) = 14.151

Hit [RTN] to continue  [RTN]

 A  N  O  V  A
---------------

Regression
df = 1
SS = 1.7231

Residual
df = 5
SS = .043024
MS = .0086048

Total, corrected
df = 6
SS = 1.7661

F = 200.24
S^2 = .0086048

Hit [RTN] to continue [RTN]
```

(8) The main menu will appear showing all the options. We want to re-select the variables to perform a straightforward linear regression. When the main menu appears select option number two.

```
Enter choice by number 2
```

(9) The program will again display information that defines the variables involved. This relieves the user from having to remember this information everytime this option is chosen. We want TEMPF to be the dependent variable and TEMPC to be the independent variable. No transformation is needed for either variable.

```
Title : DEMO1
Variable # 1 is TEMPF
Variable # 2 is TEMPC

Enter # of dependent var.  1
Enter # of independent var.  2

TRANSFORMATION MENU
0) None
1) (aX+b)^c+d
2) LOG(aX+b)+c
3) User defined

Trnsf # for depnd. var. (Y)  0
Trnsf # for indepnd. var. (X)  0
All information correct ? Y
```

(10) The main menu appears showing options zero to three. We will select option number three and watch the regression results.

```
Enter choice by number  3

Wait Please ...

X is TEMPC
Y is TEMPF
Y = A+ B*X

Number of points = 7
X BAR= 49.286
Y BAR= 119.71
SDEV X= 37.575
SDEV Y= 67.391
R^2 = .9987
A = 31.378
B = 1.7465
S.E.(A) = .028951
S.E.(B) = 1.7465
t(A) = 17.966
t(B) = 61.91

Hit [RTN] to continue  [RTN]

 A  N  O  V  A
---------------
```

```
Regression
df = 1
SS = 27214

Residual
df = 5
SS = 35.501


MS =  7.1003

Total, corrected
df = 6
SS = 27249

F = 3832.8
S^2 = 7.1003

Hit [RTN] to continue  [RTN]
```

The main menu will appear showing all the options.

(11) We want to perform some addtional statistics based on the regression results that we obtained. We will select option number four.

```
MAIN MENU

 0) Quit
 1) Data Files
 2) Variable Selection
 3) Regression
 4) More Stat
 5) Additional Data

Enter choice by number 4
```

(12) The option number chosen will lead us to another menu,

```
ADDITIONAL STAT MENU
0) Quit
1) Project Y^
2) Calc. Y^ & C.I.
3) C.I. for A & B
4) t-test for A
5) t-test for B
```

We will perform some calculations.

(13) First we will calculate Y^ (i.e. TEMPF^), for TEMPC=5, as well as the confidence interval at 95% confidence. Thus option number two is selected.

```
Enter choice by number 2

Enter % probability  95

Enter TEMPC  5



TEMPC= 5 & 95 % Prob.
TEMPF^ = 40.34
TEMPF^ Upper Limit = 46.035
TEMPF^ Lower Limit = 34.645

Want another Y^ & C.I. ?  N
```

(14) We are back at the additional stat menu. We want to evaluate, at 90% confidence, the interval for A and B, the regression coefficients. Select option number three.

```
Enter choice by number 3

Enter % probability  90

@ 90 % Prob.
A Upper Limit= 34.896
A            = 31.378
A Lower Limit= 27.86

B Upper Limit= 1.8506
B            = 1.7923
B Lower Limit= 1.734

Hit [RTN] to continue  [RTN]
```

The additional stat menu will appear again.

(15) We will select option number four to test the hypothesis that A=32 , at 95 % probability.

```
Enter choice by number 4

Enter % probability  95

From regression
A = 31.378
Enter test value  32

t calc. = -.35605
Student-t = 2.5712
@ 95 % Probability & 5 df
H0 : 32 = 31.378  ACCEPTED
```

```
Want to test another value ? N
```

We can see that the assumed hypothesis is accepted.

(16) Select option number five to test, at 95 % probability, the hypothesis that B=1.8 .

```
Enter choice by number 5

Enter % probability  95


From regression
B = 1.7923
Enter test value  1.8

t calc. = -.26503
Student-t = 2.7512
@ 95 % Probability & 5 df
H0 : 1.8 = 1.7923  ACCEPTED

Want to test another value ? N
```

(17) Select option zero from the menu to get back to the main menu, where option number zero will also be selected to exit the program.

