# PalmStats™ *Version 2.1*
## User Reference

## Introduction

**PalmStats™** is a graphically-oriented statistical analysis program written specifically for the Palm Platform. **PalmStats™** calculates over 40 statistical values on two data sets of up to 100 points each. The results are displayed both numerically and graphically to provide the most complete answers.

## Features and Functions:

- *Statistical calculations include*: Mean, Standard Deviation, Standard error of the Mean, Variance, Minimum, Maximum, Range, Sum, and Skewness.
- *Histogram and Scatter Plots* may be done quickly to characterize data.
- *Coordinate Tracking* is available for scatter and regression plots to quickly locate individual data points.
- *Hypothesis testing (1 and 2 tailed t-testing)* may be done to at 1, 2.5, 5 and 10% significance levels. Test results are graphically represented.
- *Confidence Intervals* may be calculated for the means of the data sets.
- *Probability* calculations may be done on all data or parameters may be entered individually.
- *Regression analysis* includes linear and exponential curve fitting, plotting, and interpolation.
- Data may be entered using the *Data editor* or imported from a comma-delimited memo file.
- Data may be saved as comma-delimited memo files for transfer to a PC.

## Testing and Compatibility

All statistical calculations have been verified by comparing the calculated output to either textbook examples, hand calculations, or spreadsheet outputs.

**PalmStats™** was developed on a Handspring Visor and has been extensively tested on a Palm III*xe*. During the development of **PalmStats™**, it was used and tested with the following programs and/or hacks without any conflicts:
HackMaster, ClipHack, ClearHack, PhlemHack, MagicTextHack, MenuHack, SelectHack, RPNpopUpCalculator, and PopUpNote.

# Files included with *PalmStats*™

| Filename | Description |
|----------|-------------|
| PalmStats.Prc | program module |
| Module1.Prc | program extension module |
| Module2.Prc | program extension module |
| Module3.Prc | program extension module |
| DataEdit.Prc | program extension module |
| Licence.txt | user license agreement |
| PalmStats.PDF | this document |

## Installation

Copy the Zip file to a storage directory on the PC, and extract all of the files. Use the Palm Desktop Install function to install all PRC files onto the Palm device.

## Upgrading from a previous version

**Users upgrading from <u>any</u> previous version must first delete the existing program files on their handheld device prior to installing this software.  Failure to do so may cause unpredictable behavior and or fatal errors which may  result in data loss.**

## Registration

A registration code, along with instructions regarding how to enter it, will be e-mailed after registering this software through one of the sales sites listed below.  This will usually take less than 24 hours. The registration cost is $13.95.

Currently, *PalmStats*™ may be registered at:
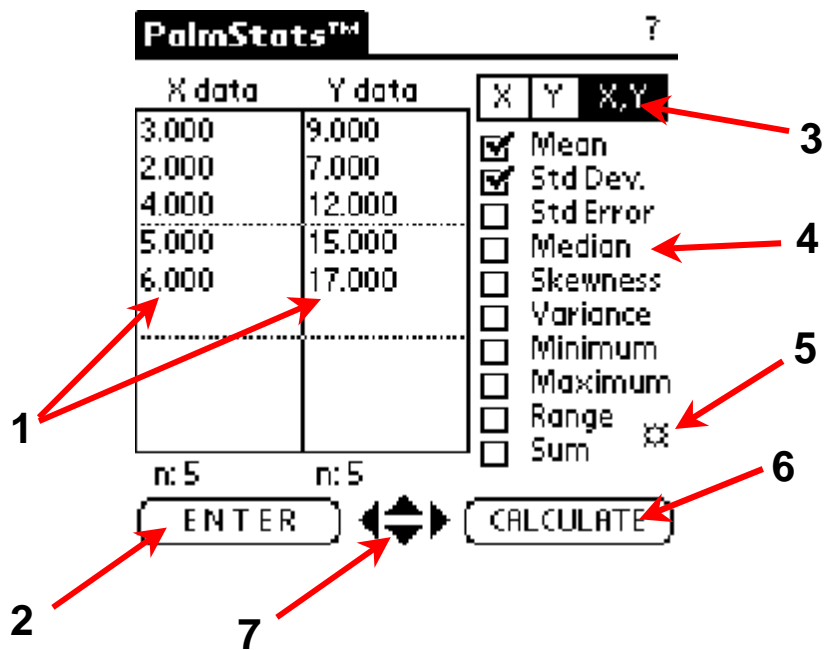
**PalmGear HQ**      www.palmgear.com

**Handango**      www.handango.com

2

# Controls and Functions

Refer to the following figures for information regarding the user interface of **PalmStats™** .



