# Contents

PREFACE ........................................................................................................................................... vii

UNDERSTANDING THE DIFFERENCES BETWEEN UNDERSTANDING BASIC STATISTICS 6/E AND UNDERSTANDABLE STATISTICS 10/E ........................................................................................................ viii

## TI-83 PLUS, TI-84 PLUS AND TI-NSPIRE GUIDE

### CHAPTER 1: GETTING STARTED

*About the TI-83 Plus, TI-84 Plus, and TI-Nspire Graphing Calculators* ........................................ 3

*Using the TI-83 Plus, TI-84 Plus, and TI-Nspire* ................................................................. 4

*Computations on the TI-83 Plus, TI-84 Plus, and TI-Nspire* .................................................... 9

**Entering Data**

*TI-83 Plus and TI-84 Plus* ........................................................................................................ 12

*TI-Nspire* .................................................................................................................................. 14

**Comments on Entering and Correcting Data**

*TI-83 Plus and TI-84 Plus* ........................................................................................................ 17

*TI-Nspire* .................................................................................................................................. 18

*Random Samples* .................................................................................................................... 19

*Lab Activities for Random Samples* .......................................................................................... 23

### CHAPTER 2: ORGANIZING DATA

*Histograms*

*TI-83 Plus and TI-84 Plus* ........................................................................................................ 25

*TI-Nspire* .................................................................................................................................. 28

*Lab Activities for Histograms* .................................................................................................... 32

*Bar Graphs and Pie Charts*

*TI-83 Plus and TI-84 Plus* ........................................................................................................ 34

*TI-Nspire* .................................................................................................................................. 34

*Lab Activities for Bar Graphs and Pie Charts* ........................................................................... 37

### CHAPTER 3: AVERAGES AND VARIATION

*One-Variable Statistics*

*TI-83 Plus and TI-84 Plus* ........................................................................................................ 39

*TI-Nspire* .................................................................................................................................. 44

*Lab Activities for One-Variable Statistics* .................................................................................. 48
Grouped Data

TI-83 Plus and TI-84 Plus ................................................................. 51
TI-Nspire ....................................................................................... 53

Lab Activities for Grouped Data ....................................................... 56

Box-and-Whisker Plots

TI-83 Plus and TI-84 Plus ................................................................. 56
TI-Nspire ....................................................................................... 59

Lab Activities for Box-and-Whisker Plots ......................................... 62

CHAPTER 4: CORRELATION AND REGRESSION

Linear Regression

TI-83 Plus and TI-84 Plus ................................................................. 65
TI-Nspire ....................................................................................... 71

Lab Activities for Linear Regression ................................................ 76

CHAPTER 5: ELEMENTARY PROBABILITY THEORY

TI-83 Plus and TI-84 Plus ................................................................. 78
TI-Nspire ....................................................................................... 79

CHAPTER 6: THE BINOMIAL PROBABILITY DISTRIBUTION AND RELATED TOPICS

Discrete Probability Distributions

TI-83 Plus and TI-84 Plus ................................................................. 80
TI-Nspire ....................................................................................... 81

Lab Activities for Discrete Probability Distributions .............................. 82

Binomial Probabilities

TI-83 Plus and TI-84 Plus ................................................................. 82
TI-Nspire ....................................................................................... 83

Lab Activities for Binomial Probabilities .............................................. 87

CHAPTER 7: NORMAL DISTRIBUTIONS

The Area Under any Normal Curve

TI-83 Plus and TI-84 Plus ................................................................. 88
TI-Nspire ....................................................................................... 89

Drawing the Normal Distribution ......................................................... 90
Lab Activities for the Area Under any Normal Curve ................................................................. 92

Sampling Distributions

TI-83 Plus and TI-84 Plus ........................................................................................................ 92
TI-Nspire .................................................................................................................................. 93

CHAPTER 8: ESTIMATION

Confidence Intervals for a Population Mean

TI-83 Plus and TI-84 Plus ........................................................................................................ 95
TI-Nspire .................................................................................................................................. 98

Lab Activities for Confidence Intervals for a Population Mean ........................................... 101

Confidence Intervals for the Probability of Success $p$ in a Binomial Distribution

TI-83 Plus and TI-84 Plus ........................................................................................................ 101
TI-Nspire .................................................................................................................................. 102

Lab Activities for Confidence Intervals for the Probability of Success $p$ in a Binomial Distribution .. 103

CHAPTER 9: HYPOTHESIS TESTING

Testing a Single Population Mean

TI-83 Plus and TI-84 Plus ........................................................................................................ 105
TI-Nspire .................................................................................................................................. 107

Lab Activities for Testing a Single Population Mean .......................................................... 109

Testing Involving a Single Proportion

TI-83 Plus and TI-84 Plus ........................................................................................................ 110
TI-Nspire .................................................................................................................................. 110

CHAPTER 10: INFERENCE ABOUT DIFFERENCES

Tests Involving Paired Differences (Dependent Samples)

TI-83 Plus and TI-84 Plus ........................................................................................................ 111
TI-Nspire .................................................................................................................................. 112

Lab Activities Using Tests Involving Paired Differences (Dependent Samples) ....................... 114

Tests of Difference of Means (Independent Samples)

TI-83 Plus and TI-84 Plus ........................................................................................................ 115
TI-Nspire .................................................................................................................................. 116

Testing a Difference of Proportions

TI-83 Plus and TI-84 Plus ........................................................................................................ 117
TI-Nspire .................................................................................................................................. 118
Lab Activities for Testing Difference of Means (Independent Samples) or Proportions .................. 118

CHAPTER 11: ADDITIONAL TOPICS USING INference

Chi-Square Test of Independence

TI-83 Plus and TI-84 Plus .................................................................................................................. 119

TI-Nspire .......................................................................................................................................... 121

Lab Activities for Chi-Square Test of Independence ...................................................................... 124

APPENDIX

PREFACE ........................................................................................................................................... A-3

SUGGESTIONS FOR USING THE DATA SETS ............................................................................... A-3

DESCRIPTIONS OF DATA SETS .................................................................................................... A-4
Preface

The use of computing technology can greatly enhance a student’s learning experience in statistics. *Understanding Basic Statistics* is accompanied by four Technology Guides, which provide basic instructions, examples, and lab activities for four different tools:

- TI-83 Plus, TI-84 Plus and TI-Nspire
- Microsoft Excel ©2010 with Analysis ToolPak for Windows ®
- MINITAB Version 15
- SPSS Version 18

The TI-83 Plus, TI-84 Plus and TI-Nspire are versatile, widely available graphing calculators made by Texas Instruments. The calculator guide shows how to use their statistical functions, including plotting capabilities.

Excel is an all-purpose spreadsheet software package. The Excel guide shows how to use Excel’s built-in statistical functions and how to produce some useful graphs. Excel is not designed to be a complete statistical software package. In many cases, macros can be created to produce special graphs, such as box-and-whisker plots. However, this guide only shows how to use the existing, built-in features. In most cases, the operations omitted from Excel are easily carried out on an ordinary calculator. The Analysis ToolPak is part of Excel and can be installed from the same source as the basic Excel program (normally, a CD-ROM) as an option on the installer program’s list of Add-Ins. Details for getting started with the Analysis ToolPak are in Chapter 1 of the Excel guide. No additional software is required to use the Excel functions described.

MINITAB is a statistics software package suitable for solving problems. It can be packaged with the text. Contact Cengage Learning for details regarding price and platform options.

SPSS is a powerful tool that can perform many statistical procedures. The SPSS guide shows how to manage data and perform various statistical procedures using this software.

The lab activities that follow accompany the text *Understanding Basic Statistics*, 6th edition by Brase and Brase. On the following page is a table to coordinate this guide with *Understandable Statistics*, 10th edition by Brase and Brase. Both texts are published by Cengage Learning.

In addition, over one hundred data files from referenced sources are described in the Appendix. These data files are available via download from the Cengage Learning Web site:

http://www.cengage.com/statistics/brase
Understanding the Differences Between *Understanding Basic Statistics* 6/e and *Understandable Statistics* 10/e

*Understandable Basic Statistics* is the brief, one-semester version of the larger book. It is currently in its Sixth Edition.

*Understandable Statistics* is the full, two-semester introductory statistics textbook, which is now in its Tenth Edition.

Unlike other brief texts, *Understanding Basic Statistics* is not just the first six or seven chapters of the full text. Rather, topic coverage has been shortened in many cases and rearranged, so that the essential statistics concepts can be taught in one semester.

The major difference between the two tables of contents is that Regression and Correlation are covered much earlier in the brief textbook. In the full text, these topics are covered in Chapter 9. In the brief text, they are covered in Chapter 4.

Analysis of Variance (ANOVA) is not covered in the brief text.

*Understanding Statistics* has 11 chapters and *Understanding Basic Statistics* has 11. The full text is a hard cover book, while the brief is softcover.

The same pedagogical elements are used throughout both texts.

The same supplements package is shared by both texts.

Following are the two Tables of Contents, side-by-side:

<table>
<thead>
<tr>
<th>Chapter</th>
<th><em>Understandable Statistics</em> (full)</th>
<th><em>Understanding Basic Statistics</em> (brief)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1</td>
<td>Getting Started</td>
<td>Getting Started</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>Organizing Data</td>
<td>Organizing Data</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>Averages and Variation</td>
<td>Averages and Variation</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>Elementary Probability Theory</td>
<td>Correlation and Regression</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>The Binomial Probability Distribution and Related Topics</td>
<td>Elementary Probability Theory</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>Normal Curves and Sampling Distributions</td>
<td>The Binomial Probability Distribution and Related Topics</td>
</tr>
<tr>
<td>Chapter 7</td>
<td>Estimation</td>
<td>Normal Curves and Sampling Distributions</td>
</tr>
<tr>
<td>Chapter 8</td>
<td>Hypothesis Testing</td>
<td>Estimation</td>
</tr>
<tr>
<td>Chapter 9</td>
<td>Correlation and Regression</td>
<td>Hypothesis Testing</td>
</tr>
<tr>
<td>Chapter 10</td>
<td>Chi-Square and F Distributions</td>
<td>Inference About Differences</td>
</tr>
<tr>
<td>Chapter 11</td>
<td>Nonparametric Statistics</td>
<td>Additional Topics Using Inference</td>
</tr>
</tbody>
</table>
TI-83 PLUS, TI-84 PLUS and TI-NSPIRE Graphing Calculators Guide
CHAPTER 1: GETTING STARTED

ABOUT THE TI-83 PLUS, TI-84 PLUS AND TI-NSPIRE GRAPHING CALCULATORS

Calculators with built-in statistical support provide tremendous aid in performing the calculations required in statistical analysis. The Texas Instruments TI-83 Plus and TI-84 Plus graphing calculators have many features that are particularly useful in an introductory statistics course. Among the features are:

(a) Data entry in a spreadsheet-like format

The TI-83 Plus and TI-84 Plus have six columns (called lists: L1, L2, L3, L4, L5, L6) in which data can be entered. There are also additional unnamed lists that allow the user to enter a list with the name of their choice. The TI-NPSIRE has a similar format to inputting lists. Here list names MUST be provided by the user. The data in a list can be edited, and new lists can be created using arithmetic that uses existing lists.

Sample Data Screen from a TI-83 Plus or TI-84 Plus

![Sample Data Screen from a TI-83 Plus or TI-84 Plus](image)

(b) Single-variable statistics: mean, standard deviation, median, maximum, minimum, quartiles 1 and 3, sums

(c) Graphs for single-variable statistics: histograms, box-and-whisker plots
(d) Estimation: confidence intervals using the normal distribution, or Student’s t distribution, for a single mean and for a difference of means; intervals for a single proportion and for the difference of two proportions

(e) Hypothesis testing: single mean (z or t); difference of means (z or t); proportions, difference of proportions, chi-square test of independence; two variances; linear regression, one-way ANOVA

(f) Two-variable statistics: linear regression

(g) Graphs for two-variable statistics: scatter diagrams, graph of the least-squares line

In this Guide we will show how to use many features on the TI-83 Plus, TI-84 Plus and TI-Nspire graphing calculators to aid you as you study particular concepts in statistics. Lab Activities coordinated to the text Understanding Basic Statistics, Sixth Edition, are also included.

**USING THE TI-83 PLUS, TI-84 PLUS AND TI-NSPIRE**

**TI-83 Plus and TI-84 Plus**

The TI-83 Plus and TI-84 Plus have several functions associated with each key, and both calculators use the same keystroke sequences to perform those functions. With the TI-83 Plus, all the yellow items on the keypad are accessed by first pressing the yellow 2nd key. On the TI-84 Plus these keys are blue in color. For both calculators, the items in green are accessed by first pressing the green ALPHA key. The four arrow keys enable you to move through a menu screen or along a graph. On the following page is the image of a TI-84 Plus calculator. The keypad of the TI-83 Plus is exactly the same as that of the TI-84 Plus.
The calculators each use screen menus to access additional operations. For instance, to access the statistics menu on the TI-83 Plus or TI-84 Plus, press STAT. Your screen should look like this:

```
       CALC TESTS
   1:Edit...
   2:SortA(  
   3:SortD(  
   4:ClrList  
   5:SetUpEditor
```

Use the arrow keys to highlight your selection. Pressing ENTER selects the highlighted item.

To leave a screen, press either 2nd [QUIT] or CLEAR, or select another menu from the keypad.

Now press MODE. When you use your TI-83 Plus or TI-84 Plus for statistics, the most convenient settings are as shown. (The TI-83 Plus does not have the SET CLOCK line; otherwise, it is identical to the TI-84 Plus).

```
NORMAL SCI ENG
FLOAT 0123456789
RADIANT DEGREE
FUNC PAR POL SEQ
CONNECTED DOT
SEQUENTIAL SIMUL
REAL a+bi re^@i
FULL HORIZ G-T
SET CLOCK:05/18/07 7:02AM
```

You can enter the settings by using the arrow keys to highlight selections and then pressing ENTER.

**TI-Nspire**

The TI-Nspire also has several functions associated with some of its keys. Options are listed in light blue above each key and can be accessed by first pressing the light blue ctrl button. In addition, the touchpad located underneath the viewing screen can be to navigate through menu options or along a graph. On the following page is the image of a TI-Nspire calculator.
The TI-Nspire also uses on-screen menus to access additional operations. For instance, to access the statistics menu, press \( \text{on} \). Using the touchpad, highlight the \( \text{on} \) icon then press the middle of the touchpad where you see the \( \text{on} \) icon. To move within menus you can also use the tab key. Your screen should look like:

![Menu Screen]

To leave a menu, press the esc key.

To change settings, press the \( \text{on} \) key, and choose 5:Settings and Status. Then, select 2:Settings and finally select 1:General. Here you can modify the number of display digits, the angle measure that you are using and other settings.

![Settings Screen]

You can enter settings by choosing options from the drop-down menus by using your touchpad.
COMPUTATIONS ON THE TI-83 PLUS, TI-84 PLUS AND TI-NSPIRE

TI-83 Plus and TI-84 Plus

In statistics, you will be evaluating a variety of expressions. The following examples demonstrate some basic keystroke patterns for the TI-83 Plus and TI-84 Plus.

Example

Evaluate \(-2(3) + 7\).

Use the following keystrokes: \([-][2][\times][3][+]7[\text{ENTER}]\) The result is 1.

To enter a negative number, be sure to use the key (-) rather than the subtract key. Notice that the expression \(-2\times3 + 7\) appears on the screen. When you press ENTER, the result 1 is shown on the far right side of the screen.

Example

(a) Evaluate \(\frac{10-7}{3.1}\) and round the answer to three places after the decimal.

A reliable approach to evaluating fractions is to enclose the numerator in parentheses, and if the denominator contains more than a single number, enclose it in parentheses as well.

\[\frac{10-7}{3.1} = (10-7) \div 3.1\]

Use the keystrokes \([-][1][0][-][7][\div][3.1][\text{ENTER}]\)

The result is 0.9677419355, which rounds to 0.968.

(b) Evaluate \(\frac{10-7}{2}\) and round the answer to three places after the decimal.

Place both numerator and denominator in parentheses: \((10 - 7) \div (3.1 \div 2)\)

Use the keystrokes \([-][1][0][-][7][\div][3.1][-][\div][2][\text{ENTER}]\)

The result is 1.935483871, or 1.935 rounded to three places after the decimal.
Example

Several formulas in statistics require that we take the square root of a value. Note that a left parenthesis, (, is automatically placed next to the square root symbol when 2nd [√] is pressed. Be sure to close the parentheses after typing in the radicand.

(a) Evaluate $\sqrt{10}$ and round the result to three places after the decimal.

$$2\text{nd}[\sqrt{10}]\text{ ENTER}$$

The result is 3.16227766, which rounds to 3.162.

(b) Evaluate $\sqrt{\frac{10}{3}}$ and round the result to three places after the decimal.

$$2\text{nd}[\sqrt{\frac{10}{3}}]\text{ ENTER}$$

The result rounds to 1.054.

Be careful to close the parentheses. If you do not close the parentheses, you will get the result of $\sqrt{\frac{10}{3}} \approx 1.826$

Example

Some expressions require us to use powers.

(a) Evaluate $3.2^2$.

In this case, we can use the $x^2$ key.

$$3,2 \text{ ENTER}$$

The result is 10.24.

(b) Evaluate $0.4^3$.

In this case, we use the $^\text{k}$ key.

$$0.4 \text{ ENTER}$$

The result is 0.064.

TI-Nspire

The following demonstrates basic keystrokes for the TI-Nspire. In order to evaluate expression, start by pressing $\text{on}$ and selecting $\text{A:Calculator}$.
Example

Evaluate \(-2(3) + 7\).

Use the following keystrokes: \([-2 \times 3 + 7 \text{ ENTER}\] The result is 1.

To enter a negative number, be sure to use the key (-) rather than the subtract key. Notice that the expression -2*3 + 7 appears on the screen. When you press enter, the result 1 is shown on the far right side of the screen.

Example

(c) Evaluate \(\frac{10-7}{3.1}\) and round the answer to three places after the decimal.

A reliable approach to evaluating fractions is to enclose the numerator in parentheses, and if the denominator contains more than a single number, enclose it in parentheses as well.

\[\frac{10-7}{3.1} = (10-7) \div 3.1\]

Use the keystrokes \([10-7 \div 3.1 \text{ ENTER}\]

The result is 0.967419355, which rounds to 0.968.

(b) Evaluate \(\frac{10-7}{3.1}\) and round the answer to three places after the decimal.

Place both numerator and denominator in parentheses: \((10 - 7) \div (3.1 \div 2)\)

Use the keystrokes \([10-7 \div (3.1 \div 2) \text{ ENTER}\]

The result is 1.935483871, or 1.935 rounded to three places after the decimal.

Example

Several formulas in statistics require that we take the square root of a value. Note that a left parenthesis, (, is automatically placed next to the square root symbol when ctrl [\(\sqrt{\quad}\)] is pressed. Unlike with the TI-83 Plus and TI-84 Plus, both parentheses are provided when a function is selected.

(a) Evaluate \(\sqrt{10}\) and round the result to three places after the decimal.
ctrl $\sqrt{\frac{1}{3}}$ 1 0 enter

The result is 3.16227766, which rounds to 3.162.

(d) Evaluate $\sqrt{\frac{1}{3}}$ and round the result to three places after the decimal.

ctrl $\sqrt{\frac{1}{3}}$ 1 0 ÷ 3 enter

The result rounds to 1.054.

Example

Some expressions require us to use powers.

(c) Evaluate $3.2^2$.

In this case, we can use the $x^2$ key.

3 $\div$ 2 $\times$ 3 ENTER

The result is 10.24.

(d) Evaluate $0.4^3$.

In this case, we use the $^3$ key.

4 $\div$ 3 ENTER

The result is 0.064.

ENTERING DATA

TI-83 Plus and TI-84 Plus

To use the statistical processes built into the TI-83 Plus and TI-84 Plus, we first enter data into lists. Press the STAT key. Next we will clear any existing data lists. With EDIT highlighted, select 4:ClrList using the arrow keys and press ENTER.
Then type in the six lists separated by commas as shown. (The lists are found using the ` key, followed by 1, 2, 3, 4, 5, and 6.) Press ENTER.

Next press STAT again and select item 1:Edit. You will see the data screen and are now ready to enter data.

Let’s enter the numbers.

2 5 7 9 11 in list L₁
3 6 9 11 1 in list L₂

Press ENTER after each number, and use the arrow keys to move between L₁ and L₂.

To correct a data entry error, highlight the entry that is wrong and enter the correct data value.

We can also create new lists by doing arithmetic on existing lists. For instance, let’s create L₃ by using the formula L₃ = 2L₁ + L₂.
Highlight \( L_3 \). Then type in \( 2L_1+L_2 \) and press ENTER.

The final result is shown on the next page.

To leave the data screen, press 2nd [QUIT] or press another menu key, such as STAT.

**TI-Nspire**

To use the statistical processes built into the TI-Nspire, we first enter data into lists. Press the \( \text{on} \) key and select \( \text{on} \). Next we will clear any existing data lists. If there is existing data in any column, simply navigate to the top of the column and press the up arrow one more time to highlight the column. Press ctrl then del to clear all entries.
You should now see a blank spreadsheet similar to a Microsoft Excel spreadsheet.

Let’s enter the numbers

2  5  7  9  11  in the first column
3  6  9  11  1  in the second column

Recall that we will need to name our lists, let’s call the first list L1 and the second list L2. To name a list, arrow to the top of the column and enter the name of the list using the keypad.
In order to enter data, arrow to the first row in the spreadsheet and type your data. Press enter after each number.

To correct a data entry error, highlight the entry that is wrong and enter the correct data value.

We can also create new lists by doing arithmetic on existing lists. For instance, let’s create L3 by using the formula $L3 = 2L1 + L2$.

Name the third column L3. Then, in the box directly under the name, type in $=2*L1+L2$ and press enter.

If a Conflict Detected window displays on the screen, select Variable Reference and press OK for each list.
The final result is shown below.

To leave the data screen, press the \text{\textbf{on}} key.

**COMMENTS ON ENTERING AND CORRECTING DATA**

**TI-83 Plus and TI-84 Plus**

(a) To create a new list from old lists, the lists must all have the same number of entries.

(b) To delete a data entry, highlight the data value you wish to correct and press the \text{DEL} key.

(c) To insert a new entry into a list, highlight the position directly below the place you wish to insert the new value. Press 2nd [INS] and then enter the data value.
TI-Nspire

(a) To create a new list from old lists, the lists must all have the same number of entries.

(b) To delete a data entry, highlight the cell and press ctrl then menu. Then select 1:Cut.

(c) To insert a new entry into a list, highlight the position where you wish to insert a new value. Press ctrl then menu. Then select 4:Insert Cell.

LAB ACTIVITIES TO GET STARTED USING THE TI-83 PLUS AND TI-84 PLUS

1. Practice doing some calculations using the TI-83 Plus and TI-84 Plus. Be sure to use parentheses around a numerator or denominator that contains more than a single number. Round all answers to three places after the decimal.

   (a) \( \frac{5-2.3}{1.3} \) Ans. 2.077  
   (b) \( \frac{8-3.3}{2} \) Ans. 3.133  
   (c) \(-2(3.4) + 5.8\) Ans. -1  
   (d) \(-4(-1.7) - 2.1\) Ans. 4.7  
   (e) \(\sqrt{5.3}\) Ans. 2.302  
   (f) \(\sqrt{6 + 3(2)}\) Ans. 3.464  
   (g) \(\sqrt{\frac{8-2.7}{5-1}}\) Ans. 1.151  
   (h) \(1.5^2\) Ans. 2.25  
   (i) \((5 - 7.2)^2\) Ans. 4.84  
   (j) \((0.7)^3(0.3)^2\) Ans. 0.031

2. Enter the following data into the designated list. Be sure to clear all lists first.

   L1: 3 7 9.2 12 -4
   L2: -2 9 4.3 16 10
   L3: L1 - 2
   L4: -2L2-L1
   L1: Change the data value in the second position to 6.

LAB ACTIVITIES TO GET STARTED USING THE TI-Nspire

2. Practice doing some calculations using the TI-Nspire. Be sure to use parentheses around a numerator or denominator that contains more than a single number. Round all answers to three places after the decimal.

   (b) \( \frac{5-2.3}{1.3} \) Ans. 2.077  
   (b) \( \frac{8-3.3}{2} \) Ans. 3.133  
   (c) \(-2(3.4) + 5.8\) Ans. -1  
   (d) \(-4(-1.7) - 2.1\) Ans. 4.7
(e) \( \sqrt{5.3} \)  
\[ \text{Ans. 2.302} \]

(f) \( \sqrt{6 + 3(2)} \)  
\[ \text{Ans. 3.464} \]

(g) \( \frac{8-2.7}{\sqrt{5-1}} \)  
\[ \text{Ans. 1.151} \]

(h) \( 1.5^2 \)  
\[ \text{Ans. 2.25} \]

(i) \( (5 - 7.2)^2 \)  
\[ \text{Ans. 4.84} \]

(j) \( (0.7)^3(0.3)^2 \)  
\[ \text{Ans. 0.031} \]

2. Enter the following lists and data. Be sure to clear all lists first.

L1: 3 7 9.2 12 -4

L2: -2 9 4.3 16 10

L3: L1 – 2

L4: -2L2-L1

L1: Change the data value in the second position to 6.

**RANDOM SAMPLES (SECTION 1.2 OF UNDERSTANDING BASIC STATISTICS)**

**TI-83 Plus and TI-84 Plus**

The TI-83 Plus and TI-84 Plus graphics calculators have a random number generator, which can be used in place of a random number table. Press MATH. Use the arrow keys to highlight PRB. Notice that **1:rand** is selected.