```
Enter choice by number 0

Enter choice by number 0    <--- main menu

End Program
```

This is the second regression program. It performs multiple linear(ized) regression involving one or more independent variables and one dependent variable. This program operates with other "satellite" programs, namely, REG22 and TRNSF. A utility data file, PASS, is also used. These programs form an extremely powerful tool for researchers who are seeking to find adequate models that can explain and predict data.

Before we give the instructions for using REG21A the user should be aware that while REG21A bears some similarities with the other two regression programs, it has a truly unique feature concerning the power and flexibility to transform data. While REG11A gave a menu for the mathematical transformations (user definable included) that enables a large number of models to be examined, REG21A uses no such menu, instead it takes a different route. It allows the user to set up the number of terms involved in the model, to dictate the type of transformation, by inputing the equation that describes the term in question, in a manner that pleases the heart! The program will convert the user's input into part of REG22 (a transformation subroutine). This means that there can be billions of models tested!

An example for the above is the following linearized model,

```
X5 = a + b*LOG(X2) + c*(1/X3+1/X1) + d*LOG((X1+X2)/(X1*X2))

          [----Y----]  [-----Y------]  [--------Y---------]

Term # ----->      1             2                  3
```

The input transformation for the above three terms will be:

Term #	Input transformation
1	LOG(X2)
2	1/X3+1/X1
3	LOG((X1+X2)/(X1*X2))

The above represents three inputs the user will make to define the independent variable terms. A similar input is made for the dependent variable term (X5 in the above case). It should become clear to the user that some parts of the programs are changing without his direct intervention. THIS IS THE "PROGRAM GENERATION" FEATURE THAT REG175 HAS!

REG21A will calculate the regression coefficients (B0, B1, B2, etc.) with their standard error values and critical student-t (except for B0). The ANOVA table will also be shown. To determine the goodness of fit the coefficient of determination and its adjusted value are also given. The latter will help to

compare model with a different number of terms, since adding more terms seems to give an advantage to the model.

To load REG21A and its acompanying programs press the following ending each line with a [RTN],

```
COPY "REG21A:C1" TO "REG21A"
COPY "REG22:C1" TO "REG22"
COPY "TRNSF:C1" TO "TRNSF"
```

To run the program key in,

```
RUN "REG21A"
```

MESSAGE : Enter no. of digit display

INPUT : Key in the number of digits used for rounding the results.

MESSAGE : All data in HP75 memory ?

The above message is asking if there is one data file to process which happens to be already in memory, probably from a previous activity. If this is the case answer be a 'Y'.

INPUT : Y for yes, N for no.

The main menu in its entirety is,

```
MAIN MENU
 0) Quit
 1) Data Files.
 2) Variable Selection
 3) Regression

 Enter choice by number
```

However not all the above options will always appear depending on which stage of the regression one is in. This is to inform the user of the available options at the status quo. Here is a complete description of what the options offer:

1) Data Files: This allows to load at least one data file stored on a mass storage medium. One can also tell the program that a number of files are to be processed, which may or may not be retrieved from the same tape or disc.

The first prompt is,

MESSAGE : Enter number of files

INPUT : Enter the number of datafiles to be recalled from mass storage and then processed.

If more than one data file is to be loaded the following meassage will appear,

MESSAGE : Need to swap storage media?

INPUT : Enter a Y if you want the program to stop every time a new data file is loaded to enable placing the appropriate mass storage medium in the mass storage device. If all the data files are on the same medium answer with an N.

The program will ask the data file name(s).

MESSAGE : Enter filename # (i)

INPUT : Enter name for file number (i). A device code (':dc') is allowed, when more than one mass storage device is used.

Upon entering all the filenames the main menu will appear showing options zero to two.

2) Variable Selection : This option will allow the user to specify the particular variables entering the regression as well as the transformations. Up to ten terms are allowed (not counting the constant coefficient). Unlike REG11A the user is not limited by the functions presented in the transformation menu. The user can enter any equation as long as it has a correct syntax and refers to available variables.

This program portion will display,

TRANSFORMATION MENU
0) Quit
1) View transformations
2) Set up new transformations

Enter choice by number

Here is what each option will do.

If option (0) is chosen the user will return back to the main menu. This step allows the "escape" from variable selections if already chosen.

If option (1) is chosen you will see,

Depend. Var. term = <equation for dependent variable term>
Term # 1 = <equation for term #1>

```
Term # 2 = <equation for term #2>
. . . . . .
```

The main menu will appear showing options zero to three.

If option (2) is chosen the program is now ready to be informed about the number of terms involved in a new model as well as the participating variables. The program first displays the variable names for the currently active data file.

