Onlinevarsity
your e-way to learning

Decision Making using MS Excel

Are you registered with **Onlinevarsity.com**?	Yes	No
Did you download this book from **Onlinevarsity.com**?	Yes	No

Scores

For each **YES** you score **50**
For each **NO** you score **0**
If you score less than 100 this book is illegal.
Register on **www.onlinevarsity.com**

Decision Making using MS Excel

© 2008 Aptech Limited

All rights reserved.

No part of this book may be reproduced or copied in any form or by any means – graphic, electronic or mechanical, including photocopying, recording, taping, or storing in information retrieval system or sent or transferred without the prior written permission of copyright owner Aptech Limited.

All trademarks acknowledged.

APTECH LIMITED
Contact E-mail: ov-support@onlinevarsity.com

First Edition - 2008

INTRODUCTION TO THE BOOK

MS Excel 2007 is one of the components of MS Office 2007 suite. Excel has been used as a data analysis tool. Most of the data analysis features from previous versions of Excel are still supported in MS Excel 2007. Data analysis using Excel 2007 is much simpler as compared to previous versions.

This book covers some important topics such as Pivot Tables, Pivot Charts, Data Tables, Scenarios, Goal Seek, Solver, and Analysis ToolPak. Students will learn to summarize data using PivotTable reports, and to use Data Tables to view and compare results of different variations of a formula. They will use Scenario Manager to predict the outcome of particular set of input data. Students will also be exposed to the various Goal Seek features to determine the input value for known output values. The Solver tool is similar to Goal Seek, but is used when you want to apply restrictions on the problem. Analysis ToolPak, which is a set of tools to carry out statistical analysis, is also covered.

This book is the result of a concentrated effort of the Design Team, which is continuously striving to bring you the best and the latest in Information Technology. The process of design has been a part of the ISO 9001 Certification for Aptech-IT Division, Education Support Services. As part of Aptech's quality drive, this team does intensive research and curriculum enrichment to keep it in line with industry trends.

We will be glad to receive your suggestions.

APTECH COMPUTER EDUCATION
(A Division of Aptech Ltd)
GLOBAL INFOTECH 56, RAJPUR ROAD DEHRADUN 248001 UTTRANCHAL India Tele. - 2744102 , 2744103 Fax -
EMail - dehradu2@acentres.com

❶ **Receipt No. :** 7045 **Date :** 19/May/2006 ❸ **Address :** H.No 24, DL Road UTTRANCHAL India
Invoice No. :
Roll No. :
Received a Sum of INR Six Thousand Six Hundred Twelve Rupees and No Paise
Cheque/Draft No. : **Dated :**
Drawn On :
❷ **From :** VISHAL SRIVASTAVA **Cheque/Draft Amount**
Cash Amount : 6,612.00
Credit Card Amount 0.00 **Authorization No.**

COURSE FEE	6,000.00
SERVICE TAX	600.00
EDUCATION CESS TAX	12.00
Total :	**6,612.00**

*CHEQUES SUBJECT TO REALISATION
THIS RECEIPT MUST BE PRODUCED WHEN DEMANDED
FEES ONCE PAID ARE NOT REFUNDABLE EXCEPT CAUTION DEP
SUBJECT TO TERMS AND CONDITION PRINTED OVERLEAF THE INVOICE

For - GLOBAL INFOTECH
(Franchisee of Aptech Limited)

❹ **Authorised Signatory**

Registered Office : Aptech Ltd. ELITE AUTO HOUSE 54-A SIR M. VASANJI ROAD CHAKALA,ANDHERI (E) MUMBAI 400093 MAHARASHTRA India Fax - 022-56974455
SERVICE TAX REG. NO. : APPLIED FOR
E. &. O.E.
eCAS Version 1.1.6
Serial No : 0717028

SAMPLE

Dear Student,

Please ensure that you receive a computer generated receipt complete with all the details as given above. This is w.e.f. 1/1/2006.

❶ Ensure that the receipt number is printed

❷ & ❸Ensure that your address and name is currently printed.

❹ Ensure the receipt is signed by the centre staff.

APTECH COMPUTER EDUCATION
(A Division of Aptech Ltd)
GLOBAL INFOTECH 56, RAJPUR ROAD DEHRADUN 248001 UTTRANCHAL India Tele. - 2744102 , 2744103 Fax -
EMail - dehradu2@acentres.com

INVOICE (DEVIATED FEE PLAN)

❶ **Name :** Mr. VISHAL SRIVASTAVA
❷ **Address** H.No 24, DL Road ,DEHRADUN ,Pin Code:248001.
❺ **Roll No. :** 06.208/200605/1217
Curriculum 1060
Curr. Desc. ICAI
Course : ICAI
Batch : 0605 11:00 To 13:00

❸ **Invoice No :** 1217
Invoice Date : 19/May/2006

Gross Inv. Amt. : 6,000.00
Total Service Tax: 600.00
Total Education Cess Tax: 12.00
❹ **Net Inv. Amt. :** 6,612.00

CAUTION DEPOSIT

EXAM FEE

INSTALLMENT NO.	INSTALLMENT DATE	AMOUNT DUE (Amount In INR)	Service Tax Amount Amount (Amount In INR)	Education Cess Tax (Amount In INR)
1	19/May/2006	6,000.00	600.00	12.00

Total Amount (In Words) : **Six Thousand Six Hundred Twelve Rupees and No Paise**

RECEIPT NUMBER(S (For First Installment Only	RECEIPT DATE (For First Installment Only	RECEIPT AMOUNT (For First Installment Only)	Service Tax (For First Insatallment Only)	Education Cess Tax (For First Insatallment Only)	RECEIPT AMOUNT
7045	19/May/2006	6,612.00	600	12	

I have read and understood the code of conduct and payment terms/installment plan mentioned above and agree to abide by them and also the terms and conditions printed overleaf

For : **GLOBAL INFOTECH**
(Franchisee of Aptech Limited)

Signature of Student

❻ **AUTHORISED SIGNATORY**

Report Generated Using e-CAS 19/May/2006 eCAS Version 1.1.6 1
Aptech Ltd. ELITE AUTO HOUSE 54-A SIR M. VASANJI ROAD CHAKALA,ANDHERI (E) MUMBAI 400093 MAHARASHTRA India Fax - 022-56974455
SERVICE TAX REG NO APPLIED FOR
E. &. O.E.

SAMPLE

Dear Student,

Please ensure that you receive a computer generated invoice complete with all the details as given above. This is w.e.f. 1/1/2006.

❶ & ❷ Please ensure that your name and address is correctly mentioned.

❸ & ❹Please ensure that Invoice Number is present along with the Net Invoice amount.

❺ Please ensure that the Roll No. provided is 15 digits in length.

❻Ensure that the Invoice is signed by the Centre Staff as well as by you.

Contents

1.	Pivot Tables and Pivot Charts	1
2.	Pivot Tables and Pivot Charts (Lab)	37
3.	Data Tables and Scenario Manager	59
4.	Data Tables and Scenario Manager (Lab)	77
5.	Goal Seek and Solver	91
6.	Goal Seek and Solver (Lab)	107
7.	Analysis ToolPak	125
8.	Analysis ToolPak (Lab)	165
9.	Functions and Macros	175
10.	Functions and Macros (Lab)	217
	Glossary	i

This page has been intentionally left blank

Session 1

Pivot Tables and Pivot Charts

Session Objectives:

At the end of this session, you will be able to -

- Describe pivot table
- Explain sorting and filtering in pivot tables
- Describe pivot chart
- Explain the options to customize pivot charts

1.1 Pivot Table

MS Excel is one of the easiest tools to maintain records of information. These records could be a few thousand in number. Figure 1.1 shows a sample of such data. To keep the example simple, the number of records is reduced to twenty.

	A	B	C	D	E	F	G	H
1	Subject	Instructor	Mode	Level	Students	Male	Female	
2	Programming	IT	Full-Time	Undergraduate	96	45	51	
3	Programming	IT	Part-Time	Undergraduate	112	60	52	
4	Programming	IT	Full-Time	Postgraduate	28	20	8	
5	Programming	IT	Part-Time	Postgraduate	48	28	20	
6	Physics	Science	Full-Time	Undergraduate	56	40	16	
7	Physics	Science	Part-Time	Undergraduate	47	22	25	
8	Physics	Science	Full-Time	Postgraduate	85	55	30	
9	Physics	Science	Part-Time	Postgraduate	78	30	48	
10	Zoology	Science	Full-Time	Undergraduate	63	42	21	
11	Zoology	Science	Part-Time	Undergraduate	82	40	42	
12	Zoology	Science	Full-Time	Postgraduate	25	12	13	
13	Zoology	Science	Part-Time	Postgraduate	79	55	24	
14	Botany	Science	Full-Time	Undergraduate	56	30	26	
15	Botany	Science	Part-Time	Undergraduate	86	43	43	
16	Botany	Science	Full-Time	Postgraduate	49	19	30	
17	Botany	Science	Part-Time	Postgraduate	93	53	40	
18	Robotics	Electronics	Full-Time	Undergraduate	66	40	26	
19	Robotics	Electronics	Part-Time	Undergraduate	51	21	30	
20	Robotics	Electronics	Full-Time	Postgraduate	85	40	45	
21								

Sheet1 Sheet2 Sheet3

Figure 1.1: Source Data

Figure 1.1 lists the data about the students of a college, offering postgraduate and undergraduate full-time and part-time courses. This data contains lot of information; however, it lacks summarized data. Typically, you would like to know the following from this data:

- How many students have enrolled for a postgraduate and how many for an undergraduate course?
- How many postgraduate and undergraduate students have enrolled for a part-time course and how many for a full-time course?
- Which subjects are taken by postgraduate and undergraduate part-time and full-time students?

The above questions require data to be summarized in an easy to understand format. This can be achieved using formulas and functions of MS Excel. However, this requires a lot of calculations and is time consuming. A quick and easy alternative is the PivotTable report. Figure 1.2 shows a PivotTable report displaying summarized data about the total number of students enrolled for the postgraduate and undergraduate courses. Note that the PivotTable report uses the data in figure 1.1 as the source data.

	A	B
1		
2		
3	**Row Labels**	**Sum of Students**
4	Postgraduate	570
5	Undergraduate	715
6	**Grand Total**	**1285**

Figure 1.2: Number of Postgraduate and Undergraduate Students

You can further customize the PivotTable report to display the total number of postgraduate and undergraduate students enrolled for a part-time course or a full-time course, as shown in figure 1.3.

	A	B
1		
2		
3	**Row Labels**	**Sum of Students**
4	**Postgraduate**	**570**
5	Full-Time	272
6	Part-Time	298
7	**Undergraduate**	**715**
8	Full-Time	337
9	Part-Time	378
10	**Grand Total**	**1285**

Figure 1.3: Number of Students Enrolled for Part-time and Full-time Course

The PivotTable report in figure 1.3 can be further customized to include details about how many students have enrolled for the subjects, namely Botany, Physics, Programming, Robotics, and Zoology. Figure 1.4 shows the modified PivotTable report.

V 1.0 © 2008 Aptech Limited

	A	B	C	D	E	F	G
1							
2							
3	Sum of Students	Column Labels					
4	Row Labels	Botany	Physics	Programming	Robotics	Zoology	Grand Total
5	Postgraduate	142	163	76	85	104	570
6	Full-Time	49	85	28	85	25	272
7	Part-Time	93	78	48		79	298
8	Undergraduate	142	103	208	117	145	715
9	Full-Time	56	56	96	66	63	337
10	Part-Time	86	47	112	51	82	378
11	Grand Total	284	266	284	202	249	1285

Figure 1.4: Number of Students Enrolled for Botany, Physics, Programming, Robotics, and Zoology

Therefore, a PivotTable report can be used to summarize almost any kind of data and up to any level of detail. Besides summarizing data in a PivotTable report, you can do the following:

- Query data in a user-friendly manner
- Subtotal and aggregate numeric data
- Expand or collapse level of detail of data displayed
- Sort, filter, group and conditionally format a subset of data
- Print PivotTable report

1.1.1 Creating a PivotTable Report

Before you create a PivotTable report, the source data for the PivotTable report should be organized for effective analysis. The source data for a PivotTable report must satisfy the following requirements:

- First row should contain column headings. The heading can be a single word or a combination of multiple words. All characters are allowed in a heading.
- The second and subsequent rows should contain data.
- No blank rows or columns should exist; however, blank cells are allowed.
- Source data cannot contain any Excel-generated subtotals and grand totals. If present, you can remove them by selecting Data, Outline, Subtotal, and then Remove All.
- Source data should not contain any total calculation formulas at the top or bottom of any column.

After the source data is organized in accordance with the requirements mentioned above, you can create a PivotTable report using the following steps:

1. Select a cell in the source data.
2. On the **Ribbon**, click the **Insert** tab, and in the **Tables** group, click **PivotTable** and then click **PivotTable**. The **Create PivotTable** dialog box is displayed, as shown in figure 1.5.

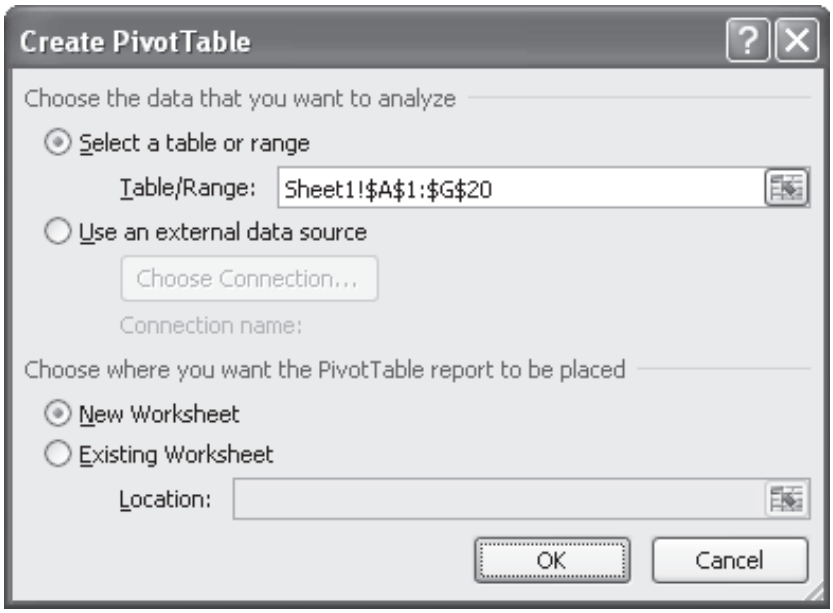


Figure 1.5: Create PivotTable Dialog Box

Note that the table range A1 to G20 is selected by default. This range of cells contains the source data in **`Sheet1`**. The **New Worksheet** option is selected by default. This causes a PivotTable report to be created in a new worksheet. An empty PivotTable report is created in a new worksheet (**`Sheet4`**), as shown in figure 1.6.

V 1.0 © 2008 Aptech Limited

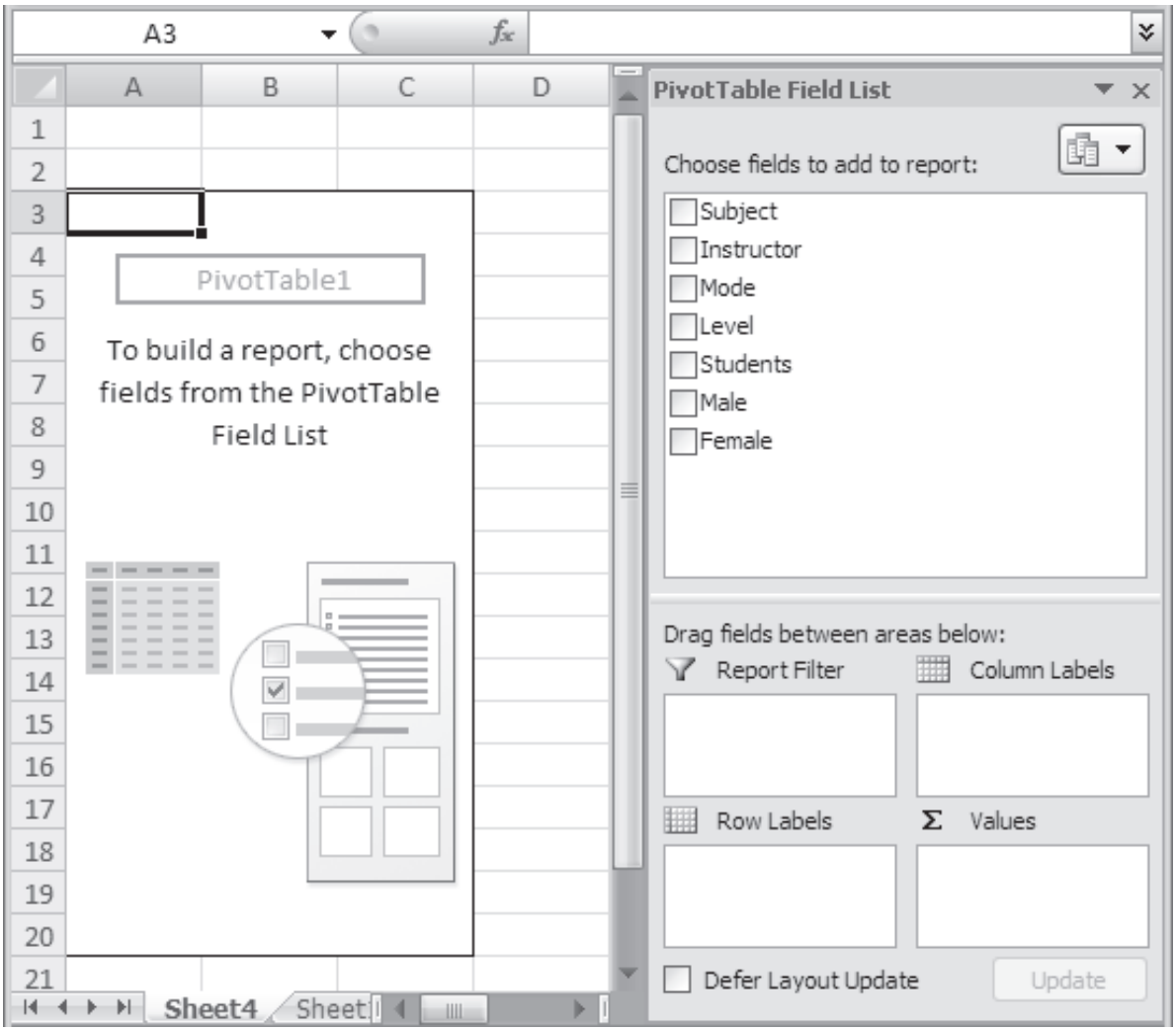


Figure 1.6: Empty PivotTable Report in a New Worksheet

The PivotTable report begins in cell A3. The **PivotTable Field List** pane is displayed on the right side of the Excel window. The **PivotTable Field List** pane comprises two sections. The upper section lists the various columns of source data as fields. The order of these fields is the same as that of the columns in source data. The bottom section of the **PivotTable Field List** pane comprises four areas, namely Report Filter, Column Labels, Row Labels, and Values. Table 1.1 describes all four areas of a **PivotTable Field List** pane.

Area	Description
Report Filter	PivotTable report data is filtered on the basis of fields added to this area.
Column Labels	Fields added to this area are displayed as columns at the top of the report. If this area contains more than one field, then the next field is nested within the field above it.
Row Labels	Fields added to this area are displayed as rows on the side of the report. If this area contains more than one field, then the next field is nested within the field above it.
Values	Fields added to this area are displayed as a single summarized value in the PivotTable report.

Table 1.1: PivotTable Field List Areas

The fields from the upper section of **PivotTable Field List** pane, are placed into one of the areas listed in Table 1.1. The data of the fields added to Report Filter, Column Labels, Row Labels, and Values area is displayed, as shown in figure 1.7. For example, a field added to the Row Labels area of **PivotTable Field List** pane is displayed in the PivotTable report in the Row Labels area.