**1 – X and Y data fields** are used to enter data into **PalmStats™**.   Use Graffiti characters followed by a return character**.   Data Markers** help in aligning data point pairs visually and may be turned off in the preferences screen.

**2 – ENTER** is used to enter the data from the X and Y data sets into **PalmStats™** .  The n: field will display the number of data points which have been entered.  **PalmStats™** uses IEEE floating point calculations internally.  Data contained in the field is parsed and converted to this format by the Enter key.  (See the next section regarding the parsing algorithm)

**3 – Data Set Selector** this control specifies what data set the calculation are performed on (X, Y or Both).  This control has the same meaning and function throughout the program.

**4 – Descriptive Statistical calculations** are chosen by checking the appropriate box.

**5 –** The **Data Editor Hotkey** is used to quickly switch to the data editor screen.  A similar key is provided on the data editor screen for a quick return to the main screen.

## Descriptive Statistics

|  | X data | Y data |
|---|---|---|
| Mean: | 57.282 | 83.313 |
| Std Dev: | 4.968 | 40.796 |
| Std Err: | 1.059 | 8.697 |
| Median: | 56.235 | 72.099 |
| Skewness: | 0.857 | 0.630 |
| Variance: | 24.684 | 1664.388 |
| Minimum: | 49.860 | 22.799 |
| Maximum: | 68.550 | 160.100 |
| Range: | 18.690 | 137.300 |
| Sum: | 1260.210 | 1832.899 |
| n: | 22 | 22 |

**6 – CALCULATE** runs the calculations and displays the results on the results screen as shown above.

**7 – Scroll Arrows**  The *up* and *down* arrows scroll the data up and down one line. The *left* arrow scrolls to the *beginning* of the data field and the *right* arrow scrolls to the *end* of the data field.   The data field to be scrolled is selected using the *Data Set Selector* ( X, Y, or both).   For example, to scroll just the X data field, select X and use the scroll controls to move through the data.

## The Data Editor

*PalmStats™* provides a data editor screen for quick and easy data entry. Data may be entered using the keypad (2) or Graffiti.
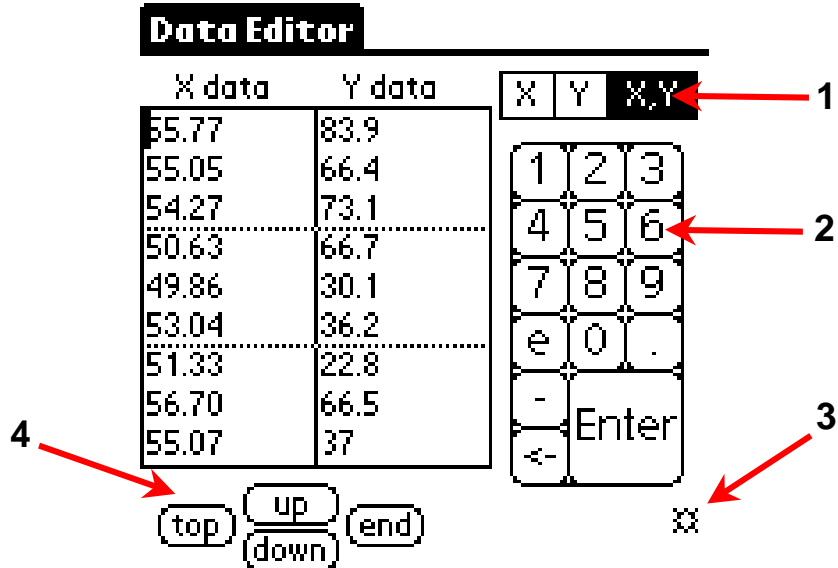


To use the data editor, tap the stylus inside one of the fields and use the keypad to enter data points followed by **Enter**. The **Keypad (2)** is equipped with a backspace key which deletes the character immediately to the left of the cursor. The **Return hotkey (3)** is used to quickly return to the main screen. Data may be scrolled using the **Scroll Controls (4)** in conjunction with the **Data Field Selector (1)**. For example, to scroll just the X data field, select X and use the scroll controls to move through the data.