When you press ENTER, rand appears on the screen. Press ENTER again and a random number between 0 and 1 appears.
To generate a random whole number with up to 3 digits, multiply the output of \texttt{rand} by 1000 and take the integer part. The integer-part command is found in the MATH menu under NUM, selection 3:iPart. Press ENTER to display the command on the main screen.

Now let’s put all commands together to generate random whole numbers with up to 3 digits. First display \texttt{iPart}. Then enter \texttt{1000*rand} and close the parentheses. Now, each time you press ENTER, a new random number will appear. Notice that the numbers might be repeated.
To generate random whole numbers with up to 2 digits, multiply `rand` by 100 instead of 1000. For random numbers with 1 digit, multiply `rand` by 10.

Simulating experiments in which outcomes are equally likely is an important use of random numbers.

**Example**

Use the TI-83 Plus or TI-84 Plus to simulate the outcomes of tossing a die six times. Record the results. In this case, the possible outcomes are the numbers 1 through 6. If we generate a random number outside of this range we will ignore it. Since we want random whole numbers with 1 digit, we will follow the method prescribed earlier and multiply each random number by 10. Press ENTER 6 times.

Of the first six values listed, we only use the outcomes 3, 2, and 2. To get the remaining outcomes, keep pressing ENTER until three more digits in the range 1 through 6 appear. When we did this, the next three such digits were 1, 4, and 3. Your results will be different, because each time the random number generator is used, it gives a different sequence of numbers.
TI-Nspire

The TI-NSpire graphics calculators have a random number generator, which can be used in place of a random number table. From the Calculator page, press the menu key and select 5:Probability. Use the arrow keys to select 4:Random.

To produce random integers, choose 2:Integers. Enter the smallest value followed by the largest value and press enter to generate a random integer.

To generate multiple random numbers, follow the procedure above and add a third argument to the randInt function indicating the number of random integers that you would like to produce.
We can also use the \texttt{rand} command to produce random numbers between 0 and 1 that are not integers. From the Random menu listing (from Calculator page: menu $\rightarrow$ Random), choose 1:\textbf{Number}. The \texttt{rand()} command will create a number between 0 and 1. There are no values needed in the parentheses. If you place a number in the function, it will produce that number of random numbers between 0 and 1.

\begin{itemize}
  \item \texttt{rand(1)}\hspace{1cm} 24
  \item \texttt{rand(1,25)}\hspace{1cm} \{23,4,13\}
  \item \texttt{rand(1,25,3)}\hspace{1cm} \{23,4,13\}
  \item \texttt{rand()}\hspace{1cm} 0.40581
  \item \texttt{rand(3)}\hspace{1cm} \{0.733812,0.043992,0.339363\}
\end{itemize}

**LAB ACTIVITIES FOR RANDOM SAMPLES**

**TI-83 Plus and TI-84 Plus**

1. Out of a population of 800 eligible county residents, select a random sample of 20 people for prospective jury duty. By what value should you multiply \texttt{rand} to generate random whole numbers with 3 digits? List the first 20 numbers corresponding to people for prospective jury duty.
2. We can simulate dealing bridge hands by numbering the cards in a bridge deck from 1 to 52. Then we draw a random sample of 13 numbers without replacement from the population of 52 numbers. A bridge deck has 4 suits: hearts, diamonds, clubs, and spades. Each suit contains 13 cards: those numbered 2 through 10,
a jack, a queen, a king, and an ace. Decide how to assign the numbers 1 through 52 to the cards in the deck. Use the random number generator on the TI-83 Plus or TI-84 Plus to get the numbers of the 13 cards in one hand. Translate the numbers to specific cards and tell what cards are in the hand. For a second game, the cards would be collected and reshuffled. Using random numbers from the TI-83 Plus or TI-84 Plus, determine the hand you might get in a second game.

**TI-Nspire**

1. Out of a population of 800 eligible county residents, select a random sample of 20 people for prospective jury duty.
2. We can simulate dealing bridge hands by numbering the cards in a bridge deck from 1 to 52. Then we draw a random sample of 13 numbers without replacement from the population of 52 numbers. A bridge deck has 4 suits: hearts, diamonds, clubs, and spades. Each suit contains 13 cards: those numbered 2 through 10, a jack, a queen, a king, and an ace. Decide how to assign the numbers 1 through 52 to the cards in the deck. Use the random number generator on the TI-Nspire to get the numbers of the 13 cards in one hand. Translate the numbers to specific cards and tell what cards are in the hand. For a second game, the cards would be collected and reshuffled. Using random numbers from the TI-Nspire, determine the hand you might get in a second game.
CHAPTER 2: ORGANIZING DATA

HISTOGRAMS (SECTION 2.1 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

The TI-83 Plus and TI-84 Plus graphing calculators draw histograms for data entered in the data lists. The calculators follow the convention that if a data value falls on a class limit, then it is tallied in the frequency of the bar to the right. However, if you specify the lower class boundary of the first class and the class width (see Understanding Basic Statistics for procedures to find these values), then no data will fall on a class boundary. The following example shows you how to draw histograms.

Example

Throughout the day from 8 A.M. to 11 P.M., Tiffany counted the number of ads occurring every hour on one commercial T.V. station. The 15 data values are as follows:

\[
\begin{array}{cccccccc}
10 & 12 & 8 & 7 & 15 & 6 & 5 & 8 \\
8 & 10 & 11 & 13 & 15 & 8 & 9 & \\
\end{array}
\]

To make a histogram with 4 classes, first enter the data in L₁.

Next, we will set the graphing window. However, we need to know both the lower class boundary for the first class and the class width. Use the techniques in Section 2.1 of Understanding Basic Statistics to find these values.

The smallest data value is 5, so the lower class boundary of the first class is 4.5. The class width is

\[
\frac{\text{largest data value} - \text{smallest data value}}{\text{Number of classes}}, \text{increased to the next integer}
\]

\[
\frac{15 - 5}{4} = 2.5, \text{increased to 3}
\]

Now we set the graphing window. Press the WINDOW key.
Select \textit{Xmin} and enter the \textit{lower class boundary} of the first class. For this example, \textbf{Xmin} = 4.5.

Setting \textbf{Xmax} = 17 ensures that the histogram fits on the screen, since our highest data value is 15.

Select \textbf{Xscl} and enter the \textit{class width}. For this example, \textbf{Xscl} = 3.

Setting \textbf{Ymin} = -5 and \textbf{Ymax} = 15 makes sure the histogram fits comfortably on the screen.

![Window Settings]

The next step is to select histogram style for the plot. Press 2nd \textbf{[STAT PLOT]}.

Highlight 1:Plot1 and press \textbf{ENTER}.

![STAT PLOTS]

Then on the next screen select \textbf{On} and the histogram shape.

In this example we stored the data in \textit{L1}. Set \textit{Xlist} to \textit{L1} by pressing 2nd [\textit{L1}]. Each data value is listed once for each time it occurs, so 1 is entered for \textit{Freq}. 

![Data List]
We are ready to graph the histogram. Press GRAPH.

Press TRACE and notice that a blinking cursor appears over the first bar. The class boundaries of the bar are given as min and max. The frequency is given by n. Use the right and left arrow keys to move from bar to bar.
After you have graphed the histogram, you can use TRACE and the arrow keys to find the class boundaries and frequency of each class. The calculator has done the work of sorting the data and tallying the data for each class.

**TI-Nspire**

The TI-Nspire draws histograms for data entered in the data lists. The calculators follow the convention that if a data value falls on a class limit, then it is tallied in the frequency of the bar to the right. However, if you specify the lower class boundary of the first class and the class width (see *Understanding Basic Statistics* for procedures to find these values), then no data will fall on a class boundary. The following example shows you how to draw histograms.

**Example**

Throughout the day from 8 A.M. to 11 P.M., Tiffany counted the number of ads occurring every hour on one commercial T.V. station. The 15 data values are as follows:

10 12 8 7 15 6 5 8
8 10 11 13 15 8 9

To make a histogram with 4 classes, first enter the data into a list named L1.

To create the histogram, while in the Spreadsheet View, press menu then choose 3:Data. From the Data submenu, choose 6:Quick Graph. This will display a dotplot of the data. In order to change the graph to a histogram, press menu and select 1:Plot Type. From the Plot Type submenu, select 3:Histogram.
You can set your own class width by pressing menu while the graphing window is selected. Scroll and select to 2:Plot Properties. From the Plot Properties menu, select 2:Histogram Properties. From the Histogram Properties submenu, choose 2:Bin Settings.

We can determine the class width using the formula:

\[
\frac{\text{largest data value} - \text{smallest data value}}{\text{Number of classes}}, \text{increased to the next integer}
\]

\[
\frac{15 - 5}{4} = 2.5, \text{increased to 3}
\]
Let's change the class width to 3. Access the bin settings as described above and change the Width to 3 then press OK.

The graphing window will be automatically configured to appropriately display the graph and choose the number of classes and class width. However, after we changed the class width, the histogram no longer fits in the graph window. Now we will display how to set the graphing window to appropriately display the histogram. Select the graphing window (recall, you can switch between windows by pressing ctrl then tab). Then press menu. Scroll to and select 5:Window/Zoom, then select 1:Window Settings.

Select Xmin and enter the lower class boundary of the first class. For this example, Xmin = 4.5.

Setting Xmax =17 ensures that the histogram fits on the screen, since our highest data value is 15.

Setting Ymax=15 makes sure the histogram fits comfortably on the screen. Notice that you cannot change Ymin, this is because there are no negative counts in a histogram.
After you have graphed the histogram, you can use the touchpad (like you would use a mouse) to hover your cursor over a bar to show the class boundaries and the number of points in the class.

LAB ACTIVITIES FOR HISTOGRAMS

TI-83 Plus, TI-84 Plus and TI-Nspire

1. A random sample of 50 professional football players produced the following data. (See Weights of Pro Football Players (data file Sv02.zip))

   The following data represents weights in pounds of 50 randomly selected pro football linebackers.


   225 230 235 238 232 227 244 222
   250 226 242 253 251 225 229 247
   239 223 233 222 243 237 230 240
   255 230 245 240 235 252 245 231
   235 234 248 242 238 240 240 240
   235 244 247 250 236 246 243 255
   241 245

   Enter this data in L1. Scan the data for the low and high values. Compute the class width for 5 classes and find the lower class boundary for the first class. Use the TI-83 Plus, TI-84 Plus, or TI-Nspire to make a histogram with 5 classes. Repeat the process for 9 classes. Is the data distribution skewed or symmetric? Is the distribution shape more pronounced with 5 classes or with 9 classes?

2. Explore the other data files found in the Appendix of this Guide, such as

   Disney Stock Volume (data file Sv01.zip)
   Heights of Pro Basketball Players (data file Sv03.zip)
   Miles per Gallon Gasoline Consumption (data file Sv04.zip)
Select one of these files and make a histogram with 7 classes. Comment on the histogram shape.

3. (a) Consider the following data:

```
1 3 7 8 10
6 5 4 2 1
9 3 4 5 2
```

Place the data in L₁. Use the TI-83 Plus, TI-84 Plus, or TI-Nspire to make a histogram with \( n = 3 \) classes. Jot down the class boundaries and frequencies so that you can compare them to part (b).

(b) Now add 20 to each data value of part (a). The results are as follows:

```
21 23 27 28 30
26 25 24 22 21
29 24 25 22
```

By using arithmetic in the data list screen, you can create \( L₂ \) and let \( L₂ = L₁ + 20 \).

Make a histogram with 3 classes. Compare the class boundaries and frequencies with those obtained in part (a). Are each of the boundary values 20 more than those of part (a)? How do the frequencies compare?

(c) Use your discoveries from part (b) to predict the class boundaries and class frequency with 3 classes for the data values below. What do you suppose the histogram will look like?

```
1001 1003 1007 1008 1010
1006 1005 1004 1002 1001
1009 1003 1004 1005 1002
```

Would it be safe to say that we simply shift the histogram of part (a) 1000 units to the right?

(d) What if we multiply each of the values of part (a) by 10? Will we effectively multiply the entries for using class boundaries by 10? To explore this question create a new list \( L₃ \) using the formula \( L₃=10L₁ \). The entries will be as follows:

```
10 30 70 80 100
60 50 40 20 10
90 30 40 50 20
```

Compare the histogram with three classes to the one from part (a). You will see that there does not seem to be an exact correspondence. To see why, look at the class width and compare it to the class width of part (a). The class width is always increased to the next integer value no matter how large the integer data values are. Consequently, the class width for the data in part (d) was increased to 31 rather than 40.

4. Histograms are not effective displays for some data. Consider the data

```
1 2 3 6 5 7
9 8 4 12 11 15
14 12 6 2 1 206
```
Use the TI-83 Plus, TI-84 Plus, or TI-Nspire to make a histogram with 2 classes. Then change to 3 classes, on up to 10 classes. Notice that all the histograms lump the first 17 data values into the first class, and the one data value, 206, in the last class. What feature of the data causes this phenomenon? Recall that

$$\text{Class width} = \frac{\text{largest value} - \text{smallest value}}{\text{number of classes}}, \text{increased to the next integer}$$

How many classes would you need before you began to see the first 17 data values distributed among several classes? What would happen if you simply did not include the extreme value 206 in your histogram?

BAR GRAPHS AND PIE CHARTS (SECTION 2.2 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

These functions are not available on the TI-83 Plus and TI-84 Plus.

TI-Nspire

The TI-Nspire is capable of producing both bar charts and pie charts for categorical data. For both of these charts, data will need to be entered into two columns, one column representing the categories (where we will need to format these specially to indicate that they are categories) and the other column representing the counts or percentages for each category.

Example

At a certain college, administration was interested to know who ate a certain café on campus. In order to determine this, the staff at the café asked each student who ate there on a certain day what year they were in college. The following data was observed:

<table>
<thead>
<tr>
<th>FRESHMAN</th>
<th>SOPHOMORE</th>
<th>JUNIOR</th>
<th>SENIOR</th>
<th>GRADUATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>67</td>
<td>32</td>
<td>15</td>
<td>28</td>
<td>35</td>
</tr>
</tbody>
</table>

Use the TI-Nspire to create a bar chart of this data.

First, enter the categories into a list named cats. In order to indicate that these are categories, we will input them as text. In order to input these as text, first press `>` and then select “ and press enter. Next type freshman and press enter. Now input the remainder of the category titles in the same manner.
Next, enter the counts into a list named counts.

Next, press menu and navigate to 5:Frequency Plot and select this option. In the Frequency Plot window that is displayed, for Data List select “cats” from the drop-down menu; for Frequency List select “counts” from the drop-down menu. Then press OK.
Example

A car company would like to know what the most popular color of car is in America. The company sends a representative to a large shopping mall parking lot. The representative observes the following data:

<table>
<thead>
<tr>
<th>Color</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silver</td>
<td>47</td>
</tr>
<tr>
<td>Red</td>
<td>21</td>
</tr>
<tr>
<td>White</td>
<td>22</td>
</tr>
<tr>
<td>Green</td>
<td>7</td>
</tr>
<tr>
<td>Tan</td>
<td>44</td>
</tr>
<tr>
<td>Other</td>
<td>10</td>
</tr>
</tbody>
</table>

Use the TI-Nspire to create a pie chart of the data.

First, we will need to input the categories into a list called colors. Recall that we must format the category data using quotations (see the section on bar charts for formatting specifics).

Then, enter the counts into a list called carcount.

Next, press menu and select 3:Data. Then, from the Data menu, select 5:Frequency Plot. In the window that appears, for Data List select colors and for Frequency List select carcount. Then press OK.

With the bar plot that is created selected, press menu. Select 1:Plot Type, then choose 9:Pie Chart.
After you have graphed the pie chart, you can use the touchpad (like you would use a mouse) to hover your areas of the pie chart to see counts and percentages for each category.

LAB ACTIVITIES FOR BAR CHARTS AND PIE CHARTS

1. According to a survey of chief information officers at large companies, the technology skills most in demand are: Networking, 33%; Internet/intranet development, 21%; Applications development, 18%; Help desk/user support, 8%; Operations, 6%; Project management, 6%; Systems analysis, 5%; Other, 3%.
   (a) Make a bar graph displaying this data.
   (b) Make a pie graph displaying this data.

2. In a survey in which respondents could name more than one choice, on-line Internet users were asked where they obtained news about current events. The results are: Search engine/directory sites, 49%; Cable news site, 41%; On-line service, 40%; Broadcast news site, 40%; Local newspapers, 30%; National newspaper site, 24%; Other, 13%; National newsweekly site, 12%; Haven’t accessed news on-line, 11%.
(a) Make a horizontal bar graph displaying this information.

(b) Is this information appropriate for a circle graph display? Why or why not?

3. What percentage of its income does the average household spend on food, and how many workdays are devoted to earning the money spent on food in an average household? The American Farm Bureau Federation gave the following information, by year: In 1930, 25% of a household’s budget went to food, and it took 91 workdays to earn the money. In 1960, 17% of the budget was for food, and the money took 64 workdays to earn. In 1990, food was 12% of the budget, earned in 43 workdays. For the year 2000, it was projected that the food budget would be 11% of total income and that it would take 40 workdays to earn the money.

   (a) Make bar charts for both the percent of budget for food, by year, and for the workdays required.
CHAPTER 3: AVERAGES AND VARIATION

ONE-VARIABLE STATISTICS (SECTIONS 3.1 AND 3.2 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

The TI-83 Plus and TI-84 Plus graphing calculators support many of the common descriptive measures for a data set. The measured supported are

- Mean $\bar{x}$
- Sample standard deviations $s_x$
- Population standard deviation $\sigma_x$
- Number of data values $n$
- Median
- Minimum data value
- Maximum data value
- Quartile 1
- Quartile 3
- $\sum x$
- $\sum x^2$

Although the mode is not provided directly, the menu choice SortA sorts the data in ascending order. By scanning the sorted list, you can find the mode fairly quickly.

Example

At Lazy River College, 15 students were selected at random from a group registering on the last day of registration. The times (in hours) necessary for these students to complete registration follow:

<table>
<thead>
<tr>
<th>1.7</th>
<th>2.1</th>
<th>0.8</th>
<th>3.5</th>
<th>1.5</th>
<th>2.6</th>
<th>2.1</th>
<th>2.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>2.1</td>
<td>1.3</td>
<td>0.5</td>
<td>2.1</td>
<td>1.5</td>
<td>1.9</td>
<td></td>
</tr>
</tbody>
</table>

Use the TI-83 Plus or TI-84 Plus to find the mean, sample standard deviation, and median. Use the SortA command to sort the data and scan for the mode, if it exists.

First enter the data into list L₁.
Press STAT again, and highlight CALC with item 1:1-Var Stats.

```
EDIT  CALC  TESTS
1:1-Var Stats
2:2-Var Stats
3:Med-Med
4:LinReg(ax+b)
5:QuadReg
6:CubicReg
7:QuartReg
```

Press ENTER. The command 1-Var Stats will appear on the screen. Then tell the calculator to use the data in list L₁ by pressing 2nd [L₁].
Press ENTER. The next two screens contain the statistics for the data in List $L_1$.

Note the arrow $\downarrow$ on the last line. This is an indication that you may use the down-arrow key to display more results.

```
1-Var Stats $L_1$

$\bar{x} = 1.973333333$
$\sum x = 29.6$
$\sum x^2 = 67.68$
$S_x = 0.8136923486$
$\sigma_x = 0.7861014919$
$n = 15$
```

Scroll down. The second full screen shows the rest of the statistics for the data in list $L_1$. 
To sort the data, first press STAT to return to the edit menu. Then highlight EDIT and item 2:SortA().

Press ENTER. The command SortA appears on the screen. Type in L₁.
When you press ENTER, the comment **Done** appears on the screen.

Press STAT and return to the data entry screen. Notice that the data in list L₁ is now sorted in ascending order.
A scan of the data shows that the mode is 2.1, since this data value occurs more than any other.

**TI-Nspire**

The TI-Nspire supports many of the common descriptive measures for a data set. The measures supported are:

- Mean $\bar{x}$
- Sample standard deviations $s_x$
- Population standard deviation $\sigma_x$
- Number of data values $n$
- Median
- Minimum data value
- Maximum data value
- Quartile 1
- Quartile 3
- $\Sigma x$
- $\Sigma x^2$

Although the mode is not provided, directly, the menu choice **Sort** sorts the data in ascending order. By scanning the sorted list, you can find the mode fairly quickly.

**Example**

At Lazy River College, 15 students were selected at random from a group registering on the last day of registration. The times (in hours) necessary for these students to complete registration follow:

<table>
<thead>
<tr>
<th>1.7</th>
<th>2.1</th>
<th>0.8</th>
<th>3.5</th>
<th>1.5</th>
<th>2.6</th>
<th>2.1</th>
<th>2.08</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>2.1</td>
<td>1.3</td>
<td>0.5</td>
<td>2.1</td>
<td>1.5</td>
<td>1.9</td>
<td></td>
</tr>
</tbody>
</table>

Use the TI-Nspire to find the mean, sample standard deviation, and median. Use the **Sort** command to sort the data and scan for the mode, if it exists.
First enter the data into list a list named times.


On the window that displays, press OK.
On the next window, for \textbf{X1 List}, select times from the drop-down menu. Then press \textbf{OK}.

The statistics appear within the second and third column of the spreadsheet.
To sort the data, highlight the column to be sorted. Then press menu and choose **1:Actions**. Next select **6:Sort**.

On the window that displays, press **OK**.

A scan of the data shows that the model is 2.1, since this data value occurs more than any other.
LAB ACTIVITIES FOR ONE-VARIABLE STATISTICS

TI-83 Plus and TI-84 Plus

1. A random sample of 50 professional football players produced the following data. (See Weights of Pro Football Players (data file Sv02.zip))

   The following data represents weights in pounds of 50 randomly selected pro football linebackers.

   225 230 235 238 232 227 244 222
   250 226 242 253 251 225 229 247
   239 223 233 222 243 237 230 240
   255 230 245 240 235 252 245 231
   235 234 248 242 238 240 240 240
   235 244 247 250 236 246 243 255
   241 245

   Enter this data in L1. Use 1-Var Stats to find the mean, median, and standard deviation for the weights. Use SortA to sort the data and scan for a mode if one exists.

2. Explore some of the other data sets found in the Appendix of this Guide, such as

   Disney Stock Volume (data file Sv01.zip)
   Heights of Pro Basketball Players (data file Sv03.zip)
   Miles per Gallon Gasoline Consumption (data file Sv04.zip)
   Fasting Glucose Blood Tests (data file Sv05.zip)
   Number of Children in Rural Families (data file Sv06.zip)

   Use the TI-83 Plus or TI-84 Plus to find the mean, median, standard deviation, and mode of the data.

3. In this problem we will explore the effects of changing data values by multiplying each data value by a constant, or by adding the same constant to each data value.

   (a) Consider the data

   1  8  3  5  7
   2 10  9  4  6
   3  5  2  9  1

   Enter the data into list L1 and use the TI-83 Plus or TI-84 Plus to find the mode (if it exists), mean, sample standard deviation, range, and median. Make a note of these values, since you will compare them to those obtained in parts (b) and (c).

   (b) Now multiply each data value of part (a) by 10 to obtain the data

   10 80 30 50 70
   20 100 90 40 60
   30 50 20 90 10

   Remember, you can enter these data in L2 by using the command L2=10L1.

   Again, use the TI-83 Plus or TI-84 Plus to find the mode (if it exists), mean, sample standard deviation, range, and median. Compare these results to the corresponding ones from part (a). Which values changed?
Did those that changed change by a factor of 10? Did the range or standard deviation change? Referring to the formulas for these measures (see Section 3.2 of *Understanding Basic Statistics*), can you explain why the values behaved the way they did? Will these results generalize to the situation of multiplying each data entry by 12 instead of by 10? What about multiplying each by 0.5? Predict the corresponding values that would occur if we multiplied the data set of part (a) by 1000.

(c) Now suppose we add 30 to each data value of part (a):

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>38</td>
<td>33</td>
<td>35</td>
<td>37</td>
</tr>
<tr>
<td>32</td>
<td>40</td>
<td>39</td>
<td>34</td>
<td>36</td>
</tr>
<tr>
<td>33</td>
<td>35</td>
<td>32</td>
<td>39</td>
<td>31</td>
</tr>
</tbody>
</table>

To enter this data, create a new list $L_3$ by using the command $L_3=L_1+30$.

Again, use the TI-83 Plus or TI-84 Plus to find the mode (if it exists), mean, sample standard deviation, range, and median. Compare these results with the corresponding ones of part (a). Which values changed? Of those that are different, did each change by being 30 more than the corresponding value of part (a)? Again, look at the formulas for range and standard deviation. Can you predict the observed behavior from the formulas? Can you generalize these results? What if we added 50 to each data value of part (a)? Predict the values for the mode, mean, sample standard deviation, range, and median.

**TI-Nspire**

1. A random sample of 50 professional football players produced the following data. (See Weights of Pro Football Players (data file Sv02.zip))
   The following data represents weights in pounds of 50 randomly selected pro football linebackers.
   
   | 225 | 230 | 235 | 238 | 232 | 227 | 244 | 222 |
   | 250 | 226 | 242 | 253 | 251 | 235 | 229 | 247 |
   | 239 | 223 | 233 | 222 | 243 | 237 | 230 | 240 |
   | 255 | 230 | 245 | 240 | 235 | 252 | 245 | 231 |
   | 235 | 234 | 248 | 242 | 238 | 240 | 240 | 240 |
   | 235 | 244 | 247 | 250 | 236 | 246 | 243 | 255 |
   | 241 | 245 |

Enter this data in $L_1$. Use **1-Variable Statistics** to find the mean, median, and standard deviation for the weights. Use **Sort** to sort the data and scan for a mode if one exists.

