Instat Tutorial

by Roger Stern, Sandro Leidi, Ian Dale and Colin Grayer

Contents

Installation 2

1. Introduction 3
   1.1. Tutorial purpose and overview 3
   1.2. Getting Started 3
2. Example 1: Basic Use of Instat 3
   2.1. Entering data 3
   2.2. Describing the data using summary statistics 5
   2.3. Plotting the data 7
   2.4. Entering sequenced data 9
   2.5. Calculating new variates 10
   2.6. Closing, saving and (re)opening worksheets 12
3. Example 1 (continued) 14
   3.1. Plotting lines 14
   3.2. Specifying groups using factors 16
   3.3. Deriving statistics for columns 18
   3.4. Review of example 1 19
4. Example 2: Boxes, Models and Data Manipulation 20
   4.1. Getting a feel for the data 20
   4.2. Comparison of means 22
   4.3. Manipulating data 23
   4.4. Understanding how Instat works 25
   4.5. Instat as a calculator 27
   4.6. Review of example 2 28
5. Example 3: A Simple Regression Model 29
   5.1. Simple Regression 29
   5.2. Help in Instat 33
   5.3. Leaving Instat 34
6. Review 34

Appendix: Exporting and Importing Data 36

© Statistical Services Centre
The University of Reading, UK
Installation

If Instat has not already been installed on your computer, follow the instructions on Instat’s download page on the web, or in the ‘ReadMe’ file on the CD, to install it onto your PC’s hard disk. The ‘typical’ installation includes all the documentation and example worksheet files.
1. **Introduction**

1.1. **Tutorial purpose and overview**

The aim of this tutorial is to help you to become familiar with the basics of using Instat, rather than to help you learn statistics. We first give you an overview of Instat's features, then lead you on a guided tour of an example analysis that uses descriptive statistics. This will take about half the time needed to follow the whole tutorial, and brings you to the end of Section 3.

Once you have gained some idea of what Instat can offer, we present an example that requires you to find your own way around the menus and dialogues to carry out an analysis – though rest assured, we do give plenty of hints and guidance in the text. This second example is in Section 4.

Section 5 builds on this, using an example of an analysis that involves simple regression. Further information on Instat's features and on statistical methods can be found in the documentation and help system: these also are described in this section.

Section 6 is a review of the topics covered, and the tutorial finishes with an appendix about import and export of data.

1.2. **Getting Started**

Instat is a Windows package for PCs: we assume in this tutorial that you have some experience of working in a Windows environment and of using a mouse.

If you are a beginner to statistical packages, it is important that you start by getting some practice in using Instat. You will not gain very much from just reading the tutorial without practicing.

If you have used another statistical package before, you can learn a lot by simply reading through the tutorial. You should find that trying out the examples is a fairly quick exercise, and that much of Instat will seem familiar, because all of the common statistical packages for Windows operate in a similar way.

Actions that you should take to follow the examples are given in the lines that start with ⇒.

⇒ To launch Instat,\(^1\) double-click on the **Instat** icon or follow the menu sequence **Start** ⇒ **Programs** ⇒ **Instat**.

The main components of the Instat window are shown in Fig. 1.2.1.

Other windows showing graphs, macros etc., open as you use Instat. The position of the windows can be set from the **Window** menu and saved for future Instat sessions using **Window** ⇒ **Save Window Positions**. Hence the initial appearance of Instat may not be quite as that in Fig. 1.2.1.

2. **Example 1: Basic Use of Instat**

2.1. **Entering data**

Instat opens with an empty, untitled worksheet in the **Current Worksheet** window.

⇒ Start entering data by clicking in the cell of row 1 of column X1 in the worksheet. Type a number, and then press the **Enter** key.

---

\(^1\) The early Windows versions of Instat were labelled ‘Instat+’ (i.e. ‘Instat Plus’) to make the distinction clear between Instat for DOS and Instat for Windows, and ‘Instat Plus’ is in fact the registered trademark. However, for convenience throughout the documentation, we now simply use ‘Instat’ to refer to the Windows version.
Fig. 1.2.1. Main components of the Instat windows interface

⇒ Enter the following numbers into the column headed X1:
   1330, 2094, 1851, 1470, 1557, 1932, 1184, 2452, 1347, 1792, 1488
⇒ Press [Ctrl]+[Enter] after typing the last number. The cursor will then move to the top of the next column.
⇒ Enter these numbers into the second column, X2:
   120, 161, 150, 142, 152, 155, 126, 183, 132, 167, 149

Fig. 2.1.1. Enter the data
If you have made any mistakes, these can be corrected: use the arrow keys to move the insertion point to the cell in question and enter the correct value. The values in X1 are the total annual rainfalls from a sequence of 11 years, and the values in X2 are the corresponding numbers of rainy days (Fig. 2.1.1).

⇒ Name the first column by highlighting the cell underneath X1 (the header line) and then typing in ‘total’. Similarly name the column X2 as ‘raindays’.

⇒ The entered data are written to a temporary file automatically, as soon as you click anywhere outside the worksheet window. To ensure that this happens, click in the Commands and Output window (Fig. 1.2.1). If you cannot see this window, press [F2] or use Window ⇒ Tile Vertical to tile the data and output windows.