## Entering Data and Intelligent Input Data Field Parsing

*PalmStats™ Version 2.1* is equipped with an fast, accurate and intelligent input parsing algorithm. Data points are entered for processing via the Graffiti system or the **Data Editor** previously described. These points may consist of any numeric value/character (either as shown or floating point format 1e3, 2.23e4 etc.) followed by a new line or carriage return. When **ENTER** is pressed, the parsing algorithm will attempt to convert the characters in the X and Y data fields to the internal IEEE floating-point format and store them internally as such. During this process, any errors encountered stop the parsing algorithm and are flagged as shown in the figure below.
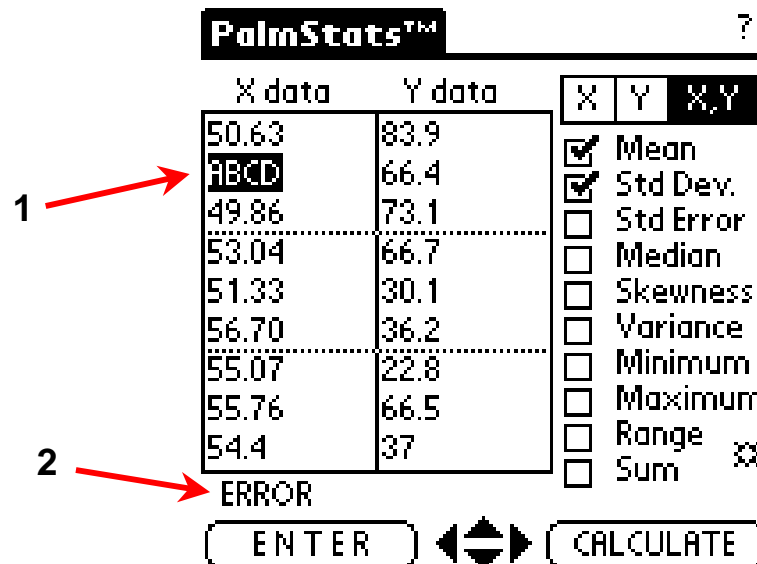


**1 –** The point containing the error is highlighted and the field is scrolled so that the point is in view. The character or characters must be edited out before parsing will continue successfully.

**2 –** The field is flagged by an ERROR label.

The parsing routine will ignore spaces (white space characters) in the data fields. This may be helpful when editing the data fields using Graffiti. Extra new line characters at the end of a data field will cause errors and are flagged by the cursor positioned after the last data point of the data field in error.

*PalmStats* will only parse 100 data points into its internal format. Any additional character data in the data fields is ignored.

# Preference Selection

**Preferences**

General:     ▼ 3 Digit(s)
  ☑ Markers     ▼ Decimal

Histogram/Plot:
  ☑ Gauss Plots   ☑ Grids
Dot Size:  ▼ Medium
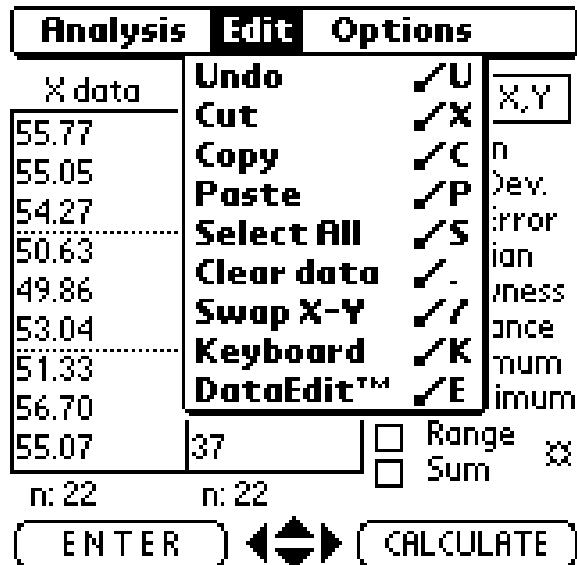# of bins:  ▼ 20
X axis:     ▼ ± 3.125 SD

[ OK ]

The **Markers check box** is used to toggle the marker lines seen in the X and Y data fields.  The numerical format may be changed using the **Digit(s)** and **Decimal/Scientific** controls.  *PalmStats™*  supports up to 6 digits after the decimal point.

The histogram and plot displays may be configured using the remaining controls.   **Gauss Plots** refer to the normal distribution curve approximated when a histogram plot is drawn.  **Grids** may be selected for scatter and regression plots.   **Dot Size** may be Small (approx. 1 mm squares), Medium (1.5 mm circles), or Large (2 mm circles). This setting also changes the dot size for the scatter plot and regression plots.    The **# of bins** is the number of bins the data will be sorted into prior to drawing the histogram.   The width of each bin is equal to the data range divided by the number of bins.  For example, if the range of a data set is 1500 and # of bins is set to 20 (as shown above) the width will be 75, or 1/20 the range.  There can be from 10 to 100 bins in increments of 5.  The **X axis** setting controls the minimum and maximum values over which the histogram will be plotted.   When set to **+/- 3.125 standard deviation**, the histogram will be centered on the screen with both tails decreasing to zero.   This is the view seen in the histogram figure.  When set to Range of Data, the ends of the plot will be the minimum and maximum data points.