```
Var # 1 is <variable #1 name>
Var # 2 is <variable #2 name>
.....
```

Then the program starts to ask for the number of terms,

MESSAGE : Number of terms ?

INPUT : Enter number of terms.

MESSAGE : Enter dep. term eqn.

INPUT : Enter the equation that will evaluate the dependent term (the one to the left hand side of the equal sign, see note below).

A loop will start to ask the equation that expresses each term.

MESSAGE : Enter eqn for term # <i>

INPUT : Enter equation for term (i).

NOTE: Each term entered above should follow these rules:

1) No equal sign is needed.
2) When referring to any variable use 'X' followed by one or two digits (e.g. variable number five is X5, variable number six is X6, variable number nine is X9 and so on).
3) All mathematical operations must observe correct syntax. Brackets can be used and must follow the syntax dictated by the BASIC language (see HP75C manual).
4) Functions can be used (e,g, LOG,SQR,SIN, etc.) but must follow the correct syntax.
5) Each term can contain any variable(s) so long they are available from the data. More than one variable can appear in each term (e.g. X3*X4, LOG((X2+X7)/X1), X2*X3^2, etc.).

When all the term are keyed in,

MESSAGE : All the information correct?

INPUT : Enter a Y to confirm, N to deny.

Upon confirmation the main menu will appear showing all options.

3) Regression : In this option the regression calculations are executed and results shown. If the user has asked to change mass storage media for multi-file operations a "Change storage medium & hit [RTN] " message will appear.

The display/printer will show the following:

```
Wait Please ...

Variable # 1 is <variable name>
Variable # 2 is <variable name>
.............................

Depnd. Var. term = <equation>

Term # 1 = <equation>
Term # 2 = <equation>
.....................

R2 = <coefficient of determination>
R2a = <adjusted coefficient of determination>
B0 = <constant coefficient>

B 1 = <value>
S.E(B 1) = <stand error for B1>
t-value = <critical student-t value for B1>

B 2 = <value>
.....................
.....................
.....................

Hit [RTN] to continue

  A  N  O  V  A
----------------
Regression
df = 1
SS = <sum of squares due to regression>

Residual
df = <degrees of freedom>
SS = <residual sum of squares>

Total, corrected
df = <degrees of freedom>
SS = <total, corrected, sum of squares>

S^2 = mean sum of squares
F = <Snedocor F statistic>
```

Hit [RTN] to continue

The main menu will appear again showing all options.

Error Handling : During program operation error messages may appear. Some of them are originated by the program itself. Consult appendix B for the the latter.

Example 4

In this example we will use the data from DEM02. The latter contains data involving three variables. We will perform a "surface" fit that involves these variables. The model involved is,

```
Z = a + b*X + c*X^2 + d*X*Y + e*Y + d*Y^2
```

The above model is a five-term model that involves two independent variables (X and Y) and a dependent variable (Z).

(1) Load programs REG21A, REG22 and TRNSF from the cassette.

```
COPY "REG21A:C1: TO "REG21A"
COPY "REG22:C1" TO "REG22"
COPY "TRNSF:C1" TO "TRNSF"
```

(2) Run program REG21A.

```
RUN "REG21A"
```

The program will ask for the number of digits displayed with the results. We will use four digits.

```
Enter no. of digits displayed 4
```

The program will now ask whether the data file which is the source of the information to be analysed is residing in the memory. This is not the case for this example.

```
All data in HP75 memory N
```

The main menu will appear showing options zero and one.

(3) We will select option number one to process file DEM02. The latter will be loaded from the cassette (device code is ':C1').

```
MAIN MENU
 0) Quit
 1) Data Files

Enter choice by number 1
Enter number of files 1
Enter filename # 1 DEM02:C1
```

The main menu will appear showing options zero to two.

(4) We will select option number two.

```
MAIN MENU
```

```
0) Quit
1) Data Files
2) Variable Selection

Enter choice by number  2
```

The transformation menu will appear. Our choice will be option number two.

```
TRANSFORMATION MENU
0) Quit
1) View transformations
2) Set up new transformations

Enter choice by number  2
```

The program will define the variables' names.

```
Title : DEMO2
Var # 1 is X
Var # 2 is Y
Var # 3 is Z
```

Next we will tell the program that our model will consist of five terms such that,

Term #	Equation	Represent
1	X1	X1
2	X1^2	X^2
3	X1*X2	X*Y
4	X2	Y
5	X2^2	Y^2
independent	X3	Z

```
Number of terms  5

Enter dep. term eqn.  X3

Enter eqn for term # 1  X1
Enter eqn for term # 2  X1^2
Enter eqn for term # 3  X1*X2
Enter eqn for term # 4  X2
Enter eqn for term # 5  X2^2
All the information correct ?  Y
```

The main menu will appear showing all options.

(5) We will select option number three to obtain the regression results.