2.2. Describing the data using summary statistics

For statistical summaries use the Statistics menu, shown in Fig. 2.2.1.

⇒ Choose Statistics ⇒ Summary ⇒ Describe

Fig. 2.2.1. The Summary menu

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Climatic</th>
<th>Window</th>
<th>Help</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple Models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis of Variance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Parametric</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time Series</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survival</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability Distributions...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample Size...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

⇒ The dialogue shown in Fig. 2.2.2 appears. Select the variates required, either by double-clicking on each in turn, or by marking both and clicking the button on the dialogue. Then click [OK].
Fig. 2.2.2. The Descriptive Statistics dialogue

By default, the Descriptive Statistics dialogue gives number of observations, minimum, maximum, range, mean and standard deviation. More statistics are available from the same dialogue.

Fig. 2.2.3. Statistics for raindays data

⇒ Select the Commands and Output window. The results from the Descriptive Statistics dialogue are shown in Fig. 2.2.3.

⇒ Find the Descriptive Statistics dialogue again and click in the Additional statistics box. Select the Median and Quartiles check boxes (Fig. 2.2.4) and click [OK].
Fig. 2.2.4. Additional descriptive statistics

From the results that appear in the **Commands and Output** window, complete the table below.

<table>
<thead>
<tr>
<th>Median</th>
<th>Lower Quartile</th>
<th>Upper Quartile</th>
<th>Inter quartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.3. Plotting the data

A range of graphs is possible from the **Graphics** menu. Graphs use any column names that have been entered (such as *total* and *raindays*) for presentation; otherwise they use X1, X2 etc.

⇒ Check that the worksheet looks similar to that shown in Fig. 2.3.1.

**Fig. 2.3.1. Worksheet showing named columns**

**Fig. 2.3.2. Graphics Menu**

⇒ Use **Graphics** ⇒ **Plot** (Fig. 2.3.2).

⇒ Complete the dialogue, as shown in Fig. 2.3.3, then click [OK].
Fig. 2.3.3. The Plot dialogue

The button on the tool bar recalls the last dialogue box used.
⇒ Click this button to recall the Plot dialogue box.
⇒ Click on the Titles tab, and add a meaningful title to your graph (see Fig. 2.3.4). The other tabs on this dialogue allow adjustments to the x- and y-axis limits, line/symbol settings, etc.
⇒ Click [OK] to produce a plot similar to that in Fig. 2.3.5.

Fig. 2.3.4. The Plot dialogue’s Titles tab
2.4. Entering sequenced data

These data values came from a sequence of 11 years. It would be useful to have this information entered too. Instat provides a flexible means of entering regular sequences of data.

⇒ Select Manage ⇒ Data ⇒ Regular Sequence (Fig. 2.4.1).

Fig. 2.4.1. The Manage menu
Fig. 2.4.2. Entering a regular sequence

- Complete the dialogue as shown in Fig. 2.4.2, to enter the numbers 1-11 into column X3. To see what the sequence will look like, click [Preview]. Click [OK] to generate the values.

- Enter yearnum into the header line as the name for column X3 (Fig. 2.4.3).

Fig. 2.4.3. Data in yearnum

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>X2</td>
<td>X3</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>total</td>
<td>raindays</td>
<td>yearnum</td>
</tr>
<tr>
<td>1</td>
<td>1330</td>
<td>120</td>
</tr>
<tr>
<td>2</td>
<td>2094</td>
<td>161</td>
</tr>
<tr>
<td>3</td>
<td>1851</td>
<td>150</td>
</tr>
<tr>
<td>4</td>
<td>1470</td>
<td>142</td>
</tr>
<tr>
<td>5</td>
<td>1557</td>
<td>152</td>
</tr>
<tr>
<td>6</td>
<td>1332</td>
<td>155</td>
</tr>
<tr>
<td>7</td>
<td>1184</td>
<td>126</td>
</tr>
<tr>
<td>8</td>
<td>2452</td>
<td>133</td>
</tr>
<tr>
<td>9</td>
<td>1347</td>
<td>132</td>
</tr>
<tr>
<td>10</td>
<td>1792</td>
<td>157</td>
</tr>
<tr>
<td>11</td>
<td>1489</td>
<td>149</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.5. Calculating new variates

It is easy to calculate new variates from those already in a worksheet. To demonstrate this, let's find the mean rainfall per rainy day.

- To calculate a new column, choose Manage ➔ Calculations (Fig. 2.5.1).

- Complete the Calculate dialogue as shown in Fig. 2.5.2. The calculation can either be typed into the top box, or you can use the mouse to click on the operator buttons and double-click on the variates as required.

- To save the results in a new column, click the check box Save result and select X4 (the next empty column) from the pull down list. Click [OK].

- A new variate will be added to the spreadsheet in column X4. Name X4 as meanrain.
Instat Tutorial

Fig. 2.5.1. Manage menu

Manage  Graphics  Statistics
Calculations...
Data  
Manipulate  
Reshape  
Column Properties...
Resize Worksheet...  Worksheet Information...