# Special Menu Functions

In addition to the normal Palm editing functions shown below, two new functions have been added to the edit menu.   **Clear Data** erases all the data contained in the field selected with the Data Selector (X, Y, or both).  The **Swap X-Y** function exchanges the X and Y datasets.

```
┌─────────────────────────────────────────┐
│  Analysis  [Edit]  Options               │
├──────────┬──────────────────┬────────────┤
│  X data  │ Undo         ╱U  │ X,Y        │
│ 55.77    │ Cut          ╱X  │ n          │
│ 55.05    │ Copy         ╱C  │ Dev.       │
│ 54.27    │ Paste        ╱P  │ rror       │
│ 50.63    │ Select All   ╱S  │ ion        │
│ 49.86    │ Clear data   ╱.  │ ness       │
│ 53.04    │ Swap X-Y     ╱/  │ ance       │
│ 51.33    │ Keyboard     ╱K  │ mum        │
│ 56.70    │ DataEdit™    ╱E  │ imum       │
│ 55.07    │ 37        □ Range     ☒       │
│          │           □ Sum              │
│  n: 22   │   n: 22                       │
│  ( ENTER )  ◄╪►  ( CALCULATE )           │
└─────────────────────────────────────────┘
```

All other editing functions behave the same as in other Palm applications. *PalmStats*™ has been tested with the following edit-enhancing operating system extensions:  ClipHack, ClearHack, MagicTextHack, MenuHack, and SelectHack.

# Analysis Functions

***PalmStats™*** provides the user very flexible data analysis capabilities.   The figure below shows the analysis function available in this version of ***PalmStats™***.

```
┌─────────────────────────────────────┐
│ Analysis  Edit   Options            │
│ Confidence Interval        /I       │
│ Hypothesis Test            /T       │
│ Probability                /B       │
│ Sample Size Calculator     /Z       │
│ Regression Analysis        /R       │
│ Histogram Plot             /H       │
│ Scatter Plot               /S       │
│ ────────────────────────────────    │
│ 51.33      22.8     □ Minimum       │
│ 56.70      66.5     □ Maximum       │
│ 55.07      37       □ Range    ☼    │
│                     □ Sum           │
│   n: 22      n: 22                  │
│ ( ENTER ) ◀♦▶ ( CALCULATE )         │
└─────────────────────────────────────┘
```

**Confidence intervals** for the means of the X and Y data sets may be calculated at 80, 90, 95, 98, and 99% confidence levels.   Calculations done on small sample sizes ($n<30$) are done with a *t distribution*.  Sample sizes of greater than 30 points are calculated using the normal gaussian distribution.

A **Sample Size Calculator** has been provided which calculates the minimum sample size required for a given confidence level, mean, standard deviation, and total error of the estimate.   The interval containing the mean is calculated along with the minimum number of samples based on a normal distribution.

**Scatter Plots** may be done to visualize the relationship between the X and Y data sets. This function treats the data sets as coordinate pairs with the X data set containing the X coordinate and the Y data set containing the Y coordinate. The figure below illustrates a typical scatter plot.
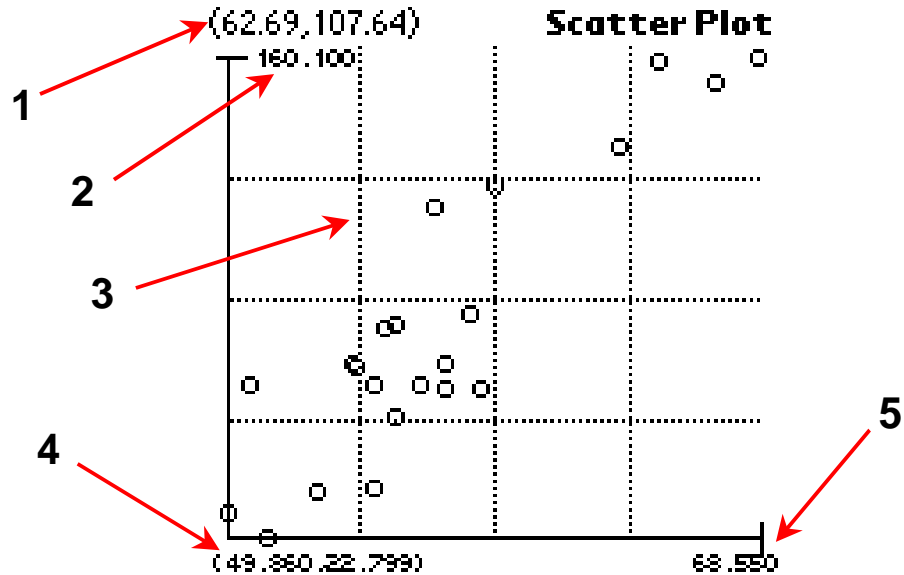


**1 – Coordinate tracking** is available by moving the stylus across the screen. The (x,y) position of the stylus is displayed.