2. Explore some of the other data sets found in the Appendix of this Guide, such as
   - Disney Stock Volume (data file Sv01.zip)
   - Heights of Pro Basketball Players (data file Sv03.zip)
   - Miles per Gallon Gasoline Consumption (data file Sv04.zip)
   - Fasting Glucose Blood Tests (data file Sv05.zip)
   - Number of Children in Rural Families (data file Sv06.zip)

Use the TI-Nspire to find the mean, median, standard deviation, and mode of the data.
3. In this problem we will explore the effects of changing data values by multiplying each data value by a constant, or by adding the same constant to each data value.

(a) Consider the data

<table>
<thead>
<tr>
<th>1</th>
<th>8</th>
<th>3</th>
<th>5</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10</td>
<td>9</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>2</td>
<td>9</td>
<td>1</td>
</tr>
</tbody>
</table>

Enter the data into list called list1 and use the TI-Nspire to find the mode (if it exists), mean, sample standard deviation, range, and median. Make a note of these values, since you will compare them to those obtained in parts (b) and (c).

(b) Now multiply each data value of part (a) by 10 to obtain the data

<table>
<thead>
<tr>
<th>10</th>
<th>80</th>
<th>30</th>
<th>50</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>100</td>
<td>90</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>30</td>
<td>50</td>
<td>20</td>
<td>90</td>
<td>10</td>
</tr>
</tbody>
</table>

Remember, you can enter these data in list2 by using the command list2=10list1.

Again, use the TI-Nspire to find the mode (if it exists), mean, sample standard deviation, range, and median. Compare these results to the corresponding ones from part (a). Which values changed? Did those that changed change by a factor of 10? Did the range or standard deviation change? Referring to the formulas for these measures (see Section 3.2 of Understanding Basic Statistics), can you explain why the values behaved the way they did? Will these results generalize to the situation of multiplying each data entry by 12 instead of by 10? What about multiplying each by 0.5? Predict the corresponding values that would occur if we multiplied the data set of part (a) by 1000.

(c) Now suppose we add 30 to each data value of part (a):

<table>
<thead>
<tr>
<th>31</th>
<th>38</th>
<th>33</th>
<th>35</th>
<th>37</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>40</td>
<td>39</td>
<td>34</td>
<td>36</td>
</tr>
<tr>
<td>33</td>
<td>35</td>
<td>32</td>
<td>39</td>
<td>31</td>
</tr>
</tbody>
</table>

To enter this data, create a new list list3 by using the command list3=list1+30.

Again, use the TI-Nspire to find the mode (if it exists), mean, sample standard deviation, range, and median. Compare these results with the corresponding ones of part (a). Which values changed? Of those that are different, did each change by being 30 more than the corresponding value of part (a)? Again, look at the formulas for range and standard deviation. Can you predict the observed behavior from the formulas? Can you generalize these results? What if we added 50 to each data value of part (a)? Predict the values for the mode, mean, sample standard deviation, range, and median.
GROUPED DATA (SECTION 3.2 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

The TI-83 Plus and TI-84 Plus support grouped data, because it allows you to specify the frequency with which each data value occurs in a separate list. The frequencies must be whole numbers. This means that numbers with decimal parts are not permitted as frequencies.

Example

A random sample of 44 automobiles registered in Dallas, Texas, shows the ages of the cars to be

<table>
<thead>
<tr>
<th>Age (in years)</th>
<th>Midpt.</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-3</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>4-6</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>7-9</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>10-12</td>
<td>11</td>
<td>3</td>
</tr>
</tbody>
</table>

Estimate the mean and standard deviation of the ages of the cars.

Put the midpoints in the list L₁ and the corresponding frequencies in the list L₂.

Press STAT and, with CALC highlighted, select item 1:1-Var Stats.
Designate the data lists, with the first list holding the data and the second list holding the frequencies.

When you press ENTER, the statistics for the grouped data will appear. Note that $n = 44$ is the correct value for the number of data values.
From the last screen, we see that the estimate for the mean is 6.16 and that the estimate for the sample standard deviation is 2.68.

**TI-Nspire**

The TI-Nspire supports grouped data, because it allows you to specify the frequency with which each data value occurs in a separate list. The frequencies must be whole numbers. This means that numbers with decimal parts are not permitted as frequencies.

**Example**

A random sample of 44 automobiles registered in Dallas, Texas, shows the ages of the cars to be

<table>
<thead>
<tr>
<th>Age (in years)</th>
<th>Midpt.</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-3</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>4-6</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>7-9</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>10-12</td>
<td>11</td>
<td>3</td>
</tr>
</tbody>
</table>

Estimate the mean and standard deviation of the ages of the cars.

Put the midpoints into a list named midpt and the corresponding frequencies into a list named freq.

Press OK.
In **X1 List**, select midpt from the drop-down menu and for **Frequency List**, select freq from the drop-down menu. Then press **OK**.

![Image of the One-Variable Statistics menu]

The statistics will appear in columns C and D.

![Image of the spreadsheet with statistics]

From the last screen, we see that the estimate for the mean is 6.16 and that the estimate for the sample standard deviation is 2.68.
LAB ACTIVITIES FOR GROUPED DATA

TI-83 Plus, TI-84 Plus and TI-Nspire

1. A teacher rating form uses the scale values -2, -1, 0, 1, 2 with -2 meaning “strongly disagree” and 2 meaning “strongly agree.” The statement “The professor grades fairly” was answered as follows for all courses in the sociology department.

<table>
<thead>
<tr>
<th>Score</th>
<th>Number of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>30</td>
</tr>
<tr>
<td>-1</td>
<td>125</td>
</tr>
<tr>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>2</td>
<td>180</td>
</tr>
<tr>
<td>1</td>
<td>72</td>
</tr>
</tbody>
</table>

Use the TI-83 Plus, TI-84 Plus or TI-Nspire to calculate the mean response and the standard deviation for this statement.

2. A stock market analyst looked at 80 key stocks and recorded the total gains and losses (in points) per share for each stock over a one-month period. The results follow with negative point changes indicating losses and positive point changes indicating gains.

<table>
<thead>
<tr>
<th>Point Change</th>
<th>Frequency</th>
<th>Point Change</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>-40 to -31</td>
<td>3</td>
<td>0 to 9</td>
<td>20</td>
</tr>
<tr>
<td>-30 to -21</td>
<td>2</td>
<td>10 to 19</td>
<td>14</td>
</tr>
<tr>
<td>-20 to -11</td>
<td>10</td>
<td>20 to 29</td>
<td>7</td>
</tr>
<tr>
<td>-10 to -1</td>
<td>19</td>
<td>30 to 39</td>
<td>5</td>
</tr>
</tbody>
</table>

First find the class midpoints. Then use the TI-83 Plus, TI-84 Plus or TI-Nspire to estimate the mean point change and standard deviation for this population of stocks.

BOX-AND-WHISKER PLOTS (SECTION 3.3 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

The box-and-whisker plot is based on the five-number summary values found under 1-Var Stats:

- Lowest value
- Quartile 1, or Q₁
- Median
- Quartile 3, or Q₃
- Highest value
Example

Let’s make a box-and-whisker plot using the data about the time it takes to register for Lazy River College students who waited till the last day to register. The times (in hours) are:

1.7  2.1  0.8  3.5  1.5  2.6  2.1  2.08
3.1  2.1  1.3  0.5  2.1  1.5  1.9

In a previous example we found the 1-Var Stats for this data. First enter the data into list L₁.

Press WINDOW to set the graphing window. Use Xmin = 0.5, since that is the smallest value. Use Xmax = 3.5, since that is the largest value. Use Xscl = 1. Use Ymin = -5 and Ymax = 5 to position the axis in the window.
Press 2nd [STAT PLOT] and highlight 1:Plot. Press ENTER.

Highlight On, and select the box plot. The Xlist should be L₁ and Freq should be 1.
Press GRAPH.

![Box and Whisker Plot]

Press TRACE and use the arrow keys to display the five-number summary.

![Five-Number Summary]

You now have many descriptive tools available: histograms, box-and-whisker plots, averages, and measures of variation such as the standard deviation. When you use all of these tools, you get a lot of information about the data.

**TI-Nspire**

The box-and-whisker plot is based on the five-number summary values found under **1-Var Stats**:

- Lowest value
Quartile 1, or $Q_1$

Median

Quartile 3, or $Q_3$

Highest value

Example

Let’s make a box-and-whisker plot using the data about the time it takes to register for Lazy River College students who waited till the last day to register. The times (in hours) are:

$$
\begin{align*}
1.7 & \quad 2.1 & \quad 0.8 & \quad 3.5 & \quad 1.5 & \quad 2.6 & \quad 2.1 & \quad 2.08 \\
3.1 & \quad 2.1 & \quad 1.3 & \quad 0.5 & \quad 2.1 & \quad 1.5 & \quad 1.9 
\end{align*}
$$

In the previous example we found the **1 Variable Statistics** for this data. First enter the data into a list named \texttt{l1}.

Press menu and choose **3:Data**. Then choose **6:Quick Graph**.
While the dotplot is selected, press menu. Choose 1:Plot Type then choose 2:Box Plot.

You can view the five-number summary by using the touchpad to navigate over points of the boxplot.

You now have many descriptive tools available: histograms, box-and-whisker plots, averages, and measures of variation such as the standard deviation. When you use all of these tools, you get a lot of information about the data.
LAB ACTIVITIES FOR BOX-AND-WHISKER PLOTS

TI-83 Plus and TI-84 Plus

1. One of the data sets included in the Appendix gives the miles per gallon gasoline consumption for a random sample of 55 makes and models of passenger cars. (See Miles per Gallon Gasoline Consumption (data file S04.zip))

The following data represents miles per gallon gasoline consumption (highway) for a random sample of 55 makes and models of passenger cars.

Source: Environmental Protection Agency

| 30 | 27 | 22 | 25 | 24 | 25 | 24 | 15 |
| 33 | 35 | 33 | 52 | 49 | 10 | 27 | 18 |
| 20 | 23 | 24 | 25 | 30 | 24 | 24 | 24 |
| 18 | 20 | 25 | 27 | 24 | 32 | 29 | 27 |
| 24 | 27 | 26 | 25 | 24 | 28 | 33 | 30 |
| 13 | 13 | 21 | 28 | 37 | 35 | 32 | 33 |
| 29 | 31 | 28 | 28 | 25 | 29 | 31 |

Enter these data into list L1, and then make a histogram and a box-and-whisker plot, and compute the mean, median, and sample standard deviation. Based on the information you obtain, response to the following questions:

(a) Is the distribution skewed or symmetric? How is this shown in both the histogram and the box-and-whisker plot?
(b) Look at the box-and-whisker plot. Are the data more spread out above the median or below the median?
(c) Look at the histogram and estimate the location of the mean on the horizontal axis. Are these data more spread out above the mean or below the mean?
(d) Do there seem to be any data values that are unusually high or unusually low? If so, how do these show up on a histogram or on a box-and-whisker plot?
(e) Pretend that you are writing a brief article for a newspaper. Describe the information about the data in non-technical terms. Be sure to make some comments about the “average” of the data values and some comments about the spread of the data.

2. (a) Consider the test scores of 30 students in a political science class.

| 85 | 73 | 43 | 86 | 73 | 59 | 73 | 84 | 62 | 100 |
| 75 | 87 | 70 | 84 | 97 | 62 | 76 | 89 | 90 | 83 |
| 70 | 65 | 77 | 90 | 84 | 80 | 68 | 91 | 67 | 79 |

For this population of test scores, find the mode, median, mean, range, variance, standard deviation, CV, the five-number summary, and make a box-and-whisker plot. Be sure to record all of these values so you can compare them to the results of part (b).

(b) Suppose Greg was in the political science class of part (a). Suppose he missed a number of classes because of illness, but took the exam anyway and made a score of 30 instead of 85 as listed as the first entry of the data in part (a). Again, use the calculator to find the mode, median, mean, range, variance, standard deviation, CV, the five-number summary, and make a box-and-whisker plot using the new data set. Compare these results to the corresponding results of part (a). Which average was most
affected: mode, median, or mean? What about the range, standard deviation, and coefficient of variation? How do the box-and-whisker plots compare?
(c) Write a brief essay in which you use the results of parts (a) and (b) to predict how an extreme data value affects a data distribution. What do you predict for the results if Greg’s test score had been 80 instead of 30 or 85?

**TI-Nspire**

1. One of the data sets included in the Appendix gives the miles per gallon gasoline consumption for a random sample of 55 makes and models of passenger cars. (See *Miles per Gallon Gasoline Consumption* (data file Sv04.zip))

The following data represents miles per gallon gasoline consumption (highway) for a random sample of 55 makes and models of passenger cars.

Source: Environmental Protection Agency

```
30 27 22 25 24 25 24 15
33 35 33 52 49 10 27 18
20 23 24 25 30 24 24 24
18 20 25 27 24 32 29 27
24 27 26 25 24 28 33 30
13 13 21 28 37 35 32 33
29 31 28 28 25 29 31
```

Enter these data into a list called l1, and then make a histogram and a box-and-whisker plot, and compute the mean, median, and sample standard deviation. Based on the information you obtain, response to the following questions:

(a) Is the distribution skewed or symmetric? How is this shown in both the histogram and the box-and-whisker plot?
(b) Look at the box-and-whisker plot. Are the data more spread out above the median or below the median?
(c) Look at the histogram and estimate the location of the mean on the horizontal axis. Are these data more spread out above the mean or below the mean?
(d) Do there seem to be any data values that are unusually high or unusually low? If so, how do these show up on a histogram or on a box-and-whisker plot?
(e) Pretend that you are writing a brief article for a newspaper. Describe the information about the data in non-technical terms. Be sure to make some comments about the “average” of the data values and some comments about the spread of the data.

2. (a) Consider the test scores of 30 students in a political science class.

```
85 73 43 86 73 59 73 84 62 100
75 87 70 84 97 62 76 89 90 83
70 65 77 90 84 80 68 91 67 79
```

For this population of test scores, find the mode, median, mean, range, variance, standard deviation, CV, the five-number summary, and make a box-and-whisker plot. Be sure to record all of these values so you can compare them to the results of part (b).
(b) Suppose Greg was in the political science class of part (a). Suppose he missed a number of classes because of illness, but took the exam anyway and made a score of 30 instead of 85 as listed as the first entry of the data in part (a). Again, use the calculator to find the mode, median, mean, range, variance,
standard deviation, CV, the five-number summary, and make a box-and-whisker plot using the new data set. Compare these results to the corresponding results of part (a). Which average was most affected: mode, median, or mean? What about the range, standard deviation, and coefficient of variation? How do the box-and-whisker plots compare?

(c) Write a brief essay in which you use the results of parts (a) and (b) to predict how an extreme data value affects a data distribution. What do you predict for the results if Greg’s test score had been 80 instead of 30 or 85?
CHAPTER 4: CORRELATION AND REGRESSION

LINEAR REGRESSION (SECTIONS 4.1 and 4.2 OF UNDERSTANDING BASIC STATISTICS)

TI-83 and TI-84 Plus

Important Note: Before beginning this chapter, press 2nd [CATALOG] (above 0) and scroll down to the entry DiagnosticOn. Press ENTER twice. After doing this, the regression correlation coefficient \( r \) will appear as output with the linear regression line.

The TI-83 Plus and TI-84 Plus graphing calculators have automatic regression functions built-in as well as summary statistics for two variables. To locate the regression menu, press STAT, and then select CALC. The choice \( 8:\text{LinReg}(a + bx) \) performs linear regression on the variables in \( L_1 \) and \( L_2 \). If the data are in other lists, then specify those lists after the \( \text{LinReg}(a + bx) \) command.

Let’s look at a specific example to see how to do linear regression on the TI-83 Plus or TI-84 Plus.

Example

Merchandise loss due to shoplifting, damage, and other causes is called shrinkage. Shrinkage is a major concern to retailers. The managers of H.R. Merchandise believe there is a relationship between shrinkage and number of clerks on duty. To explore this relationship, a random sample of 6 weeks was selected. During each week the staffing level of sales clerks was kept constant and the dollar value of the shrinkage was recorded.

<table>
<thead>
<tr>
<th>Number of Sales Clerks</th>
<th>12</th>
<th>11</th>
<th>15</th>
<th>9</th>
<th>13</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shrinkage:</td>
<td>15</td>
<td>20</td>
<td>9</td>
<td>25</td>
<td>12</td>
<td>31</td>
</tr>
</tbody>
</table>
(a) Find the equation of the least-squares line, with number of sales clerks as the explanatory variable. First put the numbers of sales clerks in L1 and the corresponding amount of shrinkage in L2.

Use the CALC menu and select 8:LinReg(a+bx).

Designate the lists containing the data. Then press ENTER.
Note that the linear regression equation is \( y = 54.48x - 3.16 \).

We also have the value of the Pearson correlation coefficient \( r \) and \( r^2 \).

(b) Make a scatter diagram and show the least-squares line on the plot.

To graph the least squares line, press \( Y= \). Clear all existing equations from the display. To enter the least-squares equation directly, press VARS. Select item 5:Statistics.

Press ENTER and then select EQ and then 1:RegEQ.
You will see the least-squares regression equation appear automatically in the equations list.

We want to graph the least-squares line with the scatter diagram. Press 2nd [STAT PLOT]. Select 1:Plot1. Then select On, scatter diagram picture, and enter L1 for Xlist and L2 for Ylist.

The last step before we graph is to set the graphing window. Press WINDOW. Set Xmin to the lowest x value, Xmax to the highest x value, Ymin at -5 and Ymax at above the highest y value.
Now press GRAPH.

(c) Predict the shrinkage when 10 clerks are on duty.

Press 2nd [CALC]. This is the Calculate function for graphs and is different from the CALC you highlight from the STAT menu. Highlight 1:Value.

Press ENTER. The graph will appear. Enter 10 next to X=
Press ENTER. The predicted value for y appears.

To display the summary statistics for x and y, press STAT and highlight CALC. Select 2:2-Var Stats. Press ENTER.

Enter the lists containing the data and press ENTER.
Scroll to the statistics of interest.

The TI-Nspire graphing calculators have automatic regression functions built-in as well as summary statistics for two variables. To locate the regression menu, enter the Spreadsheet and List page and press menu, and then select **1:Stat Calculations**. The choice **4:Linear Regression (a + bx)** performs linear regression on the variables in two lists.

Let’s look at a specific example to see how to do linear regression on the TI-Nspire.
Example

Merchandise loss due to shoplifting, damage, and other causes is called shrinkage. Shrinkage is a major concern to retailers. The managers of H.R. Merchandise believe there is a relationship between shrinkage and number of clerks on duty. To explore this relationship, a random sample of 6 weeks was selected. During each week the staffing level of sales clerks was kept constant and the dollar value of the shrinkage was recorded.

<table>
<thead>
<tr>
<th>Number of Sales Clerks</th>
<th>12</th>
<th>11</th>
<th>15</th>
<th>9</th>
<th>13</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shrinkage</td>
<td>15</td>
<td>20</td>
<td>9</td>
<td>25</td>
<td>12</td>
<td>31</td>
</tr>
</tbody>
</table>

(a) Find the equation of the least-squares line, with number of sales clerks as the explanatory variable. First put the numbers of sales clerks in a list called clerks and the corresponding amount of shrinkage in a list called shrink.

Use the Stat Calculations menu and select 4:Linear Regression (a+bx).
Designate the lists containing the data. Note where the regression equation is being saved to.

Then press OK.

Note that the linear regression equation is \( y = 54.48x - 3.16 \).

We also have the value of the Pearson correlation coefficient \( r \) and \( r^2 \).

(b) Make a scatter diagram and show the least-squares line on the plot.

Start by creating a scatter diagram. Select both columns containing the \( x \) and \( y \) data. Press menu and choose 3:Data then choose 6:Quick Plot.

Now, we will plot the regression line on the graph. Press menu and choose 4:Analyze then choose 6:Regression. Finally, choose 2:Show Linear (a + bx)
(c) Predict the shrinkage when 10 clerks are on duty.

To view the predicted values for the regression line, select the spreadsheet then press menu. Choose 5:Table and then choose 1:Switch to Table.
Select the equation where the regression line is stored and press enter.

The predicted values for the regression line are shown. You can modify these by pressing menu and choosing 5:Table then choosing 5:Edit Table Settings.

To display the summary statistics for x and y, return to the spreadsheet view by pressing menu and selecting 5:Table then choosing 1:Switch to Lists and Spreadsheet. Then press menu and choose 4:Statistics. Select 2:Two-Variable Statistics.
Enter the lists containing the data and press OK.
LAB ACTIVITIES FOR LINEAR REGRESSION

For each of the following data sets, do the following:

(a) Enter the data, putting x values into one list and y values in a second list.
(b) Find the equation of the least-squares line.
(c) Find the value of the correlation coefficient r.
(d) Draw a scatter diagram and show the least-squares line on the scatter diagram.
(e) Use 2 variable statistics to find all the sums necessary for computing the standard error of estimate $S_e$. 

Scroll to the statistics of interest.
1. **Cricket Chirps Versus Temperature** (data files Slr02.zip)
   In the following data pairs \((X, Y)\), \(X = \) Chirps/s for the striped ground cricket, and \(Y = \) Temperature in degrees Fahrenheit.

   Source: *The Song of Insects* by Dr. G.W. Pierce, Harvard College Press

   \[
   \begin{array}{cccc}
   (20.0, 88.6) & (16.0, 71.6) & (19.8, 93.3) & (18.4, 84.3) \\
   (15.5, 75.2) & (14.7, 69.7) & (17.1, 82.0) & (15.4, 69.4) \\
   (15.0, 79.6) & (17.2, 82.6) & (16.0, 80.6) & (17.0, 83.5) \\
   \end{array}
   \]

2. **List Price Versus Best Price for a New GMC Pickup Truck** (data files Slr01.zip)
   In the following data pairs \((X, Y)\), \(X = \) List price (in $1000) for a GMC Pickup Truck, and \(Y = \) Best price (in $1000) for a GMC Pickup Truck.

   Source: *Consumer's Digest*, February 1994

   \[
   \begin{array}{cccc}
   (12.400, 11.200) & (14.300, 12.500) & (14.500, 12.700) \\
   (14.900, 13.100) & (16.100, 14.100) & (16.900, 14.800) \\
   (16.500, 14.400) & (15.400, 13.400) & (17.000, 14.900) \\
   (17.900, 15.600) & (18.800, 16.400) & (20.300, 17.700) \\
   (22.400, 19.600) & (19.400, 16.900) & (15.500, 14.000) \\
   (16.700, 14.600) & (17.300, 15.100) & (18.400, 16.100) \\
   (19.200, 16.800) & (17.400, 15.200) & (19.500, 17.000) \\
   (19.700, 17.200) & (21.200, 18.600) \\
   \end{array}
   \]

3. **Diameter of Sand Granules Versus Slope on a Natural Occurring Ocean Beach** (data files Slr03.zip)
   In the following data pairs \((X, Y)\), \(X = \) Median diameter (mm) of granules of sand, and \(Y = \) Gradient of beach slope in degrees. The data is for naturally occurring ocean beaches.

   Source: *Physical Geography* by A.M. King, Oxford Press, England

   \[
   \begin{array}{cccc}
   (0.170, 0.630) & (0.190, 0.700) & (0.220, 0.820) \\
   (0.235, 0.880) & (0.235, 1.150) & (0.300, 1.500) \\
   (0.350, 4.400) & (0.420, 7.300) & (0.850, 11.300) \\
   \end{array}
   \]

4. **National Unemployment Rate Male Versus Female** (data files Slr04.zip)
   In the following data pairs \((X, Y)\), \(X = \) National unemployment rate for adult males, and \(Y = \) National unemployment rate for adult females

   Source: *Statistical Abstract of the United States*

   \[
   \begin{array}{cccc}
   (2.9, 4.0) & (6.7, 7.4) & (4.9, 5.0) \\
   (7.9, 7.2) & (9.8, 7.9) & (6.9, 6.1) \\
   (6.1, 6.0) & (6.2, 5.8) & (6.0, 5.2) \\
   (5.1, 4.2) & (4.7, 4.0) & (4.4, 4.4) \\
   (5.8, 5.2) \\
   \end{array}
   \]
CHAPTER 5: ELEMENTARY PROBABILITY THEORY

TI-83 Plus and TI-84 Plus

There are no specific TI-83 Plus or TI-84 Plus activities for the basic rules of probability. However, notice that the TI-83 Plus and TI-84 Plus have menu items for factorial notation, combinations $C_{n,r}$, and permutations $P_{n,r}$. To find these functions, press MATH and then highlight PRB.

To compute $C_{12,6}$, first clear the screen and type 12. Then use the MATH key to select PRB and item 3:nCr. Press ENTER. Type the number 6 and then press ENTER.

The menu option nPr works in a similar fashion and gives you the value of $P_{n,r}$.
TI-Nspire

There are no specific TI-Nspire activities for the basic rules of probability. However, notice that the TI-Nspire have menu items for factorial notation, combinations $C_{n,r}$, and permutations $P_{n,r}$. To find these functions, enter the Calculator page. Then press menu and select 5:Probability.

![TI-Nspire menu](image)

To compute $C_{12,6}$, select 3:Combinations. Then type 12,6 and press enter.