Fig. 2.5.2. Calculate dialogue

Fig. 2.5.3. Calculation

<table>
<thead>
<tr>
<th>X4</th>
<th>meanrain</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.08333</td>
<td></td>
</tr>
<tr>
<td>13.00621</td>
<td></td>
</tr>
<tr>
<td>12.34</td>
<td></td>
</tr>
<tr>
<td>10.35211</td>
<td></td>
</tr>
<tr>
<td>10.24342</td>
<td></td>
</tr>
<tr>
<td>12.45452</td>
<td></td>
</tr>
<tr>
<td>9.396826</td>
<td></td>
</tr>
<tr>
<td>13.33691</td>
<td></td>
</tr>
<tr>
<td>10.20455</td>
<td></td>
</tr>
<tr>
<td>10.73054</td>
<td></td>
</tr>
<tr>
<td>9.96577</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2.5.4. Format the data

The column is displayed with the same number of significant digits for all the values, which can lead to a variable number of decimal places. Change this by:

⇒ Manage ⇒ Column Properties ⇒ Format (Fig. 2.5.4).

Fig. 2.5.3 shows column X4 (meanrain) in the worksheet, which now holds the mean rainfall per rainy day for each of the 11 years. The column is displayed with the same number of significant digits for all the values, which can lead to a variable number of decimal places. Change this by:

⇒ As shown in Fig. 2.5.5, select the variate X4 as the column to be formatted and select the format 0.0 to have 1 decimal place displayed.

⇒ Click [OK] to effect the change. The resulting formatted data are shown in Fig. 2.5.6.
2.6. **Closing, saving and (re)opening worksheets**

⇒ Close the current worksheet. This can be done by clicking on the \( \times \) button in the upper right corner of the worksheet, or by selecting **File** → **Close Worksheet** from the menu bar (Fig. 2.6.1).

The worksheet you have been using up to this point is stored in a temporary file named **UNTITLED.WOR**. When you try to close it, Instat prompts you to save any changes to the worksheet (Fig. 2.6.2) and – if you choose ‘Yes’ – give the file a more suitable, permanent name.

⇒ Choose \[Yes\] (Fig. 2.6.2).

The **Save Worksheet As** dialogue appears (Fig. 2.6.3).

⇒ Type `raindays.wor` as the new name for the file, and click [**OK**].
To open a recently used worksheet, for example, raindays.wor, either select it from the list near the bottom of the File menu, as shown in Fig. 2.6.4, or use File ⇒ Open Worksheet and select the file name raindays.wor from the list displayed, as shown in Fig. 2.6.5.
3. Example 1 (continued)

3.1. Plotting lines

Try plotting the mean daily rainfall against the year number.

⇒ Use Graphics ⇒ Plot with meanrain as the Y-Variable and yearnum as the X-Variable, as shown in Fig. 3.1.1.

![Fig. 3.1.1. The Plot dialogue](image)

⇒ Click on the Titles tab to add an appropriate title (see Fig. 3.1.2).

![Fig. 3.1.2. The Plot dialogue’s Titles tab](image)

⇒ To get a line plot, tick the Show lines checkbox. To specify the type of lines you want, click on the Styles and Colours tab and complete the dialogue as shown in Fig. 3.1.3.
Fig. 3.1.3. Styles and colours for lines and symbols

⇒ Click [OK] to display the plot.

Complete the following:

Which year has highest mean rain per day?

Which year has the lowest mean rain per day?

Is there any obvious pattern?

⇒ Try some more plots, using different styles, colours and symbols. For example, plot the raindays variate, X2, against the yearnum variate. Our version of the resulting plot is shown in Fig. 3.1.4.

Fig. 3.1.4. Plot of raindays against yearnum
3.2. Specifying groups using factors

Four years in this dataset were ‘El Niño’ years: the second, third, eighth and tenth. The others were ordinary years. To specify these two groups, you need to define a factor with two levels, labelled ‘E’ for El Niño and ‘O’ for Ordinary (you can type a longer label of up to 8 letters if you wish).

⇒ Click in row 1 of column X5 and enter the following data: 1 2 1 1 1 1 2 1 2 1
⇒ Type type as the column name (Fig. 3.2.1).
⇒ Select Manage ⇒ Column Properties ⇒ Factor and in the resulting dialogue (Fig. 3.2.2) choose X5 (type) as the Data column.
⇒ Tick the option to Attach to new label column, then edit the labels ‘Level1’ and ‘Level2’ to be ‘O’ and ‘E’ respectively.
⇒ Click [OK].

As can be seen in Fig. 3.2.3, the Make or modify a factor column dialogue has associated the value 1 in X5 with the label ‘O’ and the value 2 with the label ‘E’. It also designates this column as a factor and shows this in the worksheet column header by ‘X5 - F’.

For simplicity, we have used the letters ‘O’ and ‘E’ as the labels. Up to 8 characters are allowed, so ‘Ordinary’ and ‘El Niño’ could be typed in full instead.

⇒ If you look in the worksheet and press the Labels tab (Fig. 3.2.4), you will see how Instat stores the labels for factors (Fig. 3.2.5).
The type column can be used to label a graph.

⇒ Choose **Graphics ⇒ Plot**. Select the Y-Variable and X-Variable as shown, tick the **By factor** checkbox, and select X5 from the pull-down list (Fig. 3.2.6).