**2 –** The **Y axis** is scaled with respect to the range of the Y data set. This value represents the **maximum Y value** in the data set.
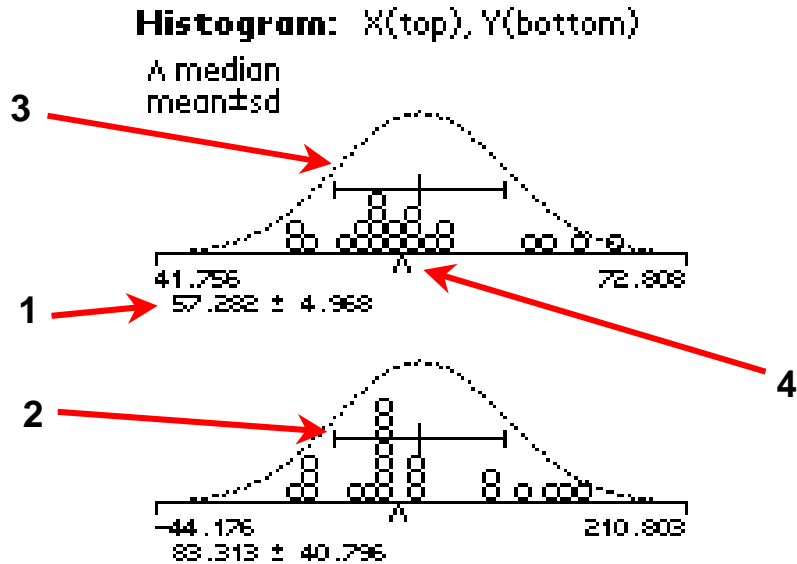
**3 –** Grids are available and may be toggled on and off in the preferences screen. The grids are located at 25, 50 and 75% points on the each axis.

**4 –** This is the **Origin** point for the X-Y axis. This point represents **the minimum x and y point** and is displayed as a coordinate.

**5 –** The **X axis** is scaled with respect to the range of the X data set. This value represents the **maximum X value** in the data set.

The number of points in the X and Y data set *must* be equal in order to do a scatter plot. *This screen is exited by tapping the **HOME** key*.

10

**Histogram Plots** may be done to quickly determine the distribution of a data set.   The figure below illustrates a typical plot of both X and Y data sets. The data set to plot is selected with the Data Set Selector on the main screen. *This screen is exited by tapping the **HOME** key.*



**1,2 –** The **Mean and standard deviation** are calculated and displayed along with corresponding error bars.

**3 – Gauss plots** are available and may be toggled on or off in the preferences menu.

**4 –** the **Median** is calculated and shown with the small **triangle marker**.

**Hypothesis Testing** may be done on all data.  These tests include both X and Y as individual data sets against a given mean,  X and Y against each other, or tests on user supplied means.   All tests may be done with 1, 2.5, 5, and 10% significance levels.  *This screen is exited by tapping the **HOME** key.* The figure below shows the test selection list.
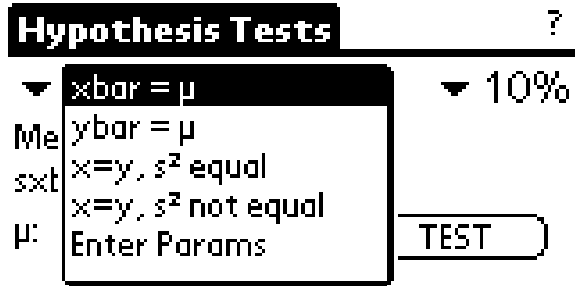


**xbar = $\mu$** and  **ybar = $\mu$**  tests the means of the respective data sets against a user supplied value ($\mu$).

**x=y, $S^2$ equal** test the means of the x and y data sets against each other assuming the sample size and variance are the same.   This test is available in most spread sheet or data analysis packages and is commonly referred to as:  *t-Test: Two-Sample Assuming Equal Variances.*

**x=y, $S^2$ equal** test the means of the x and y data sets against each other assuming the sample size and variance may not be the same.    This test is commonly referred to as:  *t-Test: Two-Sample Assuming Unequal Variances.*

**Enter Params** allows user supplied entries for the test parameters.

***PalmStats™***  is equipped with an internal t-table for use in testing data sets up to *n*=30.   Data sets containing more than 30 points are tested with respect to a normal distribution.   The following figures show the test output of two t-tests.