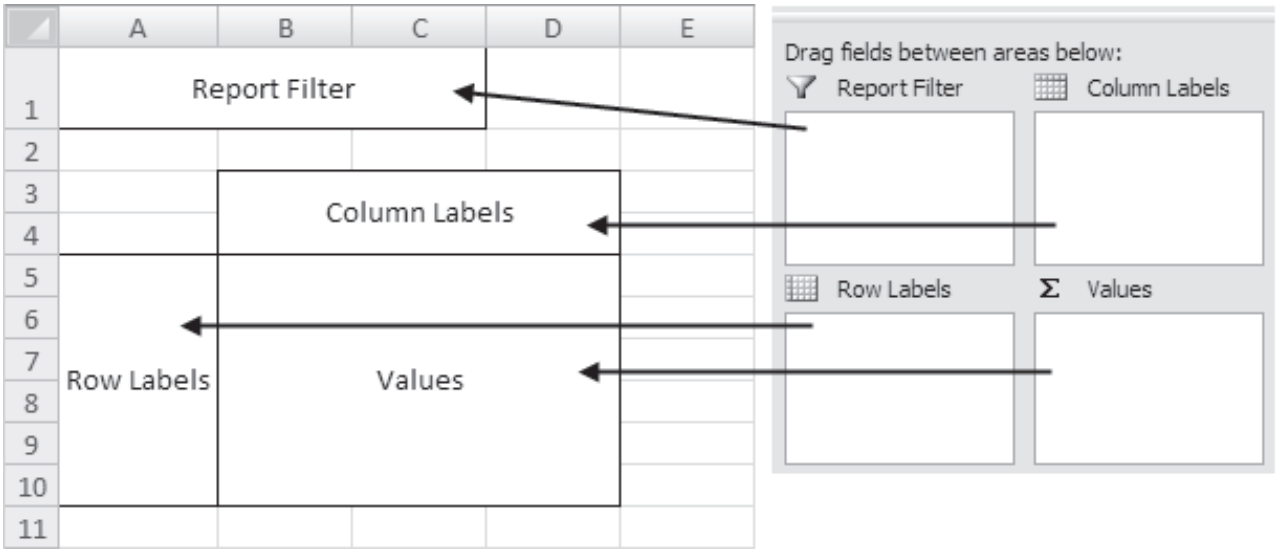


Figure 1.7: PivotTable Report Areas

Consider a case where you have to create a PivotTable report to display the total number of students enrolled for postgraduate and undergraduate courses. The **`Students`** column, in the source data, contains the number of students enrolled. The **`Level`** column, in the source data, specifies whether a student has enrolled for a postgraduate or an undergraduate course. Therefore, you select the **`Students`** field and then the **`Level`** field in the **PivotTable Field List** pane, as shown in figure 1.8.

V 1.0 © 2008 Aptech Limited

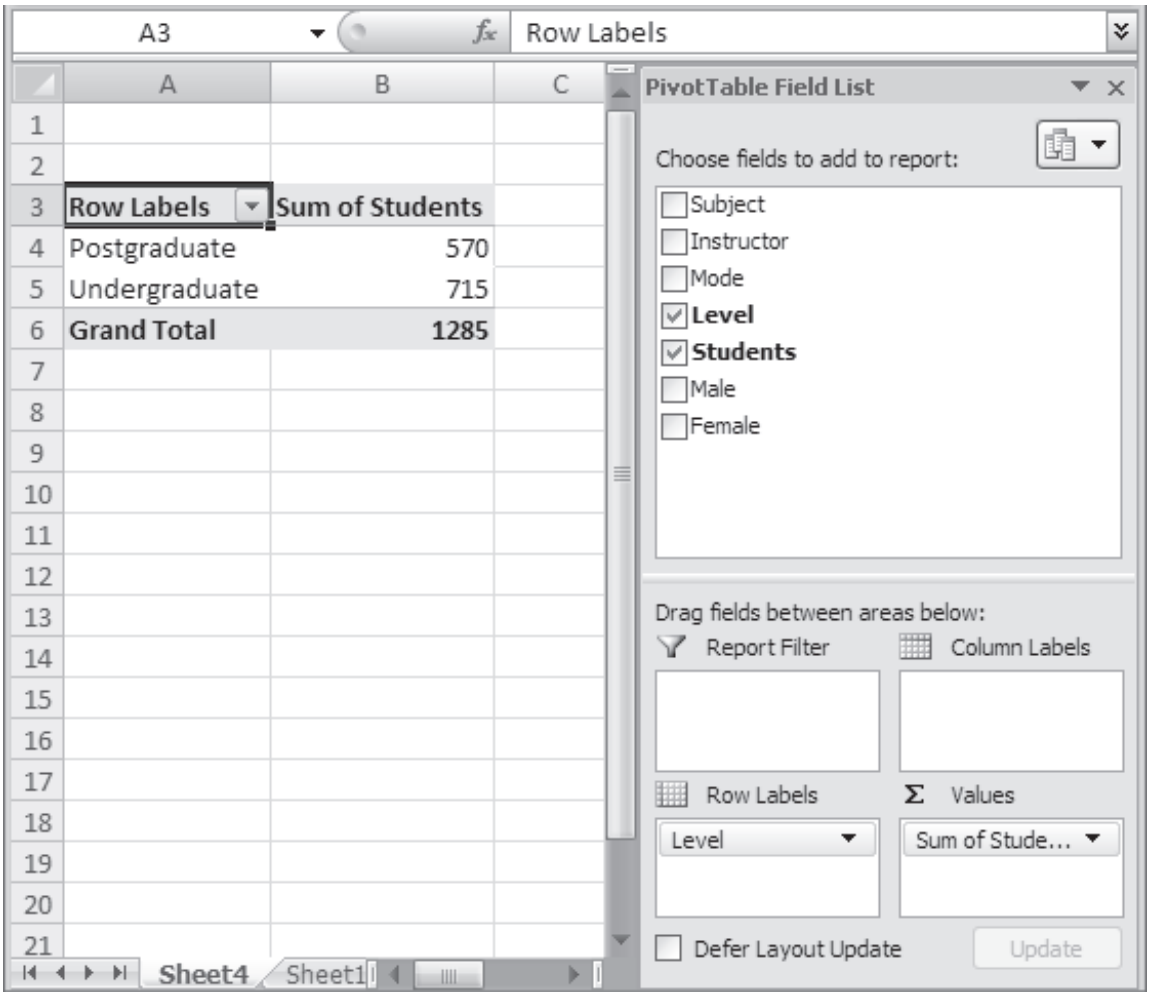


Figure 1.8: PivotTable Report – Number of Postgraduate and Undergraduate Students

Note that the **`Level`** field is added to the Row Labels area. The **`Students`** field is added to the Values area as a summation field. As the **`Level`** column contains textual data, it is added to the Row Labels area by default. The **`Students`** field contains numeric data. Hence, it is added to the Values area.

To create a summary of how many students have enrolled for full-time or part-time, postgraduate or undergraduate course, you select the **`Mode`** field. The modified PivotTable report is shown in figure 1.9.

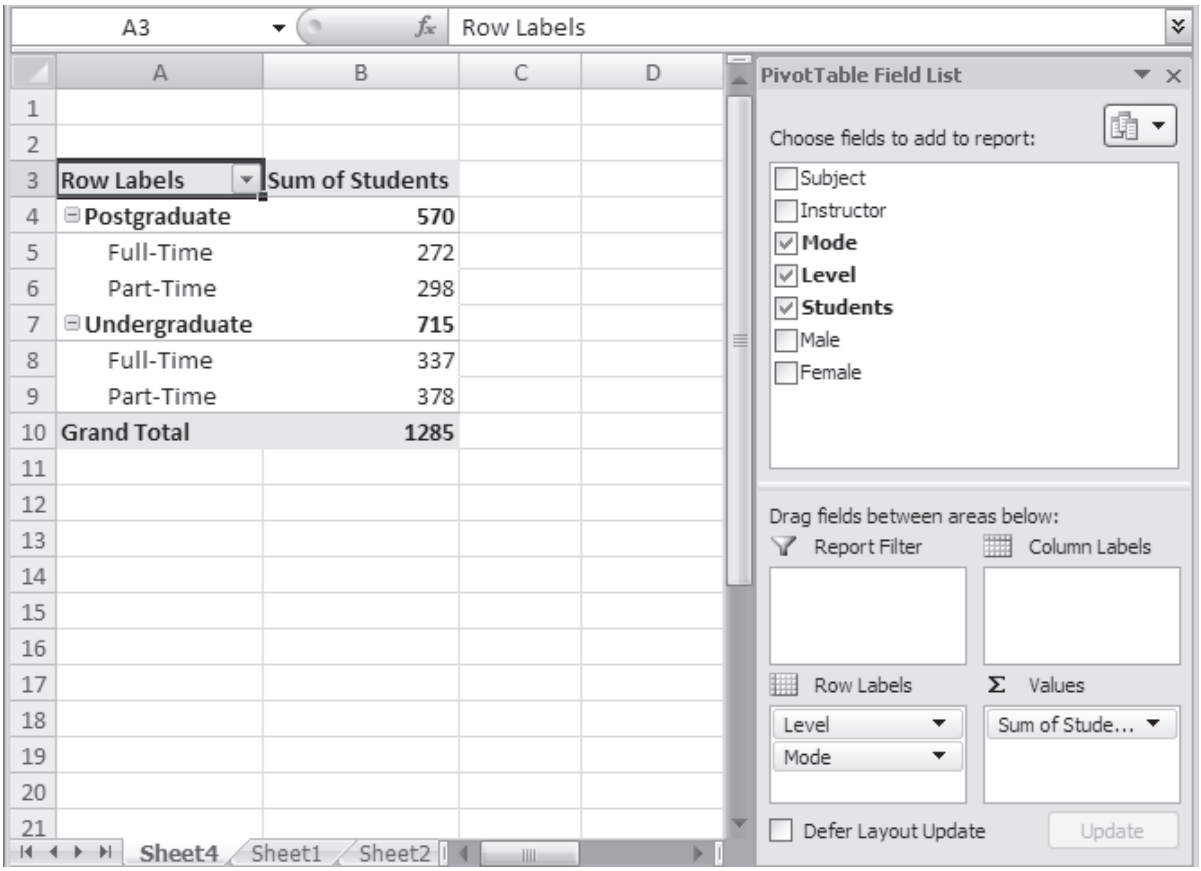


Figure 1.9: PivotTable Report with Mode Field

To change the layout of full-time and part-time data displayed in a PivotTable report, you drag the **Mode** field from the Row Labels area to the Column Labels area. The modified PivotTable report is shown in figure 1.10.

V 1.0 © 2008 Aptech Limited

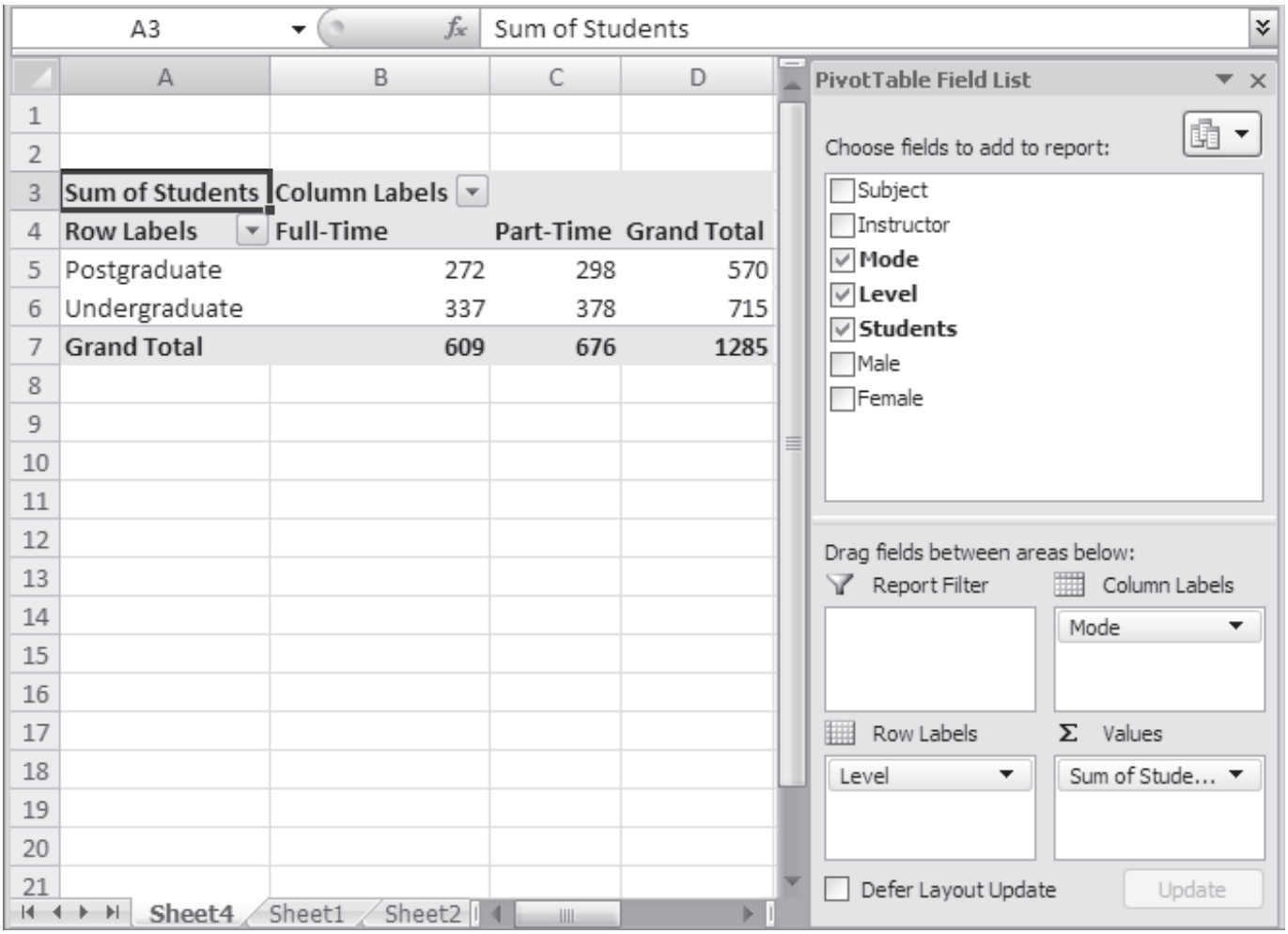
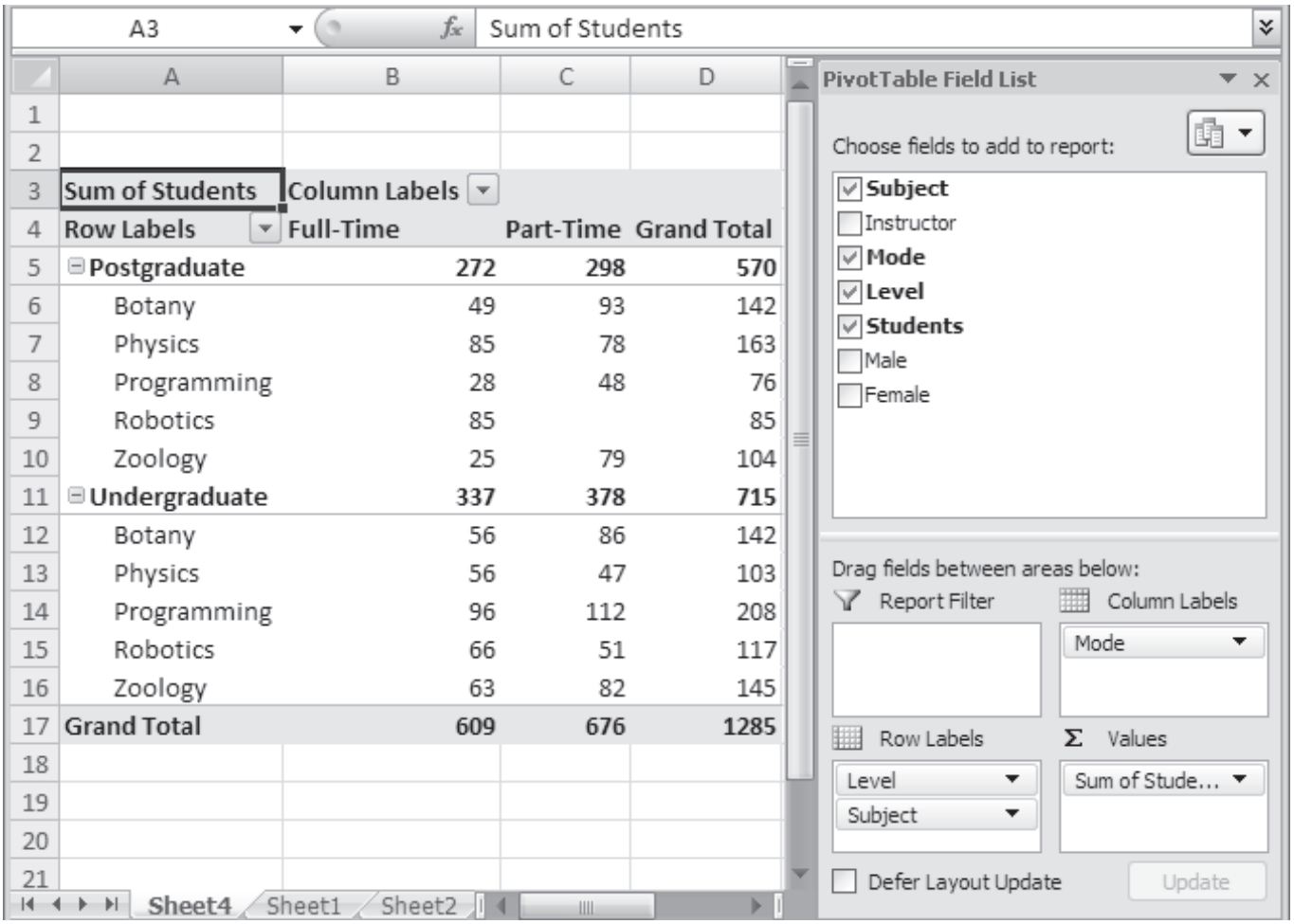


Figure 1.10: PivotTable Report with Mode Field in Column Labels Area

Note that this layout displays the PivotTable report data in tabular format.

To create a summary of the subjects selected by the postgraduate and undergraduate students, select the **`Subject`** field in the upper section of the **PivotTable Field List** pane. The PivotTable report displays the modified summarized data, as shown in figure 1.11.



Sum of Students	Column Labels		
Row Labels	Full-Time	Part-Time	Grand Total
Postgraduate	272	298	570
Botany	49	93	142
Physics	85	78	163
Programming	28	48	76
Robotics	85		85
Zoology	25	79	104
Undergraduate	337	378	715
Botany	56	86	142
Physics	56	47	103
Programming	96	112	208
Robotics	66	51	117
Zoology	63	82	145
Grand Total	609	676	1285

Figure 1.11: PivotTable Report with Subject Field

1.2 Working with PivotTable Report

After a PivotTable report is created, you can clear, refresh, or delete it.

➢ **Clearing Data in a PivotTable Report**

To clear the data in a PivotTable report, first select a cell in the PivotTable report. Next, on the **Ribbon**, under the **PivotTable Tools** tab, click the **Options** tab, click **Clear**, and then click **Clear All**, as shown in figure 1.12.