**Fig. 3.2.6. Plot dialogue**

![Plot dialogue](image)

⇒ Click on the **Styles and Colours** tab, then set the **Symbol Type** to ‘Star’, and the **Line Type** to ‘None’ (Fig. 3.2.7). Use the **Titles** tab to set a suitable title for the plot.

**Fig. 3.2.7. Specify Title, Line and Symbol Settings**

![Specify Title, Line and Symbol Settings](image)

⇒ Click [**OK**] to generate the plot, which should look similar to Fig. 3.2.8. The points from the two groups are displayed using different symbols, indicated in the legend on the right-hand side of the plot.
3.3. Deriving statistics for columns

Earlier, you used **Statistics ⇒ Summary ⇒ Describe** to give some summaries of the data. Now, with the data in two groups, it is useful to give the summaries for each group individually.

⇒ **Statistics ⇒ Summary ⇒ Column Statistics**

⇒ Complete the dialogue as shown in Fig. 3.3.1 and press [OK]. The results appear in the **Commands and Output** window, and are shown in Fig. 3.3.2.
Fig. 3.3.2. Output for level 1 (O)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Min.</th>
<th>Max.</th>
<th>SDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>1473</td>
<td>1184</td>
<td>1932</td>
<td>237.3</td>
</tr>
<tr>
<td>raindays</td>
<td>139.4</td>
<td>120</td>
<td>155</td>
<td>13.61</td>
</tr>
<tr>
<td>meanrain</td>
<td>10.53</td>
<td>9.397</td>
<td>12.46</td>
<td>0.9872</td>
</tr>
</tbody>
</table>

⇒ To close the worksheet, click on the [X] in the top right corner of the Worksheet window.
If you have made any changes to the data in the worksheet, Instat first prompts you to save them – they will be saved by default to raindays.wor. The Worksheet window then closes.
The results that appear in the Commands and Output window can also be saved if desired:
⇒ Click within the Commands and Output window, then select File ⇒ Save to bring up the Save Commands and Output dialogue (similar to Fig. 2.6.3).
⇒ Type a suitable filename for the output (e.g. raindays.out) and click on [Save].
This is the end of the first example. Leave Instat open, ready for the next exercise.

3.4. Review of example 1

You have just carried out a simple descriptive analysis, using a range of features accessed through Instat’s menus. Check that you feel confident about finding the menu options needed for these various tasks:
• Create, open, save and close worksheets; save and close output files (File menu).
• Enter regular sequences; calculate new columns; format columns; define columns as factors (Manage menu).
• Derive summaries (Statistics menu).
• Produce scatterplots, line plots, and plots that distinguish between groups (Graphics menu).

Note: when a plot is the active widow, Instat’s menu and toolbar change to provide various plot editing facilities. The plot menu and toolbar are shown in Fig. 3.5.1. If you feel like a challenge, you could explore these facilities, for instance by altering the axes or adding a legend to a plot, then exporting it in a form that can be included in a report.

Fig. 3.5.1. Revised menu and toolbar that are displayed when a plot is the active window

If you have worked through the tutorial to this point without a pause, and are happy about finding your way around Instat using the menus, now is a good time to take a break!
4. Example 2: Boxes, Models and Data Manipulation

In the tutorial so far, we have used only descriptive statistics, to summarise the data numerically and to draw some graphs. Some of Instat’s features used in this second example bring in ideas of simple statistical models.

4.1. Getting a feel for the data

This example comes from Mead, Curnow and Hasted (2002), pages 33-34 and 38-39. This compares 6 observations from a new variety of wheat with 10 observations from the standard variety. The yields (in tonnes per hectare) are as follows:

New: 2.5 2.1 2.4 2 2.6 2.2
Standard: 2.2 1.9 1.8 2.1 2.1 1.7 2.3 2.0 1.7 2.2

⇒ First create a new worksheet using File ⇒ New Worksheet
⇒ Enter the data, name the columns and give the worksheet a title, as shown in Fig. 4.1.1.

![Fig. 4.1.1. Wheat data set](image)

⇒ Save the worksheet, giving it the name wheat.wor.

An informative way to illustrate the main aspects of the data is to use a boxplot.
⇒ Find the Boxplot menu option and complete the dialogue, as shown in Fig. 4.1.2.

This gives the display shown in Fig. 4.1.3.

An additional use of boxplots is to show outliers. Fig. 4.1.4 shows (horizontal) boxplots of the wheat data, after editing the 8th value in the Standard group from 2.0 to 2.9. The general shape of the plot is the same, but the ‘odd’ large value stands out and therefore deserves closer scrutiny.
⇒ Change the value in the 8th row of X2 from 2 to 2.9 (Fig. 4.1.2), then refer to the Boxplot dialogue options and try to reproduce the plot in Fig. 4.1.4.
⇒ Once you have succeeded and feel confident about using boxplots, set the edited value back to 2.0 in the worksheet.

---

Fig. 4.1.2. Boxplot dialogue

Fig. 4.1.3. Boxplots for New and Standard wheat yields (t/ha)

Fig. 4.1.4. Horizontal boxplots for the two wheat varieties, highlighting an outlier
4.2. Comparison of means

⇒ A comparison of the means of two different samples can be made using the menu option Statistics ⇒ Simple Models ⇒ Normal, Two Samples (Fig. 4.2.1).