**Hypothesis Tests**   ?
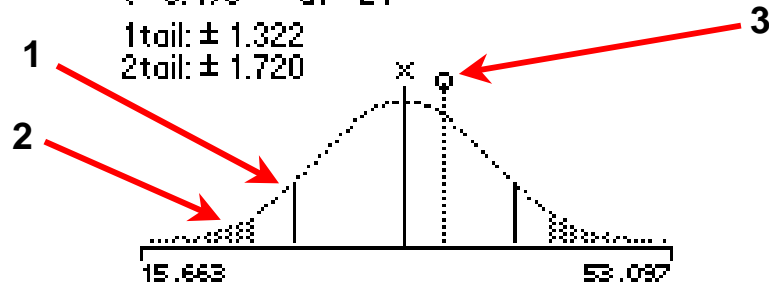
▼ xbar = μ        ▼ 10%

Mean: X: 34.196

sxbar: X: 5.879

μ: 37.1|          ( TEST )

t= 0.493     df= 21

1tail: ± 1.322
2tail: ± 1.720

**1**   **2**   **3**

15.663                53.097

**1, 2 –** The two-tail test limits (1) are shown shaded and the one-tail test limits (2) are displayed as solid lines.   The test significance may be selected with the pick list. In this case it is set to 10%.  The tail areas correspond to this value.

**3 –** The location of the test value (μ) is displayed as a circle and line.

Testing two data set against each other may be done as shown below.  Each distribution is drawn along with the respective tails and test results.

**Hypothesis Tests**   ?

▼ x=y, s² not equal   ▼ 10%

Mean: X: 54.171  Y: 83.313
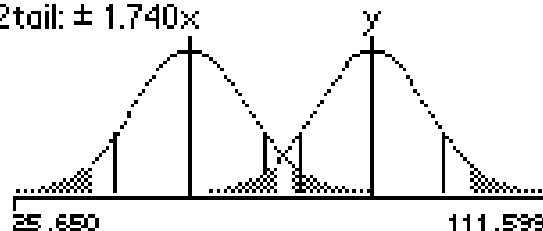
sxbar: X: 9.043  Y: 8.697

μ: |              ( TEST )
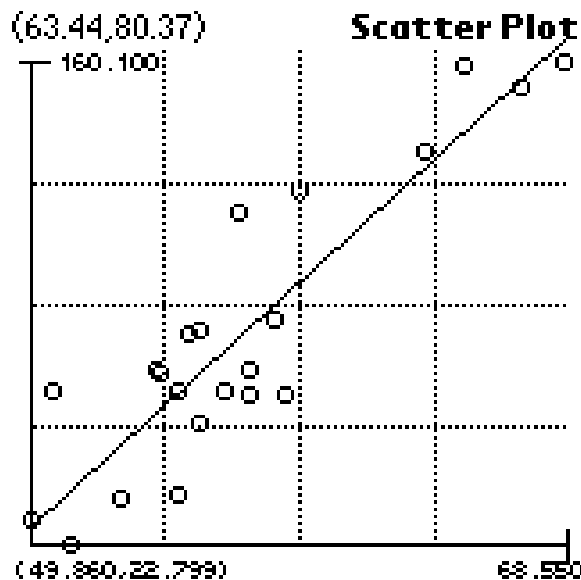
t= -2.322    df= 17

1tail: ± 1.332
2tail: ± 1.740 x          y

25.650                111.599

**Regression Analysis** may be done using three regression models: $y = bx + a$, $y = bx$, or the exponential curve $y = ae^{bx}$.    The figure below shows the regression screen.   *This screen is exited by tapping the __HOME__ key.*



```
Regression                    y=bx + a   a
                                 y = bx
   a:-337.977      Sa:         y = ae^bx

   b:7.354         Sb: .....

   Sxy: 18.592

   Pearson's r: 0.895

   Covariance: 173.298

   ☑ Reg. Line      -point calculator-
   ☐ Residuals      y value: -183.530

                    x value: 21

   ( PLOT )          ( CALCULATE )
```

After **PalmStats™** calculates the regression information, individual points may be calculated using the **point calculator**.   All error values associated with the regression constants are also calculated.    The following figures show typical regression plots of linear data.  The plot screen is exited by tapping the **HOME** key which will return to the regression screen seen above.



```
(63.44,80.37)         Scatter Plot
    160 .100

                                      o    o
                                   o


                  o

                            o
                 oo        o
             o          
           o       o o  o
                      o

              o    o

    o
       o
(49.860,22.799)              68.550
```