![TI-Nspire calculation](image)

The menu option Permutations works in a similar fashion and gives you the value of $P_{n,r}$.
CHAPTER 6: THE BINOMIAL PROBABILITY DISTRIBUTION AND RELATED TOPICS

DISCRETE PROBABILITY DISTRIBUTIONS (SECTION 6.1 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

The TI-83 Plus and TI-84 Plus will compute the mean and standard deviation of a discrete probability distribution. Just enter the values of the random variable \( x \) in list L₁ and the corresponding probabilities in list L₂. Use the techniques of grouped data and one-variable statistics to compute the mean and standard deviation.

Example

How long do we spend on hold when we call a store? One study produced the following probability distribution, with the times recorded to the nearest minute.

<table>
<thead>
<tr>
<th>Time on hold, ( x ):</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(x) ):</td>
<td>0.15</td>
<td>0.25</td>
<td>0.40</td>
<td>0.10</td>
<td>0.08</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Find the expected value of the time on hold and the standard deviation of the probability distribution.

Enter the times in L₁ and the probabilities in list L₂. Press the STAT key, choose CALC, and use 1-Var Stats with lists L₁ and L₂. The results are displayed.

![1-Var Stats](image)

The expected value is \( x = 1.77 \) minutes with standard deviation \( \sigma = 1.182 \).
The TI-Nspire will compute the mean and standard deviation of a discrete probability distribution. Just enter the values of the random variable $x$ a list and the corresponding probabilities into a second list. Use the techniques of grouped data and one-variable statistics to compute the mean and standard deviation.

**Example**

How long do we spend on hold when we call a store? One study produced the following probability distribution, with the times recorded to the nearest minute.

<table>
<thead>
<tr>
<th>Time on hold, $x$:</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(x)$:</td>
<td>0.15</td>
<td>0.25</td>
<td>0.40</td>
<td>0.10</td>
<td>0.08</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Find the expected value of the time on hold and the standard deviation of the probability distribution.

Enter the times in a list called $xval$ and the probabilities in list called $probs$. Press the menu key, choose 4:Statistics, then choose 1:Stat Calculations, then choose 1:One-Variable Statistics. Press OK. Then for X1 List choose $xval$ from the drop-down menu and for Frequency List choose $probs$ from the drop-down list. Then choose OK. The results are displayed.

The expected value is $x=1.77$ minutes with standard deviation $\sigma = 1.182$. 

![TI-Nspire Calculations](image)
LAB ACTIVITIES FOR DISCRETE PROBABILITY DISTRIBUTIONS

TI-83 Plus, TI-84 Plus and TI-Nspire

1. The probability distribution for scores on a mechanical aptitude test is as follows:

<table>
<thead>
<tr>
<th>Score, x:</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(x):</td>
<td>0.130</td>
<td>0.200</td>
<td>0.300</td>
<td>0.170</td>
<td>0.120</td>
<td>0.080</td>
</tr>
</tbody>
</table>

Use the TI-83 Plus, TI-84 Plus, TI-Nspire to find the expected value and standard deviation of the probability distribution.

2. Hold Insurance has calculated the following probabilities for claims on a new $20,000 vehicle for one year if the vehicle is driven by a single male under 25.

<table>
<thead>
<tr>
<th>$ Claims, x:</th>
<th>0</th>
<th>1000</th>
<th>5000</th>
<th>100,000</th>
<th>200,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(x):</td>
<td>0.63</td>
<td>0.24</td>
<td>0.10</td>
<td>0.02</td>
<td>0.01</td>
</tr>
</tbody>
</table>

What is the expected value of the claim in one year? What is the standard deviation of the claim distribution? What should the annual premium be to include $400 in overhead and profit as well as the expected claim?

BINOMIAL PROBABILITIES (SECTION 6.2 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

For a binomial distribution with $n$ trials, $r$ successes, and the probability of success on a single trial $p$, the formula for the probability of $r$ successes out of $n$ trials is

$$P(r) = \binom{n}{r} p^r (1-p)^{n-r}$$

To compute probabilities for specific values of $n$, $r$, and $p$, we can use the TI-83 Plus or TI-84 Plus graphing calculator.

Example

Consider a binomial experiment with 10 trials and probability of success on a single trial $p = 0.72$. Compute the probability of 7 successes out of $n$ trials.

On the TI-83 Plus or TI-84 Plus, the combinations function $nCr$ is found in the MATH menu under PRB.
To use the function, we first type the value of $n$, then $n\text{Cr}$, and finally the value of $r$. Here, $n = 10$, $r = 7$, $p = 0.72$ and $(1 - p) = 0.28$. By the formula, we raise 0.72 to the power 7 and 0.28 to the power 3. The probability of 7 successes is 0.264.

**TI-Nspire**

For a binomial distribution with $n$ trials, $r$ successes, and the probability of success on a single trial $p$, the formula for the probability of $r$ successes out of $n$ trials is

$$P(r) = C_{n,r} p^r (1 - p)^{n-r}$$

To compute probabilities for specific values of $n$, $r$, and $p$, we can use the TI-Nspire.
Example

Consider a binomial experiment with 10 trials and probability of success on a single trial \( p = 0.72 \). Compute the probability of 7 successes out of \( n \) trials.

On the TI-Nspire, the combinations function \( nCr \) is found in the Calculator page.

\[
\binom{10}{7} \cdot (0.72)^7 \cdot (0.28)^3 = 0.26423
\]

After entering Calculator page, press menu. Then choose 5:Probability, then choose 3:Combinations. Then the value of \( n \) then a comma and then the value of \( r \). Here, \( n = 10, r = 7, p = 0.72 \) and \( (1 - p) = 0.28 \). By the formula, we raise 0.72 to the power 7 and 0.28 to the power 3. The probability of 7 successes is 0.264.

USING THE PROBABILITY DISTRIBUTIONS ON THE TI-83 PLUS AND TI-84 PLUS

The TI-83 Plus and TI-84 Plus fully support the binomial distribution and has it built in as a function. To access the menu that contains the probability functions, press 2nd [DISTR]. Then scroll to item A:binompdf and press ENTER. Note on TI-83 Plus, the item will be 0:binompdf.
Type in the number of trials, followed by the probability of success on a single trial, followed by the number of successes. Separate each entry by a comma. Press ENTER.

To find the cumulative probability, \texttt{binomcdf}(n, p, r).

\textbf{USING THE PROBABILITY DISTRIBUTIONS ON THE TI-NSPIRE}

The TI-Nspire fully supports the binomial distribution and has it built in as a function. To access the menu that contains the probability functions, from the Calculator screen press menu. Choose \texttt{5:Probability} then choose \texttt{5:Distributions}. Choose \texttt{D:Binomial Pdf}. 
For Num Trials, n, input the number of trials, in Prob Success, p input the success probability on a single trial, and in X Value input the number of successes. Press OK.
To find the cumulative probability, Binomial Cdf.

LAB ACTIVITIES FOR BINOMIAL PROBABILITIES

1. Consider a binomial distribution with $n = 8$ and $p = 0.43$. Use the formula to find the probability of $r$ successes for $r$ from 0 through 8.
2. Consider a binomial distribution with $n = 8$ and $p = 0.43$. Use the built-in binomial probability distribution function to find the probability of $r$ successes for $r$ from 0 through 8.
CHAPTER 7: NORMAL CURVES AND SAMPLING DISTRIBUTIONS

THE AREA UNDER ANY NORMAL CURVE (SECTION 7.3 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

The TI-83 Plus and TI-84 Plus give the area under any normal distribution and shades the specified area. Press 2nd [DISTR] to access the probability distributions. Select 2:normalcdf( and press ENTER.

Then type in the lower bound, upper bound, \( \mu \), and \( \sigma \) in that order, separated by commas. To find the area under the normal curve with \( \mu = 10 \) and \( \sigma = 2 \) between 4 and 12, select normalcdf( and enter the values as shown. Then press ENTER.
TI-Nspire

The TI-Nspire gives the area under any normal distribution and shades the specified area. From the Calculator page, press menu and choose 5:Probability. Then choose 5:Distributions and then select 2:Normal Cdf.

Then type in the lower bound, upper bound, $\mu$, and $\sigma$ in that order. To find the area under the normal curve with $\mu = 10$ and $\sigma = 2$ between 4 and 12, select Normal Cdf and enter the values as shown. Then press OK.
DRAWING THE NORMAL DISTRIBUTION

To draw the graph, first set the window to accommodate the graph. Press the WINDOW button and enter values as shown.

Press 2nd [DISTR] again and highlight DRAW. Select 1:ShadeNorm.
Again enter the lower limit, upper limit, $\mu$, and $\sigma$, separated by commas.

Finally press ENTER.
LAB ACTIVITIES FOR THE AREA UNDER ANY NORMAL CURVE

1. Find the area under a normal curve with mean 10 and standard deviation 2, between 7 and 9. Show the shaded region.
2. To find the area in the right tail of a normal distribution, select the value of 5Å for the upper limit of the region. Find the area that lies to the right of 12 under a normal curve with mean 10 and standard deviation 2.
3. Consider a random variable x that follows a normal distribution with mean 100 and standard deviation 15.
   (a) Shade the region corresponding to P(x < 90) and find the probability.
   (b) Shade the region corresponding to P(70 < x < 100) and find the probability.
   (c) Shade the region corresponding to P(x > 115) and find the probability.
   (d) If the random variable were larger than 145, would that be an unusual event? Explain by computing P(x > 145) and commenting on the meaning of the result.

SAMPLING DISTRIBUTIONS (SECTION 7.5 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

In this chapter, we use the TI-83 Plus or TI-84 Plus graphing calculator to do computations. For example, to compute the z score corresponding to a raw score from an x distribution, we use the formula

\[ z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \]

To evaluate z, be sure to use parentheses as necessary.

Example

If a random sample of size 40 is taken from a distribution with mean \( \mu = 10 \) and standard deviation \( \sigma = 2 \), find the z score corresponding to \( x = 9 \).

We use the formula

\[ z = \frac{9 - 10}{2/\sqrt{40}} \]

Key in the expression using parentheses: \((9 - 10) ÷ (2 ÷ \sqrt{40})\) ENTER.
The result rounds to $z = -3.16$.

**TI-Nspire**

In this chapter, we use the TI-Nspire graphing calculator to do computations. For example, to compute the $z$ score corresponding to a raw score from an $x$ distribution, we use the formula

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

To evaluate $z$, be sure to use parentheses as necessary.

**Example**

If a random sample of size 40 is taken from a distribution with mean $\mu = 10$ and standard deviation $\sigma = 2$, find the $z$ score corresponding to $x = 9$.

We use the formula

$$z = \frac{9 - 10}{\frac{2}{\sqrt{40}}}$$

From the Calculator page, key in the expression using parentheses: $(9 - 10) \div (2 \div \sqrt{40})$ enter.
The result rounds to $z = -3.16$. 
CHAPTER 8: ESTIMATION

CONFIDENCE INTERVALS FOR A POPULATION MEAN (SECTIONS 8.1 AND 8.2 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

The TI-83 Plus and TI-84 Plus fully support confidence intervals. To access the confidence interval choices, press STAT and select TESTS. The confidence interval choices are found in items 7 through B.

Example (σ is known with a Large Sample)

Suppose a random sample of 250 credit card bills showed an average balance of $1200. Also assume that the population standard deviation is $350. Find a 95% confidence interval for the population mean credit card balance.

Since σ is known and we have a large sample, we use the normal distribution. Select 7:ZInterval.
In this example, we have summary statistics, so we will select the STATS option for input. We enter the value of \( \sigma \), the value of \( \bar{x} \) and the sample size \( n \). Use 0.95 for the C-Level.

Highlight **Calculate** and press ENTER to get the results. Notice that the interval is given using standard mathematical notation for an interval. The interval for \( \mu \) goes from $1156.6 to $1232.4.

Example (\( \sigma \) is unknown)

A random sample of 16 wolf dens showed the number of pups in each to be as follows

5 8 7 5 3 4 3 9
5 8 5 6 5 6 4 7

Find a 90% confidence interval for the population mean number of pups in such dens.

In this case we have raw data, so enter the data in list L₁ using the **EDIT** option of the STAT key. Since \( \sigma \) is unknown, we use the \( t \) distribution. Under **Tests** from the STAT menu, select item **8:TInterval**. Since we have raw data, select the DATA option for Input. The data is in list L₁, and occurs with frequency 1. Enter 0.90 for the C-Level.
Highlight Calculate and press ENTER. The result is the interval from 4.84 pups to 6.41 pups.
TI-Nspire

The TI-Nspire fully supports confidence intervals. To access the confidence interval choices, enter the Lists and Spreadsheet page. Press menu and choose 4:Statistics then choose 3:Confidence Intervals.

Example (σ is known with a Large Sample)

Suppose a random sample of 250 credit card bills showed an average balance of $1200. Also assume that the population standard deviation is $350. Find a 95% confidence interval for the population mean credit card balance.

Since σ is known and we have a large sample, we use the normal distribution. Select 1:z Interval.

In this example, we have summary statistics, so we will select Stats for Data Input Method and press OK.
We enter the value of $\sigma$, the value of $\bar{x}$ and the sample size $n$. Use 0.95 for the C-Level.

Then press **OK**.

Example (*$\sigma$* is unknown)

A random sample of 16 wolf dens showed the number of pups in each to be as follows

\[
\begin{align*}
5 & \quad 8 & \quad 7 & \quad 5 & \quad 3 & \quad 4 & \quad 3 & \quad 9 \\
5 & \quad 8 & \quad 5 & \quad 6 & \quad 5 & \quad 6 & \quad 4 & \quad 7
\end{align*}
\]

Find a 90% confidence interval for the population mean number of pups in such dens.

In this case we have raw data, so enter the data in a list called **dens**. Since $\sigma$ is unknown, we use the $t$ distribution. Press menu and select **4:Statistics**, select item **3:Confidence Intervals** then select **2:t Interval**. Since we have raw data, select the DATA option for Data Input Method and press **OK**.
The data is in list dens, and occurs with frequency 1. Enter 0.90 for the C-Level.

Press OK.
LAB ACTIVITIES FOR CONFIDENCE INTERVALS FOR A POPULATION MEAN

1. Markey Survey was hired to do a study for a new soft drink, Refresh. 20 people were given a can of
Refresh and asked to rate it for taste on a scale of 1 to 10 (with 10 being the highest rating). The ratings
were

5  8  3  7  5  9  10  6  6  2
9  2  1  8  10  2  5  1  4  7

Find an 85% confidence interval for the population mean rating of Refresh.

2. Suppose a random sample of 50 basketball players showed the average height to be 78 inches. Also
assume that the population standard deviation is 1.5 inches.
   (a) Find a 99% confidence interval for the population mean height.
   (b) Find a 95% confidence interval for the population mean height.
   (c) Find a 90% confidence interval for the population mean height.
   (d) Find an 85% confidence interval for the population mean height.
   (e) What do you notice about the length of the confidence interval as the confidence level goes down?
      If you used a confidence level of 80%, would you expect the confidence interval to be longer or
      shorter than that of 85%? Run the program again to verify your answer.

CONFIDENCE INTERVALS FOR THE PROBABILITY OF SUCCESS p IN A
BINOMIAL DISTRIBUTION (SECTION 8.3 OF UNDERSTANDING BASIC
STATISTICS)

TI-83 Plus and TI-84 Plus

To find a confidence interval for a proportion, press the STAT key and use option A:1-PropZInt under TESTS.
Notice that the normal distribution will be used.

Example

The public television station BPBS wants to find the percent of its viewing population that gives donations to
the station. 300 randomly selected viewers were surveyed, and it was found that 123 viewers made contributions to
the station. Find a 95% confidence interval for the probability that a viewer of BPBS selected at random contributes
to the station.

The letter x is used to count the number of successes (the letter r is used in the text). Enter 123 for x and 300 for
n. Use 0.95 for the C-level.
Highlight **Calculate** and press ENTER. The result is the interval from 0.35 to 0.47.

**TI-Nspire**

To find a confidence interval for a proportion, enter the Spreadsheet and Lists page and press the menu key and select **4:Statistics** then choose **3:Confidence Intervals** and use option **5:1-Prop z Interval**. Notice that the normal distribution will be used.

**Example**

The public television station BPBS wants to find the percent of its viewing population that gives donations to the station. 300 randomly selected viewers were surveyed, and it was found that 123 viewers made contributions to the station. Find a 95% confidence interval for the probability that a viewer of BPBS selected at random contributes to the station.
The letter $x$ is used to count the number of successes (the letter $r$ is used in the text). Enter 123 for $x$ and 300 for $n$. Use 0.95 for the C-level.

Press OK. The result is the interval from 0.35 to 0.47.

**LAB ACTIVITIES FOR CONFIDENCE INTERVALS FOR THE PROBABILITY OF SUCCESS $p$ IN A BINOMIAL DISTRIBUTION**

1. Many types of error will cause a computer program to terminate or give incorrect results. One type of error is punctuation. For instance, if a comma is inserted in the wrong place, the program might not run. A study of programs written by students in an introductory programming course showed that 75 out of 300 errors selected at random were punctuation errors. Find a 99% confidence interval for the proportion of errors made by beginning programming students that are punctuation errors. Next, find a 90% confidence interval. Is this interval longer or shorter?

2. Sam decided to do a statistics project to determine a 90% confidence interval for the probability that a student at West Plains College eats lunch in the school cafeteria. He surveyed a random sample of 12
students and found that 9 ate lunch in the cafeteria. Can Sam use the program to find a confidence interval for the population proportion of students eating in the cafeteria? Why or why not? Try using a z interval with \( n = 12 \) and \( r = 9 \). What happens? What should Sam do to complete his project?
CHAPTER 9: HYPOTHESIS TESTING

TI-83 Plus and TI-84 Plus

The TI-83 Plus and TI-84 Plus both fully support hypothesis testing. Use the STAT key, then highlight TESTS. The options used in Chapter 8 are given on the two screens.

TI-Nspire

The TI-Nspire fully supports hypothesis testing. While in the Spreadsheet and Lists page press the menu key, then choose 4:Statistics. Next choose 4:Stat Tests. The options used in Chapter 8 are given on the following screen.

TESTING A SINGLE POPULATION MEAN (SECTION 9.2 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

When the value of $\sigma$ is known, a $z$-test based on normal distribution is used to test a population mean, provided that the population has a normal distribution or the sample size is large (at least 30). If the value of $\sigma$ is unknown, then a $t$-test will be used if either the population has approximately a normal distribution or the sample size is large.
Select option **1:Z-Test** for a z-test or **2:T-Test** for a t-test. As with confidence intervals, we have a choice of entering raw data into a list using the **Data** input option or using summary statistics with the **Stats** option. The null hypothesis is \( H_0: \mu = \mu_0 \). Enter the value of \( \mu_0 \) in the spot indicated. To select the alternate hypothesis, use one of the three options: \( \mu \neq \mu_0 \), \( \mu < \mu_0 \), or \( \mu > \mu_0 \). Output consists of the value of the sample mean \( \bar{x} \) and its corresponding \( z \) value. The \( p \) value of the sample statistic is also given.

**Example (Testing a mean when \( \sigma \) is known)**

Ten years ago, State College did a study regarding the number of hours full-time students worked each week. The mean number of hours was 8.7. A recent study involving a random sample of 45 full-time students showed that the average number of hours worked per week was 10.3. Also assume that the population standard deviation was 2.8. Use a 5% level of confidence to test if the mean number of hours worked per week by full-time students has increased.

Because we have a large sample, we will use the normal distribution for our sample test statistic. Select option **1:Z-Test**. We have summary statistics rather than raw data, so use the **Stats** input option. Enter the appropriate values on the screen.

Next highlight **Calculate** and press ENTER.

We see that the \( z \) value corresponding to the sample test statistic \( \bar{x} = 10.3 \) is \( z = 3.83 \). The critical value for a 5% level of significance and a right-tailed test is \( z_0 = 1.645 \). Clearly the sample \( z \) value is to the right of \( z_0 \), and we reject
Notice that the \( p \) value = 6.325E-5. This means that we move the decimal 5 places to the left, giving a \( p \) value of 0.000063. Since the \( p \) value is less than 0.05, we reject \( H_0 \).

Graph

The TI-83 Plus and TI-84 Plus both give an option to show the sample test statistic on the normal distribution. Highlight the Draw option on the Z-Test screen. Because the sample \( z \) is so far to the right, it does not appear on this window. However, its value does, and a rounded \( p \) value shows as well.

To do hypothesis testing of the mean when \( \sigma \) is unknown, we use a Student’s \( t \) distribution. Select option 2:T-Test. The entry screens are similar to those for the Z-Test.

TI-Nspire

When the value of \( \sigma \) is known, a \( z \)-test based on normal distribution is used to test a population mean, provided that the population has a normal distribution or the sample size is large (at least 30). If the value of \( \sigma \) is unknown, then a \( t \)-test will be used if either the population has approximately a normal distribution or the sample size is large. Select option 1:z Test for a \( z \)-test or 2:t Test for a \( t \)-test. As with confidence intervals, we have a choice of entering raw data into a list using the Data input option or using summary statistics with the Stats option. The null hypothesis is \( H_0: \mu = \mu_0 \). Enter the value of \( \mu_0 \) in the spot indicated. To select the alternate hypothesis, use one of the three options: \( \mu \neq \mu_0, \mu < \mu_0, \) or \( \mu > \mu_0 \). Output consists of the value of the sample mean \( \bar{x} \) and its corresponding \( z \) value. The \( p \) value of the sample statistic is also given.

Example (Testing a mean when \( \sigma \) is known)

Ten years ago, State College did a study regarding the number of hours full-time students worked each week. The mean number of hours was 8.7. A recent study involving a random sample of 45 full-time students showed that the average number of hours worked per week was 10.3. Also assume that the population standard deviation was 2.8. Use a 5% level of confidence to test if the mean number of hours worked per week by full-time students has increased.
Because we have a large sample, we will use the normal distribution for our sample test statistic. Select option 1:z Test. We have summary statistics rather than raw data, so use the Stats input option. Enter the appropriate values on the screen.

Next press OK.

We see that the $z$ value corresponding to the sample test statistic $\bar{x} = 10.3$ is $z = 3.83$. The critical value for a 5% level of significance and a right-tailed test is $z_0 = 1.645$. Clearly the sample $z$ value is to the right of $z_0$, and we reject $H_0$. Notice that the $p$ value = 6.325E-5. This means that we move the decimal 5 places to the left, giving a $p$ value of 0.000063. Since the $p$ value is less than 0.05, we reject $H_0$.

Graph

The TI-Nspire gives an option to show the sample test statistic on the normal distribution. Check the Shade P Value box z Test screen. Because the sample $z$ is so far to the right, it does not appear on this window. However, its value does, and a rounded $p$ value shows as well.
To do hypothesis testing of the mean when $\sigma$ is unknown, we use a Student’s $t$ distribution. Select option 2: Test. The entry screens are similar to those for the Z-Test.

LAB ACTIVITIES FOR TESTING A SINGLE POPULATION MEAN

1. A random sample of 65 pro basketball players showed their heights (in feet) to be as follows:

<table>
<thead>
<tr>
<th>Height (feet)</th>
<th>Height (feet)</th>
<th>Height (feet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.50</td>
<td>6.25</td>
<td>6.33</td>
</tr>
<tr>
<td>6.17</td>
<td>7.00</td>
<td>5.67</td>
</tr>
<tr>
<td>6.00</td>
<td>6.75</td>
<td>7.00</td>
</tr>
<tr>
<td>5.92</td>
<td>6.08</td>
<td>7.00</td>
</tr>
<tr>
<td>6.00</td>
<td>6.25</td>
<td>6.75</td>
</tr>
<tr>
<td>5.92</td>
<td>6.58</td>
<td>6.13</td>
</tr>
<tr>
<td>6.67</td>
<td>6.17</td>
<td>6.25</td>
</tr>
<tr>
<td>6.00</td>
<td>6.42</td>
<td>6.92</td>
</tr>
<tr>
<td>6.08</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) Enter the data into a list. Use one variable statistics to determine the sample standard deviation.
(b) Use the $t$ Test option to test the hypothesis that the average height of the players is greater than 6.2 feet, at the 1% level of significance.

2. In this problem, we will see how the test conclusion is possibly affected by a change in the level of significance. Teachers for Excellence is interested in the attention span of students in grades 1 and 2, now as compared to 20 years ago. They believe it has decreased. Studies done 20 years ago indicate that the attention span of children in grades 1 and 2 was 15 minutes. A study sponsored by Teachers for Excellence involved a random sample of 20 students in grades 1 and 2. The average attention span of these students was (in minutes) $\bar{x} = 14.2$ with standard deviation $s = 1.5$.

(a) Conduct the hypothesis test using $\alpha = 0.05$ and a left-tailed test. What is the test conclusion? What is the $p$ value?
(b) Conduct the hypothesis test using $\alpha = 0.01$ and a left-tailed test. What is the test conclusion? How could you have predicted this result by looking at the $p$ value from part (a)? Is the $p$ value for this part the same as it was for part (a)?
3. In this problem, let’s explore the effect that sample size has on the process of testing a mean. Run $z$ test with the hypotheses $H_0: \mu = 200$, $H_1: \mu > 200$, $\alpha = 0.05$, $\bar{x} = 210$ and $\sigma = 40$.

(a) Use the sample size $n = 30$. Note the $p$ value and $z$ score of the sample test statistic and test conclusion.

(b) Use the sample size $n = 50$. Note the $p$ value and $z$ score of the sample test statistic and test conclusion.

(c) Use the sample size $n = 100$. Note the $p$ value and $z$ score of the sample test statistic and test conclusion.

(d) In general, if your sample statistic is close to the proposed population mean specified in $H_0$, and you want to reject $H_0$, would you use a smaller or a larger sample size?