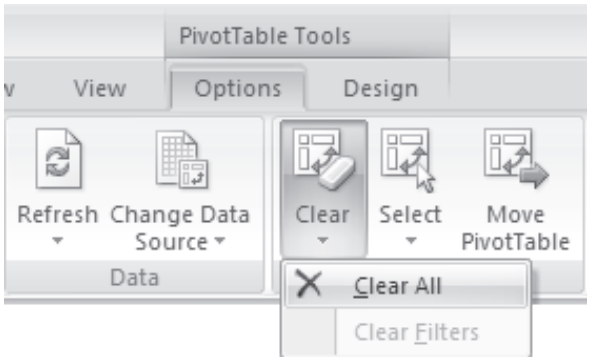


Figure 1.12: Clearing PivotTable Report

V 1.0 © 2008 Aptech Limited

➢ **Refreshing a PivotTable Report**

There are times when the source data changes, and the PivotTable report requires an update. To refresh a PivotTable report, first select a cell in the PivotTable report. Next, on the **Ribbon**, under the **PivotTable Tools** tab, click the **Options** tab, and in the **Data** group, click **Refresh** and then click **Refresh**, as shown in figure 1.13.

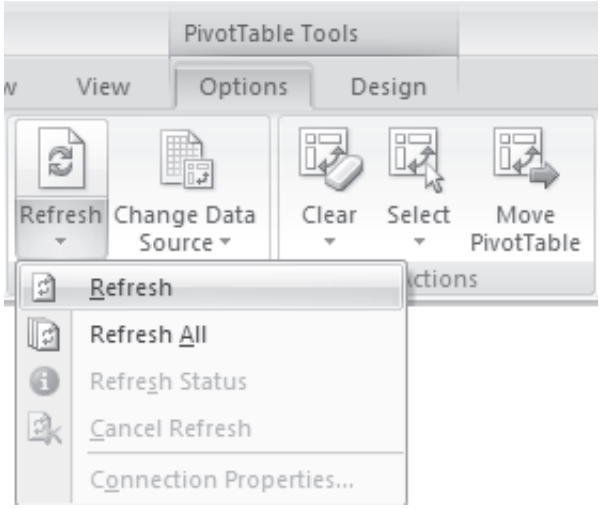


Figure 1.13: Refreshing a PivotTable Report

➢ **Deleting a PivotTable Report**

To delete a PivotTable report, first select a cell in the PivotTable report. Next, on the **Ribbon**, under the **PivotTable Tools** tab, click the **Options** tab, and in the **Actions** group, click **Select** and then click **Entire PivotTable**. This action selects the entire PivotTable report. Figure 1.14 shows the **Entire PivotTable** option. To delete the selected PivotTable report, press the DELETE key.

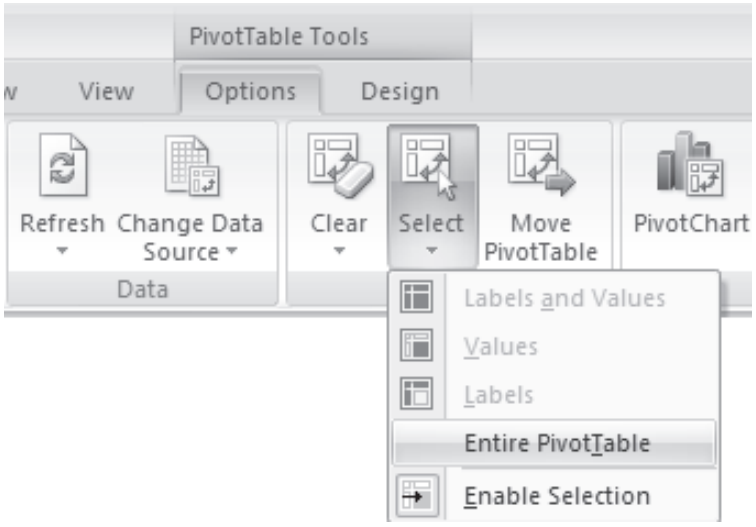


Figure 1.14: Selecting a PivotTable Report

1.3 Sorting in PivotTable Report

Consider the source data shown in figure 1.15. This data is about the various invoices issued to customers against the purchase of computer devices.

	A	B	C	D	E	F	G	H	I
1	Invoice No.	Date	Customer	Product	Category	Quantiy	Unit Price	Total Price	Executive
2	1002	4/11/2008	J & K Shipping	Samsung 750	Monitors	2	450	900	Jacob
3	1002	4/11/2008	J & K Shipping	HP Laser 28ppm	Printers	4	470	1,880.00	Jacob
4	1003	4/18/2008	Centrum Point	Dell 260 mb	Servers	5	800	4,000.00	James
5	1003	4/18/2008	Centrum Point	Lexmark G25 printer	Printers	2	150	300	James
6	1003	4/18/2008	Centrum Point	Sony 23" Flat screen	Monitors	4	650	2,600.00	James
7	1004	4/27/2008	PAFEX	HP Deskjet	Printers	4	120	480	Liza
8	1004	4/27/2008	PAFEX	Sony 23" Flat screen	Monitors	3	625	1,875.00	Liza
9	1005	5/3/2008	Stallard Tech	Canon C56	Monitors	12	325	3,900.00	Jacob
10	1005	5/3/2008	Stallard Tech	IBM M45 Server	Servers	2	950	1,900.00	Jacob
11	1006	5/10/2008	J & K Shipping	Daewoo 23"	Monitors	3	400	1,200.00	James
12	1007	5/16/2008	Indoco House	Compaq 1000mb	Servers	2	1,150.00	2,300.00	Jenny
13	1007	5/16/2008	Indoco House	IBM M45 Server	Servers	5	995	4,975.00	Jenny
14	1009	5/28/2008	Centrum Point	Canon C56	Monitors	3	390	1,170.00	Liza
15	1009	5/28/2008	Centrum Point	Lexmark G25 printer	Printers	1	150	150	Liza
16	1010	6/4/2008	J & K Shipping	Dell 260 mb	Servers	2	825	1,650.00	Jacob
17	1010	6/4/2008	J & K Shipping	Samsung 750	Monitors	10	387	3,870.00	Jacob
18	1011	6/10/2008	Stallard Tech	HP Deskjet	Printers	7	105	735	Liza
19	1011	6/10/2008	Stallard Tech	Sony 23" Flat screen	Monitors	2	600	1,200.00	Liza
20	1012	6/16/2008	Indoco House	Sony 23" Flat screen	Monitors	6	585	3,510.00	Jenny
21	1014	6/28/2008	PAFEX	Compaq H16 250mb	Servers	2	825	1,650.00	Liza
22									

Sheet5 Sheet1 Sheet2 Sheet3

Figure 1.15: Source Data

Figure 1.16 shows the PivotTable report summarizing the source data shown in figure 1.15. This report gives a summary of invoices of the products purchased by customers.

V 1.0 © 2008 Aptech Limited

A3 | Sum of Total Price

	A	B	C	D	E
1					
2					
3	Sum of Total Price	Column Labels			
4	Row Labels	Monitors	Printers	Servers	Grand Total
5	Centrum Point	3770	450	4000	8220
6	1003	2600	300	4000	6900
7	1009	1170	150		1320
8	Indoco House	3510		7275	10785
9	1007			7275	7275
10	1012	3510			3510
11	J & K Shipping	5970	1880	1650	9500
12	1002	900	1880		2780
13	1006	1200			1200
14	1010	3870		1650	5520
15	PAFEX	1875	480	1650	4005
16	1004	1875	480		2355
17	1014			1650	1650
18	Stallard Tech	5100	735	1900	7735
19	1005	3900		1900	5800
20	1011	1200	735		1935
21	Grand Total	20225	3545	16475	40245
22					

Sheet5 Sheet1 Sheet2 Sh

PivotTable Field List

Choose fields to add to report:

- [x] Invoice No.
- [ ] Date
- [x] Customer
- [ ] Product
- [x] Category
- [ ] Quantiy
- [ ] Unit Price
- [x] Total Price
- [ ] Executive

Drag fields between areas below:

Report Filter | Column Labels: Category

Row Labels: Customer, Invoice No. | Values: Sum of Total ...

Defer Layout Update | Update

Figure 1.16: PivotTable Report

A PivotTable report usually contains labels and values. Therefore, you either sort the row and column labels or the values. Sorting of the labels and the values is done using the various options provided in the **Sort** group on the **Options** tab. Figure 1.17 shows the **Sort** group highlighted.

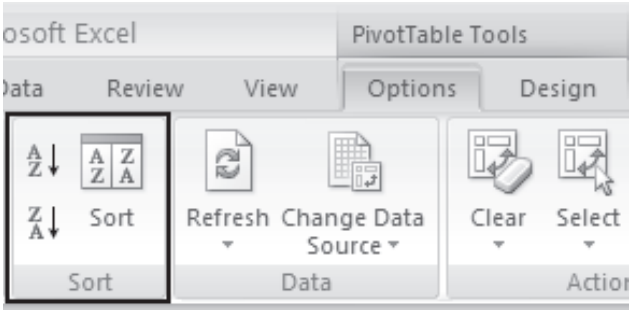


Figure 1.17: Sort Group

1.3.1 Sorting Labels

Labels usually contain textual data. However, labels can contain numbers, date and time data as well. So based on the data in the label, Excel provides corresponding sorting options.

Consider the row labels for customer name in the PivotTable report shown in figure 1.16. The customer names are sorted in ascending alphanumeric order by default. To sort the customer names in descending alphanumeric order, first select a customer row label, for example, Centrum Point (cell A5). Next, on the **Options** tab, in the **Sort** group, click **Sort Z to A**.

The modified PivotTable report is shown in figure 1.18.

	A	B	C	D	E
1					
2					
3	Sum of Total Price	Column Labels			
4	Row Labels	Monitors	Printers	Servers	Grand Total
5	Stallard Tech	5100	735	1900	7735
6	1005	3900		1900	5800
7	1011	1200	735		1935
8	PAFEX	1875	480	1650	4005
9	1004	1875	480		2355
10	1014			1650	1650
11	J & K Shipping	5970	1880	1650	9500
12	1002	900	1880		2780
13	1006	1200			1200
14	1010	3870		1650	5520
15	Indoco House	3510		7275	10785
16	1007			7275	7275
17	1012	3510			3510
18	Centrum Point	3770	450	4000	8220
19	1003	2600	300	4000	6900
20	1009	1170	150		1320
21	Grand Total	20225	3545	16475	40245
22					

Sheet5 Sheet1 Sheet2 Sh

Figure 1.18: Customer Names Sorted in Descending Order

Note: There are several ways in which a label or values can be sorted. The easiest way to sort labels containing textual data, is to right-click a label, point to Sort and then click Sort A to Z or Sort Z to A.

The invoice numbers in the PivotTable report contain numeric data. To sort these invoice numbers in descending order, right-click an invoice number (cell A6) in the report, point to **Sort**, and then click **Sort Largest to Smallest**. Figure 1.19 shows how to do this.

V 1.0 © 2008 Aptech Limited

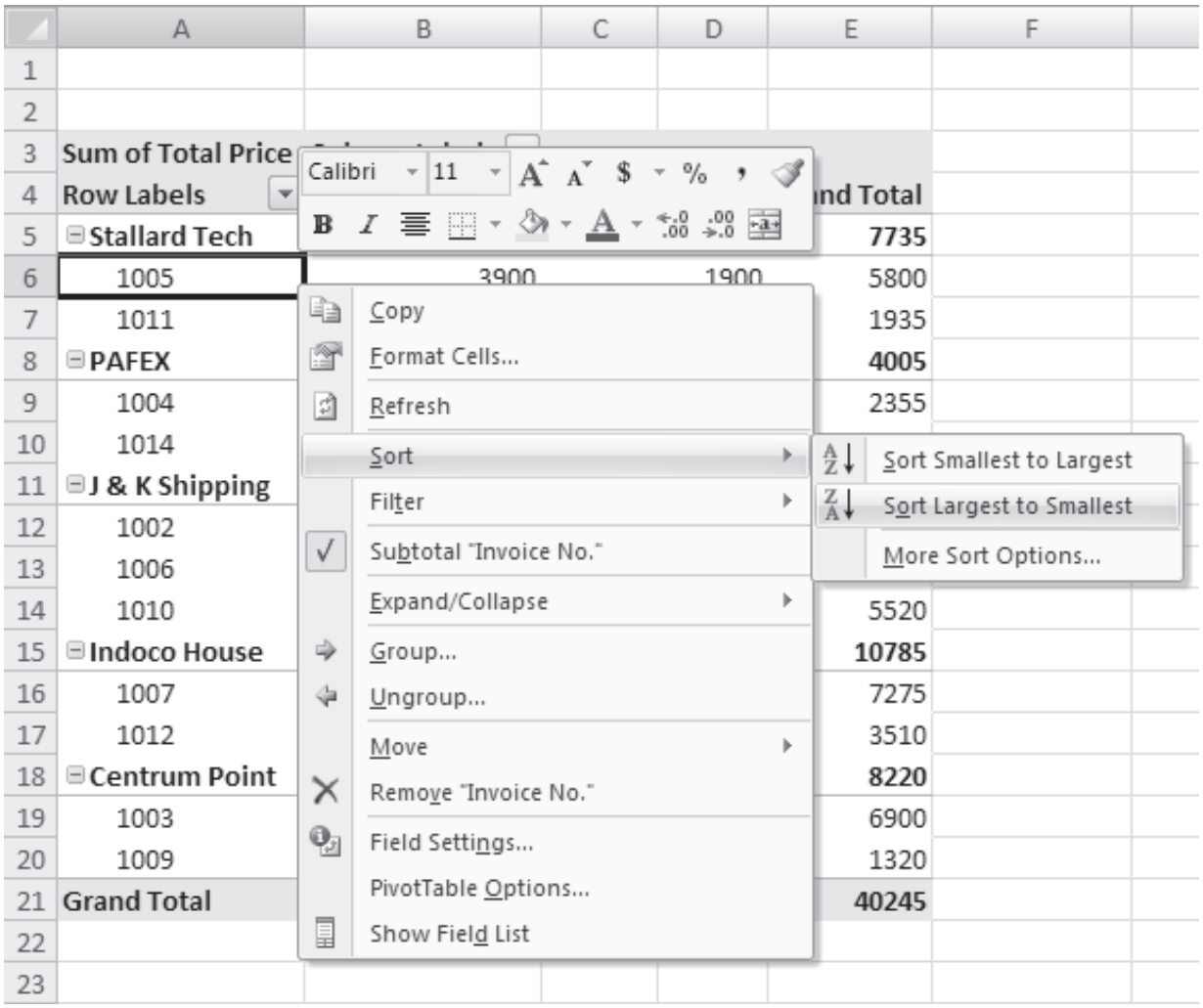


Figure 1.19: Sorting Numeric Labels

Figure 1.20 shows the PivotTable report with invoice number sorted from the largest to the smallest number.

	A	B	C	D	E
1					
2					
3	Sum of Total Price	Column Labels			
4	Row Labels	Monitors	Printers	Servers	Grand Total
5	Stallard Tech	5100	735	1900	7735
6	1011	1200	735		1935
7	1005	3900		1900	5800
8	PAFEX	1875	480	1650	4005
9	1014			1650	1650
10	1004	1875	480		2355
11	J & K Shipping	5970	1880	1650	9500
12	1010	3870		1650	5520
13	1006	1200			1200
14	1002	900	1880		2780
15	Indoco House	3510		7275	10785
16	1012	3510			3510
17	1007			7275	7275
18	Centrum Point	3770	450	4000	8220
19	1009	1170	150		1320
20	1003	2600	300	4000	6900
21	Grand Total	20225	3545	16475	40245
22					

Figure 1.20: Invoice Numbers Sorted from Largest to Smallest

Labels in a PivotTable report can contain date and time information as well. To sort such labels, right-click the label, point to **Sort** and then click **Sort Oldest to Newest** or **Sort Newest to Oldest**.

1.3.2 Sorting Values

The Values area of a PivotTable report usually contains numeric data. Consider the PivotTable report shown in figure 1.21. This report gives a summary of how many devices of each type, namely monitor, printer, or server are purchased by a customer.

V 1.0 © 2008 Aptech Limited

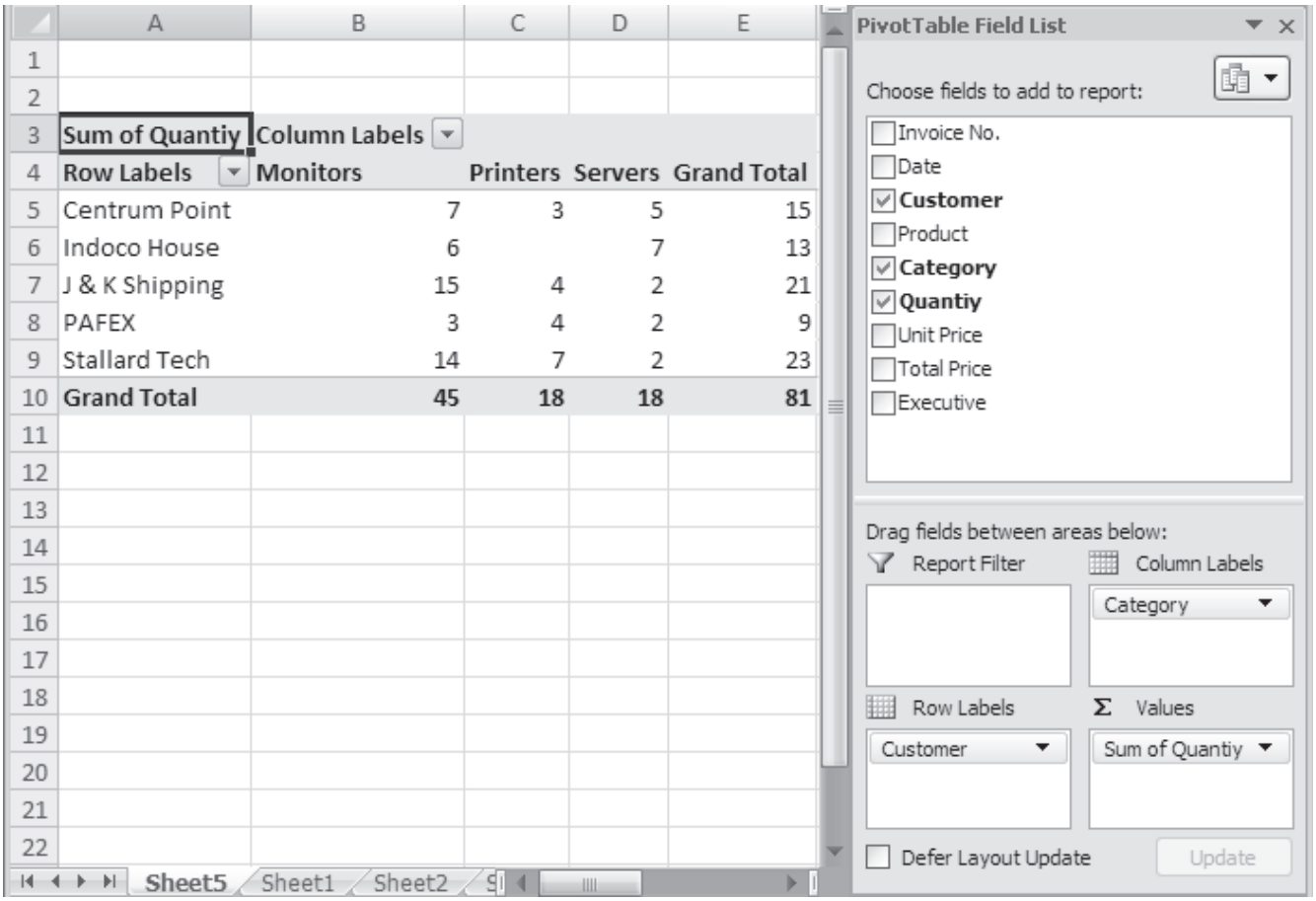


Figure 1.21: PivotTable Report Summarizing Customer Names and the Devices Purchased

Sorting can be applied to the values in this report to gain more specific information. For example, to find which customer has purchased the maximum number of monitors, you can sort the values in the `Monitors` column. To do this, right-click cell B7, point to **Sort**, and then click **Sort Largest to Smallest**. Figure 1.22 shows that J & K Shipping has purchased the maximum number of monitors.