The figure above shows a typical regression plot.  As with the scatter plot, the coordinate tracking function may be used be dragging the stylus over the screen.   The following plot shows a regression plot with the residuals plotted.
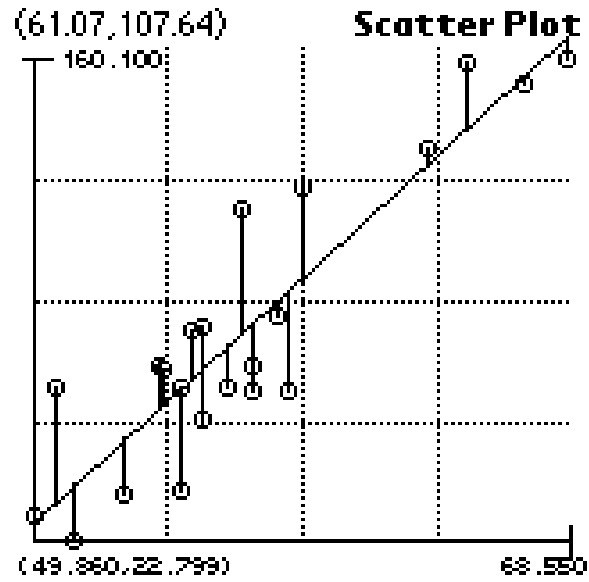
**Probability Calculations** may be done on all data.   These calculations include both X and Y with respect to one or two values, or user specified data input.   All calculations are done assuming a normal distribution with n>30.

```
┌─────────────────────────────────────────────┐
│ ▉Probability▉                          ?     │
│ ───────────────────────────────────────────  │
│ Mean:     57.282                    ┌──┬──┐  │
│                                     │▉X│ Y│  │
│ Std. Dev: 4.968                     ├──┴──┤  │
│                                     │enter│  │
│ µ1: 55                              └─────┘  │
│                                              │
│ µ2: 61                        ( CALCULATE )  │
│                                              │
│ p(x<µ1)= 0.322    p(µ1- mean)=0.177          │
│ p(x>µ1)= 0.677    p(µ1-µ2)=0.449             │
│                                              │
│                                              │
│        41.338              73.536            │
└─────────────────────────────────────────────┘
```
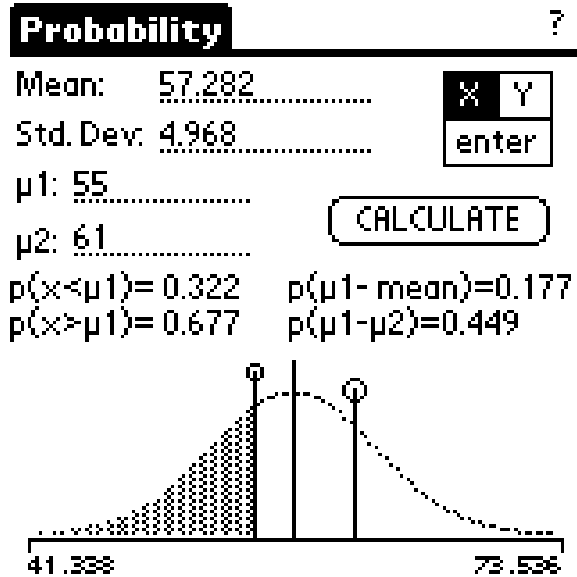
The locations of **µ1** and **µ2** are marked by  small and large circles respectively.   The probability of **µ1** being less than the mean is shaded.
*This screen is exited by tapping the **HOME** key.*

# The Filer

*PalmStats™* is capable of reading and writing comma-delimited memo files. These files are used to store datasets or transfer data to and from a PC during synchronization.
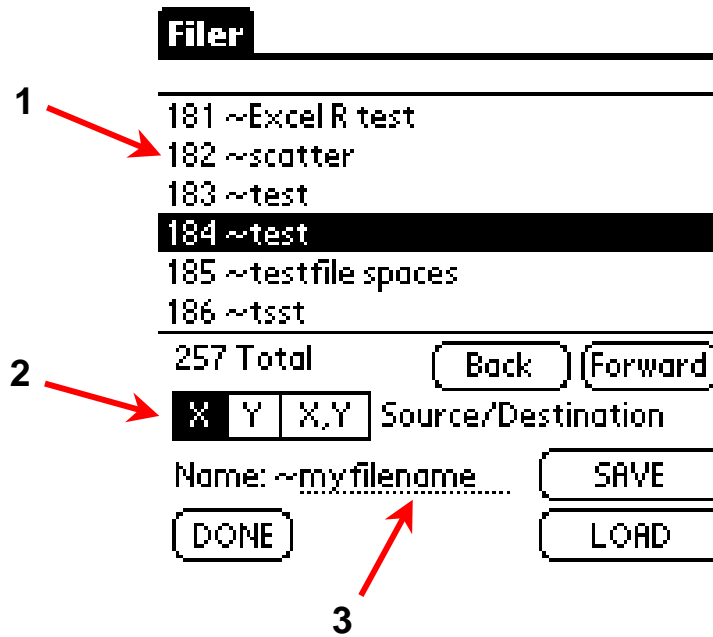


**1 – Filenames** are displayed, along with their number (left margin) as shown in above. The filename list will display <u>all</u> of the memos in the memo pad database twenty at a time. The total number of memos is shown below the list. The **Forward** and **Back** controls are used to step though the memos in increments of 20.