TESTS INVOLVING A SINGLE PROPORTION (SECTION 9.3 OF UNDERSTANDING BASIC STATISTICS)

**TI-83 Plus and TI-84 Plus**

To conduct a hypotheses test of a single proportion, select option **5:1-PropZTest**. The null hypothesis is $H_0: p = p_0$. Enter the value of $p_0$. The number of successes is designated by the value $x$. Enter the value. The sample size or number of trials is $n$. The alternate hypothesis will be prop $\neq p_0$, $< p_0$, or $> p_0$. Highlight the appropriate choice. Finally highlight **Calculate** and press ENTER. Notice that the **Draw** option is available to show the results on the standard normal distribution.

**TI-Nspire**

To conduct a hypotheses test of a single proportion, select option **5:1-Prop z Test**. The null hypothesis is $H_0: p = p_0$. Enter the value of $p_0$. The number of successes is designated by the value $x$. Enter the value. The sample size or number of trials is $n$. The alternate hypothesis will be prop $\neq p_0$, $< p_0$, or $> p_0$. Highlight the appropriate choice. Finally press **OK**. Notice that the **Draw** option is available to show the results on the standard normal distribution.
CHAPTER 10: INFERENCE ABOUT DIFFERENCES

TESTS INVOLVING PAIRED DIFFERENCES (DEPENDENT SAMPLES) (SECTION 10.1 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

To perform a paired difference test, we put our paired data into two columns, and then put the differences between corresponding pairs of values in a third column. For example, put the “before” data in list L₁ and the “after” data in L₂. Create L₃ = L₁ − L₂.

Example

Promoters of a state lottery decided to advertise the lottery heavily on television for one week during the middle of one of the lottery games. To see if the advertising improved ticket sales, the promoters surveyed a random sample of 8 ticket outlets and recorded weekly sales for one week before the television campaign and for one week after the campaign. The results follow (in ticket sales) where B stands for “before” and A for “after” the advertising campaign.

<table>
<thead>
<tr>
<th>B:</th>
<th>3201</th>
<th>4529</th>
<th>1425</th>
<th>1272</th>
<th>1784</th>
<th>1733</th>
<th>2563</th>
<th>3129</th>
</tr>
</thead>
<tbody>
<tr>
<td>A:</td>
<td>3762</td>
<td>4851</td>
<td>1202</td>
<td>1131</td>
<td>2172</td>
<td>1802</td>
<td>2492</td>
<td>3151</td>
</tr>
</tbody>
</table>

Test the claim that the television campaign increased lottery ticket sales at the 0.05 level of significance. We want to test to see if \( D = B - A \) is less than zero, since we are testing the claim that the lottery ticket sales are greater after the television campaign. We will put the “before” data in L₁ and the “after” data in L₂, and put the differences in list L₃ by highlighting the header L₃, entering L₁ − L₂, and pressing ENTER.

Next we will conduct a \( t \)-test on the differences in list L₃. Select option 2:T-Test from the TESTS menu. Select Data for Inpt. Since the null hypotheses is that \( d = 0 \), we use the null hypothesis \( H₀: d = µ₀ \) with \( µ₀ = 0 \). Since the data is in L₃, enter that as the List. Since the test is a left-tailed test, select \( µ < µ₀ \).
Highlight Calculate and press ENTER.

We see that the sample $t$ value is $-1.17$ with a corresponding $p$ value of 0.138. Since the $p$ value is greater than 0.05, we do not reject $H_0$.

**TI-Nspire**

To perform a paired difference test, we put our paired data into two columns, and then put the differences between corresponding pairs of values in a third column. For example, put the “before” data in a list called before and the “after” data in a list called after. Create $\text{diff} = \text{before} - \text{after}$.

**Example**

Promoters of a state lottery decided to advertise the lottery heavily on television for one week during the middle of one of the lottery games. To see if the advertising improved ticket sales, the promoters surveyed a random sample
of 8 ticket outlets and recorded weekly sales for one week before the television campaign and for one week after the campaign. The results follow (in ticket sales) where B stands for “before” and A for “after” the advertising campaign.

<table>
<thead>
<tr>
<th></th>
<th>B:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3201</td>
<td>4529</td>
<td>1425</td>
<td>1272</td>
<td>1784</td>
<td>1733</td>
<td>2563</td>
<td>3129</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3762</td>
<td>4851</td>
<td>1202</td>
<td>1131</td>
<td>2172</td>
<td>1802</td>
<td>2492</td>
<td>3151</td>
<td></td>
</tr>
</tbody>
</table>

Test the claim that the television campaign increased lottery ticket sales at the 0.05 level of significance. We want to test to see if $D = B - A$ is less than zero, since we are testing the claim that the lottery ticket sales are greater after the television campaign. We will put the “before” data in a list called before and the “after” data in a list called after, and put the differences in a list called diff. Then highlight the cell under the name diff and enter before – after, and pressing enter.

Next we will conduct a $t$-test on the differences in list called diff. Select option 2: $t$ Test from the Stat Tests menu. Select Data for Data Input Method. Since the null hypotheses is that $d = 0$, we use the null hypothesis $H_0: d = \mu_0$ with $\mu_0 = 0$. Since the data is in diff, enter that as the List. Since the test is a left-tailed test, select $\mu: < \mu_0$.

Press OK.
We see that the sample $t$ value is $-1.17$ with a corresponding $p$ value of 0.138. Since the $p$ value is greater than 0.05, we do not reject $H_0$.

LAB ACTIVITIES USING TESTS INVOLVING PAIRED DIFFERENCES
(DEPENDENT SAMPLES)

1. The data are pairs of values where the first entry represents average salary (in thousands of dollars/year) for male faculty members at an institution and the second entry represents the average salary for female faculty members (in thousands of dollars/year) at the same institution. A random sample of 22 U.S. colleges and universities was used (source: Academe, Bulletin of the American Association of University Professors).

   $$(34.5, 33.9) \quad (30.5, 31.2) \quad (35.1, 35.0) \quad (35.7, 34.2) \quad (31.5, 32.4)$$
   $$\quad (34.4, 34.1) \quad (32.1, 32.7) \quad (30.7, 29.9) \quad (33.7, 31.2) \quad (35.3, 35.5)$$
   $$\quad (30.7, 30.2) \quad (34.2, 34.8) \quad (39.6, 38.7) \quad (30.5, 30.0) \quad (33.8, 33.8)$$
   $$\quad (31.7, 32.4) \quad (32.8, 31.7) \quad (38.5, 38.9) \quad (40.5, 41.2) \quad (25.3, 25.5)$$
   $$\quad (28.6, 28.0) \quad (35.8, 35.1)$$

   (a) Put the first entries in a list and the second in another list and create a third list to be the difference of the first two lists.

   (b) Use the $t$ test to test the hypothesis that there is a difference in salary. What is the $p$ value of the sample test statistic? Do we reject or fail to reject the null hypothesis at the 5% level of significance? What about at the 1% level of significance?

   (c) Use the $t$ Test option to test the hypothesis that female faculty members have a lower average salary than male faculty members. What is the test conclusion at the 5% level of significance? At the 1% level of significance?

2. An audiologist is conducting a study on noise and stress. Twelve subjects selected at random were given a stress test in a room that was quiet. Then the same subjects were given another stress test, this time in a room with high-pitched background noise. The results of the stress tests were scores 1 through 20, with 20 indicating the greatest stress. The results, where $B$ represents the score of the test administered in the quiet room and $A$ represents the scores of the test administered in the room with the high-pitched background noise, are shown below.
Test the hypothesis that the stress level was greater during exposure to noise. Look at the $p$ value. Should you reject the null hypotheses at the 1% level of significance? At the 5% level?

**TESTS OF DIFFERENCE OF MEANS (INDEPENDENT SAMPLES) (SECTION 10.2 OF UNDERSTANDING BASIC STATISTICS)**

TI-83 Plus and TI-84 Plus

We consider the $x_1 \sim x_2$ distribution. The null hypothesis is that there is no difference between means, so $H_0: \mu_1 = \mu_2$ or $H_0: \mu_1 = \mu_2$.

**Example**

Sellers of microwave French fry cookers claim that their process saves cooking time. McDougle Fast Food Chain is considering the purchase of these new cookers, but wants to test the claim. Six batches of French fries were cooked in the traditional way. These times (in minutes) are as follows:

15 17 14 15 16 13

Six batches of French fries of the same weight were cooked using the new microwave cooker. These cooking times (in minutes) are as follows:

11 14 12 10 11 15

Test the claim that the microwave process takes less time. Use $\alpha = 0.05$.

Put the data for traditional cooking in list L1 and the data for the new method in List L2. Since we have small samples, we want to use the Student’s $t$ distribution. Select option 4:2-SampTTest. Select Data for Inpt, use the alternate hypothesis $\mu_1 > \mu_2$, select Yes for Pooled:
The output is on two screens.

\[
\begin{array}{l}
\text{2-SampTTTest} \\
\nu_1 > \mu_2 \\
t = 2.890086704 \\
p = .0080523896 \\
df = 10 \\
\bar{x}_1 = 15 \\
\downarrow \bar{x}_2 = 12.16666667 \\
\end{array}
\]

\[
\begin{array}{l}
\text{2-SampTTTest} \\
\nu_1 > \mu_2 \\
\uparrow s\bar{x}_1 = 1.41421356 \\
s\bar{x}_2 = 1.94079022 \\
s_{xp} = 1.69803808 \\
n_1 = 6 \\
n_2 = 6 \\
\end{array}
\]

Since the \( p \) value is 0.008, which is less than 0.05, we reject \( H_0 \) and conclude that the new method cooks food faster.

**TI-Nspire**

We consider the \( \bar{x}_1 - \bar{x}_2 \) distribution. The null hypothesis is that there is no difference between means, so \( H_0: \mu_1 = \mu_2 \) or \( H_0: \mu_1 - \mu_2. \)

**Example**

Sellers of microwave French fry cookers claim that their process saves cooking time. McDougle Fast Food Chain is considering the purchase of these new cookers, but wants to test the claim. Six batches of French fries were cooked in the traditional way. These times (in minutes) are as follows:

\[
15 \quad 17 \quad 14 \quad 15 \quad 16 \quad 13
\]

Six batches of French fries of the same weight were cooked using the new microwave cooker. These cooking times (in minutes) are as follows:

\[
11 \quad 14 \quad 12 \quad 10 \quad 11 \quad 15
\]

Test the claim that the microwave process takes less time. Use \( \alpha = 0.05. \)

Put the data for traditional cooking in list called trad and the data for the new method in a list called new. Since we have small samples, we want to use the Student’s \( t \) distribution. Select option **4:2-Sample t Test**. Select **Data** for **Data Input Method**, use the alternate hypothesis \( \mu_1 > \mu_2 \), select **Yes** for **Pooled**.
The output is on two screens.

Since the $p$ value is 0.008, which is less than 0.05, we reject $H_0$ and conclude that the new method cooks food faster.

**TESTING A DIFFERENCE OF PROPORTIONS (SECTION 10.3 OF *UNDERSTANDING BASIC STATISTICS*)

**TI-83 Plus and TI-84 Plus**

To conduct a hypothesis test for a difference of proportions, select option **6:2-PropZTest** and enter appropriate values.
**TI-Nspire**

To conduct a hypothesis test for a difference of proportions, select option 6:2-Prop z Test and enter appropriate values.

**LAB ACTIVITIES FOR TESTING DIFFERENCE OF MEANS (INDEPENDENT SAMPLES) OR PROPORTIONS**

1. Calm Cough Medicine is testing a new ingredient to see if its addition will lengthen the effective cough relief time of a single dose. A random sample of 15 doses of the standard medicine was tested, and the effective relief times were (in minutes):

   \[
   42 \quad 35 \quad 40 \quad 32 \quad 30 \quad 26 \quad 51 \quad 39 \quad 33 \quad 28 \\
   37 \quad 22 \quad 36 \quad 33 \quad 41
   \]

   A random sample of 20 doses was tested when the new ingredient was added. The effective relief times were (in minutes):

   \[
   43 \quad 51 \quad 35 \quad 49 \quad 32 \quad 29 \quad 42 \quad 38 \quad 45 \quad 74 \\
   31 \quad 31 \quad 46 \quad 36 \quad 33 \quad 45 \quad 30 \quad 32 \quad 41 \quad 25
   \]

   Assume that the standard deviations of the relief times are equal for the two populations. Test the claim that the effective relief time is longer when the new ingredient is added. Use \( \alpha = 0.01 \).

2. Publisher’s Survey did a study to see if the proportion of men who read mysteries is different from the proportion of women who read them. A random sample of 402 women showed that 112 read mysteries regularly (at least six books per year). A random sample of 365 men showed that 92 read mysteries regularly. Is the proportion of mystery readers different between men and women? Use a 1% level of significance.
   (a) Find the \( p \) value of the test conclusion.
   (b) Test the hypothesis that the proportion of women who read mysteries is greater than the proportion of men. Use a 1% level of significance. Is the \( p \) value for a right-tailed test half that of a two-tailed test? If you know the \( p \) value for a two-tailed test, can you draw conclusions for a one-tailed test?
CHAPTER 11: ADDITIONAL TOPICS USING INFERENCE

CHI-SQUARE TEST OF INDEPENDENCE (SECTION 11.1 OF UNDERSTANDING BASIC STATISTICS)

TI-83 Plus and TI-84 Plus

The TI-83 Plus and TI-84 Plus calculators support tests for independence. Press the STAT key, select TESTS, and option C: \( \chi^2 \)-Test. The original observed values need to be entered in a matrix.

Example

A computer programming aptitude test has been developed for high school seniors. The test designers claim that scores on the test are independent of the type of school the student attends: rural, suburban, urban. A study involving a random sample of students from each of these types of institutions yielded the following information, where aptitude scores range from 200 to 500 with 500 indicating the greatest aptitude and 200 the least. The entry in each cell is the observed number of students achieving the indicated score on the test.

<table>
<thead>
<tr>
<th>School Type</th>
<th>Score</th>
<th>Rural</th>
<th>Suburban</th>
<th>Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>200-299</td>
<td>33</td>
<td>65</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>300-399</td>
<td>45</td>
<td>79</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>400-500</td>
<td>21</td>
<td>47</td>
<td>63</td>
</tr>
</tbody>
</table>

Using the option C: \( \chi^2 \)-Test ..., test the claim that the aptitude test scores are independent of the type of school attended at the 0.05 level of significance.

We need to use a matrix to enter the data. Press 2nd [MATRIX]. Highlight EDIT, and select 1:[A]. Press ENTER.
Since the contingency table has 3 rows and 3 columns, type 3 (for number of rows), press ENTER, type 3 (for number of columns), press ENTER. Then type in the observed values. Press ENTER after each entry. Notice that you enter the table by rows. When the table is entered completely, press 2nd [MATRX] again.

![Matrix A 3x3 with values]

This time, select 2:[B]. Set the dimensions to be the same as for matrix [A]. For this example, [B] should be 3 rows × 3 columns.

Now press the STAT key, highlight TESTS, and use option C:χ²-Test. This tells you that the observed values of the table are in matrix [A]. The expected values will be placed in matrix [B].

Highlight Calculate and press ENTER. We see that the sample value of χ² is 1.32. The p value is 0.858. Since the p value is larger than 0.05, we do not reject the null hypothesis.
If you want to see the expected values in matrix \([B]\), type 2nd [MATRX] and select \([B]\) under NAMES.

**Graph**

Notice that one of the options of the \(\chi^2\) test is to graph the \(\chi^2\) distribution and show the sample test statistic on the graph. Highlight **Draw** and press ENTER. Before you do this, be sure that all STAT PLOTS are set to Off, and that you have cleared all the entries in the \(Y=\) menu.

![Graph](image.png)

**TI-Nspire**

The TI-Nspire calculators support tests for independence. While in the Spreadsheet and Lists page, press the menu key. Select **4:Stat Tests** then choose **8: \(\chi^2\) 2-way test**. The original observed values need to be entered into a matrix.

**Example**

A computer programming aptitude test has been developed for high school seniors. The test designers claim that scores on the test are independent of the type of school the student attends: rural, suburban, urban. A study involving a random sample of students from each of these types of institutions yielded the following information, where aptitude scores range from 200 to 500 with 500 indicating the greatest aptitude and 200 the least. The entry in each cell is the observed number of students achieving the indicated score on the test.

<table>
<thead>
<tr>
<th>School Type</th>
<th>Score</th>
<th>Rural</th>
<th>Suburban</th>
<th>Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>200-299</td>
<td>33</td>
<td>65</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>300-399</td>
<td>45</td>
<td>79</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>400-500</td>
<td>21</td>
<td>47</td>
<td>63</td>
</tr>
</tbody>
</table>

Using the option **8: \(\chi^2\) 2-way test**, test the claim that the aptitude test scores are independent of the type of school attended at the 0.05 level of significance.
Enter the data into a matrix by first entering the Calculator page. Press the menu key and select 7:Matrix & Vector. Choose 1:Create then choose 1:Matrix.

Input the number of rows and columns. In this case, there are 3 rows and 3 columns. Press OK.

Now input the values into the matrix. Press tab after entering each value and press enter when you are finished entering values. Now to store the matrix, press ctrl and then var. Type in a name for the matrix and press enter.
Press the menu key select **6:Statistics**, then **7:Stat Tests** and select option **8:χ² 2-way test**. For Observed Matrix, select amat from the drop-down list.

Press **OK**. We see that the sample value of $\chi^2$ is 1.32. The $p$ value is 0.858. Since the $p$ value is larger than 0.05, we do not reject the null hypothesis.
If you want to see the expected values in the ExpMatrix, type stat. and then select expmatrix and press enter twice.

LAB ACTIVITIES FOR CHI-SQUARE TEST OF INDEPENDENCE

1. We Care Auto Insurance had its staff of actuaries conduct a study to see if vehicle type and loss claim are independent. A random sample of auto claims over six months gives the information in the contingency table.

<table>
<thead>
<tr>
<th>Type of vehicle</th>
<th>$0-999</th>
<th>$1000-2999</th>
<th>$3000-5999</th>
<th>$6000+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sports car</td>
<td>20</td>
<td>10</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>Family Sedan</td>
<td>40</td>
<td>68</td>
<td>17</td>
<td>7</td>
</tr>
<tr>
<td>Compact</td>
<td>52</td>
<td>73</td>
<td>48</td>
<td>12</td>
</tr>
</tbody>
</table>

Test the claim that car type and loss claim are independent. Use $\alpha = 0.05$. 

124
2. An educational specialist is interested in comparing three methods of instruction:

- SL - standard lecture with discussion
- TV - video taped lectures with no discussion
- IM - individualized method with reading assignments and tutoring, but no lectures

The specialist conducted a study of these three methods to see if they are independent. A course was taught using each of the three methods and a standard final exam was given at the end. Students were put into the different method sections at random. The course type and test results are shown in the next contingency table.

<table>
<thead>
<tr>
<th>Course Type</th>
<th>&lt;60</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
<th>90-100</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL</td>
<td>10</td>
<td>4</td>
<td>70</td>
<td>31</td>
<td>25</td>
</tr>
<tr>
<td>TV</td>
<td>8</td>
<td>3</td>
<td>62</td>
<td>27</td>
<td>23</td>
</tr>
<tr>
<td>IM</td>
<td>7</td>
<td>2</td>
<td>58</td>
<td>25</td>
<td>22</td>
</tr>
</tbody>
</table>

Test the claim that the instruction method and final test scores are independent, using $\alpha = 0.01$. 
Appendix: Descriptions of Data Sets on the Student Website
Preface

There are over 70 data sets saved in Excel, Minitab Portable, SPSS, TI-83 Plus, TI-84 Plus, TI-Nspire/ASCII formats to accompany Understanding Statistics, 6th edition. These files can be found on the Brase/Brase statistics site at http://www.cengage.com/statistics/brase. The data sets are organized by category.

A. The following are provided for each data set:
   1. The category
   2. A brief description of the data and variables with a reference when appropriate
   3. File names for Excel, Minitab, SPSS, and TI-83 Plus, TI-84 Plus, TI-Nspire/ASCII formats

B. The categories are
   1. **Single variable large and small sample**
      File name prefix Sv followed by the data set number
      42 data sets……………………………………………………………………page A-4
   2. **Two variable independent samples** (large and small sample)
      File name prefix Tvis followed by the data set number
      10 data sets……………………………………………………………………page A-21
   3. **Two variable dependent samples** appropriate for \( t \)-tests
      File name prefix Tvcs followed by the data set number
      10 data sets……………………………………………………………………page A-26
   4. **Simple linear regression**
      File name prefix Slr followed by the data set number
      12 data sets……………………………………………………………………page A-30

C. The formats are
   1. Excel files in subdirectory Excel_10e. These files have suffix .xls
   2. Minitab portable files in subdirectory Minitab_10e. These files have suffix .zip
   3. TI-83 Plus and TI-84 Plus/ASCII files in subdirectory TI8384_10e. These files have suffix .zip
   4. SPSS files in subdirectory SPSS_10e. These files have suffix .zip

Suggestions for Using the Data Sets

1. Single variable large and small sample (file name prefix Sv)
   These data sets are appropriate for:
   - Graphs: Histograms, box plots
   - Descriptive statistics: Mean, median, mode, variance, standard deviation, coefficient of variation, 5 number summary
   - Inferential statistics: Confidence intervals for the population mean, hypothesis tests of a single mean
2. Two independent data sets (file name prefix Tvis)
   Graphs: Histograms, box plots for each data set
   Descriptive statistics: Mean, median, mode, variance, standard deviation, coefficient of variation, 5-number summary for each data set
   Inferential statistics: Confidence intervals for the difference of means, hypothesis tests for the difference of means

3. Paired data, dependent samples (file name prefix Tvcds)
   Descriptive statistics: Mean, median, mode, variance, standard deviation, coefficient of variation, 5 number summary for the difference of the paired data values.
   Inferential statistics: Hypothesis tests for the difference of means (paired data)

4. Data pairs for simple linear regression (file name prefix Slr)
   Graphs: Scatter plots, for individual variables histograms and box plots
   Descriptive statistics:
   • Mean, median, mode, variance, standard deviation, coefficient of variation, 5 number summary for individual variables.
   • Least squares line, sample correlation coefficient, sample coefficient of determination
   Inferential statistics: Testing $\rho$, confidence intervals for $\beta$, testing $\beta$

Descriptions of Data Sets

SINGLE VARIABLE LARGE AND SMALL SAMPLE

File name prefix: Sv followed by the number of the data file

01. Disney Stock Volume (Single Variable Large Sample $n \geq 30$)
The following data represents the number of shares of Disney stock (in hundreds of shares) sold for a random sample of 60 trading days
Reference: The Denver Post, Business section

```
12584  9441  18960  21480  10766  13059  8589  4965
4803  7240  10906  8561  6389  14372  18149  6309
13051 12754  10860  9574  19110  29585  21122  14522
17330 18119  10902  29158  16065  10376  10999  17950
15418 12618  16561  8022  9567  12245  8172  13708
11259 10518  9301  5197  13708  10518  9301  13708
7628  7304  7628  14265  10518  9301  5197  13708
6758  7304  7628  14265  13054  15336  14682  27804
16022 24009  32613  19111
```

File names
Excel: Sv01.xls
Minitab: Sv01.zip
SPSS: Svspss01.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Svls01.zip
02. Weights of Pro Football Players (Single Variable Large Sample \( n \geq 30 \))

The following data represents weights in pounds of 50 randomly selected pro football linebackers.

Reference: The Sports Encyclopedia Pro Football

<table>
<thead>
<tr>
<th>Weight (lbs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>225</td>
</tr>
<tr>
<td>230</td>
</tr>
<tr>
<td>235</td>
</tr>
<tr>
<td>238</td>
</tr>
<tr>
<td>232</td>
</tr>
<tr>
<td>227</td>
</tr>
<tr>
<td>244</td>
</tr>
<tr>
<td>222</td>
</tr>
<tr>
<td>250</td>
</tr>
<tr>
<td>226</td>
</tr>
<tr>
<td>242</td>
</tr>
<tr>
<td>253</td>
</tr>
<tr>
<td>251</td>
</tr>
<tr>
<td>225</td>
</tr>
<tr>
<td>229</td>
</tr>
<tr>
<td>247</td>
</tr>
<tr>
<td>239</td>
</tr>
<tr>
<td>223</td>
</tr>
<tr>
<td>233</td>
</tr>
<tr>
<td>222</td>
</tr>
<tr>
<td>243</td>
</tr>
<tr>
<td>237</td>
</tr>
<tr>
<td>230</td>
</tr>
<tr>
<td>240</td>
</tr>
<tr>
<td>255</td>
</tr>
<tr>
<td>230</td>
</tr>
<tr>
<td>245</td>
</tr>
<tr>
<td>240</td>
</tr>
<tr>
<td>246</td>
</tr>
<tr>
<td>243</td>
</tr>
<tr>
<td>255</td>
</tr>
<tr>
<td>241</td>
</tr>
<tr>
<td>245</td>
</tr>
</tbody>
</table>

File names

Excel: Sv02.xls
Minitab: Sv02.zip
SPSS: Svspss02.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv02.zip

03. Heights of Pro Basketball Players (Single Variable Large Sample \( n \geq 30 \))

The following data represents heights in feet of 65 randomly selected pro basketball players.