	A	B	C	D	E
1					
2					
3	Sum of Quantiy	Column Labels			
4	Row Labels	Monitors	Printers	Servers	Grand Total
5	J & K Shipping	15	4	2	21
6	Stallard Tech	14	7	2	23
7	Centrum Point	7	3	5	15
8	Indoco House	6		7	13
9	PAFEX	3	4	2	9
10	**Grand Total**	**45**	**18**	**18**	**81**

Figure 1.22: Values in Monitor column Sorted from Largest to Smallest

Similarly, you can sort the grand total values as well. For example, to find the least sold devices, the grand total values of row 10 can be sorted. To do this, right-click the cell B10, point to **Sort**, and then click **Sort Smallest to Largest**. Figure 1.23 shows printers and servers are the least sold devices.

	A	B	C	D	E
1					
2					
3	Sum of Quantiy	Column Labels			
4	Row Labels	Printers	Servers	Monitors	Grand Total
5	J & K Shipping	4	2	15	21
6	Stallard Tech	7	2	14	23
7	Centrum Point	3	5	7	15
8	Indoco House		7	6	13
9	PAFEX	4	2	3	9
10	**Grand Total**	**18**	**18**	**45**	**81**

Figure 1.23: Values in Grand Total Row Sorted from Smallest to Largest

Consider another case where you want to find which type of device was most ordered by a particular customer. For example, to find which device was most ordered by Stallard Tech, you can sort the values in the `Stallard Tech` row. To do this, right-click cell B6, point to **Sort** and then click **More Sort Options**. The **Sort By Value** dialog box is displayed. In this dialog box, select **Largest to Smallest** and **Left to Right**, as shown in figure 1.24, and then click **OK**. On doing this, the values in the `Stallard Tech` row are sorted from left to right with the largest value on the left.

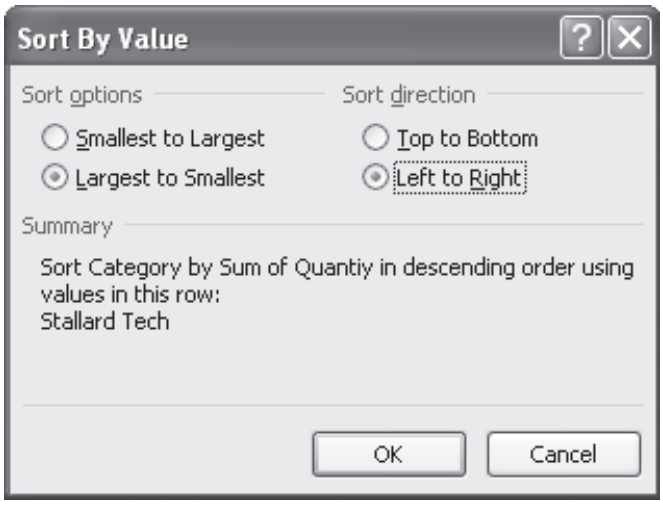


Figure 1.24: Sort By Value Dialog Box

Figure 1.25 shows that the most ordered devices by Stallard Tech are monitors.

V 1.0 © 2008 Aptech Limited

	A	B	C	D	E
1					
2					
3	Sum of Quantiy	Column Labels			
4	Row Labels	Monitors	Printers	Servers	Grand Total
5	J & K Shipping	15	4	2	21
6	Stallard Tech	14	7	2	23
7	Centrum Point	7	3	5	15
8	Indoco House	6		7	13
9	PAFEX	3	4	2	9
10	**Grand Total**	**45**	**18**	**18**	**81**

Figure 1.25: Sorting from Left to Right

1.4 Filtering in PivotTable Report

Filtering in PivotTable reports involves filtering labels, values, and the report. A PivotTable report using a filter displays only subset of data. Excel provides several options to filter labels and values. Figure 1.26 shows the various label filter options.

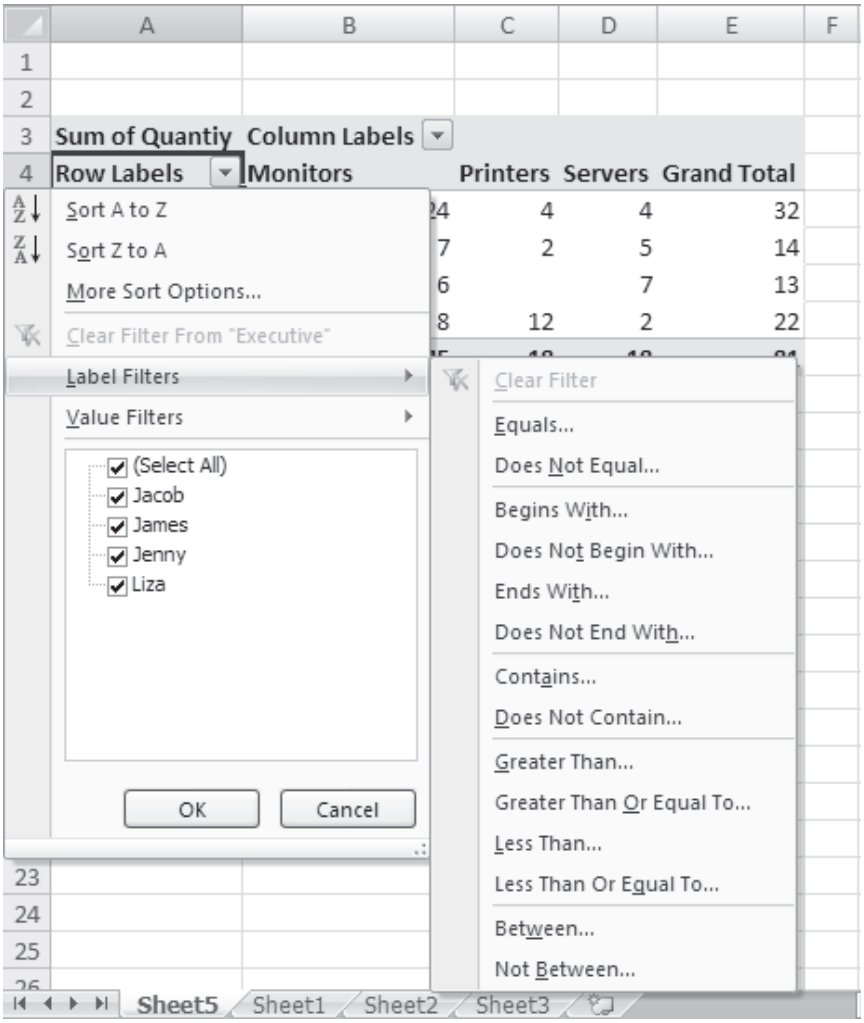


Figure 1.26: Label Filter Options

Note that the menu in figure 1.26 is displayed by clicking the arrow of the label in cell A4 or cell B3. An alternate way to display this menu is to right-click a row or a column label such as either Monitors, Printers, or Servers.

Figure 1.27 shows the various value filters options.

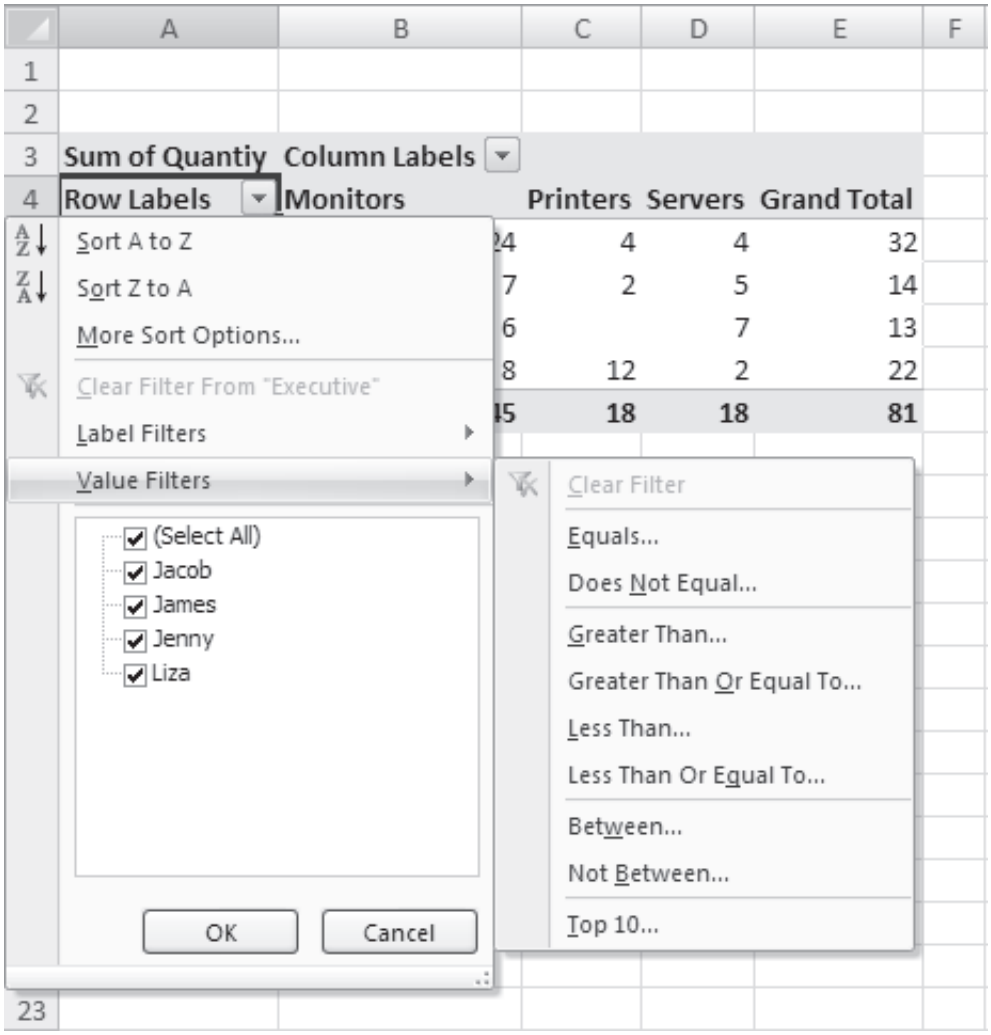


Figure 1.27: Value Filters

Figure 1.28 shows the various date filter options.

V 1.0 © 2008 Aptech Limited

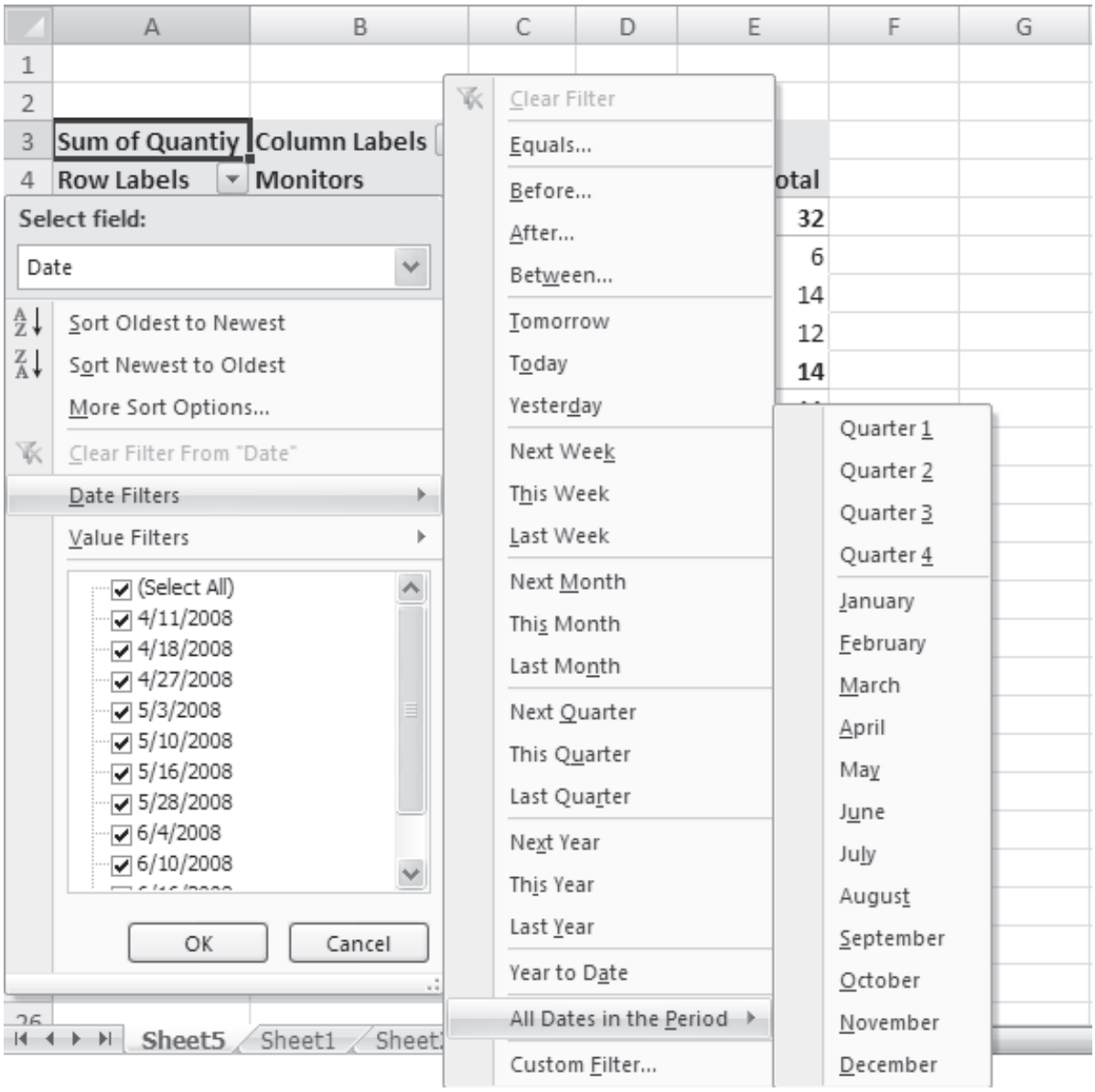


Figure 1.28: Date Filters

Note: The Date Filter option is displayed in the menu only if a label containing date or time data is right-clicked.

1.4.1 Filtering Labels

Consider the PivotTable report shown in figure 1.29. This report gives a summary about the number of devices and the date on which an executive sold a device to a customer.

	A	B	C	D	E
1					
2					
3	Sum of Quantiy	Column Labels			
4	Row Labels	Monitors	Printers	Servers	Grand Total
5	Jacob	24	4	4	32
6	4/11/2008	2	4		6
7	5/3/2008	12		2	14
8	6/4/2008	10		2	12
9	James	7	2	5	14
10	4/18/2008	4	2	5	11
11	5/10/2008	3			3
12	Jenny	6		7	13
13	5/16/2008			7	7
14	6/16/2008	6			6
15	Liza	8	12	2	22
16	4/27/2008	3	4		7
17	5/28/2008	3	1		4
18	6/10/2008	2	7		9
19	6/28/2008			2	2
20	Grand Total	45	18	18	81
21					
22					
23					
24					
25					

Sheet5 Sheet1 Sheet2

PivotTable Field List

Choose fields to add to report:

- Invoice No.
- Date
- Customer
- Product
- Category
- Quantiy
- Unit Price
- Total Price
- Executive

Drag fields between areas below:

Report Filter | Column Labels: Category

Row Labels: Executive, Date | Σ Values: Sum of Quantiy

Defer Layout Update | Update

Figure 1.29: PivotTable Report Summarizing the Devices Sold by an Executive

To apply a label filter, right-click a row or a column label, point to **Filters**, and then select one of the label filters. For example, to display the data of executives whose names begin with 'J', right-click the row label (cell A5), point to **Label Filters**, and then click **Begins With**. The **Label Filter (Executive)** dialog box is displayed. In the box, type J, as shown in figure 1.30, and then click **OK**.

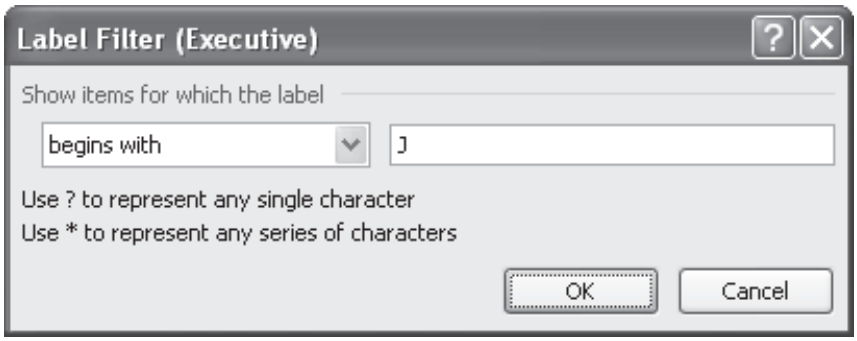


Figure 1.30: Label Filter (Executive) Dialog Box

The modified report containing data of executives whose names begin with the letter 'J' is shown in figure 1.31.

V 1.0 © 2008 Aptech Limited

	A	B	C	D	E
1					
2					
3	Sum of Quantiy	Column Labels			
4	Row Labels	Monitors	Printers	Servers	Grand Total
5	Jacob	24	4	4	32
6	4/11/2008	2	4		6
7	5/3/2008	12		2	14
8	6/4/2008	10		2	12
9	James	7	2	5	14
10	4/18/2008	4	2	5	11
11	5/10/2008	3			3
12	Jenny	6		7	13
13	5/16/2008			7	7
14	6/16/2008	6			6
15	Grand Total	37	6	16	59

Figure 1.31: PivotTable Report with Label Filter

Similarly, you can apply a date filter. For example, to filter data shown in figure 1.31, so as to display data only for the month of May, click the arrow in the row label (cell A4). Next, select the **Date** field from the **Select field** drop-down box, as shown in figure 1.32, point to **Date Filters**, point to **All Dates in the Period** and then click **May**.

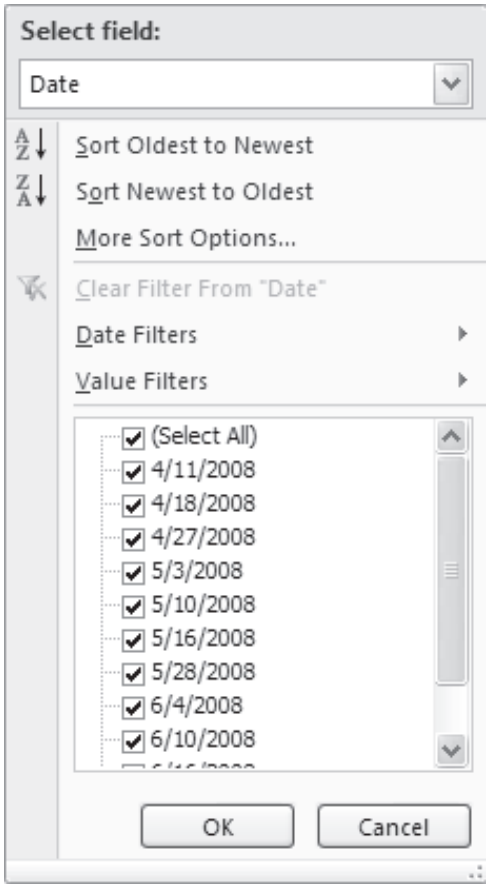


Figure 1.32: Selecting Date Field

The modified report containing data only for the month of May is shown in figure 1.33.