**2 – Source/Destination** controls the place the data is coming from during a **SAVE** operation or the place the data will be going to during a **LOAD**. This control functions the same as the Data Selector on the main screen. For example to **LOAD** the data file into the X data set - - choose a file, select X and tap LOAD.

**3 –** The filename (**Name:**) may be specified during a SAVE to add a descriptive name to the data set. *PalmStats™* adds a **~** character to the front of all the files it saves to distinguish data files from other memos. The filer will only load files beginning with the ~ character. If a **Name** is not specified during a SAVE, the filer will create a unique name based on the date, time, and data sets specified with Source/Destination.

# Data File Format

The figure below shows the comma-delimited format of a *PalmStats™* data file.  Line one is the title and can be up to 40 characters long.  Each remaining line may contain any numeric value separated by a comma.  Spaces are allowed between numbers.  Each line in the data file, including the title line, must be followed by a new line character.

```
Memo 7 of 45          ▼ Unfiled
~Data01
55.77,83.9,
55.05,66.4,
54.27,73.1,
50.63,66.7,
49.86,30.1,
53.04,36.2,
51.33,22.8,
56.70,66.5,
55.07,37,
55.76,58,

( Done ) ( Details )
```

```
Memo 4 of 45          ▼ Unfiled
~409x
6.900,
7.300,
7.800,
7.400,
7.199,
6.599,
6.199,
8.199,
7.599,
5.699,

( Done ) ( Details )
```

**Note:** When data is loaded from a data file into **BOTH X and Y data sets**, data will be alternately put into X and Y.   The first, third, fifth, etc… go into X and the second, fourth, sixth, etc… go into Y.

## References

The calculations used in *PalmStats™* have been derived from the following texts:

Glantz, S. A., <u>Primer of Biostatistics</u>, 4th ed. New York: McGraw-Hill, 1997.

Bevington, R., and Robertson, D., <u>Data Reduction and Error Analysis for the Physical Sciences</u>, 2nd ed. New York: McGraw-Hill, 1992.

Crow, E., Davis, F., and Maxfield, M., <u>Statistics Manual</u>, New York: Dover, 1960.

Taylor, John R., <u>An Introduction to Error Analysis:  The Study of Uncertainties in Physical Measurements</u>, 2nd ed.
Sausalito, CA: University Science Books, 1982.

## Contacting Us:

Email may be sent to **seisupport@home.com**

Our Web Page is  **http://members.home.net/seisupport/index.html**

Any comments or observations are greatly welcome.  *PalmStats™* will be updated, improved and features added based on user's comments. Currently,  *PalmStats™* is being used by students, engineers, medical professionals, and scientists world wide.  It is our goal to produce the finest statistical analysis program available for the Palm Platform and value for the registration cost.

**License Agreement**
You should carefully read the following terms and conditions
before using this software.  Unless you have a different license
agreement signed by Software Engineering Inc., your use of this
software indicates your acceptance of this license agreement and
warranty.

**Registered Version**
One registered copy of PalmStats™ may only be used by a single person
on one Palm Computing$^{(R)}$ Platform device.  Since the registration
code is derived from the PalmOS user name, a new registration code
is required for each Palm Computing Platform device.

**Governing Law**
This agreement shall be governed by the laws of the State of
Indiana.

**Disclaimer of Warranty**

THIS SOFTWARE AND THE ACCOMPANYING FILES ARE SOLD "AS IS" AND
WITHOUT WARRANTIES AS TO PERFORMANCE OR MERCHANTABILITY OR ANY
OTHER WARRANTIES WHETHER EXPRESSED OR IMPLIED.  Because of the
various hardware and software environments into which PalmStats™ may be utilized, NO
WARRANTY OF FITNESS FOR
A PARTICULAR PURPOSE IS OFFERED.

**Good data processing procedure dictates that any program be
thoroughly tested with non-critical data before relying on it.**
The user must assume the entire risk of using the program.  ANY
LIABILITY OF THE SELLER WILL BE LIMITED EXCLUSIVELY TO PRODUCT
REPLACEMENT OR REFUND OF PURCHASE PRICE.

**Evaluation and Registration**

This is not free software.  Subject to the terms below, you are hereby licensed to use this software for evaluation purposes for 30 days only.  If you use this software after the 30 day evaluation period a registration fee of $13.95 is required.  Registration is available through one of the purchasing sites listed.   When payment is received you  will be sent a registration code which should be entered into PalmStats™.

Unregistered use PalmStats™ after the 30-day evaluation period is in violation of U.S. and international copyright laws.