Reference: All-Time Player Directory, The Official NBA Encyclopedia

<table>
<thead>
<tr>
<th>Height (ft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.50</td>
</tr>
<tr>
<td>6.25</td>
</tr>
<tr>
<td>6.33</td>
</tr>
<tr>
<td>6.50</td>
</tr>
<tr>
<td>6.42</td>
</tr>
<tr>
<td>6.67</td>
</tr>
<tr>
<td>6.83</td>
</tr>
<tr>
<td>6.82</td>
</tr>
<tr>
<td>6.75</td>
</tr>
<tr>
<td>6.50</td>
</tr>
<tr>
<td>6.50</td>
</tr>
<tr>
<td>6.75</td>
</tr>
<tr>
<td>6.50</td>
</tr>
<tr>
<td>6.54</td>
</tr>
<tr>
<td>6.42</td>
</tr>
<tr>
<td>6.58</td>
</tr>
<tr>
<td>6.00</td>
</tr>
<tr>
<td>7.00</td>
</tr>
<tr>
<td>6.58</td>
</tr>
<tr>
<td>6.50</td>
</tr>
<tr>
<td>6.67</td>
</tr>
<tr>
<td>6.17</td>
</tr>
<tr>
<td>6.17</td>
</tr>
<tr>
<td>6.25</td>
</tr>
<tr>
<td>6.00</td>
</tr>
<tr>
<td>6.25</td>
</tr>
<tr>
<td>6.75</td>
</tr>
<tr>
<td>6.17</td>
</tr>
<tr>
<td>6.17</td>
</tr>
<tr>
<td>6.25</td>
</tr>
<tr>
<td>6.00</td>
</tr>
<tr>
<td>6.17</td>
</tr>
<tr>
<td>6.75</td>
</tr>
<tr>
<td>6.63</td>
</tr>
<tr>
<td>6.58</td>
</tr>
<tr>
<td>6.58</td>
</tr>
<tr>
<td>6.50</td>
</tr>
<tr>
<td>6.33</td>
</tr>
<tr>
<td>6.92</td>
</tr>
<tr>
<td>6.67</td>
</tr>
<tr>
<td>6.33</td>
</tr>
<tr>
<td>6.00</td>
</tr>
<tr>
<td>6.08</td>
</tr>
</tbody>
</table>

File names

Excel: Sv03.xls
Minitab: Sv03.zip
SPSS: Svspss03.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv03.zip

04. Miles per Gallon Gasoline Consumption (Single Variable Large Sample \( n \geq 30 \))

The following data represents miles per gallon gasoline consumption (highway) for a random sample of 55 makes and models of passenger cars.

Reference: Environmental Protection Agency

<table>
<thead>
<tr>
<th>Miles/Gallon</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
</tr>
<tr>
<td>27</td>
</tr>
<tr>
<td>22</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>35</td>
</tr>
<tr>
<td>35</td>
</tr>
<tr>
<td>33</td>
</tr>
<tr>
<td>52</td>
</tr>
<tr>
<td>49</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>27</td>
</tr>
<tr>
<td>18</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>23</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>25</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>24</td>
</tr>
</tbody>
</table>
05. **Fasting Glucose Blood Tests (Single Variable Large Sample $n \geq 30$)**

The following data represents glucose blood level (mg/100mL) after a 12-hour fast for a random sample of 70 women.


| 45 | 66 | 83 | 71 | 76 | 64 | 59 | 59 |
| 76 | 82 | 80 | 81 | 85 | 77 | 82 | 90 |
| 87 | 72 | 79 | 69 | 83 | 71 | 87 | 69 |
| 81 | 76 | 96 | 83 | 67 | 94 | 101 | 94 |
| 89 | 94 | 73 | 99 | 93 | 85 | 83 | 80 |
| 78 | 80 | 85 | 83 | 84 | 74 | 81 | 70 |
| 65 | 89 | 70 | 80 | 84 | 77 | 65 | 46 |
| 80 | 70 | 75 | 45 | 101 | 71 | 109 | 73 |
| 73 | 80 | 72 | 81 | 63 | 74 |

06. **Number of Children in Rural Canadian Families (Single Variable Large Sample $n \geq 30$)**

The following data represents the number of children in a random sample of 50 rural Canadian families.


| 11 | 13 | 4 | 14 | 10 | 2 | 5 | 0 |
| 0 | 3 | 9 | 2 | 5 | 2 | 3 | 3 |
| 3 | 4 | 7 | 1 | 9 | 4 | 3 | 3 |
| 2 | 6 | 0 | 2 | 6 | 5 | 9 | 5 |
| 4 | 3 | 2 | 5 | 2 | 2 | 3 | 5 |
| 14 | 7 | 6 | 6 | 2 | 5 | 3 | 4 |
| 6 | 1 |
07. Children as a % of Population (Single Variable Large Sample \( n \geq 30 \))

The following data represent percentage of children in the population for a random sample of 72 Denver neighborhoods.
Reference: The Piton Foundation, Denver, Colorado

30.2  18.6  13.6  36.9  32.8  19.4  12.3  39.7  22.2  31.2
36.4  37.7  38.8  28.1  18.3  22.4  26.5  20.4  37.6  23.8
22.1  53.2  6.8  20.7  31.7  10.4  21.3  19.6  41.5  29.8
14.7  12.3  17.0  16.7  20.7  34.8  7.5  19.0  27.2  16.3
24.3  39.8  31.1  34.3  15.9  24.2  20.3  31.2  30.0  33.1
29.1  39.0  36.0  31.8  32.9  26.5  4.9  19.5  21.0  24.2
12.1  38.3  39.3  20.2  24.0  28.6  27.1  30.0  60.8  39.2
21.6  20.3

File names
Excel: Sv07.xls
Minitab: Sv07.zip
SPSS: Svspss07.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv07.zip

08. Percentage Change in Household Income (Single Variable Large Sample \( n \geq 30 \))

The following data represent the percentage change in household income over a five-year period for a random sample of \( n = 78 \) Denver neighborhoods.
Reference: The Piton Foundation, Denver, Colorado

27.2  25.2  25.7  80.9  26.9  20.2  25.4  26.9  26.4  26.3
27.5  38.2  20.9  31.3  23.5  26.0  35.8  30.9  15.5  24.8
29.4  11.7  32.6  32.2  27.6  27.5  28.7  28.0  15.6  20.0
21.8  18.4  27.3  13.4  14.7  21.6  26.8  20.9  32.7  29.3
21.4  29.0  7.2  25.7  25.5  39.8  26.6  24.2  33.5  16.0
29.4  26.8  32.0  24.7  24.2  29.8  25.8  18.2  26.0  26.2
21.7  27.0  23.7  28.0  11.2  26.2  21.6  23.7  28.3  34.1
40.8  16.0  50.5  54.1  3.3  23.5  10.1  14.8

File names
Excel: Sv08.xls
Minitab: Sv08.zip
SPSS: Svspss08.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv08.zip

09. Crime Rate per 1,000 Population (Single Variable Large Sample \( n \geq 30 \))

The following data represent the crime rate per 1,000 population for a random sample of 70 Denver neighborhoods.
Reference: The Piton Foundation, Denver, Colorado
10. Percentage Change in Population (Single Variable Large Sample $n \geq 30$)

The following data represent the percentage change in population over a nine-year period for a random sample of 64 Denver neighborhoods.

Reference: The Piton Foundation, Denver, Colorado

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2</td>
<td>5.4</td>
<td>8.5</td>
<td>1.2</td>
<td>5.6</td>
<td>28.9</td>
<td>6.3</td>
<td>10.5</td>
<td>-1.5</td>
</tr>
<tr>
<td>21.6</td>
<td>-2.0</td>
<td>-1.0</td>
<td>3.3</td>
<td>2.8</td>
<td>3.3</td>
<td>28.5</td>
<td>-0.7</td>
<td>8.1</td>
</tr>
<tr>
<td>68.6</td>
<td>56.0</td>
<td>19.8</td>
<td>7.0</td>
<td>38.3</td>
<td>41.2</td>
<td>4.9</td>
<td>7.8</td>
<td>7.8</td>
</tr>
<tr>
<td>5.5</td>
<td>21.6</td>
<td>32.5</td>
<td>-0.5</td>
<td>2.8</td>
<td>4.9</td>
<td>8.7</td>
<td>-1.3</td>
<td>4.0</td>
</tr>
<tr>
<td>2.0</td>
<td>6.4</td>
<td>7.1</td>
<td>8.8</td>
<td>3.0</td>
<td>5.1</td>
<td>-1.9</td>
<td>-2.6</td>
<td>1.6</td>
</tr>
<tr>
<td>10.8</td>
<td>4.8</td>
<td>1.4</td>
<td>19.2</td>
<td>2.7</td>
<td>71.4</td>
<td>2.5</td>
<td>6.2</td>
<td>2.3</td>
</tr>
<tr>
<td>1.9</td>
<td>2.3</td>
<td>-3.3</td>
<td>2.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

File names
Excel: Sv09.xls
Minitab: Sv09.zip
SPSS: Svspss09.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv09.zip

11. Thickness of the Ozone Column (Single Variable Large Sample $n \geq 30$)

The following data represent the January mean thickness of the ozone column above Arosa, Switzerland (Dobson units: one milli-centimeter ozone at standard temperature and pressure). The data is from a random sample of years from 1926 on.

Reference: Laboratorium fuer Atmosphaerensphysik, Switzerland

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>324</td>
<td>332</td>
<td>362</td>
<td>383</td>
<td>335</td>
<td>349</td>
<td>354</td>
<td>319</td>
<td>360</td>
</tr>
<tr>
<td>400</td>
<td>341</td>
<td>315</td>
<td>368</td>
<td>361</td>
<td>336</td>
<td>349</td>
<td>347</td>
<td>338</td>
</tr>
<tr>
<td>341</td>
<td>352</td>
<td>342</td>
<td>361</td>
<td>318</td>
<td>337</td>
<td>300</td>
<td>352</td>
<td>340</td>
</tr>
<tr>
<td>327</td>
<td>357</td>
<td>320</td>
<td>377</td>
<td>338</td>
<td>361</td>
<td>301</td>
<td>331</td>
<td>334</td>
</tr>
<tr>
<td>336</td>
<td>378</td>
<td>369</td>
<td>332</td>
<td>344</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
12. Sun Spots (Single Variable Large Sample \( n \geq 30 \))

The following data represent the January mean number of sunspots. The data is taken from a random sample of Januarys from 1749 to 1983.


\[
\begin{array}{cccccccccc}
12.5 & 14.1 & 37.6 & 48.3 & 67.3 & 70.0 & 43.8 & 56.5 & 59.7 & 24.0 \\
12.0 & 27.4 & 53.5 & 73.9 & 104.0 & 54.6 & 4.4 & 177.3 & 70.1 & 54.0 \\
28.0 & 13.0 & 6.5 & 134.7 & 114.0 & 72.7 & 81.2 & 24.1 & 20.4 & 13.3 \\
9.4 & 25.7 & 47.8 & 50.0 & 45.3 & 61.0 & 39.0 & 12.0 & 7.2 & 11.3 \\
22.2 & 26.3 & 34.9 & 21.5 & 12.8 & 17.7 & 34.6 & 43.0 & 52.2 & 47.5 \\
30.9 & 11.3 & 4.9 & 88.6 & 188.0 & 35.6 & 50.5 & 12.4 & 3.7 & 18.5 \\
115.5 & 108.5 & 119.1 & 101.6 & 59.9 & 40.7 & 26.5 & 23.1 & 73.6 & 165.0 \\
202.5 & 217.4 & 57.9 & 38.7 & 15.3 & 8.1 & 16.4 & 84.3 & 51.9 & 58.0 \\
74.7 & 96.0 & 48.1 & 51.1 & 31.5 & 11.8 & 4.5 & 78.1 & 81.6 & 68.9 \\
\end{array}
\]

13. Motion of Stars (Single Variable Large Sample \( n \geq 30 \))

The following data represent the angular motions of stars across the sky due to the stars own velocity. A random sample of stars from the M92 global cluster was used. Units are arc seconds per century.


\[
\begin{array}{cccccccccccc}
0.042 & 0.048 & 0.019 & 0.025 & 0.028 & 0.041 & 0.030 & 0.051 & 0.026 \\
0.040 & 0.018 & 0.022 & 0.048 & 0.045 & 0.019 & 0.028 & 0.029 & 0.018 \\
0.033 & 0.035 & 0.019 & 0.046 & 0.021 & 0.026 & 0.026 & 0.033 & 0.046 \\
0.023 & 0.036 & 0.024 & 0.014 & 0.012 & 0.037 & 0.034 & 0.032 & 0.035 \\
0.015 & 0.027 & 0.017 & 0.035 & 0.021 & 0.016 & 0.036 & 0.029 & 0.031 \\
0.016 & 0.024 & 0.015 & 0.019 & 0.037 & 0.016 & 0.024 & 0.029 & 0.025 \\
0.022 & 0.028 & 0.023 & 0.021 & 0.020 & 0.020 & 0.016 & 0.016 & 0.016 \\
0.040 & 0.029 & 0.025 & 0.025 & 0.042 & 0.022 & 0.037 & 0.024 & 0.046 \\
0.016 & 0.024 & 0.028 & 0.027 & 0.060 & 0.045 & 0.037 & 0.027 & 0.028 \\
0.022 & 0.048 & 0.053 \\
\end{array}
\]
14. Arsenic and Ground Water (Single Variable Large Sample $n \geq 30$)

The following data represent (naturally occurring) concentration of arsenic in ground water for a random sample of 102 Northwest Texas wells. Units are parts per billion.

Reference: Nichols, C.E. and Kane, V.E., Union Carbide Technical Report K/UR-1

7.6 10.4 13.5 4.0 19.9 16.0 12.0 12.2 11.4 12.7
3.0 10.3 21.4 19.4 9.0 6.5 10.1 8.7 9.7 6.4
9.7 63.0 15.5 10.7 18.2 7.5 6.1 6.7 6.9 0.8
73.5 12.0 28.0 12.6 9.4 6.2 15.3 7.3 10.7 15.9
5.8 1.0 8.6 1.3 13.7 2.8 2.4 1.4 2.9 13.1
15.3 9.2 11.7 4.5 1.0 1.2 0.8 1.0 2.4 4.4
2.2 2.9 3.6 2.5 1.8 5.9 2.8 1.7 4.6 5.4
3.0 3.1 1.3 2.6 1.4 2.3 1.0 5.4 1.8 2.6
3.4 1.4 10.7 18.2 7.7 6.5 12.2 10.1 6.4 10.7
6.1 0.8 12.0 28.1 9.4 6.2 7.3 9.7 62.1 15.5
6.4 9.5

File names  Excel: Sv14.xls
            Minitab: Sv14.zip
            SPSS: Svspss14.zip
            TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv14.zip

15. Uranium in Ground Water (Single Variable Large Sample $n \geq 30$)

The following data represent (naturally occurring) concentrations of uranium in ground water for a random sample of 100 Northwest Texas wells. Units are parts per billion.

Reference: Nichols, C.E. and Kane, V.E., Union Carbide Technical Report K/UR-1

8.0 13.7 4.9 3.1 78.0 9.7 6.9 21.7 26.8
56.2 25.3 4.4 29.8 22.3 9.5 13.5 47.8 29.8
13.4 21.0 26.7 52.5 6.5 15.8 21.2 13.2 12.3
5.7 11.1 16.1 11.4 18.0 15.5 35.3 9.5 2.1
10.4 5.3 11.2 0.9 7.8 6.7 21.9 20.3 16.7
2.9 124.2 58.3 83.4 8.9 18.1 11.9 6.7 9.8
15.1 70.4 21.3 58.2 25.0 5.5 14.0 6.0 11.9
15.3 7.0 13.6 16.4 35.9 19.4 19.8 6.3 2.3
1.9 6.0 1.5 4.1 34.0 17.6 18.6 8.0 7.9
56.9 53.7 8.3 33.5 38.2 2.8 4.2 18.7 12.7
3.8 8.8 2.3 7.2 9.8 7.7 27.4 7.9 11.1
24.7

File names  Excel: Sv15.xls
            Minitab: Sv15.zip
16. **Ground Water pH (Single Variable Large Sample \( n \geq 30 \))**

A pH less than 7 is acidic, and a pH above 7 is alkaline. The following data represent pH levels in ground water for a random sample of 102 Northwest Texas wells.

Reference: Nichols, C.E. and Kane, V.E., Union Carbide Technical Report K/UR-1

```
7.6 7.7 7.4 7.7 7.1 8.2 7.4 7.5 7.2 7.4
7.2 7.6 7.4 7.8 8.1 7.5 7.1 8.1 7.3 8.2
7.6 7.0 7.3 7.4 7.8 8.1 7.3 8.0 7.2 8.5
7.1 8.2 8.1 7.9 7.2 7.1 7.0 7.5 7.2 7.3
8.6 7.7 7.5 7.8 7.6 7.1 7.8 7.3 8.4 7.5
7.1 7.4 7.2 7.4 7.3 7.7 7.0 7.3 7.6 7.2
8.1 8.2 7.4 7.6 7.3 7.1 7.0 7.0 7.4 7.2
8.2 8.1 7.9 8.1 8.2 7.7 7.5 7.3 7.9 8.8
7.1 7.5 7.9 7.5 7.6 7.7 8.2 8.7 7.9 7.0
8.8 7.1 7.2 7.3 7.6 7.1 7.0 7.0 7.3 7.2
7.8 7.6
```

File names

- Excel: Sv16.xls
- Minitab: Sv16.zip
- SPSS: Svspss16.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv16.zip

17. **Static Fatigue 90% Stress Level (Single Variable Large Sample \( n \geq 30 \))**

Kevlar Epoxy is a material used on the NASA space shuttle. Strands of this epoxy were tested at 90% breaking strength. The following data represent time to failure in hours at the 90% stress level for a random sample of 50 epoxy strands.

Reference: R.E. Barlow University of California, Berkeley

```
0.54 1.80 1.52 2.05 1.03 1.18 0.80 1.33 1.29 1.11
3.34 1.54 0.08 0.12 0.60 0.72 0.92 1.05 1.43 3.03
1.81 2.17 0.63 0.56 0.03 0.09 0.18 0.34 1.51 1.45
1.52 0.19 1.55 0.02 0.07 0.65 0.40 0.24 1.51 1.45
1.60 1.80 4.69 0.08 7.89 1.58 1.64 0.03 0.23 0.72
```

File names

- Excel: Sv17.xls
- Minitab: Sv17.zip
- SPSS: Svspss17.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv17.zip

18. **Static Fatigue 80% Stress Level (Single Variable Large Sample \( n \geq 30 \))**

Kevlar Epoxy is a material used on the NASA space shuttle. Strands of this epoxy were tested at 80% breaking strength. The following data represent time to failure in hours at the 80% stress level for a random sample of 50 epoxy strands.

Reference: R.E. Barlow University of California, Berkeley

```
0.54 1.80 1.52 2.05 1.03 1.18 0.80 1.33 1.29 1.11
3.34 1.54 0.08 0.12 0.60 0.72 0.92 1.05 1.43 3.03
1.81 2.17 0.63 0.56 0.03 0.09 0.18 0.34 1.51 1.45
1.52 0.19 1.55 0.02 0.07 0.65 0.40 0.24 1.51 1.45
1.60 1.80 4.69 0.08 7.89 1.58 1.64 0.03 0.23 0.72
```

File names

- Excel: Sv17.xls
- Minitab: Sv17.zip
- SPSS: Svspss17.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv17.zip
level for a random sample of 54 epoxy strands.
Reference: R.E. Barlow University of California, Berkeley

| 152.2   166.9  183.8   8.5  1.8   118.0  125.4  132.8  10.6 |
| 29.6  50.1  202.6  177.7  160.0  87.1  112.6  122.3  124.4 |
| 131.6  140.9  7.5  41.9  59.7  80.5  83.5  149.2  137.0 |
| 301.1  329.8  461.5  739.7  304.3  894.7  220.2  251.0  269.2 |
| 130.4  77.8   64.4  381.3  329.8  451.3  451.3  663.0  49.1 |
| 31.7  116.8  140.2  334.1  285.9  59.7  44.1  351.2  93.2 |

File names
Excel: Sv18.xls
Minitab: Sv18.zip
SPSS: Svspss18.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv18.zip

19. Tumor Recurrence (Single Variable Large Sample \( n \geq 30 \))
Certain kinds of tumors tend to recur. The following data represents the length of time in months for a tumor to recur after chemotherapy (sample size: 42).

| 19  18  17  1  21  22  54  46  25  49 |
| 50  1  59  39  43  39  5  9  38  18 |
| 14  45  54  59  46  50  29  12  19  36 |
| 38  40  43  41  10  50  41  25  19  39 |
| 27  20 |

File names
Excel: Sv19.xls
Minitab: Sv19.zip
SPSS: Svspss19.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv19.zip

20. Weight of Harvest (Single Variable Large Sample \( n \geq 30 \))
The following data represent the weights in kilograms of maize harvest from a random sample of 72 experimental plots on the island of St Vincent (Caribbean).

| 24.0  27.1  26.5  13.5  19.0  26.1  23.8  22.5  20.0 |
| 23.1  23.8  24.1  21.4  26.7  22.5  22.8  25.2  20.9 |
| 23.1  24.9  26.4  12.2  21.8  19.3  18.2  14.4  22.4 |
| 16.0  17.2  20.3  23.8  24.5  13.7  11.1  20.5  19.1 |
| 20.2  24.1  10.5  13.7  16.0  7.8  12.2  12.5  14.0 |
| 22.0  16.5  23.8  13.1  11.5  9.5  22.8  21.1  22.0 |
| 11.8  16.1  10.0  9.1  15.2  14.5  10.2  11.7  14.6 |
| 15.5  23.7  25.1  29.5  24.5  23.2  25.5  19.8  17.8 |
21. Apple Trees (Single Variable Large Sample \( n \geq 30 \))

The following data represent the trunk girth (mm) of a random sample of 60 four-year-old apple trees at East Malling Research Station (England)

Reference: S.C. Pearce, University of Kent at Canterbury

\[
\begin{array}{cccccccccccc}
108 & 99 & 106 & 102 & 115 & 120 & 120 & 117 & 122 & 142 \\
103 & 114 & 101 & 99 & 112 & 120 & 108 & 91 & 115 & 109 \\
114 & 105 & 99 & 122 & 106 & 113 & 114 & 75 & 96 & 124 \\
91 & 102 & 108 & 110 & 83 & 90 & 69 & 117 & 84 & 142 \\
122 & 113 & 105 & 112 & 117 & 122 & 129 & 100 & 138 & 117 \\
\end{array}
\]

22. Black Mesa Archaeology (Single Variable Large Sample \( n \geq 30 \))

The following data represent rim diameters (cm) of a random sample of 40 bowls found at Black Mesa archaeological site. The diameters are estimated from broken pot shards.