Fig. 4.2.1. Simple Models menu

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Climatic</th>
<th>Window</th>
<th>Help</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple Models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis of Variance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Parametric</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time Series</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survival</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability Distributions...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample Size...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal, One Sample...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion, One Sample...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poisson, One Sample...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal, Two Samples...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportions, Two Samples...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poisson, Two Samples...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extremes...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-square Test...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goodness of Fit...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

⇒ Complete the dialogue as in Fig. 4.2.2, ticking both the Significance Test and Plot boxes.

Fig. 4.2.2. Two sample confidence interval and t-test

The resulting output is shown in Fig. 4.2.3 and the plot in Fig. 4.2.4. The 'confidence interval' and 'null hypothesis' are indicated in the plot, to highlight issues that could be discussed when teaching statistics. In this tutorial the main concern is to show how to use Instat to produce results.
Fig. 4.2.3. Comparison of means

<table>
<thead>
<tr>
<th>Column</th>
<th>New</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Minimum</td>
<td>2</td>
<td>1.7</td>
</tr>
<tr>
<td>Maximum</td>
<td>2.6</td>
<td>2.3</td>
</tr>
<tr>
<td>Range</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Mean</td>
<td>2.3</td>
<td>2</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>0.23664</td>
<td>0.21602</td>
</tr>
</tbody>
</table>

Pooled standard deviation = 0.223607

Difference between means = 0.3 s.e. of difference = 0.11547 with 14 d.f.

95% confidence interval for the difference between means

0.05234 to 0.54766

$t$ value testing mean difference=0 is 2.60

Significance level is 2.11% for 2 sided test

Fig. 4.2.4. Plot from two sample $t$-test

4.3. Manipulating data

Data often need reorganising before analysis. Many of Instat's data reorganisation facilities are to be found on the Manage menu.
⇒ Examine the various options offered by the Manage menu.

In the two-sample example that was used for the \( t \)-test, the data were put into two separate columns, \textit{New} and \textit{Standard}. We now wish to append the data from the two columns into \textit{X3} and add a factor column, \textit{X4}, to indicate from which set each observation has come.

⇒ Find the Stack option from the menu, and complete the dialogue as shown in Fig. 4.3.1. Note that ‘stacking’, or appending one column onto the end of another, is an example of reshaping the worksheet.

⇒ Give the name \textit{Yield} to the new column in \textit{X3}.

The stacked column, \textit{X3}, and the corresponding factor column, \textit{X4}, are shown in Fig. 4.3.2.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{stack_columns.png}
\caption{Stack columns}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{stacked_data.png}
\caption{Stacked data}
\end{figure}

⇒ Return to the \( t \)-test dialogue. Change the dialogue options to those shown in Fig. 4.3.3 and click [\text{OK}].

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{t_testDialogue.png}
\caption{t-test dialogue}
\end{figure}
Fig. 4.3.3. Analysis for a single column with two factor levels

You should find that the output and the plot show essentially the same results as the method using two columns in Section 4.2.

The stacked form of the data, with a factor column indicating the groups, is more common. This ‘list’ format, as it is called, is the standard way of arranging data for statistical work.

4.4. Understanding how Instat works

Before the Windows version of Instat was developed, you could use it only if you knew its ‘command language’. You typed commands that you ‘submitted’ to perform an action. Instat may still be used in this way, as shown next. Instat may also be used as a calculator. The Windows version of Instat gives you several ways of preparing instructions for submission to the Instat server (i.e. the program at the heart of Instat), as follows.

1. Type commands directly into the **Commands and Output** window

2. Open an **Editing** window, type commands, and submit them (e.g. using the toolbar button)

3. Use the menus and dialogues, as described in previous sections.

Here these three alternative ways of working are illustrated by performing a simple calculation.

**Method 1**

⇒ Click in the **Commands and Output** window. At the ‘?’ prompt, type `3 + 4` and press `<Enter>`. The answer 7 is displayed, as shown in Fig. 4.4.1.

**N.B.** If the **Commands and Output** window is not visible (e.g. it may be hidden by other windows), you can make it appear by selecting it from the list near the bottom of the **Window** menu.

**Method 2**

⇒ Use **Edit** ⇒ **Edit Text** ⇒ **New** (Fig. 4.4.2). This gives you an **Editing** window.
In this window, type \(3 + 4\) as shown Fig. 4.4.3.

Click on the Submit to Instat button (Fig. 4.4.4).

You have now submitted your ‘program’ of commands — a single line of instructions — to the Instat server. Instat normally echoes the command in the Commands and Output window, and the results appear exactly as before (as in Fig. 4.4.1). Note that this answer is only displayed; it is not saved in the worksheet.

Method 3

The third method is to use Manage \(\Rightarrow\) Calculations. Type \(3 + 4\) as the function, then click on [OK]. If you look in the Commands and Output window, notice that \(3 + 4\) still equals 7!

So that is how Instat works. You prepare commands, which are submitted to the Instat server. The Windows version simply gives you a variety of ways to prepare commands for Instat. Instat obeys the commands and (usually) puts the results in the Output Window.