	A	B	C	D
1				
2				
3	Sum of Quantiy	Column Labels		
4	Row Labels	Monitors	Servers	Grand Total
5	Jacob	12	2	14
6	5/3/2008	12	2	14
7	James	3		3
8	5/10/2008	3		3
9	Jenny		7	7
10	5/16/2008		7	7
11	Grand Total	15	9	24

Figure 1.33: PivotTable Report with Date Filter

Excel also enables filtering at the report level. For example, to view data only about monitors, printers, or servers, the **Category** field is placed in the Report Filter area of the **Pivot Table Field List** pane. Figure 1.34 shows how the layout of the PivotTable report changes after the **Category** field is moved from the Column Labels area to the Report Filter area.

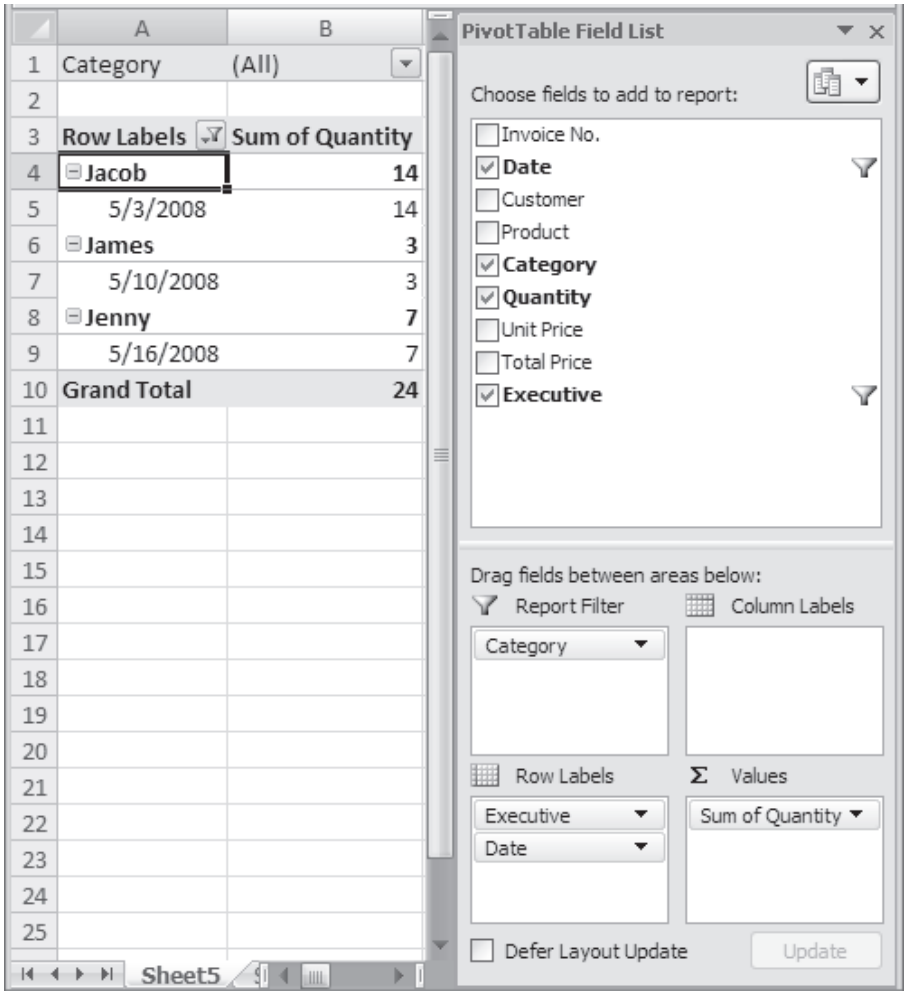


Figure 1.34: PivotTable Report with Report Filter

V 1.0 © 2008 Aptech Limited

Note that in the PivotTable report, **Category** field is displayed in the top-left corner. This area is specifically for fields added to the Report Filter area. To filter the report to display data only about servers, click the arrow in cell B1, select **Servers** and then click **OK**. The filtered report is shown in figure 1.35.

	A	B
1	Category	Servers
2		
3	Row Labels	Sum of Quantity
4	Jacob	2
5	5/3/2008	2
6	Jenny	7
7	5/16/2008	7
8	Grand Total	9

Figure 1.35: PivotTable Report with Servers Category Filter

The report shows that Jacob has sold two servers on 3rd May and Jenny has sold seven servers on 16th May.

1.4.2 Filtering Values

PivotTables also support Value filters. For example, to find out if an executive has sold more than five servers, a value filter can be applied. To do this, click the arrow in the row label (cell A4), point to **Value Filters**, and then click **Greater Than**. The V**alue Filter (Executive**) dialog box is displayed. In this dialog box, type **5** in the empty box, as shown in figure 1.36, and then click **OK**.

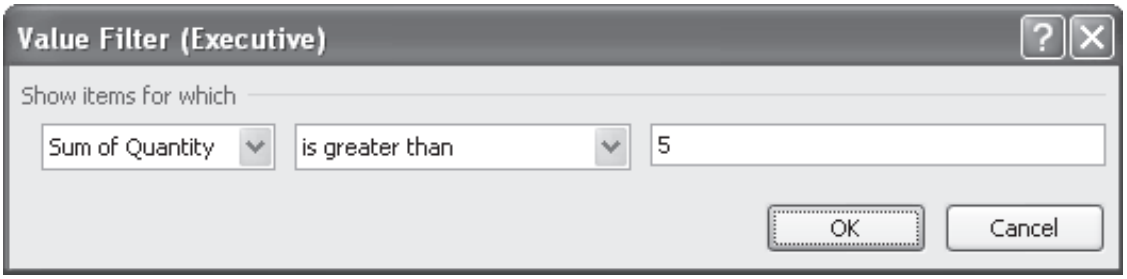


Figure 1.36: Value Filter (Executive) Dialog Box

Figure 1.37 shows that Jenny has sold more than five servers on 16th May 2008.

	A	B
1	Category	Servers
2		
3	Row Labels	Sum of Quantity
4	Jenny	7
5	5/16/2008	7
6	Grand Total	7

Figure 1.37: PivotTable Report with Value Filter

1.5 PivotChart Report

PivotChart report is a graphical representation of the data displayed in a PivotTable report. A PivotChart report cannot exist on its own. PivotTable report and PivotChart report are linked in that any changes made to the PivotTable report are automatically reflected in the PivotChart report.

1.5.1 Creating PivotChart Report

Consider the PivotTable report in figure 1.38

	A	B
1		
2		
3	Row Labels	Sum of Students
4	Postgraduate	570
5	Undergraduate	715
6	Grand Total	1285

Figure 1.38: PivotTable Report with one Row Label

To create a PivotChart report based on the PivotTable report in figure 1.38, first select a cell in the PivotTable report. Next, one of the following techniques can be used to create a PivotChart report:

- On the **Ribbon**, on the **PivotTable Tools** tab, click **Options** tab, and in the **Tools** group, click **PivotChart**. This creates a PivotChart report in the same worksheet as the PivotTable report.
- Press the F11 key on the keyboard. This creates a new worksheet named **`Chart1`** and adds the PivotChart report to **`Chart1`**. Figure 1.39 shows the **`Chart1`** worksheet created in this manner.

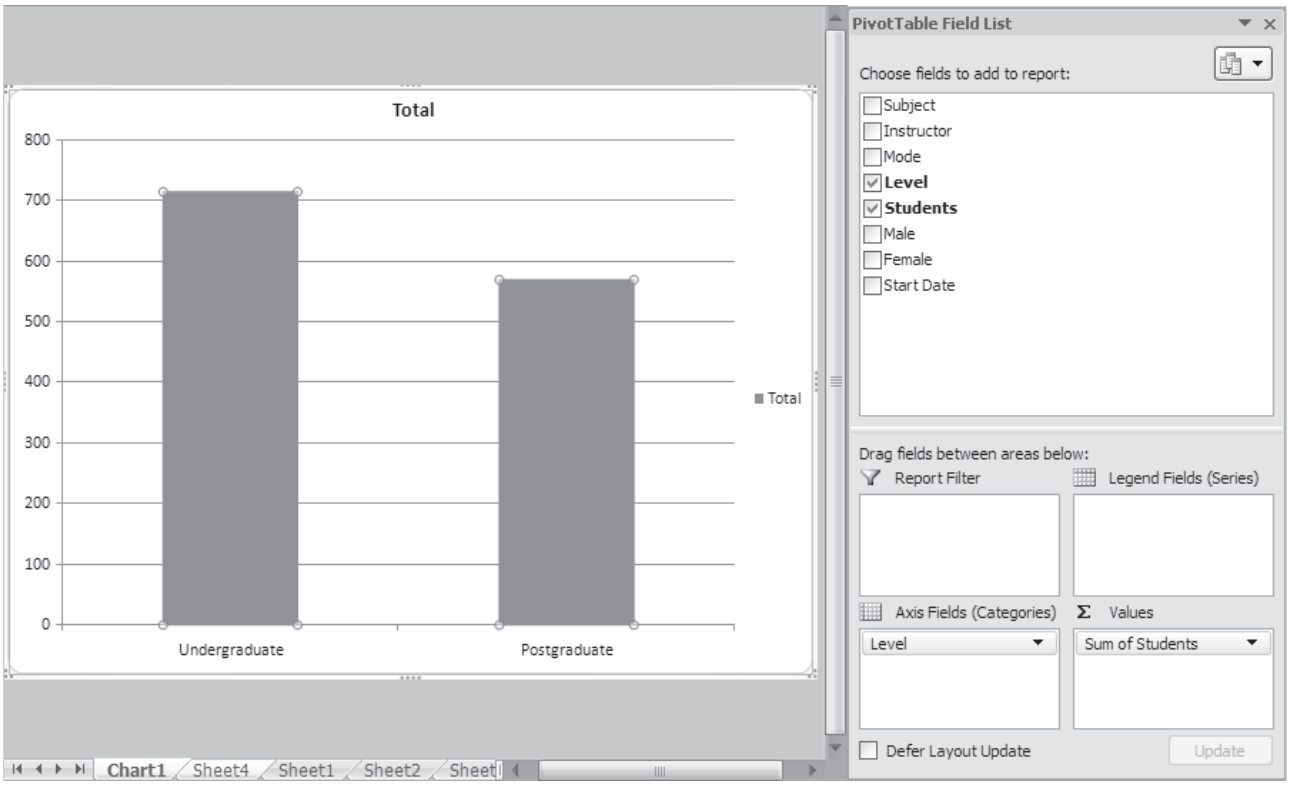


Figure 1.39: PivotChart Report

V 1.0 © 2008 Aptech Limited

The PivotChart report comprises mainly of two sections, namely the PivotChart and the **PivotTable Field List** pane. The PivotTable Field List pane comprises of the fields from the source data and the four areas listed in Table 1.2.

Area	Description
Report Filter	Fields added to this area filter the data of the PivotChart report.
Legend Fields	Fields in this area are plotted as data series in the PivotChart report on the y-axis. In figure 1.39, the total number of postgraduate and undergraduate students is plotted as data series. Column Labels fields of a PivotTable report are added to this area.
Axis Fields	Fields in this area appear as items on the x-axis. In figure 1.39, the values Postgraduate and Undergraduate of field Level are plotted on the x-axis. Row Label fields of the PivotTable report are added to this area.
Values	Fields from the PivotTable report's Values area are placed in this area.

Table 1.2: PivotChart Report's PivotTable Field List Pane's Areas

To include information about the mode (full-time or part-time) of the course, select the `Mode` field in the **PivotTable Field List** pane. Figure 1.40 shows the modified PivotChart report.

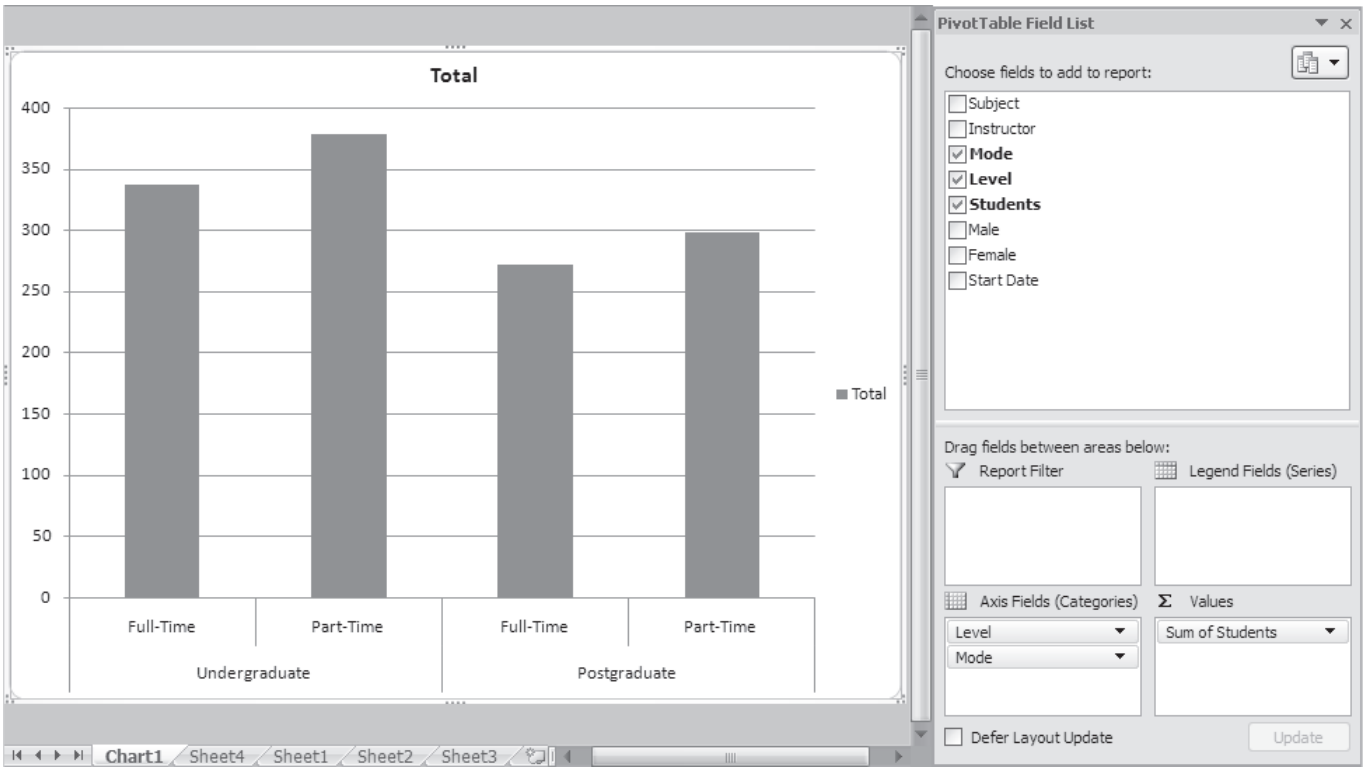


Figure 1.40: PivotChart Report with Mode Field

Note that the Axis Fields area contains two fields, namely Level and Mode. Since the `Mode` field appears below the `Level` field in the Axis Fields area, the `Mode` field is nested within the level field.

Similarly, the PivotChart report can be modified to include information about the subjects selected by the students. Figure 1.41 shows the PivotChart report with the `Subject` field's information.

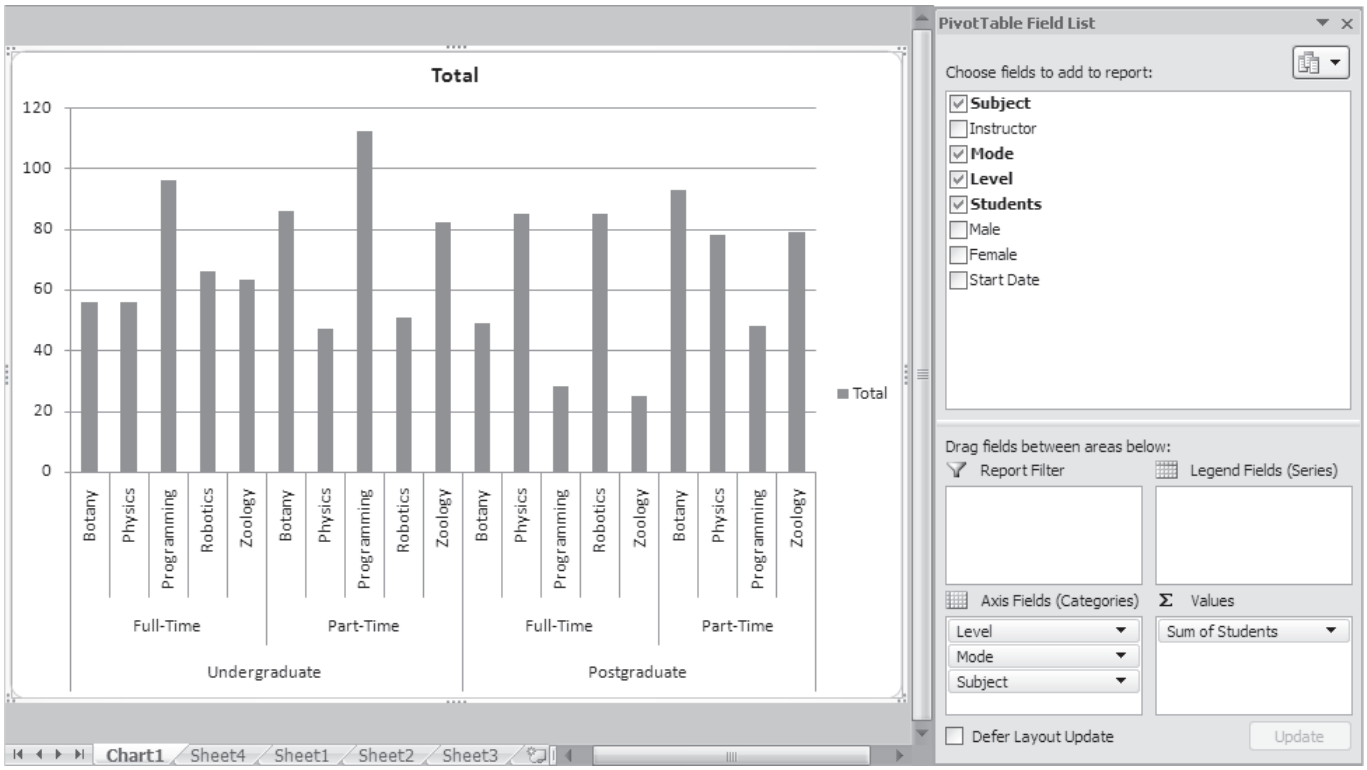


Figure 1.41: PivotChart Report with Subject Field

Note that the layout of the PivotChart report is not easy to understand. The layout of the report can be modified by dragging one or more fields to other areas in the report. For example, the **Subject** field can be dragged to the Legend Fields area to create a PivotChart report, as shown in figure 1.42.