Reference: Michelle Hegmon, Crow Canyon Archaeological Center, Cortez, Colorado

\[
\begin{array}{cccccccccccc}
17.2 & 15.1 & 13.8 & 18.3 & 17.5 & 11.1 & 7.3 & 23.1 & 21.5 & 19.7 \\
17.6 & 15.9 & 16.3 & 25.7 & 27.2 & 33.0 & 10.9 & 23.8 & 24.7 & 18.6 \\
16.9 & 18.8 & 19.2 & 14.6 & 8.2 & 9.7 & 11.8 & 13.3 & 14.7 & 15.8 \\
17.4 & 17.1 & 21.3 & 15.2 & 16.8 & 17.0 & 17.9 & 18.3 & 14.9 & 17.7 \\
\end{array}
\]

23. Wind Mountain Archaeology (Single Variable Large Sample \( n \geq 30 \))

The following data represent depth (cm) for a random sample of 73 significant archaeological artifacts at the Wind Mountain excavation site.


\[
\begin{array}{cccccccccccc}
85 & 45 & 75 & 60 & 90 & 90 & 115 & 30 & 55 & 58 \\
78 & 120 & 80 & 65 & 65 & 140 & 65 & 50 & 30 & 125 \\
\end{array}
\]
24. Arrow Heads (Single Variable Large Sample \( n \geq 30 \))

The following data represent the lengths (cm) of a random sample of 61 projectile points found at the Wind Mountain Archaeological site.

\[
\begin{array}{cccccccccccccccccccc}
75 & 137 & 80 & 120 & 15 & 45 & 70 & 65 & 50 & 45 \\
95 & 70 & 70 & 28 & 40 & 125 & 105 & 75 & 80 & 70 \\
90 & 68 & 73 & 75 & 55 & 70 & 95 & 65 & 200 & 75 \\
15 & 90 & 46 & 33 & 100 & 65 & 60 & 55 & 85 & 50 \\
10 & 68 & 99 & 145 & 45 & 75 & 45 & 95 & 85 & 65 \\
65 & 52 & 82 \\
\end{array}
\]

File names
- Excel: Sv23.xls
- Minitab: Sv23.zip
- SPSS: Svspss23.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv23.zip

25. Anasazi Indian Bracelets (Single Variable Large Sample \( n \geq 30 \))

The following data represent the diameter (cm) of shell bracelets and rings found at the Wind Mountain archaeological site.

\[
\begin{array}{cccccccccccccccccccc}
3.1 & 4.1 & 1.8 & 2.1 & 2.2 & 1.3 & 1.7 & 3.0 & 3.7 & 2.3 \\
2.6 & 2.2 & 2.8 & 3.0 & 3.2 & 3.3 & 2.4 & 2.8 & 2.8 & 2.9 \\
2.9 & 2.2 & 2.4 & 2.1 & 3.4 & 3.1 & 1.6 & 3.1 & 3.5 & 2.3 \\
3.1 & 2.7 & 2.1 & 2.0 & 4.8 & 1.9 & 3.9 & 2.0 & 5.2 & 2.2 \\
2.6 & 1.9 & 4.0 & 3.0 & 3.4 & 4.2 & 2.4 & 3.5 & 3.1 & 3.7 \\
3.7 & 2.9 & 2.6 & 3.6 & 3.9 & 3.5 & 1.9 & 4.0 & 4.0 & 4.6 \\
1.9 \\
\end{array}
\]

File names
- Excel: Sv24.xls
- Minitab: Sv24.zip
- SPSS: Svspss24.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv24.zip

24. Arrow Heads (Single Variable Large Sample \( n \geq 30 \))

The following data represent the lengths (cm) of a random sample of 61 projectile points found at the Wind Mountain Archaeological site.

\[
\begin{array}{cccccccccccccccccccc}
5.0 & 5.0 & 8.0 & 6.1 & 6.0 & 5.1 & 5.9 & 6.8 & 4.3 & 5.5 \\
7.2 & 7.0 & 5.0 & 5.6 & 5.3 & 7.0 & 3.4 & 8.2 & 4.3 & 5.2 \\
1.5 & 6.1 & 4.0 & 6.0 & 5.5 & 5.2 & 5.2 & 5.2 & 5.5 & 7.2 \\
6.0 & 6.2 & 5.2 & 5.0 & 4.0 & 5.7 & 5.1 & 6.1 & 5.7 & 7.3 \\
7.3 & 6.7 & 4.2 & 4.0 & 6.0 & 7.1 & 7.3 & 5.5 & 5.8 & 8.9 \\
7.5 & 8.3 & 6.8 & 4.9 & 4.0 & 6.2 & 7.7 & 5.0 & 5.2 & 6.8 \\
6.1 & 7.2 & 4.4 & 4.0 & 5.0 & 6.0 & 6.2 & 7.2 & 5.8 & 6.8 \\
7.7 & 4.7 & 5.3 \\
\end{array}
\]
26. Pizza Franchise Fees  (Single Variable Large Sample $n \geq 30$)
The following data represent annual franchise fees (in thousands of dollars) for a random sample of 36 pizza franchises.
Reference: *Business Opportunities Handbook*

\[
\begin{array}{cccccccccccc}
25.0 & 15.5 & 7.5 & 19.9 & 18.5 & 25.5 & 15.0 & 5.5 & 15.2 & 15.0 \\
14.9 & 18.5 & 14.5 & 29.0 & 22.5 & 10.0 & 25.0 & 35.5 & 22.1 & 89.0 \\
17.5 & 33.3 & 17.5 & 12.0 & 15.5 & 25.5 & 12.5 & 17.5 & 12.5 & 35.0 \\
30.0 & 21.0 & 35.5 & 10.5 & 5.5 & 20.0 \\
\end{array}
\]

27. Pizza Franchise Start-up Requirement  (Single Variable Large Sample $n \geq 30$)
The following data represent annual the start-up cost (in thousands of dollars) for a random sample of 36 pizza franchises.
Reference: *Business Opportunities Handbook*

\[
\begin{array}{cccccccccccc}
40 & 25 & 50 & 129 & 250 & 128 & 110 & 142 & 25 & 90 \\
75 & 100 & 500 & 214 & 275 & 50 & 128 & 250 & 50 & 75 \\
30 & 40 & 185 & 50 & 175 & 125 & 200 & 150 & 150 & 120 \\
95 & 30 & 400 & 149 & 235 & 100 \\
\end{array}
\]

28. College Degrees  (Single Variable Large Sample $n \geq 30$)
The following data represent percentages of the adult population with college degrees. The sample is from a random sample of 68 Midwest counties.
Reference: *County and City Data Book* 12th edition, U.S. Department of Commerce

\[
\begin{array}{cccccccccccc}
9.9 & 9.8 & 6.8 & 8.9 & 11.2 & 15.5 & 9.8 & 16.8 & 9.9 & 11.6 \\
9.2 & 8.4 & 11.3 & 11.5 & 15.2 & 10.8 & 16.3 & 17.0 & 12.8 & 11.0 \\
6.0 & 16.0 & 12.1 & 9.8 & 9.4 & 9.9 & 10.5 & 11.8 & 10.3 & 11.1 \\
12.5 & 7.8 & 10.7 & 9.6 & 11.6 & 8.8 & 12.3 & 12.2 & 12.4 & 10.0 \\
10.0 & 18.1 & 8.8 & 17.3 & 11.3 & 14.5 & 11.0 & 12.3 & 9.1 & 12.7 \\
\end{array}
\]
<table>
<thead>
<tr>
<th>Poverty Level (Single Variable Large Sample ( n \geq 30 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>The following data represent percentages of all persons below the poverty level. The sample is from a random collection of 80 cities in the Western U.S.</td>
</tr>
<tr>
<td>Reference: <em>County and City Data Book</em> 12th edition, U.S. Department of Commerce</td>
</tr>
<tr>
<td>12.1</td>
</tr>
<tr>
<td>9.4</td>
</tr>
<tr>
<td>21.6</td>
</tr>
<tr>
<td>19.4</td>
</tr>
<tr>
<td>30.0</td>
</tr>
<tr>
<td>21.0</td>
</tr>
<tr>
<td>17.9</td>
</tr>
<tr>
<td>16.6</td>
</tr>
<tr>
<td>28.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Working at Home (Single Variable Large Sample ( n \geq 30 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>The following data represent percentages of adults whose primary employment involves working at home. The data is from a random sample of 50 California cities.</td>
</tr>
<tr>
<td>Reference: <em>County and City Data Book</em> 12th edition, U.S. Department of Commerce</td>
</tr>
<tr>
<td>4.3</td>
</tr>
<tr>
<td>4.3</td>
</tr>
<tr>
<td>7.0</td>
</tr>
<tr>
<td>2.4</td>
</tr>
<tr>
<td>3.8</td>
</tr>
</tbody>
</table>

File names |
Excel: Sv28.xls |
Minitab: Sv28.zip |
SPSS: Svspss28.zip |
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv28.zip |

29. Poverty Level (Single Variable Large Sample \( n \geq 30 \))

30. Working at Home (Single Variable Large Sample \( n \geq 30 \))
31. **Number of Pups in Wolf Den (Single Variable Small Sample $n < 30$)**

The following data represent the number of wolf pups per den from a random sample of 16 wolf dens.


5 8 7 5 3 4 3 9 5 8 5 6 5 6 4 7

File names  
Excel: Sv31.xls  
Minitab: Sv31.zip  
SPSS: Svspss31.zip  
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv31.zip

32. **Glucose Blood Level (Single Variable Small Sample $n < 30$)**

The following data represent glucose blood level (mg/100ml) after a 12-hour fast for a random sample of 6 tests given to an individual adult female.


83 83 86 86 78 88

File names  
Excel: Sv32.xls  
Minitab: Sv32.zip  
SPSS: Svspss32.zip  
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv32.zip

33. **Length of Remission (Single Variable Small Sample $n < 30$)**

The drug 6-mP (6-mercaptopurine) is used to treat leukemia. The following data represent the length of remission in weeks for a random sample of 21 patients using 6-mP.

Reference: E.A. Gehan, University of Texas Cancer Center

10 7 32 23 22 6 16 34 32 25 11 20 19 6 17 35 6 13 9 6 10

File names  
Excel: Sv33.xls  
Minitab: Sv33.zip  
SPSS: Svspss33.zip  
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv33.zip
34. **Entry Level Jobs (Single Variable Small Sample $n < 30$)**  
The following data represent percentage of entry-level jobs in a random sample of 16 Denver neighborhoods.  
Reference: The Piton Foundation, Denver, Colorado

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8.9</td>
<td>22.6</td>
<td>18.5</td>
<td>9.2</td>
<td>8.2</td>
<td>24.3</td>
<td>15.3</td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>9.2</td>
<td>14.9</td>
<td>4.7</td>
<td>16.5</td>
<td>11.6</td>
<td>15.3</td>
<td>9.7</td>
<td>8.0</td>
<td></td>
</tr>
</tbody>
</table>

File names  
Excel: Sv34.xls  
Minitab: Sv34.zip  
SPSS: Svspss34.zip  
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv34.zip

35. **Licensed Child Care Slots (Single Variable Small Sample $n < 30$)**  
The following data represents the number of licensed childcare slots in a random sample of 15 Denver neighborhoods.  
Reference: The Piton Foundation, Denver, Colorado

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>523</td>
<td>106</td>
<td>184</td>
<td>121</td>
<td>357</td>
<td>319</td>
<td>656</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>241</td>
<td>226</td>
<td>741</td>
<td>172</td>
<td>266</td>
<td>423</td>
<td>212</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

File names  
Excel: Sv35.xls  
Minitab: Sv35.zip  
SPSS: Svspss35.zip  
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv35.zip

36. **Subsidized Housing (Single Variable Small Sample $n < 30$)**  
The following data represent the percentage of subsidized housing in a random sample of 14 Denver neighborhoods.  
Reference: The Piton Foundation, Denver, Colorado

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10.2</td>
<td>11.8</td>
<td>9.7</td>
<td>22.3</td>
<td>6.8</td>
<td>10.4</td>
<td>11.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.4</td>
<td>6.6</td>
<td>13.7</td>
<td>13.6</td>
<td>6.5</td>
<td>16.0</td>
<td>24.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

File names  
Excel: Sv36.xls  
Minitab: Sv36.zip  
SPSS: Svspss36.zip  
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv36.zip

37. **Sulfate in Ground Water (Single Variable Small Sample $n < 30$)**  
The following data represent naturally occurring amounts of sulfate SO4 in well water. Units: parts per million. The data is from a random sample of 24 water wells in Northwest Texas.  
38. **Earth’s Rotation Rate (Single Variable Small Sample $n < 30$)**

The following data represent changes in the earth’s rotation (i.e. day length). Units: 0.00001 second. The data is for a random sample of 23 years.


\[ \begin{array}{cccccccccccc}
-12 & 110 & 78 & 126 & -35 & 104 & 111 & 22 & -31 & 92 \\
51 & 36 & 231 & -13 & 65 & 119 & 21 & 104 & 112 & -15 \\
137 & 139 & 101 \\
\end{array} \]

File names
- Excel: Sv38.xls
- Minitab: Sv38.zip
- SPSS: Svspss38.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv38.zip

39. **Blood Glucose (Single Variable Small Sample $n < 30$)**

The following data represent glucose levels (mg/100ml) in the blood for a random sample of 27 non-obese adult subjects.


\[ \begin{array}{cccccccccccc}
80 & 85 & 75 & 90 & 70 & 97 & 91 & 85 & 90 & 85 \\
105 & 86 & 78 & 92 & 93 & 90 & 80 & 102 & 90 & 90 \\
99 & 93 & 91 & 86 & 98 & 86 & 92 \\
\end{array} \]

File names
- Excel: Sv39.xls
- Minitab: Sv39.zip
- SPSS: Svspss39.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv39.zip

40. **Plant Species (Single Variable Small Sample $n < 30$)**

The following data represent the observed number of native plant species from random samples of study plots on different islands in the Galapagos Island chain.


\[ \begin{array}{cccccccccccc}
23 & 26 & 33 & 73 & 21 & 35 & 30 & 16 & 3 & 17 \\
9 & 8 & 9 & 19 & 65 & 12 & 11 & 89 & 81 & 7 \\
23 & 95 & 4 & 37 & 28 \\
\end{array} \]

File names
- Excel: Sv40.xls
41. **Apples (Single Variable Small Sample \( n < 30 \))**

The following data represent mean fruit weight (grams) of apples per tree for a random sample of 28 trees in an agricultural experiment.


<table>
<thead>
<tr>
<th>85.3</th>
<th>86.9</th>
<th>96.8</th>
<th>108.5</th>
<th>113.8</th>
<th>87.7</th>
<th>94.5</th>
<th>99.9</th>
<th>92.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>67.3</td>
<td>90.6</td>
<td>129.8</td>
<td>48.9</td>
<td>117.5</td>
<td>100.8</td>
<td>94.5</td>
<td>94.4</td>
<td>98.9</td>
</tr>
<tr>
<td>96.0</td>
<td>99.4</td>
<td>79.1</td>
<td>108.5</td>
<td>84.6</td>
<td>117.5</td>
<td>70.0</td>
<td>104.4</td>
<td>127.1</td>
</tr>
<tr>
<td>135.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

File names
- Excel: Sv41.xls
- Minitab: Sv41.zip
- SPSS: Svspss41.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv41.zip

**Space Shuttle: Epoxy (Single Variable Large Sample \( n \geq 30 \))**

The following data represent time to failure (in hours) for a random sample of 50 epoxy strands.

Reference: R. E. Barlow, University of California, Berkeley

<table>
<thead>
<tr>
<th>0.54</th>
<th>1.80</th>
<th>1.52</th>
<th>2.05</th>
<th>1.03</th>
<th>1.18</th>
<th>0.80</th>
<th>1.33</th>
<th>1.29</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.11</td>
<td>3.34</td>
<td>1.54</td>
<td>0.08</td>
<td>0.12</td>
<td>0.60</td>
<td>0.72</td>
<td>0.92</td>
<td>1.05</td>
</tr>
<tr>
<td>1.43</td>
<td>3.03</td>
<td>1.81</td>
<td>2.17</td>
<td>0.63</td>
<td>0.56</td>
<td>0.03</td>
<td>0.09</td>
<td>0.18</td>
</tr>
<tr>
<td>0.34</td>
<td>1.51</td>
<td>1.45</td>
<td>1.52</td>
<td>0.19</td>
<td>1.55</td>
<td>0.02</td>
<td>0.07</td>
<td>0.65</td>
</tr>
<tr>
<td>0.40</td>
<td>0.24</td>
<td>1.51</td>
<td>1.45</td>
<td>1.60</td>
<td>1.80</td>
<td>4.69</td>
<td>0.08</td>
<td>7.89</td>
</tr>
<tr>
<td>1.58</td>
<td>1.64</td>
<td>0.03</td>
<td>0.23</td>
<td>0.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

File names
- Excel: Sv42.xls
- Minitab: Sv42.zip
- SPSS: Svspss42.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Sv42.zip
TWO VARIABLE INDEPENDENT SAMPLES

File name prefix: Tvis followed by the number of the data file

01. Heights of Football Players Versus Heights of Basketball Players (Two variable independent large samples)
The following data represent heights in feet of 45 randomly selected pro football players and 40 randomly selected pro basketball players.
Reference: Sports Encyclopedia of Pro Football and Official NBA Basketball Encyclopedia

X1 = heights (ft.) of pro football players
6.42 6.58 6.08 6.58 6.50 6.42 6.25 6.67 5.91 6.00
5.83 6.00 5.83 5.08 6.75 5.83 6.17 5.75 6.00 5.75
6.50 5.83 5.91 5.67 6.00 6.08 6.17 6.58 6.50 6.25
6.33 5.25 6.67 6.50 5.83

X2 = heights (ft.) of pro basketball players
6.08 6.58 6.25 6.58 6.25 5.92 7.00 6.41 6.75 6.25
6.00 6.92 6.83 6.58 6.41 6.67 6.67 5.75 6.25 6.25
6.50 6.00 6.92 6.25 6.42 6.58 6.58 6.08 6.75 6.50
6.83 6.08 6.92 6.00 6.33 6.50 6.58 6.83 6.50 6.58

File names
Excel: Tvis01.xls
Minitab: Tvis01.zip
SPSS: Tvisssps01.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Tvis01.zip

02. Petal Length for Iris Virginica Versus Petal Length for Iris Setosa (Two variable independent large samples)
The following data represent petal length (cm.) for a random sample of 35 iris virginica and a random sample of 38 iris setosa

X1 = petal length (c.m.) iris virginica
5.1 5.8 6.3 6.1 5.1 5.5 5.3 5.5 6.9 5.0 4.9 6.0 4.8 6.1 5.6 5.1
5.6 4.8 5.4 5.1 5.1 5.9 5.2 5.7 5.4 4.5 6.1 5.3 5.5 6.7 5.7 4.9
4.8 5.8 5.1

X2 = petal length (c.m.) iris setosa
1.5 1.7 1.4 1.5 1.5 1.6 1.4 1.1 1.2 1.4 1.7 1.0 1.7 1.9 1.6 1.4
1.5 1.4 1.2 1.3 1.5 1.3 1.6 1.9 1.4 1.6 1.5 1.4 1.2 1.9 1.5
1.6 1.4 1.3 1.7 1.5 1.7

File names
Excel: Tvis02.xls
Minitab: Tvis02.zip
03. Sepal Width Of Iris Versicolor Versus Iris Virginica (Two variable independent large samples)

The following data represent sepal width (cm.) for a random sample of 40 iris versicolor and a random sample of 42 iris virginica.


\[
\begin{align*}
X1 &= \text{sepal width (cm.) iris versicolor} \\
&= 3.2 \ 3.2 \ 3.1 \ 2.3 \ 2.8 \ 2.8 \ 3.3 \ 2.4 \ 2.9 \ 2.7 \ 2.0 \ 3.0 \ 2.2 \ 2.9 \ 2.9 \ 3.1 \\
&= 3.0 \ 2.7 \ 2.2 \ 2.5 \ 3.2 \ 2.8 \ 2.5 \ 2.8 \ 2.9 \ 3.0 \ 2.8 \ 3.0 \ 2.9 \ 2.6 \ 2.4 \ 2.4 \\
&= 2.7 \ 2.7 \ 3.0 \ 3.4 \ 3.1 \ 2.3 \ 3.0 \ 2.5 \\

X2 &= \text{sepal width (cm.) iris virginica} \\
&= 3.3 \ 2.7 \ 3.0 \ 2.9 \ 3.0 \ 3.0 \ 2.5 \ 2.9 \ 2.5 \ 3.6 \ 3.2 \ 3.2 \ 2.7 \ 3.0 \ 2.5 \ 2.8 \ 3.2 \\
&= 3.0 \ 3.8 \ 2.6 \ 2.2 \ 3.2 \ 2.8 \ 2.8 \ 2.7 \ 3.3 \ 3.2 \ 2.8 \ 3.0 \ 2.8 \ 3.0 \ 2.8 \ 3.8 \\
&= 2.8 \ 2.8 \ 2.6 \ 3.0 \ 3.4 \ 3.1 \ 3.0 \ 3.1 \ 3.1 \ 3.1 \\

\end{align*}
\]

File names
- Excel: Tvis03.xls
- Minitab: Tvis03.zip
- SPSS: Tvisspss03.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Tvis03.spss

04. Archaeology, Ceramics (Two variable independent large samples)

The following data represent independent random samples of shard counts of painted ceramics found at the Wind Mountain archaeological site.

Reference: Woosley and McIntyre, Mimbres Mogollon Archaeology, Univ. New Mexico Press

\[
\begin{align*}
X1 &= \text{count Mogollon red on brown} \\
&= 52 \ 10 \ 8 \ 71 \ 7 \ 31 \ 24 \ 20 \ 17 \ 5 \\
&= 16 \ 75 \ 25 \ 17 \ 14 \ 33 \ 13 \ 17 \ 12 \ 19 \\
&= 67 \ 13 \ 35 \ 14 \ 3 \ 7 \ 9 \ 19 \ 16 \ 22 \\
&= 7 \ 10 \ 9 \ 49 \ 6 \ 13 \ 24 \ 45 \ 14 \ 20 \\
&= 3 \ 6 \ 30 \ 41 \ 26 \ 32 \ 14 \ 33 \ 1 \ 48 \\
&= 44 \ 14 \ 16 \ 15 \ 13 \ 8 \ 61 \ 11 \ 12 \ 16 \\
&= 20 \ 39 \\

X2 &= \text{count Mimbres black on white} \\
&= 61 \ 21 \ 78 \ 9 \ 14 \ 12 \ 34 \ .54 \ 10 \ 15 \\
&= 43 \ 9 \ 7 \ 67 \ 18 \ 18 \ 24 \ .54 \ 8 \ 10 \\
&= 16 \ 6 \ 17 \ 14 \ 25 \ 22 \ 25 \ 13 \ 23 \ 12 \\
&= 36 \ 10 \ 56 \ 35 \ 79 \ 69 \ 41 \ 36 \ 18 \ 25 \\
&= 27 \ 27 \ 11 \ 13 \\
\end{align*}
\]
05. Agriculture, Water Content of Soil (Two variable independent large samples)
The following data represent soil water content (% water by volume) for independent random samples of soil from two experimental fields growing bell peppers.

$X_1$ = soil water content from field I
15.1 11.2 10.3 10.8 16.6 8.3 9.1 12.3 9.1 14.3 10.7 16.1 10.2 15.2 8.9 9.5 9.6 11.3 14.0 11.3 15.6 11.2 13.8 9.0 8.4 8.2 12.0 13.9 11.6 16.0 9.6 11.4 8.4 8.0 14.1 10.9 13.2 13.8 14.6 10.2 11.5 13.1 14.7 12.5 10.2 11.8 11.0 12.7 10.3 10.8 11.0 12.6 10.8 9.6 11.5 10.6 11.7 10.1 9.7 9.7 11.2 9.8 10.3 11.9 9.7 11.3 10.4 12.0 11.0 10.7 8.8 11.1

$X_2$ = soil water content from field II
12.1 10.2 13.6 8.1 13.5 7.8 11.8 7.7 8.1 9.2 14.1 8.9 13.9 7.5 12.6 7.3 14.9 12.2 7.6 8.9 13.9 8.4 13.2 7.3 11.3 7.5 9.7 12.3 6.9 7.6 13.8 7.5 13.3 8.0 11.3 6.8 7.4 11.7 11.8 7.7 12.6 7.7 13.2 13.9 10.4 12.8 7.6 10.7 10.7 10.9 12.5 11.3 10.7 13.2 8.9 12.9 7.7 9.7 9.7 11.4 11.9 13.4 9.2 13.4 8.8 11.9 7.1 8.5 14.0 14.2

06. Rabies (Two variable independent small samples)
The following data represent the number of cases of red fox rabies for a random sample of 16 areas in each of two different regions of southern Germany.

$X_1$ = number cases in region I
10 2 2 5 3 4 3 3 4 0 2 6 4 8 7 4
X2 = number cases in region 2
1  1  2  1  3  9  2  2  4  5  4  2  2  0  0  2

File names
Excel: Tvis06.xls
Minitab: Tvis06.zip
SPSS: Tvissp06.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Tvis06.zip

07. Weight of Football Players Versus Weight of Basketball Players (Two variable independent small samples)
The following data represent weights in pounds of 21 randomly selected pro football players, and 19 randomly selected pro basketball players.
Reference: *Sports Encyclopedia of Pro Football and Official NBA Basketball Encyclopedia*

X1 = weights (lb) of pro football players
245  262  255  251  244  276  240  265  257  252  282
256  250  264  270  275  245  275  253  265  270

X2 = weights (lb) of pro basketball
205  200  220  210  191  215  221  216  228  207
225  208  195  191  207  196  181  193  201

File names
Excel: Tvis07.xls
Minitab: Tvis07.zip
SPSS: Tvissp07.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Tvis07.zip

08. Birth Rate (Two variable independent small samples)
The following data represent birth rate (per 1000 residential population) for independent random samples of counties in California and Maine.
Reference: *County and City Data Book* 12th edition, U.S. Dept. of Commerce

X1 = birth rate in California counties
14.1  18.7  20.4  20.7  16.0  12.5  12.9  9.6  17.6
18.1  14.1  16.6  15.1  18.5  23.6  19.9  19.6  14.9
17.7  17.8  19.1  22.1  15.6

X2 = birth rate in Maine counties
15.1  14.0  13.3  13.8  13.5  14.2  14.7  11.8  13.5  13.8
16.5  13.8  13.2  12.5  14.8  14.1  13.6  13.9  15.8

File names
Excel: Tvis08.xls
Minitab: Tvis08.zip
SPSS: Tvissp08.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Tvis08.zip
09. Death Rate (Two variable independent small samples)
The following data represents death rate (per 1000 resident population) for independent random samples of counties in Alaska and Texas.
Reference: *County and City Data Book* 12th edition, U.S. Dept. of Commerce

\[X_1 = \text{death rate in Alaska counties}\]

<table>
<thead>
<tr>
<th>1.4</th>
<th>4.2</th>
<th>7.3</th>
<th>4.8</th>
<th>3.2</th>
<th>3.4</th>
<th>5.1</th>
<th>5.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.7</td>
<td>3.3</td>
<td>1.9</td>
<td>8.3</td>
<td>3.1</td>
<td>6.0</td>
<td>4.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

\[X_2 = \text{death rate in Texas counties}\]

<table>
<thead>
<tr>
<th>7.2</th>
<th>5.8</th>
<th>10.5</th>
<th>6.6</th>
<th>6.9</th>
<th>9.5</th>
<th>8.6</th>
<th>5.9</th>
<th>9.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>8.8</td>
<td>6.1</td>
<td>9.5</td>
<td>9.6</td>
<td>7.8</td>
<td>10.2</td>
<td>5.6</td>
<td>8.6</td>
</tr>
</tbody>
</table>

File names Excel: Tvis09.xls
Minitab: Tvis09.zip
SPSS: Tvisspss09.zip

10. Pickup Trucks (Two variable independent small samples)
The following data represent the retail price (in thousands of dollars) for independent random samples of models of pickup trucks.
Reference: *Consumer Guide* Vol.681

\[X_1 = \text{prices for different GMC Sierra 1500 models}\]

<table>
<thead>
<tr>
<th>17.4</th>
<th>23.3</th>
<th>29.2</th>
<th>19.2</th>
<th>17.6</th>
<th>19.2</th>
<th>23.6</th>
<th>19.5</th>
<th>22.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.0</td>
<td>26.4</td>
<td>23.7</td>
<td>29.4</td>
<td>23.7</td>
<td>26.7</td>
<td>24.0</td>
<td>24.9</td>
<td></td>
</tr>
</tbody>
</table>

\[X_2 = \text{prices for different Chevrolet Silverado 1500 models}\]

<table>
<thead>
<tr>
<th>17.5</th>
<th>23.7</th>
<th>20.8</th>
<th>22.5</th>
<th>24.3</th>
<th>26.7</th>
<th>24.5</th>
<th>17.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.4</td>
<td>29.7</td>
<td>20.1</td>
<td>21.1</td>
<td>22.1</td>
<td>24.2</td>
<td>27.4</td>
<td>28.1</td>
</tr>
</tbody>
</table>

File names Excel: Tvis10.xls
Minitab: Tvis10.zip
SPSS: Tvisspss10.zip

TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Tvis09.zip
TWO VARIABLE DEPENDENT SAMPLES

File name prefix: Tvds followed by the number of the data file

01. Average Faculty Salary, Males vs Female (Two variable dependent samples)
In following data pairs, A = average salaries for males ($1000/yr) and B = average salaries for females ($1000/yr) for assistant professors at the same college or university. A random sample of 22 US colleges and universities was used.
Reference: Academe, Bulletin of the American Association of University Professors

A: 34.5 30.5 35.1 35.7 31.5 34.4 32.1 30.7 33.7 35.3
B: 33.9 31.2 35.0 34.2 32.4 34.1 32.7 29.9 31.2 35.5

A: 30.7 34.2 39.6 30.5 33.8 31.7 32.8 38.5 40.5 25.3
B: 30.2 34.8 38.7 30.0 33.8 32.4 31.7 38.9 41.2 25.5

A: 28.6 35.8
B: 28.0 35.1

File names
Excel: Tvcds01.xls
Minitab: Tvcds01.zip
SPSS: Tvcdssps01.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Tvcds01.zip

02. Unemployment for College Graduates Versus High School Only
(Two variable dependent samples)
In the following data pairs, A = Percent unemployment for college graduates and B = Percent unemployment for high school only graduates. The data are paired by year.
Reference: Statistical Abstract of the United States

A: 2.8 2.2 2.2 1.7 2.3 2.3 2.4 2.7 3.5 3.0 1.9 2.5
B: 5.9 4.9 4.8 5.4 6.3 6.9 6.9 7.2 10.0 8.5 5.1 6.9

File names
Excel: Tvcds02.xls
Minitab: Tvcds02.zip
SPSS: Tvcdssps02.zip
TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Tvcds02.zip

03. Number of Navajo Hogans versus Modern Houses (Two variable dependent samples)
In the following data pairs, A = Number of traditional Navajo hogans in a given district and B = Number of modern houses in a given district. The data are paired by district of the Navajo reservation. A random sample of 8 districts was used.