If the commands produce graphs, then Instat puts the graphs in a Graphics Window. If you make a mistake in the command, it prints an error message in the status line at the bottom of the Instat screen, and also in the Output window.
4.5. Instat as a calculator

The example above (3 + 4 = 7) showed how Instat may be used as a simple calculator. Here we build on this idea. Suppose you want to calculate the difference between 4.35 and 2.37 expressed as a percentage of 4.35.

⇒ Open the calculator using Manage ⇒ Calculations, construct the calculation shown in Fig. 4.5.1 (i.e. 100 * (4.35 - 2.37) / 4.35)

It is important that parentheses – ‘(’ and ‘)’ – are included where appropriate in the formula to ensure that the calculation you are trying to do has only one meaning.

Fig. 4.5.1. Instat as a calculator: the Calculate dialogue

Fig. 4.5.2. Output from calculation

⇒ Click [OK] to get the result, namely 45.517 (Fig. 4.5.2): i.e. the difference is 45.52% of 4.35.

⇒ To try some more calculations, click the Edit last dialogue button on the toolbar (Fig. 4.5.3) to recall the Calculations dialogue.

Notice that Instat remembers the previous calculation (Fig. 4.5.4).

This time we will find the mean of X1 and save the result from the calculation in a constant, K1.

⇒ Make sure that the worksheet wheat.wor is open. In the Calculate dialogue, construct the expression mea(X1) in the formula field, click the Save Result box, choose the option to save in a constant and select K1 as the result (Fig. 4.5.5). Click [OK].

Fig. 4.5.5. Save result in a constant

⇒ To see the result, either click the Constant tab in the Current Worksheet window (Fig. 4.5.6) or type ? K1 <Enter> in the Commands and Output window (Fig. 4.5.7).
The examples so far show that Instat, in common with all other statistics packages, is basically a column calculator. The main component of an Instat worksheet comprises the columns X1, X2 etc. (See, for example, Fig. 4.3.2.) Occasionally, as in the calculation above, it is useful to store single values, for which the 'constants' (K1, K2 ...) are provided.

Next we show how Instat can be used as a scientific column calculator. It allows a wide range of data transformations, which are sometimes required before the data are analysed.

⇒ Bring up the Calculate dialogue. In the formula field box, either type or use the calculator keys to enter `sqr(X1)` (as shown in Fig. 4.5.8). Choose the option to save the result in a column, and select X5 from the list of available columns. This will give the square root of the data contained in the column New, and save the results in a new column, X5. Click [OK].

⇒ Name X5 as SqrNew and format the result to display 3 decimal places (Fig. 4.5.9).

4.6. Review of example 2

This second example introduced some more of Instat's facilities, showing several ways to compare the means of two samples. In the process, some reorganisation of the data columns was needed. As you read the points below, refer book back at the example to remind yourself about the menus and dialogues used.

- Boxplots allow a visual comparison to be made of the average (mean, median) and spread of data in a single column or in several columns.
- The means of data from two samples can be compared by assuming that a normal model is suitable, and the difference then tested by calculating the \( t \)-value. (More on the underlying statistical ideas is given in the Instat Introductory Guide.)
- Columns can be joined or stacked, and a new, classifying (factor) column generated: this type of manipulation is frequently necessary to put data in a suitable form for analysis.
- Instat can be used via the menus and dialogues. You can also type and submit commands.
• For statistical analysis, the data are arranged in columns. So statistical packages can be thought of as ‘column calculators’.

5. Example 3: A Simple Regression Model

The final example in this tutorial is of simple regression and it reviews some of the key elements of data analysis using Instat. This example is taken from pages 193–196 of Mead, Curnow and Hasted (2002), and deals with the uptake of CO\textsubscript{2} by leaves (in cm\textsuperscript{3}/dm\textsuperscript{2}/hour) at various concentrations of CO\textsubscript{2} in the air.

5.1. Simple Regression

⇒ Start a new worksheet using File ⇒ New Worksheet
⇒ Enter the concentration and uptake data and name the two columns, as shown in Fig. 5.1.1.
⇒ Give the worksheet the title MCH Example, p193-196.

**Fig. 5.1.1 Data for simple linear regression example**

![Data for simple linear regression example](image)

⇒ Click outside the worksheet and when prompted, save it as leaves.wor.
⇒ Calculate some summary statistics for each of the columns of data.
⇒ Produce a scatterplot by completing the appropriate dialogue as shown in Fig. 5.1.2.

Your plot should look similar to Fig. 5.1.3. There is clearly a linear relationship between the two variables.

---

Fig. 5.1.2. Plot dialogue showing options available using the tabs

Fig. 5.1.3. Plot as specified by the dialogue options in Fig. 5.1.2

⇒ Find the **Correlation** dialogue and complete it, as shown in Fig. 5.1.4, to give the correlation between *Uptake* and *Conc*.

The result appears in the output window: you should find that the value is 0.984.
Find the dialogue for Simple Linear Regression and complete it, as shown in Fig. 5.1.5.