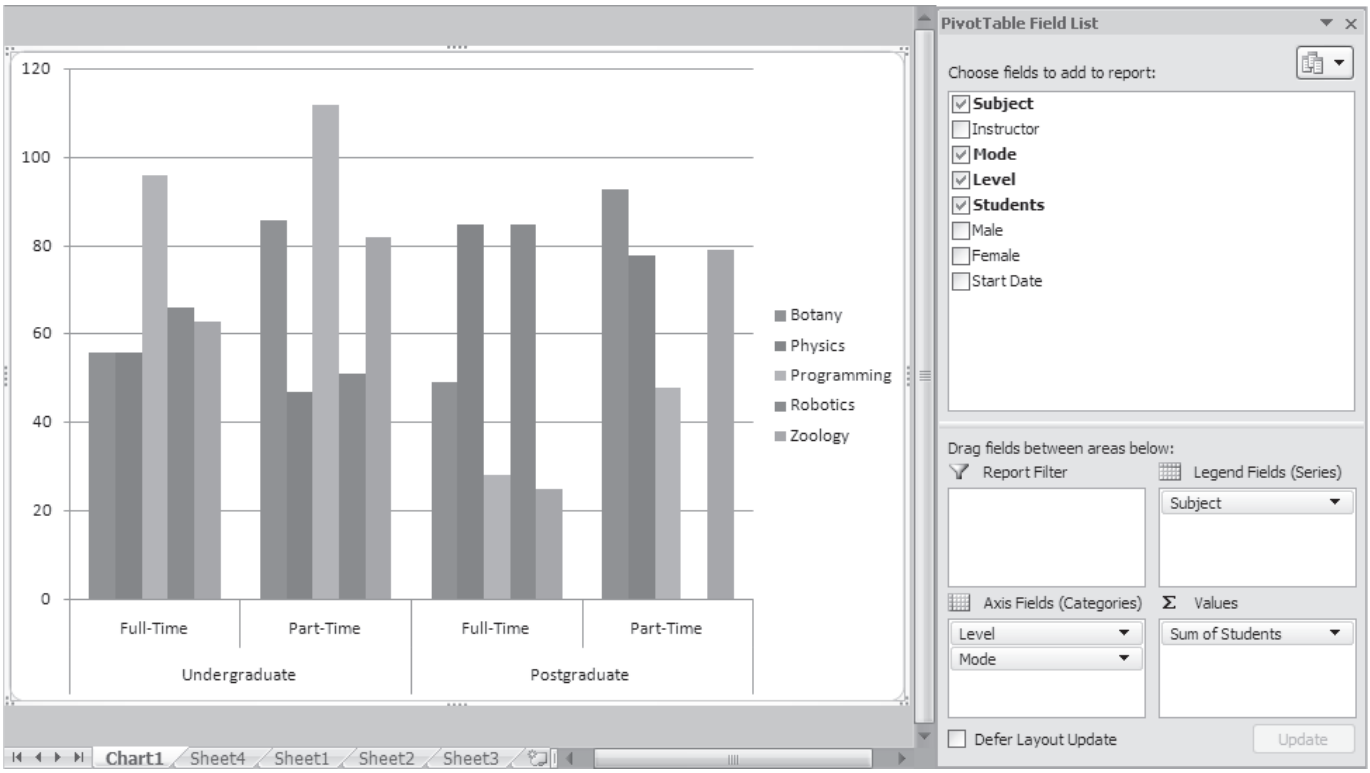


Figure 1.42: Edited PivotChart Report

V 1.0 © 2008 Aptech Limited

1.6 Customizing PivotChart Report

A PivotChart report can be customized to make it as appealing as possible. For example, you can change the chart type, chart tile, chart legend, and so on.

1.6.1 Changing the Type of Chart of the PivotChart Report

Excel provides various types of charts to select from. To change the type of chart, right-click on the chart in the PivotChart report, and then click **Change Chart Type**. The **Change Chart Type** dialog box is displayed. In this dialog box, in the **Bar** pane, click **Clustered Bar in 3-D** icon, as shown in figure 1.43, and then click **OK**. Note that the chart types XY (Scatter), Bubble, and Stock cannot be used with PivotChart reports.

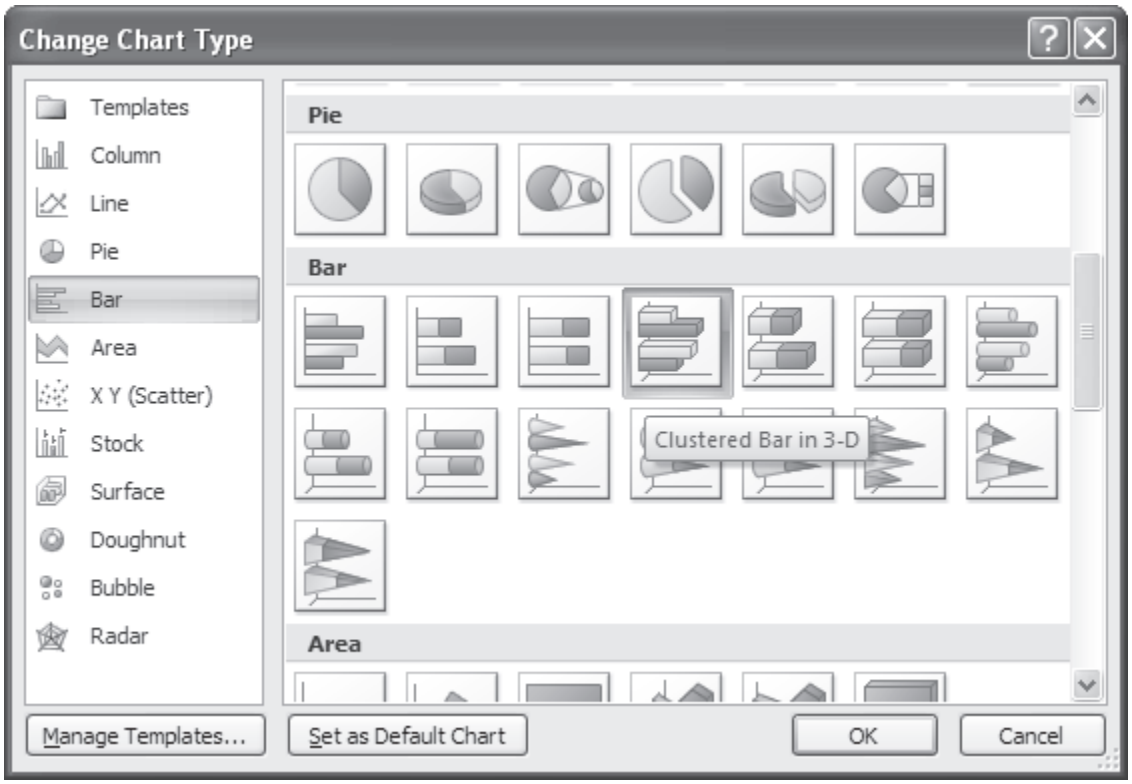


Figure 1.43: Selecting Clustered Bar in 3-D

The modified PivotChart report is displayed, as shown in figure 1.44.

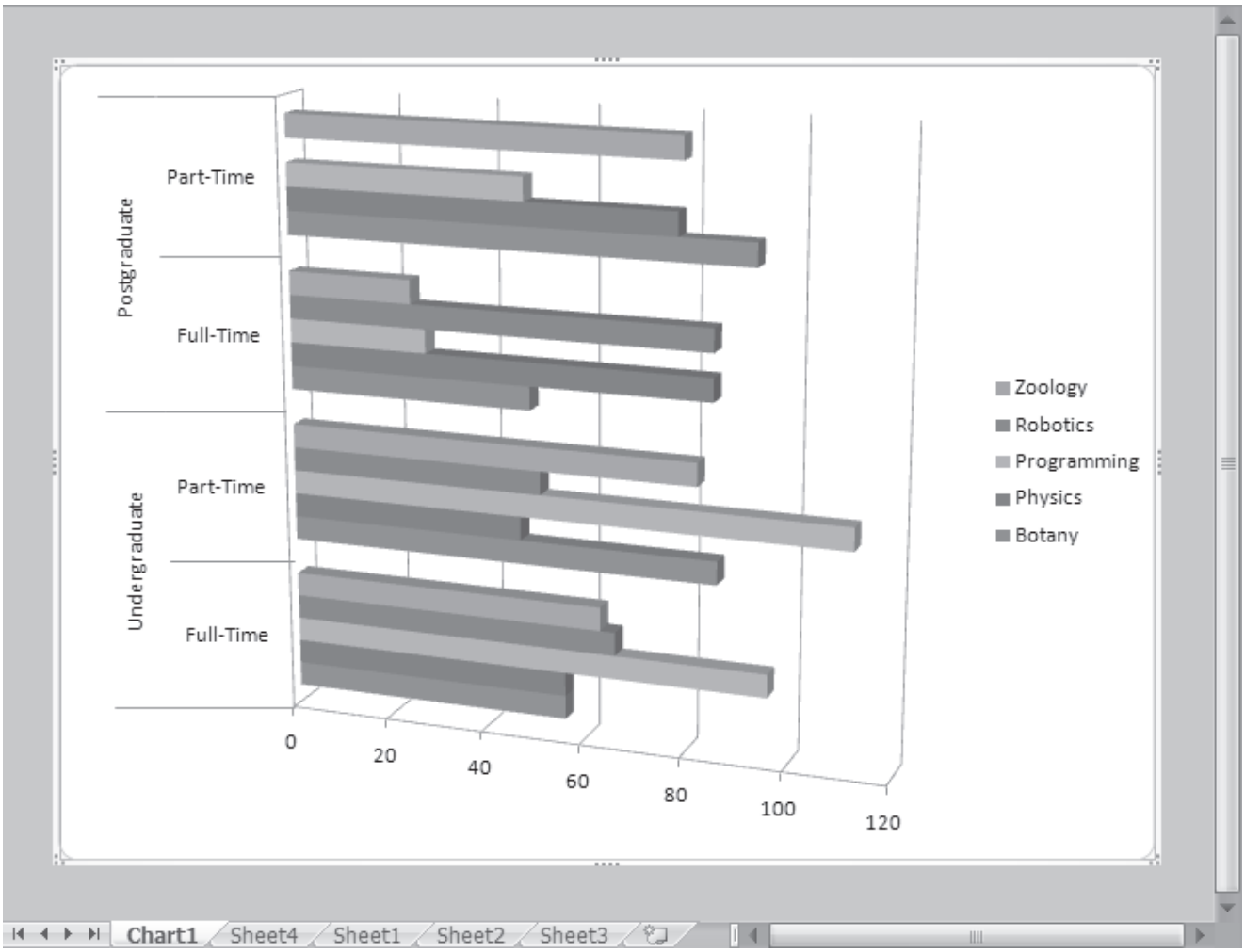


Figure 1.44: PivotChart Report with Chart Type as Clustered Bar in 3-D

Note that changing the chart type causes the chart layout, as well as the chart style, to be modified. However, you can change the chart layout and style separately, using the **Chart Layouts** and **Chart Styles** groups on the **Design** tab of the **PivotChart Tools** tab. Figure 1.45 shows the **Chart Layouts** and **Chart Styles** groups.

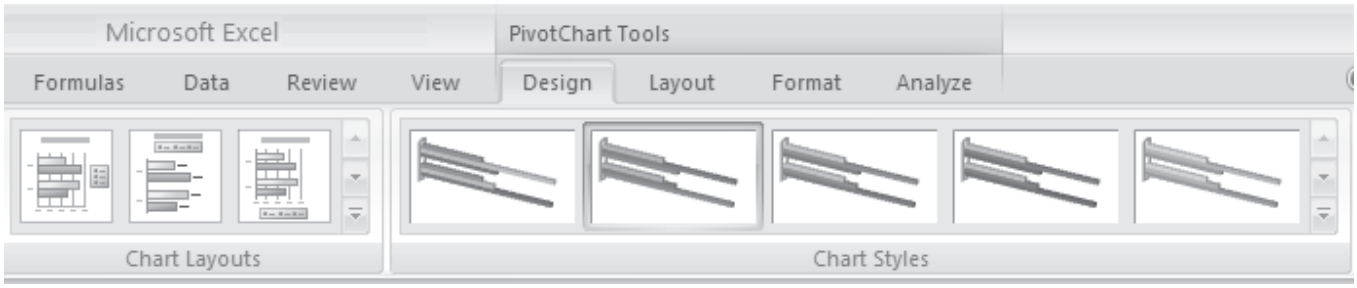


Figure 1.45: Chart Layouts and Chart Styles Groups

1.6.2 Adding Labels to the Chart

The **Labels** group of **Layout** tab provides several options to change the labels displayed in the chart. For example, you can add a chart title, data labels and also display the data table associated with the chart. Figure 1.46 shows the **Labels** group highlighted.

V 1.0 © 2008 Aptech Limited

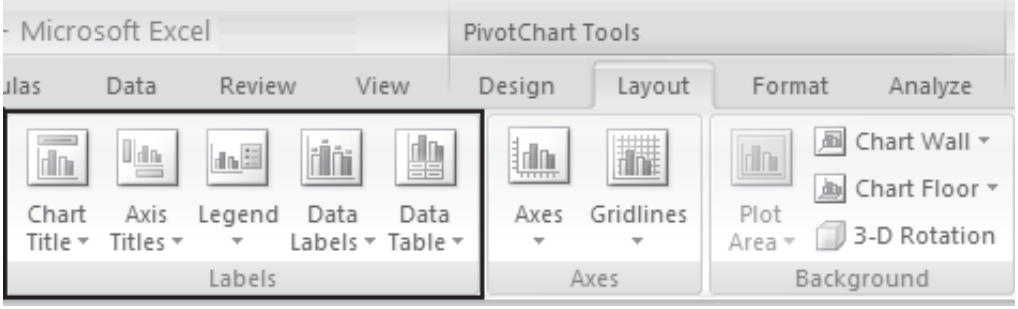


Figure 1.46: Labels Group

To add a title to the PivotChart report, in the **Labels** group, click **Chart Title** and then click **Above Chart**. A box with the text **Chart Title** is added at the top of the chart. In the box, type `Course Data`. Figure 1.47 shows the PivotChart report with the title `Course Data`.

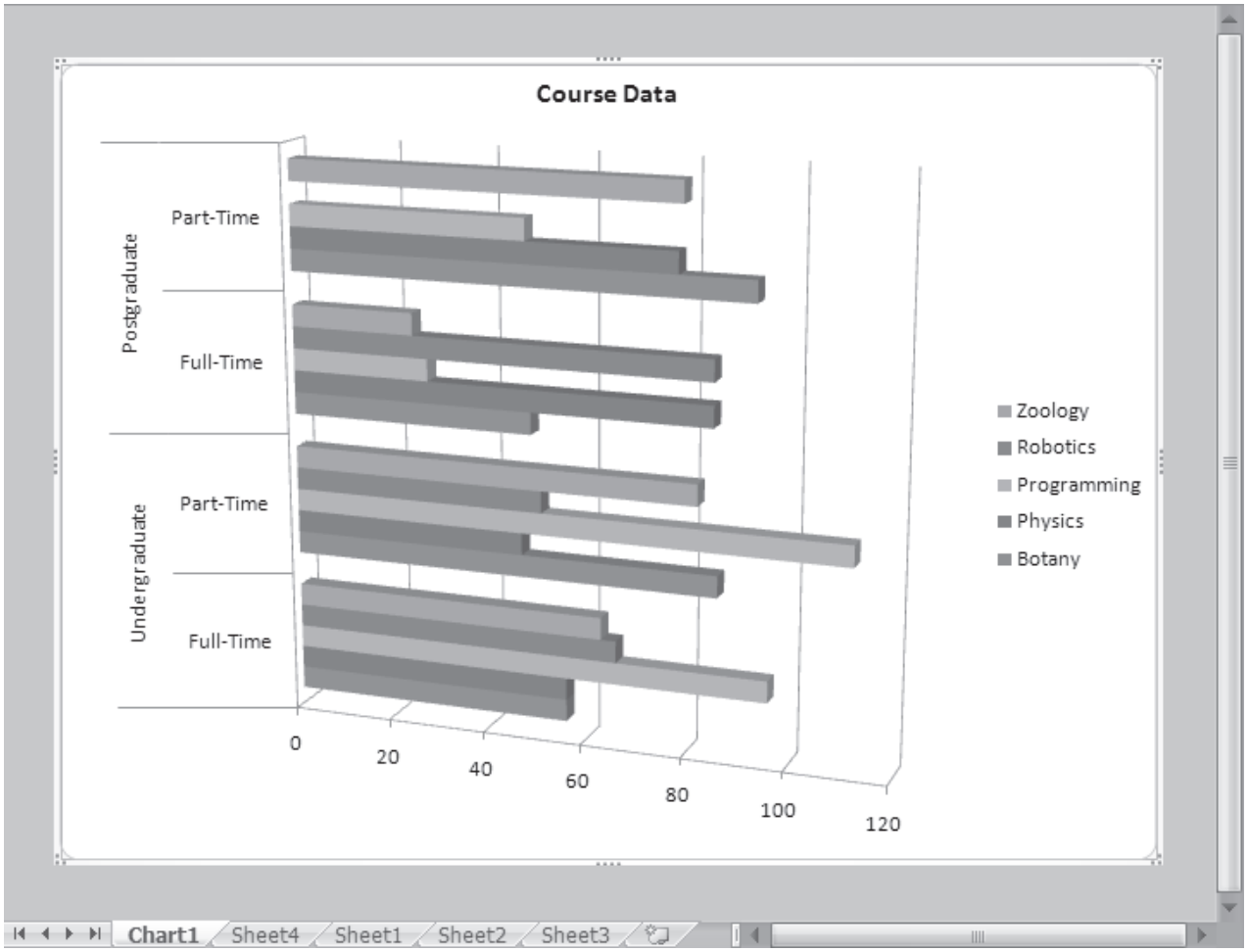


Figure 1.47: PivotChart Report with Chart Title

To add data labels representing the values of each colored bar in the chart, you can add data labels. To add a data label, in the **Labels** group, click **Data Labels**, and then click **Show**. Figure 1.48 shows the PivotChart report with data labels.

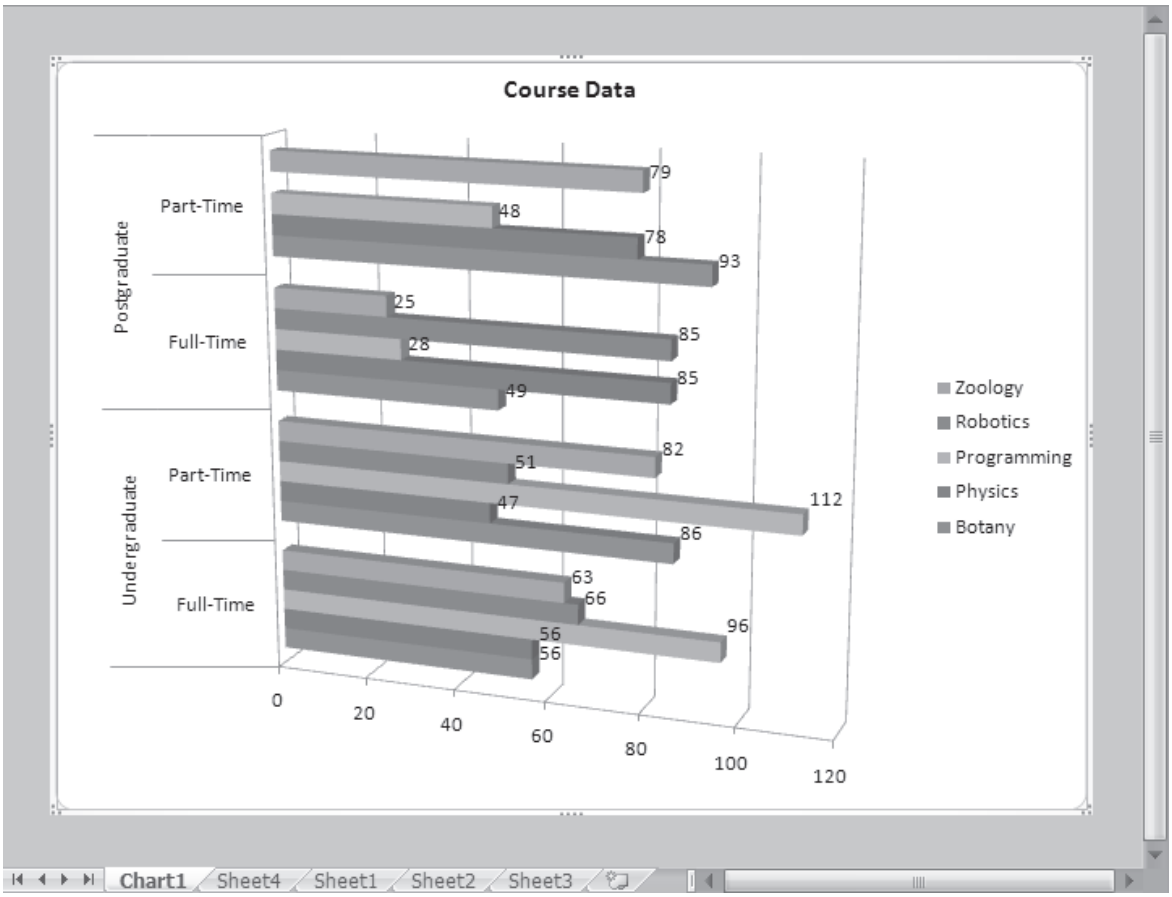


Figure 1.48: PivotChart Report with Data Labels

Similarly, the location of legends can be changed to provide more space to the chart. For example, legends can be placed horizontally at the top. To display legends horizontally at the top, in the **Labels** group, click **Legend**, and then click **Show Legend at Top**. Figure 1.49 shows legends at the top.