```
MAIN MENU
 0) Quit
 1) Data Files
 2) Variable Selection
 3) Regression


Enter choice by number  3

Wait Please ...

Title : DEM03
Variable # 1 is X
Variable # 2 is Y
Variable # 3 is Z

Depnd. Var. term = X3

Term # 1 = X1
Term # 2 = X1^2
Term # 3 = X1*X2
Term # 4 = X2
Term # 5 = X2^2
R2 = .99897

R2 adj = .99769

B0 = 2.0596

B  1 = 1.0983
S.E(B  1) = .18911
t-value = 5.8076

B  2 = .01148
S.E(B  2) = .074298
t-value = .15451

B  3 = -1.0523
S.E(B  3) = .09227
t-value = -11.405

B  4 = .99847
S.E(B  4) = .16799
t-value = 5.9438

B  5 = -.99072
S.E(B  5) = .043271
t-value = -22.896

Hit [RTN] to continue  [RTN]

 A  N  O  V  A
```

```
---------------

Total, corrected
df = 6
SS = 239.82

Regression
df= 5
SS = 239.76

Residual
df = 4
SS = .24635

Sey^2 = .061586
F = 778.01

Hit [RTN] to continue  [RTN]
```

(6) When the main menu appears select option number zero to exit.

REG31A

This program is very similar to REG21A except it studies the relationship between two variables only using a polynomial equation. Again the user can select particular variables if the data table has more than two. The program will calculate the regression coefficients (B0, B1, B2, etc.) along with their standard errors and critical student-t values (except for B0). The coefficient of determination is also evaluated together with its adjusted value. The ANOVA table is also shown. The operation of REG31A is more straight-forward since no transformations are required. This program is a special case of REG21A.

To load REG31A key in,

```
COPY "REG31A:C1" TO "REG31A"
```

To run the program press the [RUN] key. The first prompt will come as,

MESSAGE : Enter no. of digit display

INPUT : Key in the number of digits used for rounding the results.

MESSAGE : All data in HP75 memory ?

The above message is asking if there is one data file to process which happens to be already in memory, probably from a previous activity. If this is the case answer be a 'Y'.

INPUT : Y for yes, N for no.

The main menu in its entirety is,

```
MAIN MENU
 0) Quit
 1) Data Files.
 2) Regression

Enter choice by number
```

However not all the above options will always appear depending on which stage of the regression one is in. This is to inform the user of the available options at the status quo. Here is a complete description of what the options offer:

1) Data Files: This allows to load at least one data file stored on a mass storage meduim. One can also tell the program that a number of files are to be processed, which may or may not be retreived from the same tape or disc.

The first prompt is,

MESSAGE : Enter number of files

INPUT : Enter the number of datafiles to be recalled from mas storage and then processed.

If more than one data file is to be loaded the following meassage will appear,

MESSAGE : Need to swap storage media?

INPUT : Enter a Y if you want the program to stop every time a new data file is loaded to enable placing the appropriate mass storage medium in the mass storage device. If all the data files are on the same meduim answer with an N.

The program will ask the data file name(s).

MESSAGE : Enter filename # (i)

INPUT : Enter name for file number (i). A device code (':dc') is allowed, when more than one mass storage device is used.

Upon entering all the filenames the main menu will appear showing options zero to two.

3) Regression : In this option the regression calculations are executed and results shown. If the user has asked to change mass storage media for multi-file operations a "Change storage medium & hit [RTN] " message will appear.

The program will show some data file information,

```
Title : <title>
Var # 1 is <variable name>
Var # 2 is <variable name>
.........................
```

The program now will ask about the variables entering the polynomial fit and its degree.

MESSAGE : Enter polynomial degree

INPUT : Enter polynomial degree (less than ten).

MESSAGE : Enter var. no. for Y

INPUT : Enter the variable number that will be the dependent variable.

MESSAGE : Enter var. no. for X

INPUT : Enter the variable number that will be the independent variable.

The display/printer will then show the following:

```
Wait Please ...

Title : <title>
Variable # 1 is <variable name>
Variable # 2 is <variable name>
.............................

Y is <variable name>
X is <variable name>

R2 = <coefficient of determination>
R2a = <adjusted coefficient of determination>
B0 = <constant coefficient>

B 1 = <value>
S.E(B 1) = <stand error for B1>
t-value = <critical student-t value for B1>

B 2 = <value>
.......................
.......................
.......................