The results are in the output window (Fig. 5.1.6). They show that the fitted equation is:

\[ \text{Uptake} = -2.043 + 0.0249 \times \text{Conc} \]
Fig. 5.1.6. Output from the regression

![Simple Linear Regression](image)

A plot of the regression data and the fitted line are shown in Fig. 5.1.7. The line is labelled ‘S1’ in the figure: this is an abbreviated form of ‘String 1’.

**Fig. 5.1.7. Plot of points and fitted line**

![Plot of Uptake vs Concentration](image)

⇒ Return to the *leaves* worksheet and press the **Strings** tab (Fig. 5.1.8).

**Fig. 5.1.8. The contents of S1**

<table>
<thead>
<tr>
<th>String values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S1</strong></td>
</tr>
</tbody>
</table>

Fig. 5.1.8 shows that Instat has stored, in S1, the equation of the fitted line plotted in Fig. 5.1.7. This indicated that with Instat you are able to plot functions, as well as points and lines.
5.2. Help in Instat

There are many ways to get help. Some users will look for help from the dialogues. Every Instat dialogue has a [Help] button, which provides information about the purpose of the dialogue and usually includes three links labelled 'Commands', 'Examples' and 'Data'. An example can be seen (Fig. 5.2.1) if you click on [Help] in the Plot dialogue (shown in Fig. 3.2.6).

**Commands** lists the Instat commands that perform the same function(s) as the dialogue. Each listed item links directly to the appropriate place in the Reference Guide.

**Examples** describes where the facilities in the dialogue are used. Most items link to relevant sections in Instat's Introductory Guide or Climatic Guide, where more information is provided.

**Data** tells you which of Instat's sample datasets can be used to illustrate the working of the dialogue (e.g. Fig. 5.2.1). From there you can jump to a full description.

![Fig. 5.2.1. Help for the Plot dialogue](image)

The main alternative to using the button on the dialogues is the Help menu (Fig. 5.2.2), which gives access to general information and to the contents of all Instat's documentation (which runs to more than 1000 pages). There are four ways to access the help using this menu:

⇒ Choose **Help** ⇒ **Contents** (or press the <F1> shortcut key, which has the same effect).
⇒ For help on how to use Instat, select **Help** ⇒ **Getting Started**.
⇒ To get help on the syntax of an Instat command, choose **Help** ⇒ **Reference Guide**.
⇒ To see, for example, how to analyse your data using ANOVA, choose **Help** ⇒ **Introductory Guide** and go to Chapter 16 (Analysis of Variance).
5.2.2. The Help menu

If you are typing an Instat command and need help on its use, then press <F1> while on the same line as the command. You will be taken directly to the reference information for that command.

5.3. Leaving Instat

⇒ To end an Instat session, choose File ⇒ Exit. You will be asked if you want to save any of the open windows or spreadsheets. For the purposes of this tutorial, select [No] for all windows, and [Exit] from Instat.

6. Review

Here is your chance to review some of the tasks you have undertaken in the tutorial. Make brief notes in the Comment boxes to remind yourself how to carry out each of the tasks.

<table>
<thead>
<tr>
<th>Task</th>
<th>Comment</th>
<th>Hint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open a set of data that you entered earlier (for example, the file wheat.wor).</td>
<td></td>
<td>Figs 2.6.4 and 2.6.5</td>
</tr>
<tr>
<td>Enter a new set of data that has 3 columns and 6 rows.</td>
<td></td>
<td>Figs 1.2.1 and 2.1.1</td>
</tr>
<tr>
<td>Derive a new column that contains the squares of the values in an existing column.</td>
<td></td>
<td>Fig. 4.5.8</td>
</tr>
<tr>
<td>Examine data in two columns by means of boxplots.</td>
<td></td>
<td>Figs 4.1.2–4</td>
</tr>
<tr>
<td>Produce a line plot.</td>
<td></td>
<td>Figs 3.1.1–4</td>
</tr>
<tr>
<td>Summarise a column of data.</td>
<td></td>
<td>Figs 2.2.1–4 and 3.3.1–2</td>
</tr>
<tr>
<td>Mention three ways to submit commands to the Instat server.</td>
<td></td>
<td>Section 4.4</td>
</tr>
<tr>
<td>Explain what is meant by a</td>
<td></td>
<td>Section 3.2 and</td>
</tr>
<tr>
<td>Question</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>factor column.</td>
<td>Fig. 4.3.2</td>
<td></td>
</tr>
<tr>
<td>Make an existing column into a factor column.</td>
<td>Figs 3.2.1–3</td>
<td></td>
</tr>
</tbody>
</table>
| Describe the roles of the four types of data stored in an Instat worksheet? | Fig. 2.1.1 (X)  
                              | Fig. 3.2.4 (L)  
                              | Fig. 4.5.6 (K)  
                              | Fig. 5.1.8 (S) |
| Leave Instat? (If you cannot do this, just keep practising!)           | Section 5.3 |
| Why are statistics packages mainly ‘column calculators’? How does this compare with a spreadsheet package like Excel, or a database package like Access? |           |
Appendix: Exporting and Importing Data

You might wish to export Instat worksheets to other applications, like Genstat, that have more powerful statistical techniques, or that have more immediate data management and plotting facilities, like Excel.

Exporting from Instat

As a demonstration, we will export data from Instat to Excel, do some work in Excel, and then import back from Excel to Instat.