V 1.0 © 2008 Aptech Limited

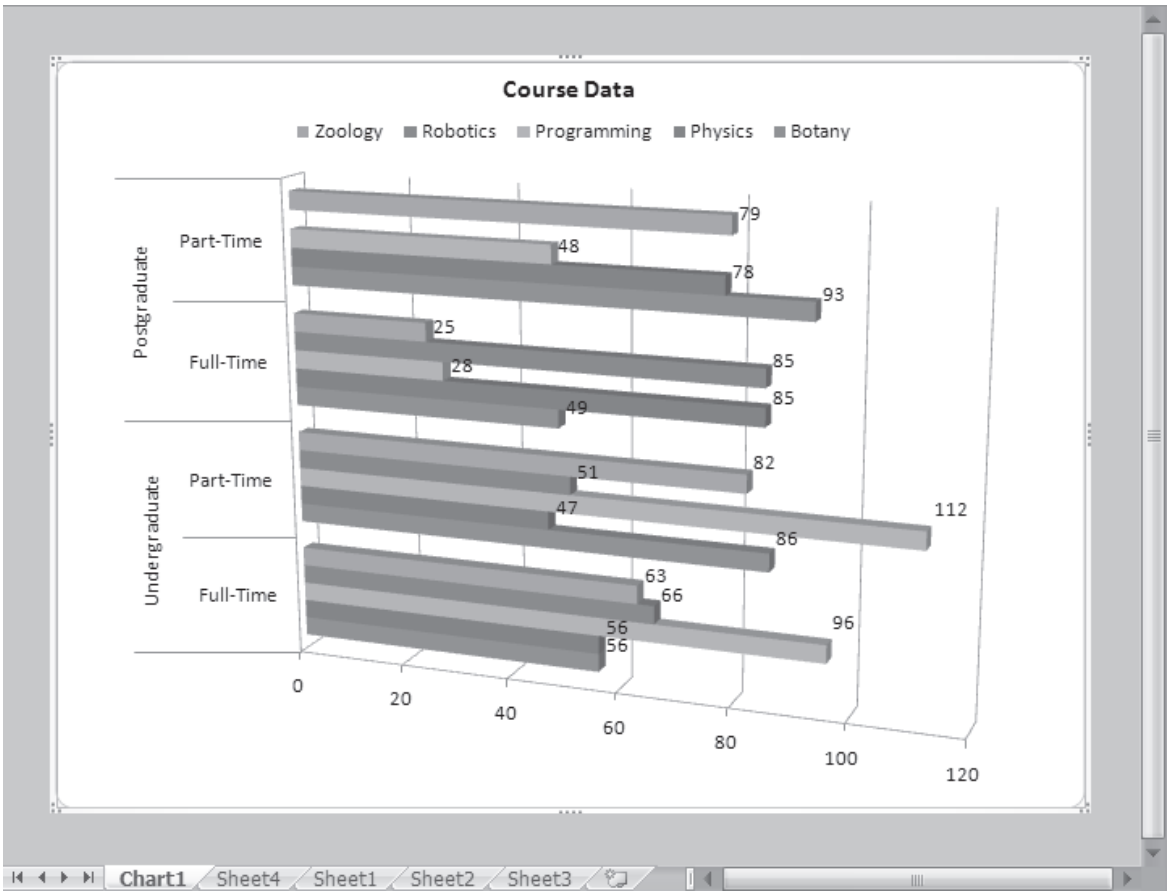


Figure 1.49: PivotChart Report with Legends at the Top

The Session In Brief

- Pivot Tables summarize, sort, and filter data.
- The PivotTable report is displayed in four areas, namely Report Filter, Column Labels, Row Labels, and Values.
- A PivotTable report is sorted using either a row label, column label, value in Values area, or a grand total.
- A PivotTable report supports label, value, date and report filter.
- A PivotChart report provides a graphical representation of the PivotTable report.
- PivotChart reports do not support the chart types XY (Scatter), Bubble, and Stock.

V 1.0 © 2008 Aptech Limited

Check Your Progress

1. A Pivot Table is used to _____ data.

   a. process

   b. print

   c. summarize

   d. display

2. The first row of source data for a pivot table should contain a _____.

   a. label

   b. formula

   c. subtotal

   d. heading

3. Fields in the Row Labels area of the PivotTable Field List pane are added to the _____ area of the PivotTable report.

   a. Report Filter

   b. Row Labels

   c. Column Labels

   d. Values

4. The option _____ sorts numeric data in descending order.

   a. Sort Largest to Smallest

   b. Sort Smallest to Largest

   c. Sort A to Z

   d. Sort Z to A

5. The label filter _____ displays labels that start with the specified character or text.

   a. Contains

   b. Begins With

   c. Does Not Contain

   d. Starts With

6. To create a PivotChart, first select a cell in the PivotTable report and then press the _____ key on the keyboard.

   a. F5

   b. F10

   c. F12

   d. F11

This page has been intentionally left blank

Session 2

Pivot Tables and Pivot Charts (Lab)

Session Objectives:

At the end of this session, you will be able to -

- Create PivotTable reports
- Create PivotChart reports

Part I - 60 Minutes

Exercise 1: PivotChart Report - MoonLight Chocolates

Problem

A chocolate store named MoonLight Chocolates has been selling chocolates since 1990. They prepare various types of chocolates such as Milk Chocolate with Almonds, Caramel Cookies Bean and so on. They have branches in San Diego, San Francisco, San Leandro, and Vacaville. The head office is in Vacaville. The head office has a data entry operator, who maintains an Excel sheet, as shown in figure 2.1.

	A	B	C	D	E	F	G
1	Year	Month	Location	Category	Name	Units	Revenue
2	2007	August	Vacaville	Chocolate Bars	Milk Chocolate with Almonds	805	$ 3,187.80
3	2007	March	Vacaville	Hot Cocoa	Milk Chocolate Hazlenut Hot Cocoa	992	$ 3,412.48
4	2007	January	San Francisco	Chocolate Bars	Caramel Cookies Bean	712	$ 1,808.48
5	2006	March	San Leandro	Hot Cocoa	Milk Chocolate Hazlenut Hot Cocoa	904	$ 2,260.00
6	2006	January	Vacaville	Hot Cocoa	Milk Chocolate Hazlenut Hot Cocoa	647	$ 2,076.87
7	2005	September	San Diego	Intense Dark Milk Chocolate	Mint Bliss	739	$ 1,707.09
8	2006	March	San Francisco	Chocolate Bars	Milk Chocolate with Almonds	974	$ 2,181.76
9	2007	June	San Leandro	Intense Dark Milk Chocolate	Twilight Delight	615	$ 1,894.20
10	2007	July	San Diego	Intense Dark Milk Chocolate	Citrus Sunset	714	$ 1,856.40
11	2006	May	Vacaville	SQUARES Milk Chocolates	Sublime White Vanila Dream	703	$ 1,553.63
12	2005	June	San Diego	Hot Cocoa	Milk Chocolate Hazlenut Hot Cocoa	528	$ 2,064.48
13	2006	October	San Francisco	SQUARES Milk Chocolates	Dark Milk Chocolate with Mint	644	$ 1,809.64
14	2005	June	Vacaville	Intense Dark Milk Chocolate	Toffee Interlude	919	$ 2,196.41
15	2007	May	San Diego	Chocolate Bars	Caramel Cookies Bean	767	$ 1,932.84
16	2007	June	San Diego	SQUARES Milk Chocolates	Dark Milk Chocolate with Mint	984	$ 1,987.68
17	2005	March	Vacaville	SQUARES Milk Chocolates	Dark Milk Chocolate with Mint	744	$ 2,217.12
18	2007	September	San Francisco	Hot Cocoa	Milk Chocolate Hazlenut Hot Cocoa	693	$ 2,189.88
19	2006	October	San Francisco	Chocolate Bars	Milk Chocolate	658	$ 1,895.04
20	2005	November	San Francisco	Hot Cocoa	Milk Chocolate Mocha Hot Cocoa	878	$ 3,274.94
21	2005	May	Vacaville	Hot Cocoa	Milk Chocolate Mocha Hot Cocoa	848	$ 3,281.76
22	2007	August	Vacaville	Chocolate Bars	Milk Chocolate with Almonds	547	$ 1,247.16
23	2005	June	San Diego	Hot Cocoa	Milk Chocolate Hazlenut Hot Cocoa	664	$ 1,513.92
24	2007	April	Vacaville	Hot Cocoa	Milk Chocolate Hazlenut Hot Cocoa	882	$ 2,716.56
25	2007	February	San Diego	Intense Dark Milk Chocolate	Mint Bliss	930	$ 2,408.70

Sheet1 Sheet2 Sheet3

Figure 2.1: Source Data

The Excel sheet contains information about the revenue generated from the various chocolates sold in the four branches.

The CEO of MoonLight Chocolates wants a 3-D Column chart providing the following information about the San Diego branch, for the year 2007:

- Chocolate category and names sold in the branch
- Number of units sold for each chocolate
- Total number of units for each chocolate category
- Revenue generated by each chocolate and its corresponding chocolate category

Additionally, the CEO would like the legends in the chart to appear at the top along with a chart title 'San Diego – 2007 Sales' at the top. In addition, the 3-D columns in the chart should display their respective values.

Analysis

The source data has several columns, namely Year, Month, Location, Category, Name, Units and Revenue. The CEO wants a 3-D column chart. Therefore, the first step is to create a PivotTable report and then use it to create the PivotChart report. The CEO wants to know the sales data of the San Diego branch for the year 2007. Therefore, the report will have the fields, **`Location`** and **`Year`** as Report Filters. The chocolate category and chocolate name is available in the **`Category`** and **`Name`** fields respectively. Therefore, these fields can be added to the Row Labels area. The number of units sold for a chocolate is stored in the **`Units`** field and the revenue information is stored in the **`Revenue`** field. Therefore, these two fields can be added to the Values area to provide information about the number of units and revenue generated by each chocolate. The PivotChart report can then be created from the PivotTable report. The PivotChart can then be customized to include 3-D column, legends at the top, chart title and data labels.

Solution

The steps to a create PivotTable report are as follows:

1. **Select a cell in the source data present in Sheet1.**
2. **On the Ribbon, click Insert → PivotTable → PivotTable.**

The **Create PivotTable** dialog box is displayed, as shown in figure 2.2.

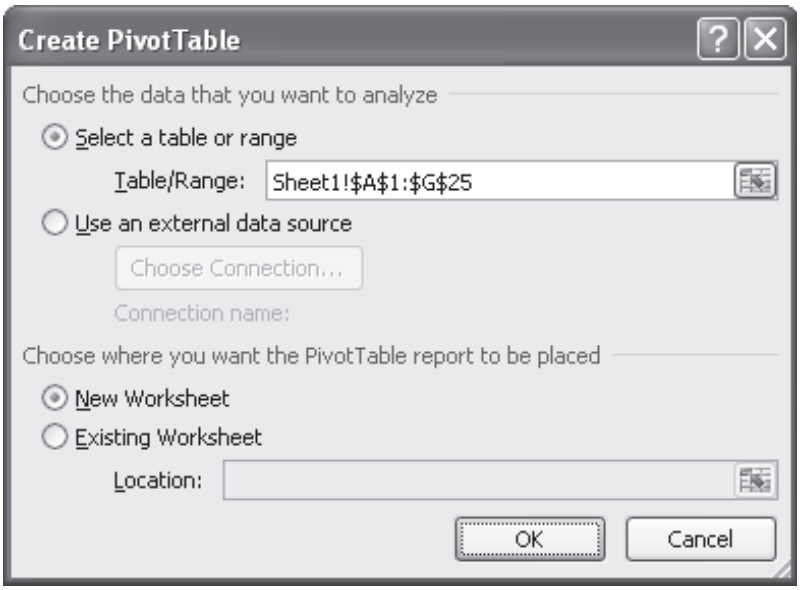


Figure 2.2: Create PivotTable Dialog Box

3. **Click OK on the Create PivotTable dialog box.**

V 1.0 © 2008 Aptech Limited

A new worksheet named **Sheet4** is automatically created and an empty PivotTable report is added to the worksheet. Figure 2.3 shows the empty PivotTable report created.

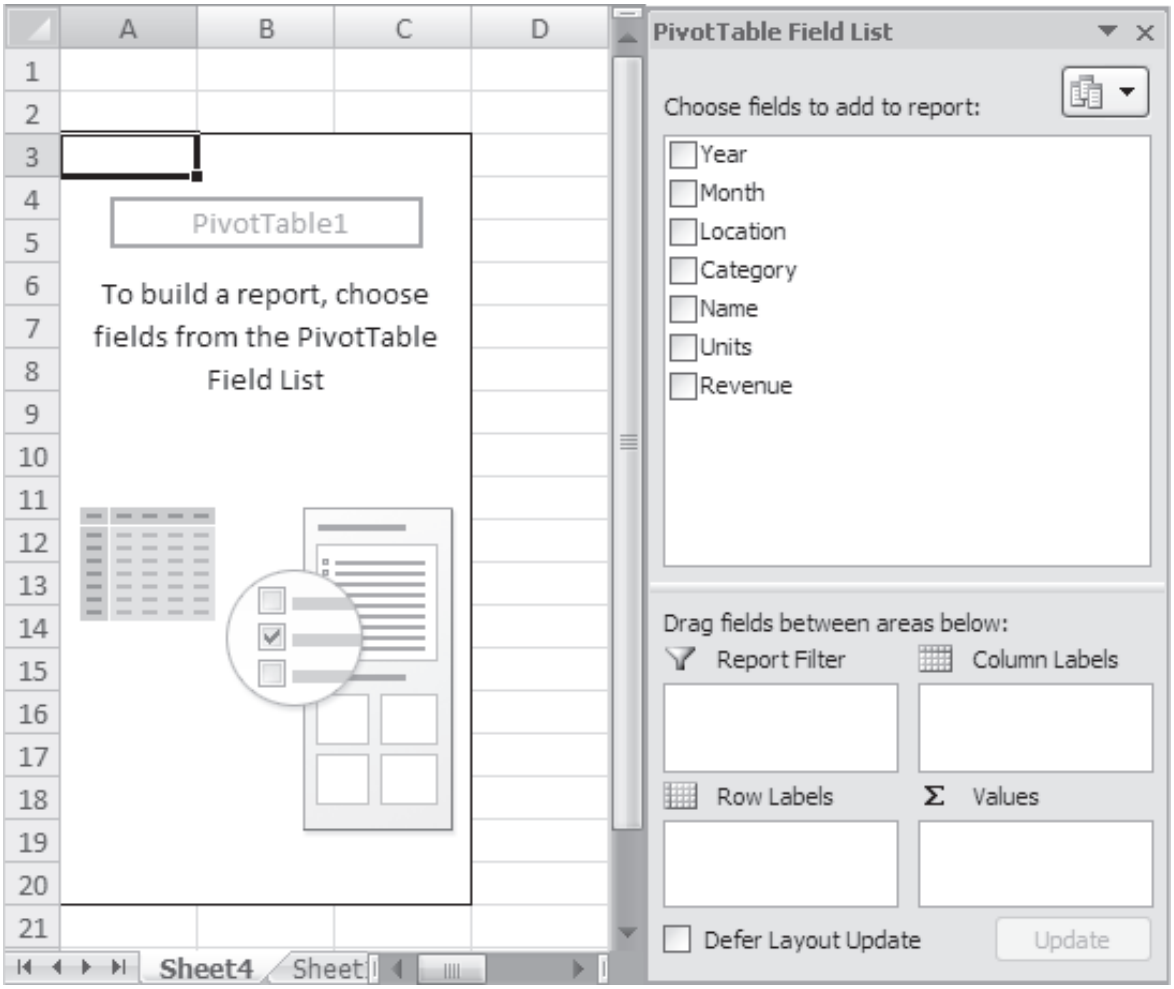


Figure 2.3: Empty PivotTable Report

4. To display information only about specific branches, in the PivotTable Field List pane, drag the Location field to the Report Filter area.

Figure 2.4 shows the **Location** field added to the Report Filter area in the PivotTable report.

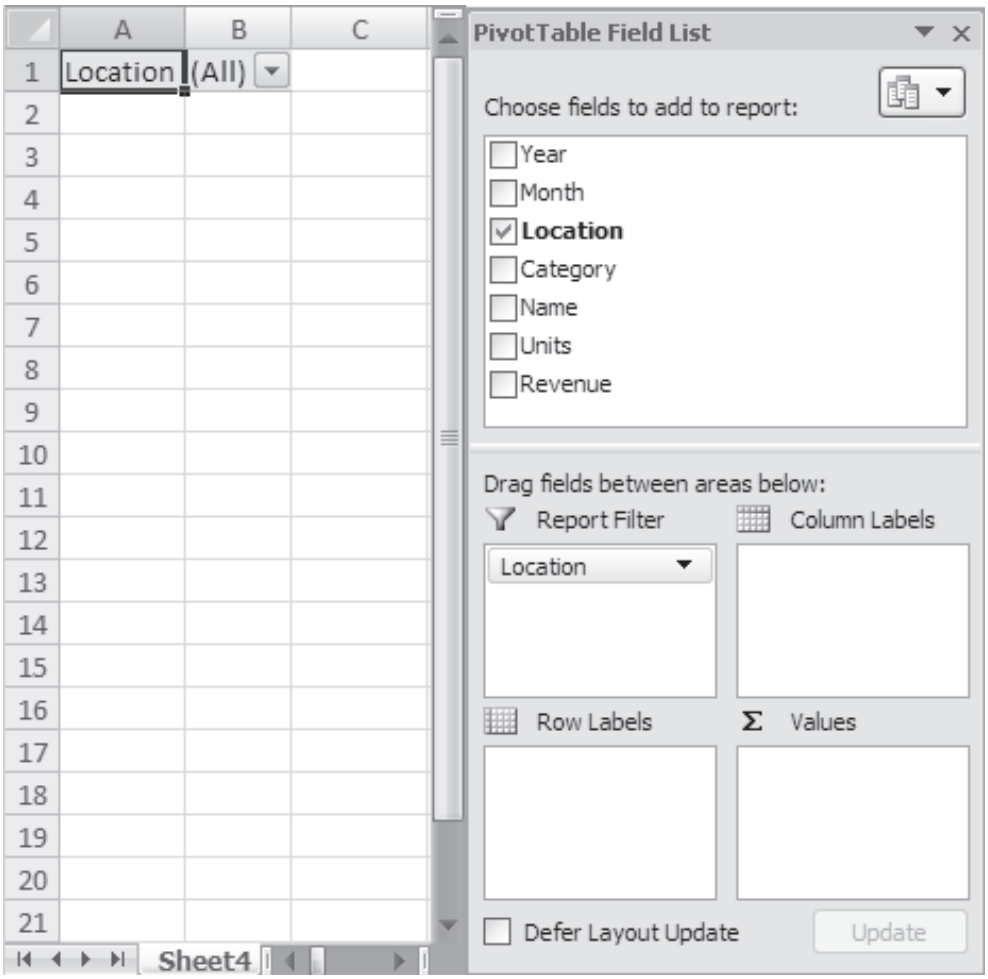


Figure 2.4: Location Field added to the Report Filter area

5. To display information only about the San Diego branch, click the arrow in cell B1.

A list displaying all the branches is displayed as shown in figure 2.5.