Hit [RTN] to continue

  A  N  O  V  A
----------------
Regression
df = 1
SS = <sum of squares due to regression>

Residual
df = <degrees of freedom>
SS = <residual sum of squares>

Total, corrected
df = <degrees of freedom>
SS = <total, corrected, sum of squares>

S^2 = mean sum of squares
F = <Snedocor F statistic>


Hit [RTN] to continue
```

The main menu will appear again.

Error Handling : During program operation error messages may appear. Some of them are originated by the program itself. Consult appendix B for the the latter.

Example 5

In this example we will use the data from file DEM04 which contains information relating two variables, namely TEMPF and TEMPC. A quadratic fit (i.e. polynomial of second degree) will be performed.

(1) Load the program REG31A after purging any large program that is in the memory.

```
COPY "REG31A:C1" TO "REG31A"
```

(2) To start running the program key in,

```
RUN "REG31A"
```

The program will ask for the number of digits displayed with the results. We will use four digits.

```
Enter no. of digits displayed 4
```

The program will now ask whether the data file which is the source of the information to be analysed is residing in the memory. This is not the case for this example.

```
All data in HP75 memory N
```

The main menu will appear showing options zero and one.

(3) We will select option number one to process file DEM02. The latter will be loaded from the cassette (device code is ':C1').

```
MAIN MENU
 0) Quit
 1) Data Files

Enter choice by number 1
Enter number of files 1
Enter filename # 1 DEM04:C1
```

The main menu will appear showing options zero to two.

(4) We will select option number two and proceed with the regression.

```
MAIN MENU
 0) Quit
 1) Data Files
 2) Regression

Enter choice by number 2
```

The program will display information defining the variables available in the data file.

```
Title : DEMO1
Variable # 1 is TEMPF
Variable # 2 is TEMPC
```

Next we will tell the program that we want a second order polynomial fit with TEMPF (i.e. X1) as the dependent variable and TEMPC (i.e. X2) as the independent variable.

```
Enter polynomial degree  2
Enter var. no. for Y  1
Enter var. no. for X  2

Wait Please ...

Title : DEMO1
Variable # 1 is TEMPF
Variable # 2 is TEMPC

Y is X1
X is X2

R2= .99873

R2 adj = .99809

B0 = 30.111

B  1 = 1.8577
S.E(B  1) = .2029
t-value = 9.1459

B  2 = -.00051022
S.E(B  2) = .0016126
t-value = -.31639

Hit [RTN] to continue [RTN]

 A  N  O  V  A
---------------

Total, corrected
df = 3
SS = 27249

Regression
df = 2
```

```
SS = 27241



Residual
df = 4
SS = 34.635

Sey^2 = 8.6586
F = 1571.5

Hit [RTN] to continue [RTN]
```

(5) The main menu will appear again. Select option number zero to exit.

Appendix A

Data File Configuration

The following is the structure of a data file. This includes some provisions that allow other software developed by Shammas Software Services.

Record #	Content
1	DATA1 (a type check code)
2	Title string.
3	Total number of data.
4	Actual number of data, ND.
5	Number of variables, NV.
6	Variable names string.
7	Number of data subfiles.
8	Subfiles names string.
9	X(1,1)
10	X(2,1)
..	.....
9+ND	X(ND,1)
10+ND	X(1,2)
.....	.....
9+ND*NV	X(ND,NV)

Appendix B

Error Messages

Here are the error messages for the programs in the REG175 package.

DB11A

Error	Cause
FILE FULL	There are already 100 data points. No more data can be stored.
File mismatch	The first record is not "DATA1" used to identify the data file as of appropriate type.
OUT OF RANGE	The user is attempting to alter an unavailable data element.
No data in memory	Attempt to perform data manipulations when none exist.
No data to delete	Attempt to delete data when no more is available.
No vars to delete	Same as above.
No. of vars=10	Attempt to add a new variable when there are already ten variables.

DB12A

Error	Cause
Bad input	Error in inputing test symbols for a screen-copy
File#1 has error type	First data file does not start with the 'DATA1' type check code.
File#2 has error type	Second data file does not start with the 'DATA1' type check code
File type error	Similar to above.
No data passed screening	All the data in the source file failed to meet the screening criteria.
No space	The total number of data points of the two files to be merged exceeds 100.
No. of vars mismatch	The files to be merged are incompatible regarding the number of table columns.

REG1A

Error	Cause
FATAL ERROR	First data file does not contain the 'DATA1' type check code.
INSUFFICIENT DATA	The total number of points is less than three.
Transformation error	Mathematical error in transformation of data.

REG21A/REG31A

Error	Cause
FATAL ERROR	First data file does not contain the 'DATA1' type check code.
INSUFFICIENT DATA	The total number of points is less than three.