A: 13 14 46 32 15 47 17 18
B: 18 16 68 9 11 28 50 50
04. Temperatures in Miami versus Honolulu (Two variable dependent samples)
In the following data pairs, A = Average monthly temperature in Miami and B = Average monthly temperature in Honolulu. The data are paired by month.
Reference: U.S. Department of Commerce Environmental Data Service

A:  67.5  68.0  71.3  74.9  78.0  80.9  82.2  82.7  81.6  77.8  72.3  68.5  
B:  74.4  72.6  73.3  74.7  76.2  78.0  79.1  79.8  79.5  78.4  76.1  73.7

05. January/February Ozone Column (Two variable dependent samples)
In the following pairs, the data represent the thickness of the ozone column in Dobson units: one milli-centimeter ozone at standard temperature and pressure.
   A = monthly mean thickness in January
   B = monthly mean thickness in February
The data are paired by year for a random sample of 15 years.
Reference: Laboratorium für Atmosphärenphysik, Switzerland

A:  360    324    377    336    383    361    369    349
B:  365    325    359    352    397    351    367    397
A:  301    354    344    329    337    387    378
B:  335    338    349    393    370    400    411

06. Birth Rate/Death Rate (Two variable dependent samples)
In the following data pairs, A = birth rate (per 1000 resident population) and B = death rate (per 1000 resident population). The data are paired by county in Iowa
Reference: County and City Data Book, 12th edition, U.S. Dept. of Commerce

A:  12.7   13.4   12.8   12.1   11.6   11.1   14.2
B:   9.8   14.5   10.7   14.2   13.0   12.9   10.9
A:  12.5   12.3   13.1   15.8   10.3   12.7   11.1
B:  14.1   13.6   9.1   10.2   17.9   11.8   7.0
07. Democrat/Republican (Two variable dependent samples)
In the following data pairs A = percentage of voters who voted Democrat and B = percentage of voters who voted Republican in a recent national election. The data are paired by county in Indiana.
Reference: County and City Data Book, 12th edition, U.S. Dept. of Commerce

A: 42.2 34.5 44.0 34.1 41.8 40.7 36.4 43.3 39.5
B: 35.4 45.8 39.4 40.0 39.2 40.2 44.7 37.3 40.8

A: 35.4 44.1 41.0 42.8 40.8 36.4 40.6 37.4
B: 39.3 36.8 35.5 33.2 38.3 47.7 41.1 38.5

08. Santiago Pueblo Pottery (Two variable dependent samples)
In the following data, A = percentage of utility pottery and B = percentage of ceremonial pottery found at the Santiago Pueblo archaeological site. The data are paired by location of discovery.
Reference: Laboratory of Anthropology, Notes 475, Santa Fe, New Mexico

A: 41.4 49.6 55.6 49.5 43.0 54.6 46.8 51.1 43.2 41.4
B: 58.6 50.4 44.4 59.5 57.0 45.4 53.2 48.9 56.8 58.6

09. Poverty Level (Two variable dependent samples)
In the following data pairs, A = percentage of population below poverty level in 1998 and B = percentage of population below poverty level in 1990. The data are grouped by state and District of Columbia.

A: 14.5 9.4 16.6 14.8 15.4 9.2 9.5 10.3 22.3 13.1
B: 19.2 11.4 13.7 19.6 13.9 13.7 6.0 6.9 21.1 14.4

A: 13.6 10.9 13.0 10.1 9.4 9.1 9.6 13.5 19.1 10.4
B: 15.8 11.0 14.9 13.7 13.0 10.4 10.3 17.3 23.6 13.1
### 10. Cost of Living Index (Two variable dependent samples)

The following data pairs represent cost of living index for A = grocery items and B = health care. The data are grouped by metropolitan areas.


<table>
<thead>
<tr>
<th>Grocery</th>
<th>A: 96.6 97.5 113.9 88.9 108.3 99.0 97.3 87.5 96.8</th>
<th>B: 91.6 95.9 114.5 93.6 112.7 93.6 99.2 93.2 105.9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A: 102.1 114.5 100.9 100.0 100.7 99.4 117.1 111.3 102.2</td>
<td>B: 110.8 127.0 91.5 100.5 104.9 104.8 124.1 124.6 109.1</td>
</tr>
<tr>
<td></td>
<td>A: 95.3 91.1 95.7 87.5 91.8 97.9 97.4 102.1 94.0</td>
<td>B: 98.7 95.8 99.7 93.2 100.7 96.0 99.6 98.4 94.0</td>
</tr>
<tr>
<td></td>
<td>A: 115.7 118.3 101.9 88.9 100.7 99.8 101.3 104.8 100.9</td>
<td>B: 121.2 122.4 110.8 81.2 104.8 109.9 103.5 113.6 94.6</td>
</tr>
<tr>
<td></td>
<td>A: 102.7 98.1 105.3 97.2 105.2 108.1 110.5 99.3 99.7</td>
<td>B: 109.8 97.6 109.8 107.4 97.7 124.2 110.9 106.8 94.8</td>
</tr>
</tbody>
</table>

**File names**

- Excel: Tvcds09.xls
- Minitab: Tvcds09.zip
- SPSS: Tvcdsspss09.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Tvcds09.zip

<table>
<thead>
<tr>
<th></th>
<th>A: 7.2 8.7 11.0 10.4 17.6 9.8 16.6 12.3 10.6 9.8</th>
<th>B: 9.9 10.7 14.3 12.0 25.7 13.4 16.3 10.3 9.8 6.3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A: 8.6 20.4 16.7 14.0 15.1 11.2 14.1 15.0 11.2 11.6</td>
<td>B: 9.2 20.9 14.3 13.0 13.7 11.5 15.6 9.2 11.0 7.5</td>
</tr>
<tr>
<td></td>
<td>A: 13.7 10.8 13.4 15.1 9.0 9.9 8.8 8.9 17.8 8.8 10.6</td>
<td>B: 16.2 13.3 16.9 15.9 8.2 10.9 11.1 8.9 18.1 9.3 11.0</td>
</tr>
</tbody>
</table>

**File names**

- Excel: Tvcds10.xls
- Minitab: Tvcds10.zip
- SPSS: Tvcdsspss10.zip
- TI-83 Plus, TI-84 Plus and TI-Nspire/ASCII: Tvcds10.zip
SIMPLE LINEAR REGRESSION

File name prefix: Slr followed by the number of the data file

01. List Price versus Best Price for a New GMC Pickup Truck (Simple Linear Regression)
In the following data, X = List price (in $1000) for a GMC pickup truck and Y = Best price (in $1000) for a GMC pickup truck.
Reference: Consumer’s Digest

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.4</td>
<td>11.2</td>
</tr>
<tr>
<td>14.3</td>
<td>12.5</td>
</tr>
<tr>
<td>14.5</td>
<td>12.7</td>
</tr>
<tr>
<td>14.9</td>
<td>13.1</td>
</tr>
<tr>
<td>16.1</td>
<td>14.1</td>
</tr>
<tr>
<td>16.9</td>
<td>14.8</td>
</tr>
<tr>
<td>16.5</td>
<td>14.4</td>
</tr>
<tr>
<td>15.4</td>
<td>13.4</td>
</tr>
<tr>
<td>17.0</td>
<td>14.9</td>
</tr>
<tr>
<td>17.9</td>
<td>15.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.8</td>
<td>16.4</td>
</tr>
<tr>
<td>20.3</td>
<td>17.7</td>
</tr>
<tr>
<td>22.4</td>
<td>19.6</td>
</tr>
<tr>
<td>19.4</td>
<td>16.9</td>
</tr>
<tr>
<td>15.5</td>
<td>14.0</td>
</tr>
<tr>
<td>16.7</td>
<td>14.6</td>
</tr>
<tr>
<td>17.3</td>
<td>15.1</td>
</tr>
<tr>
<td>18.4</td>
<td>16.1</td>
</tr>
<tr>
<td>19.2</td>
<td>16.8</td>
</tr>
<tr>
<td>17.4</td>
<td>15.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.5</td>
<td>17.0</td>
</tr>
<tr>
<td>19.7</td>
<td>17.2</td>
</tr>
<tr>
<td>21.2</td>
<td>18.6</td>
</tr>
</tbody>
</table>

File names
- Excel: Slr01.xls
- Minitab: Slr01.zip
- SPSS: Slrspss01.zip
- TI-83 Plus, TI-94 Plus and TI-Nspire/ASCII: Slr01.zip

02. Cricket Chirps versus Temperature (Simple Linear Regression)
In the following data, X = chirps/sec for the striped ground cricket and Y = temperature in degrees Fahrenheit.
Reference: The Song of Insects by Dr.G.W. Pierce, Harvard College Press

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>20.0</td>
<td>88.6</td>
</tr>
<tr>
<td>16.0</td>
<td>71.6</td>
</tr>
<tr>
<td>19.8</td>
<td>93.3</td>
</tr>
<tr>
<td>18.4</td>
<td>84.3</td>
</tr>
<tr>
<td>17.1</td>
<td>80.6</td>
</tr>
<tr>
<td>15.5</td>
<td>75.2</td>
</tr>
<tr>
<td>14.7</td>
<td>69.7</td>
</tr>
<tr>
<td>17.1</td>
<td>82.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.4</td>
<td>69.4</td>
</tr>
<tr>
<td>16.2</td>
<td>83.3</td>
</tr>
<tr>
<td>15.0</td>
<td>79.6</td>
</tr>
<tr>
<td>17.2</td>
<td>82.6</td>
</tr>
<tr>
<td>16.0</td>
<td>80.6</td>
</tr>
<tr>
<td>17.0</td>
<td>83.5</td>
</tr>
<tr>
<td>14.4</td>
<td>76.3</td>
</tr>
</tbody>
</table>

File names
- Excel: Slr02.xls
- Minitab: Slr02.zip
- SPSS: Slrspss02.zip
- TI-83 Plus, TI-94 Plus and TI-Nspire/ASCII: Slr02.zip

03. Diameter of Sand Granules versus Slope on Beach (Simple Linear Regression)
In the following data pairs, X = median diameter (mm) of granules of sand and Y = gradient of beach slope in degrees.
The data is for naturally occurring ocean beaches
Reference: Physical geography by A.M King, Oxford Press, England

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.170</td>
<td>0.630</td>
</tr>
<tr>
<td>0.190</td>
<td>0.700</td>
</tr>
<tr>
<td>0.220</td>
<td>0.820</td>
</tr>
<tr>
<td>0.235</td>
<td>0.880</td>
</tr>
<tr>
<td>0.235</td>
<td>1.150</td>
</tr>
<tr>
<td>0.300</td>
<td>1.500</td>
</tr>
<tr>
<td>0.350</td>
<td>4.400</td>
</tr>
<tr>
<td>0.420</td>
<td>7.300</td>
</tr>
<tr>
<td>0.850</td>
<td>11.300</td>
</tr>
</tbody>
</table>
04. National Unemployment Male versus Female (Simple Linear Regression)
In the following data pairs, $X =$ national unemployment rate for adult males and $Y =$ national unemployment rate for adult females.
Reference: *Statistical Abstract of the United States*

<table>
<thead>
<tr>
<th>X</th>
<th>2.9</th>
<th>6.7</th>
<th>4.9</th>
<th>7.9</th>
<th>9.8</th>
<th>6.9</th>
<th>6.1</th>
<th>6.2</th>
<th>6.0</th>
<th>5.1</th>
<th>4.7</th>
<th>4.4</th>
<th>5.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>4.0</td>
<td>7.4</td>
<td>5.0</td>
<td>7.2</td>
<td>7.9</td>
<td>6.1</td>
<td>6.0</td>
<td>5.8</td>
<td>5.2</td>
<td>4.2</td>
<td>4.0</td>
<td>4.4</td>
<td>5.2</td>
</tr>
</tbody>
</table>

05. Fire and Theft in Chicago (Simple Linear Regression)
In the following data pairs, $X =$ fires per 1000 housing units and $Y =$ thefts per 1000 population within the same zip code in the Chicago metro area.
Reference: U.S. Commission on Civil Rights

<table>
<thead>
<tr>
<th>X</th>
<th>6.2</th>
<th>9.5</th>
<th>10.5</th>
<th>7.7</th>
<th>8.6</th>
<th>34.1</th>
<th>11.0</th>
<th>6.9</th>
<th>7.3</th>
<th>15.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>29</td>
<td>44</td>
<td>36</td>
<td>37</td>
<td>53</td>
<td>68</td>
<td>75</td>
<td>18</td>
<td>31</td>
<td>25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>29.1</th>
<th>2.2</th>
<th>5.7</th>
<th>2.0</th>
<th>2.5</th>
<th>4.0</th>
<th>5.4</th>
<th>2.2</th>
<th>7.2</th>
<th>15.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>34</td>
<td>14</td>
<td>11</td>
<td>11</td>
<td>22</td>
<td>16</td>
<td>27</td>
<td>9</td>
<td>29</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>16.5</th>
<th>18.4</th>
<th>36.2</th>
<th>39.7</th>
<th>18.5</th>
<th>23.3</th>
<th>12.2</th>
<th>5.6</th>
<th>21.8</th>
<th>21.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>40</td>
<td>32</td>
<td>41</td>
<td>147</td>
<td>22</td>
<td>29</td>
<td>46</td>
<td>23</td>
<td>4</td>
<td>31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>9.0</th>
<th>3.6</th>
<th>5.0</th>
<th>28.6</th>
<th>17.4</th>
<th>11.3</th>
<th>3.4</th>
<th>11.9</th>
<th>10.5</th>
<th>10.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>39</td>
<td>15</td>
<td>32</td>
<td>27</td>
<td>32</td>
<td>34</td>
<td>17</td>
<td>46</td>
<td>42</td>
<td>43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>10.8</th>
<th>4.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>34</td>
<td>19</td>
</tr>
</tbody>
</table>
06. Auto Insurance in Sweden (Simple Linear Regression)
In the following data, X = number of claims and Y = total payment for all the claims in thousands of Swedish Kronor for geographical zones in Sweden.
Reference: Swedish Committee on Analysis of Risk Premium in Motor Insurance

X: 108 19 13 124 40 57 23 14 45 10
Y: 392.5 46.2 15.7 422.2 119.4 170.9 56.9 77.5 214.0 65.3

X: 5 48 11 23 7 2 24 6 3 23
Y: 20.9 248.1 23.5 39.6 48.8 6.6 134.9 50.9 4.4 113.0

X: 6 9 9 3 29 7 4 20 7 4
Y: 14.8 48.7 52.1 13.2 103.9 77.5 11.8 98.1 27.9 38.1

X: 0 25 6 5 22 11 61 12 4 16
Y: 0.0 69.2 14.6 40.3 161.5 57.2 217.6 58.1 12.6 59.6

X: 13 60 41 37 55 41 11 27 8 3
Y: 89.9 202.4 181.3 152.8 162.8 73.4 21.3 92.6 76.1 39.9

X: 17 13 13 15 8 29 30 24 9 31
Y: 142.1 93.0 31.9 32.1 55.6 133.3 194.5 137.9 87.4 209.8

X: 14 53 26
Y: 95.5 244.6 187.5

File names
Excel: Slr06.xls
Minitab: Slr06.zip
SPSS: Slrsspss06.zip
TI-83 Plus, TI-94 Plus and TI-Nspire/ASCII: Slr06.zip

07. Gray Kangaroos (Simple Linear Regression)
In the following data pairs, X = nasal length (mm × 10) and Y = nasal width (mm × 10) for a male gray kangaroo from a random sample of such animals.

X: 609 629 620 564 645 493 606 660 630 672
Y: 241 222 233 207 247 189 226 240 215 231

X: 778 616 727 810 778 823 755 710 701 803
Y: 263 220 271 284 279 272 268 278 238 255

X: 855 838 830 864 635 565 562 580 596 597
Y: 308 281 288 306 236 204 216 225 220 219
08. Pressure and Weight in Cryogenic Flow Meters (Simple Linear Regression)

In the following data pairs, $X =$ pressure (lb/sq in) of liquid nitrogen and $Y =$ weight in pounds of liquid nitrogen passing through flow meter each second.


<table>
<thead>
<tr>
<th>X</th>
<th>75.1</th>
<th>74.3</th>
<th>88.7</th>
<th>114.6</th>
<th>98.5</th>
<th>114.8</th>
<th>62.2</th>
<th>107.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>577.8</td>
<td>577.0</td>
<td>570.9</td>
<td>578.6</td>
<td>572.4</td>
<td>411.2</td>
<td>531.7</td>
<td>563.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>90.5</th>
<th>73.8</th>
<th>115.8</th>
<th>99.4</th>
<th>93.0</th>
<th>73.9</th>
<th>65.7</th>
<th>66.2</th>
<th>77.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>507.1</td>
<td>496.4</td>
<td>505.2</td>
<td>506.4</td>
<td>510.2</td>
<td>503.9</td>
<td>506.2</td>
<td>506.3</td>
<td>510.2</td>
</tr>
</tbody>
</table>

| X     | 109.8| 105.4| 88.6 | 73.8  | 101.3| 120.0| 75.9 | 76.2  |
|-------|------|------|------|-------|------|-------|------|-------|------|
| Y     | 508.6| 510.9| 505.4| 512.8 | 502.8| 493.0| 510.8| 512.8 | 513.4|

File names
- Excel: Slr08.xls
- Minitab: Slr08.zip
- SPSS: Slrsps08.zip
- TI-83 Plus, TI-94 Plus and TI-Nspire/ASCII: Slr08.zip

09. Ground Water Survey (Simple Linear Regression)

In the following data, $X =$ pH of well water and $Y =$ Bicarbonate (parts per million) of well water.

The data is by water well from a random sample of wells in Northwest Texas.


<table>
<thead>
<tr>
<th>X</th>
<th>7.6</th>
<th>7.1</th>
<th>8.2</th>
<th>7.5</th>
<th>7.4</th>
<th>7.8</th>
<th>7.3</th>
<th>8.0</th>
<th>7.1</th>
<th>7.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>157</td>
<td>174</td>
<td>175</td>
<td>188</td>
<td>171</td>
<td>143</td>
<td>217</td>
<td>190</td>
<td>142</td>
<td>190</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>8.1</th>
<th>7.0</th>
<th>7.3</th>
<th>7.8</th>
<th>7.3</th>
<th>8.0</th>
<th>8.5</th>
<th>7.1</th>
<th>8.2</th>
<th>7.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>215</td>
<td>199</td>
<td>262</td>
<td>105</td>
<td>121</td>
<td>81</td>
<td>82</td>
<td>210</td>
<td>202</td>
<td>155</td>
</tr>
</tbody>
</table>
10. Iris Setosa (Simple Linear Regression)

In the following data, X = sepal width (cm) and Y = sepal length (cm). The data is for a random sample of the wild flower iris setosa.


\[
\begin{array}{cccccccccc}
X: & 3.5 & 3.0 & 3.2 & 3.1 & 3.6 & 3.9 & 3.4 & 3.4 & 2.9 & 3.1 \\
Y: & 5.1 & 4.9 & 4.7 & 4.6 & 5.0 & 5.4 & 4.6 & 5.0 & 4.4 & 4.9 \\
\end{array}
\]

\[
\begin{array}{cccccccccc}
X: & 3.7 & 3.4 & 3.0 & 4.0 & 4.4 & 3.9 & 3.5 & 3.8 & 3.8 & 3.4 \\
Y: & 5.4 & 4.8 & 4.3 & 5.8 & 5.7 & 5.4 & 5.1 & 5.7 & 5.1 & 5.4 \\
\end{array}
\]

\[
\begin{array}{cccccccccc}
X: & 3.7 & 3.6 & 3.3 & 3.4 & 3.0 & 3.4 & 3.5 & 3.4 & 3.2 & 3.1 \\
Y: & 5.1 & 4.6 & 5.1 & 4.8 & 5.0 & 5.0 & 5.2 & 5.2 & 4.7 & 4.8 \\
\end{array}
\]

\[
\begin{array}{cccccccccc}
X: & 3.4 & 4.1 & 4.2 & 3.1 & 3.2 & 3.5 & 3.6 & 3.0 & 3.4 & 3.5 \\
Y: & 5.4 & 5.2 & 5.5 & 4.9 & 5.0 & 5.5 & 4.9 & 4.4 & 5.1 & 5.0 \\
\end{array}
\]

\[
\begin{array}{cccccccccc}
X: & 2.3 & 3.2 & 3.5 & 3.8 & 3.0 & 3.8 & 3.7 & 3.3 \\
Y: & 4.5 & 4.4 & 5.0 & 5.1 & 4.8 & 4.6 & 5.3 & 5.0 \\
\end{array}
\]

File names
- Excel: Slr09.xls
- Minitab: Slr09.zip
- SPSS: Sltspss09.zip
- TI-83 Plus, TI-94 Plus and TI-Nspire/ASCII: Slr09.zip

11. Pizza Franchise (Simple Linear Regression)

In the following data, X = annual franchise fee ($1000) and Y = start up cost ($1000) for a pizza franchise.

Reference: *Business Opportunity Handbook*

\[
\begin{array}{ccccccccccccc}
X: & 25.0 & 8.5 & 35.0 & 15.0 & 10.0 & 30.0 & 10.0 & 50.0 & 17.5 & 16.0 \\
Y: & 125 & 80 & 330 & 58 & 110 & 338 & 30 & 175 & 120 & 135 \\
\end{array}
\]
X:  18.5  7.0  8.0  15.0  5.0  15.0  12.0  15.0  28.0  20.0  20.0  15.0  25.0  20.0  25.0  35.0  25.0  3.5  8.5  10.0  5.0  15.0  12.0  15.0  28.0  20.0
Y:  97   50   55   40   35   45   75   33   55   90   20.0  15.0  20.0  25.0  20.0  3.5  35.0  25.0  8.5  10.0

X:  20.0  15.0  20.0  25.0  20.0  3.5  35.0  25.0  8.5  10.0
Y:  85   125  150  120  95   30   400  148  135  45

X:  10.0  25.0
Y:  87   150

File names
Excel: Slr11.xls
Minitab: Slr11.zip
SPSS: Slrspss11.zip
TI-83 Plus, TI-94 Plus and TI-Nspire/ASCII: Slr11.zip

12. Prehistoric Pueblos (Simple Linear Regression)
In the following data, X = estimated year of initial occupation and Y = estimated year of end of occupation. The data are for each prehistoric pueblo in a random sample of such pueblos in Utah, Arizona, and Nevada.
Reference: Prehistoric Pueblo World, by A. Adler, Univ. of Arizona Press

X:  1000  1125  1087  1070  1100  1150  1250  1150  1100
Y:  1050  1150  1213  1275  1300  1300  1400  1400  1250

X:  1350  1275  1375  1175  1200  1175  1300  1260  1330
Y:  1830  1350  1450  1300  1300  1275  1375  1285  1400

X:  1325  1200  1225  1090  1075  1080  1080  1180  1225
Y:  1400  1285  1275  1135  1250  1275  1150  1250  1275

X:  1175  1250  1250  750  1125  700  900  900  850
Y:  1225  1280  1300  1250  1175  1300  1250  1300  1200

File names
Excel: Slr12.xls
Minitab: Slr12.zip
SPSS: Slrspss12.zip
TI-83 Plus, TI-94 Plus and TI-Nspire/ASCII: Slr12.zip