⇒ From the menu select File ⇒ Import/Export ⇒ Export as
⇒ Using the upper Browse button, select raindays.wor as the worksheet to be exported
⇒ Use the lower Browse button to specify the file name for the exported data. This lets you specify the File type from a pull-down list: select Excel, specify the filename raindays and click [Open] and then [OK].
⇒ Minimise the Instat window, load Excel and open the raindays.xls file, which should contain a single spreadsheet.
⇒ Insert a new Excel spreadsheet and call it managedata.
⇒ Copy the columns yearnum, total and type across to this new worksheet.

Figures A.1 and A.2 show how to produce boxplots in Excel. This facility is currently only available in Excel if you have installed an add-in that provides high-resolution graphics. One possibility, used here, is the SSC-Stat add-in available on the CD and described elsewhere in this handbook. If you have not installed this add-in then read the next part and continue the tutorial from after Fig. A.2.

⇒ From the menu choose SSC-Stat ⇒ Visualisation ⇒ Boxplot, then plot total by factor type and check the Display outliers box. Name the new worksheet Boxplots.

Fig. A.1. Boxplots using SSC-Stat in Excel

In the managedata spreadsheet, add new information about the predicted year (from a seasonal forecast) and the average yields, as shown in Fig. A.3.

⇒ In a column named pred type the codes (using L for low, A for average and H for High):

\[
\begin{align*}
\text{H} & \quad \text{H} & \quad \text{A} & \quad \text{H} & \quad \text{L} & \quad \text{A} & \quad \text{L} & \quad \text{A} & \quad \text{H} & \quad \text{L} & \quad \text{L}
\end{align*}
\]
In a column named *yield* type: 1.2 0.7 0.9 1.3 1.1 0.6 1.2 0.7 0.9 0.9 1.0

To save these changes use **File ⇒ Save**. This updates the Excel workbook file *raindays.xls*.

Close Excel.

**Fig. A.3. The Excel worksheet **managedata**

If you want to analyse these data it is usually quicker in Instat than in Excel. The next section describes the import of data into Instat.

**Importing datasets from external sources**

Instat can import datasets from a range of different formats from spreadsheets to databases to other statistical packages. We give an example of importing from Excel because the flexibility of spreadsheets makes Excel an obvious package for data entry.

**Fig. A.4. Instat’s Open file for input dialogue**

To import the *managedata* spreadsheet from Excel into an Instat worksheet, do the following:

In Instat, use **File ⇒ Open Worksheet** and select the **file type** to Excel (Fig. A4).
⇒ Locate the file `raindays.xls` and click [Open] to invoke the **Import** dialogue shown in Fig. A.5.

⇒ Tick **Spreadsheet options** and select **MANAGEDATA** from the list of sheet names.

⇒ To preserve the original `raindays.wor` worksheet, change the **Worksheet to which data will be saved** from the name suggested (**MANAGEDATA**) to `raindays2.wor` (Fig. A.5).

**Fig. A.5. Instat’s Import dialogue**

Text columns (such as `type` and `pred` in this example) are imported into Instat as factors.

**Fig. A.6. Instat worksheet showing imported data**

Importing data using ODBC (Open DataBase Connectivity)

Open DataBase Connectivity provides a very flexible way of importing data from external sources using features typical of database queries, like subsetting rows by specifying criteria such as numeric ranges, factor levels or logical operators.

As an example, let’s say we wish to import only the ordinary years (classified as ‘O’) into Instat.

⇒ Use Instat’s menu File ⇒ Export/Import ⇒ ODBC Query to start the ODBC process.

⇒ In the **Select Data Source** dialogue (Fig. A.7), click the **Machine Data Source** tab and select **Excel files**. Click [OK].

⇒ Locate the folder where you saved the file `raindays.xls` and click [OK] (Fig. A.8).
In the **Open Spreadsheet from ODBC Data** dialogue (Fig. A.9), the **Table** field initially displays the first worksheet in the workbook, i.e. **Table: ‘InstatData$’**.

- Use the pull-down arrow to select **SYSTEM:managedata$** – this refers to the *managedata* sheet, which holds 3 numeric columns (prefixed ‘N:’) and 2 character columns (prefixed ‘C:’).
- Select all but the *yearnum* column, as shown (Fig. A.9).
- To import only selected rows, click on the [Where] button to bring up the **ODBC Data: Select rows with Where statement** dialogue (Fig. A.10). In the top box type *type_* = ‘O’ and then click [OK] to return to the previous dialogue.

- If you want to view the SQL statement, tick the **Edit resulting SQL statement** checkbox, as shown in Fig. A.9, and click [OK]. The SQL statement appears in a new dialogue with the title **ODBC: Edit SQL statement** (Fig. A.11).
Click [**OK**] to accept the ODBC query, click [**OK**] again to perform the selection and save the results in a new Instat worksheet named *raindays3.wor*.

This action does not automatically open the Instat worksheet that is produced. To do so,

Use **File** ⇒ **Open Worksheet**, locate the file that you just saved (*raindays3.wor*), and check that the correct subset of data has been imported. If so, your worksheet should look like that in Fig. A.12.

The ODBC query option can be used in a similar way to extract data from sources of data other than Excel, such as Access.