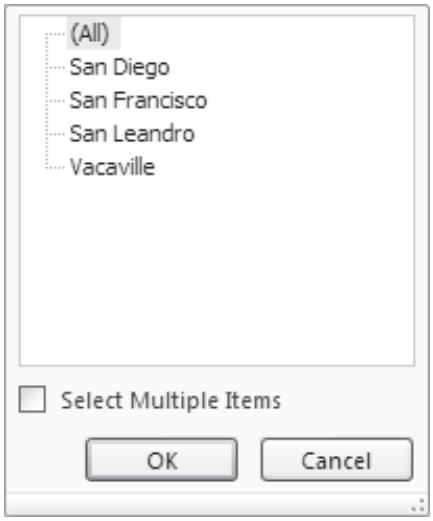


Figure 2.5: Filter displaying all the branches

V 1.0 © 2008 Aptech Limited

6. **To select the San Diego branch, in the list, click San Diego and then click OK.**
7. **To display information about a specific year, in the PivotTable Field List pane, drag the Year field to the Report Filter area.**

Figure 2.6 shows the `Year` field added to the Report Filter area in the PivotTable report.

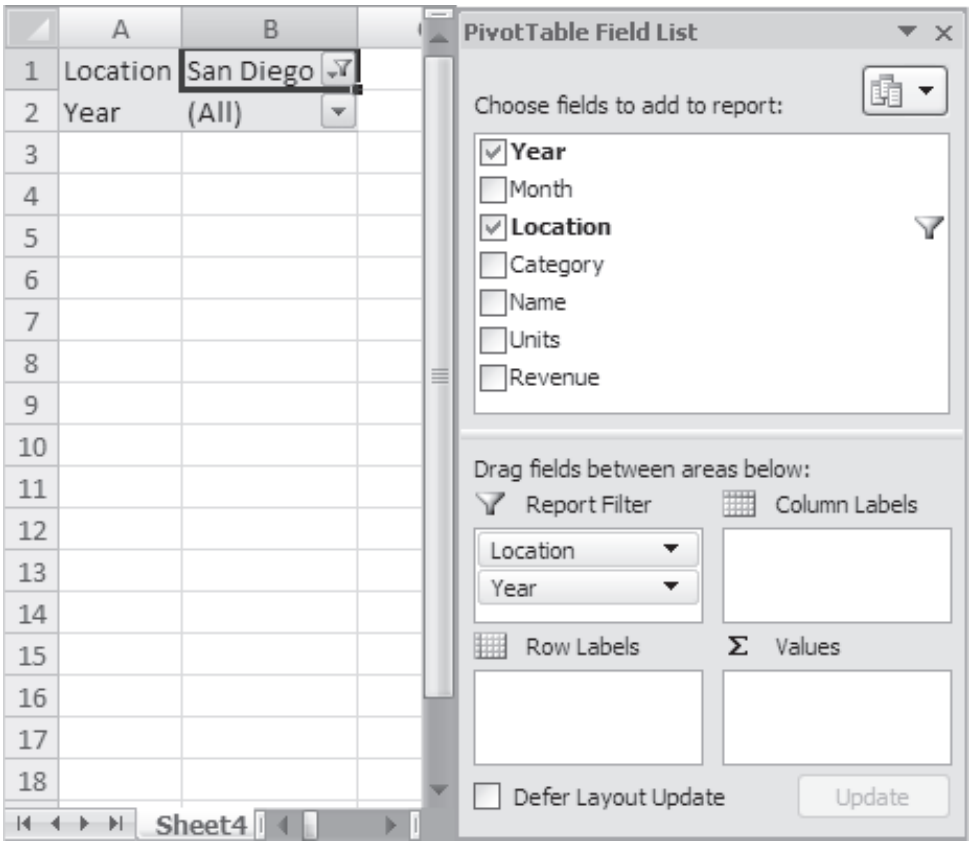


Figure 2.6: Year Field added to the Report Filter area

8. **To display information only about the year 2007, click the arrow in cell B2.**

A list displaying all the years is displayed as shown in figure 2.7.

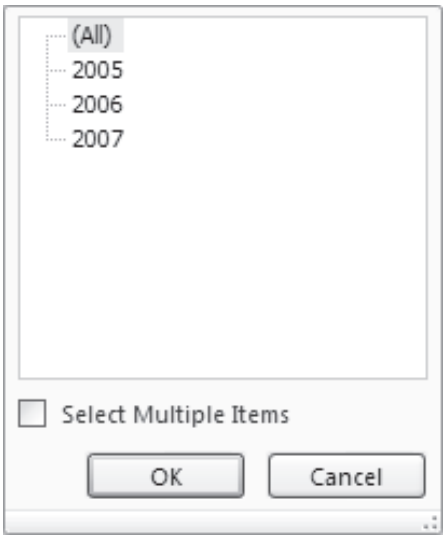


Figure 2.7: Filter displaying all the Years

9. To select the year, 2007, in the list, click 2007 and then click OK.

The two report filters added to the PivotTable report are displayed, as shown in figure 2.8.

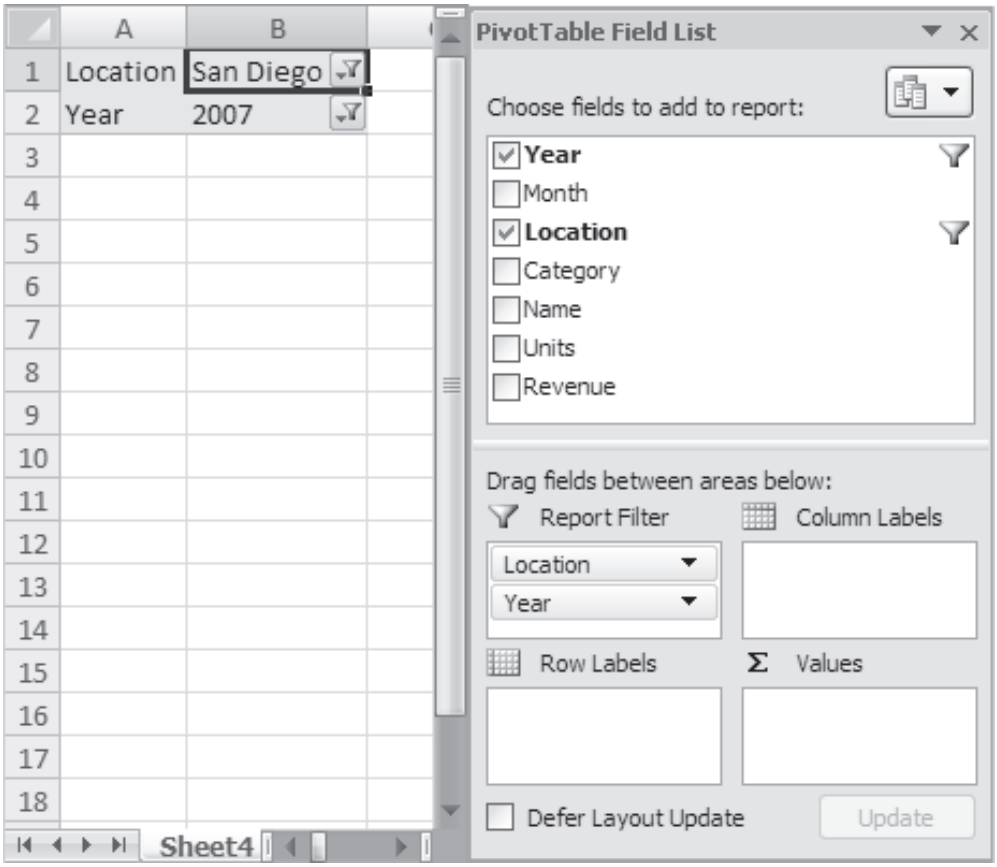


Figure 2.8: Year and Location Filters added to the Report

10. To display each chocolate category in the report, in the PivotTable Field List pane, click the Category check box.

11. To display all the chocolate names in each category, in the PivotTable Field List pane, select the Name check box.

Figure 2.9 shows the PivotTable report, listing each chocolate category and the chocolate names.

V 1.0 © 2008 Aptech Limited

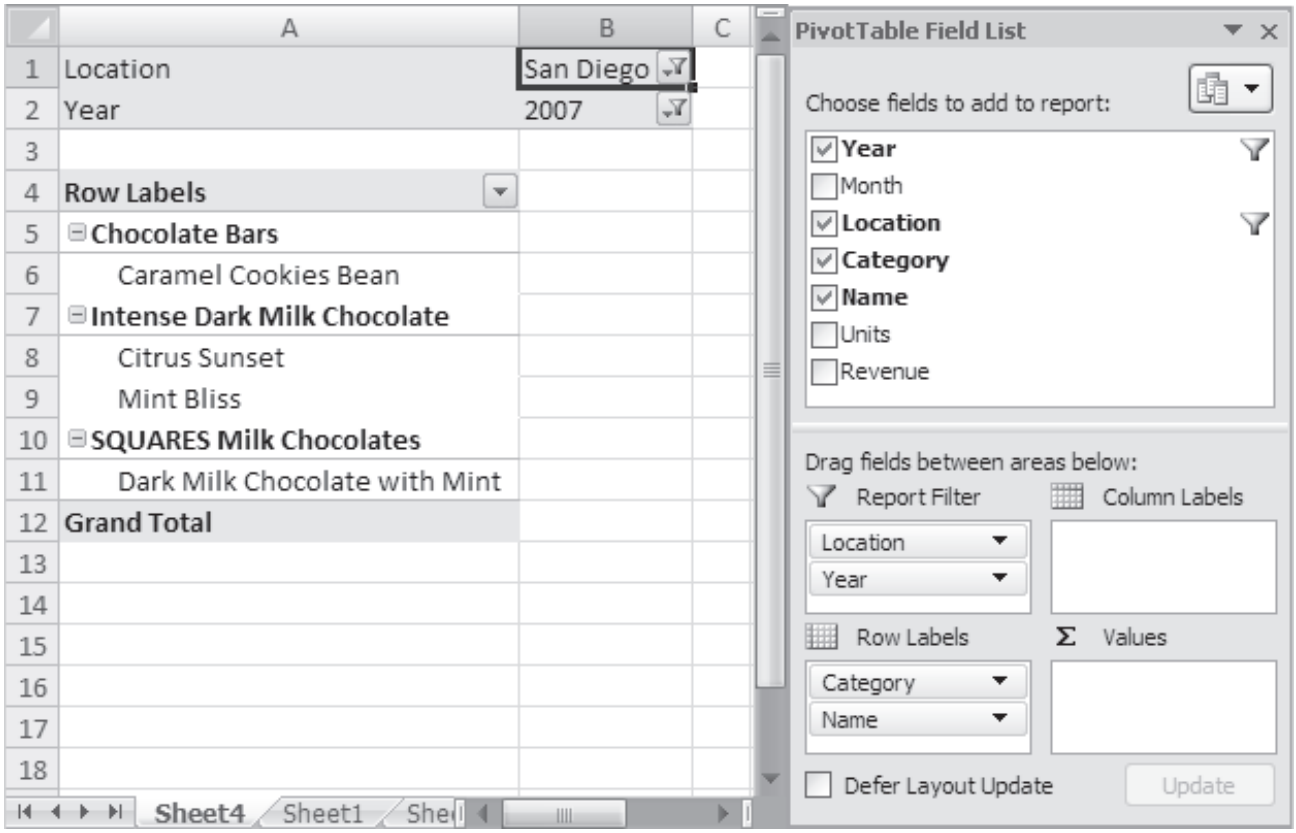
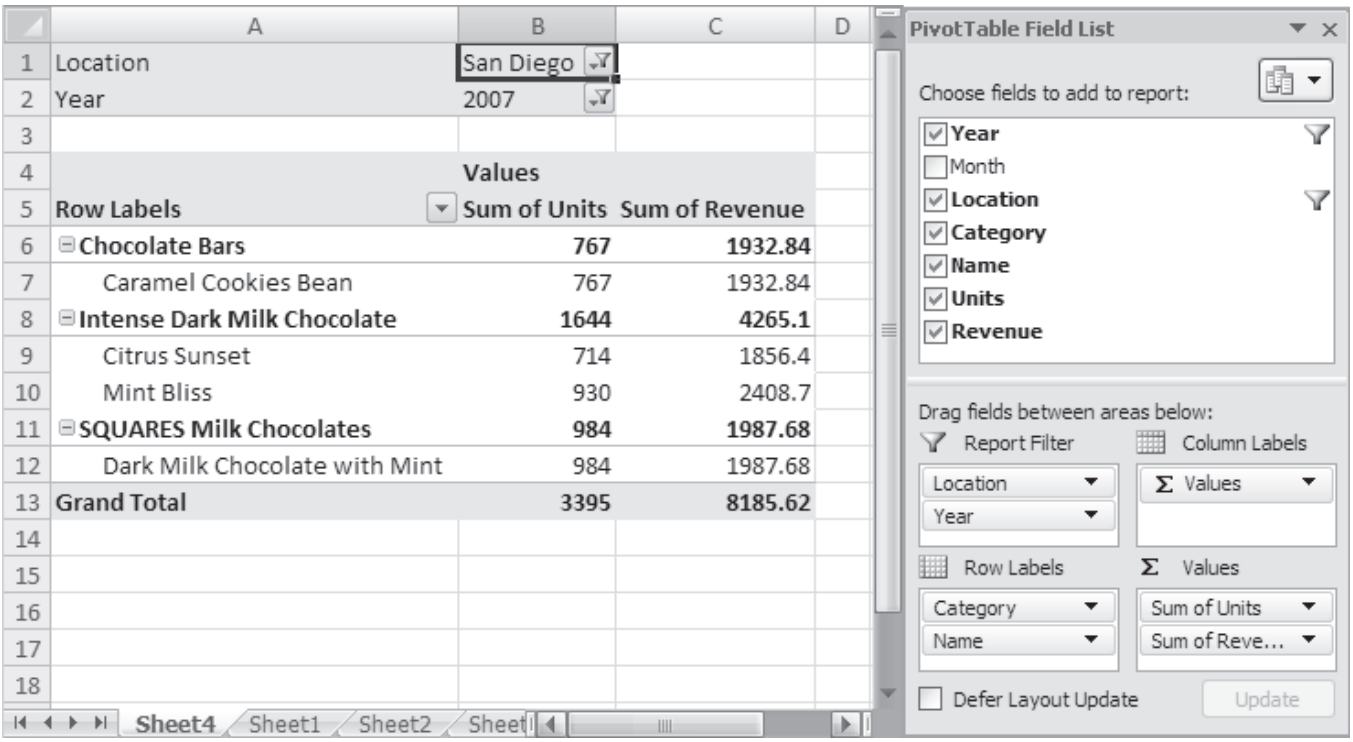


Figure 2.9: PivotTable Report displaying Chocolate Categories and Names

12. **To display the number of units sold for each chocolate, in the PivotTable Field List pane, select the Units check box.**
13. **To display the revenue generated for each chocolate, in the PivotTable Field List pane, select the Revenue check box.**

Figure 2.10 shows the report with the number of units sold for and revenue generated by each chocolate. This report displays San Diego branch's data for the year 2007.



	A	B	C
1	Location	San Diego	
2	Year	2007	
3			
4		Values	
5	Row Labels	Sum of Units	Sum of Revenue
6	Chocolate Bars	767	1932.84
7	Caramel Cookies Bean	767	1932.84
8	Intense Dark Milk Chocolate	1644	4265.1
9	Citrus Sunset	714	1856.4
10	Mint Bliss	930	2408.7
11	SQUARES Milk Chocolates	984	1987.68
12	Dark Milk Chocolate with Mint	984	1987.68
13	Grand Total	3395	8185.62

Figure 2.10: San Diego Branch – Year 2007 Data

The PivotTable report is complete with all the required data. Next, the PivotChart report is to be created. The steps to create the PivotChart report required by the CEO are as follows:

14. Select a cell in the PivotTable report.

15. Press F11 key on the keyboard.

A PivotChart report, as shown in figure 2.11, is created in new worksheet named `Chart1`.

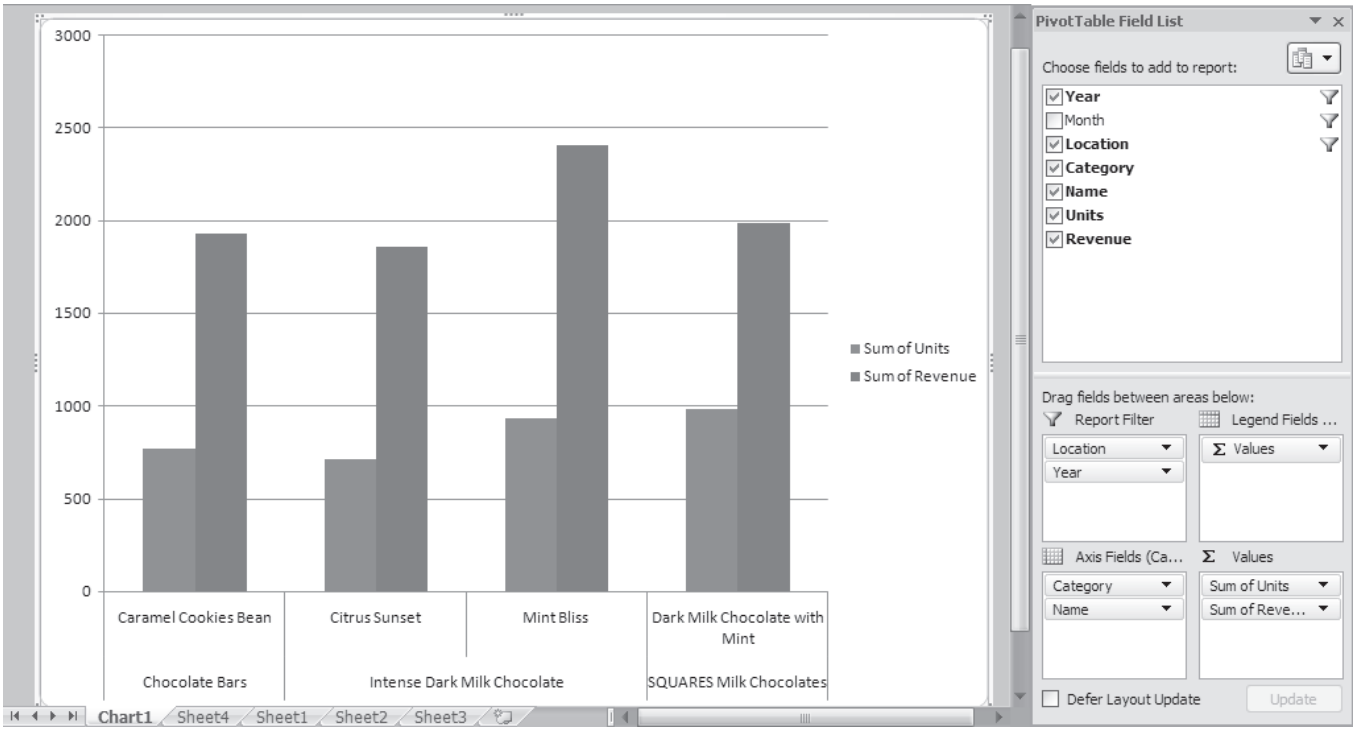


Figure 2.11: PivotChart Report

V 1.0 © 2008 Aptech Limited

16. To change the chart type to 3-D Column, right-click in the Chart Area and then click Change Chart Type.

The **Change Chart Type** dialog box is displayed.

17. To select 3-D Column chart type, click the 3-D Column icon, as shown in figure 2.12, and then click OK.

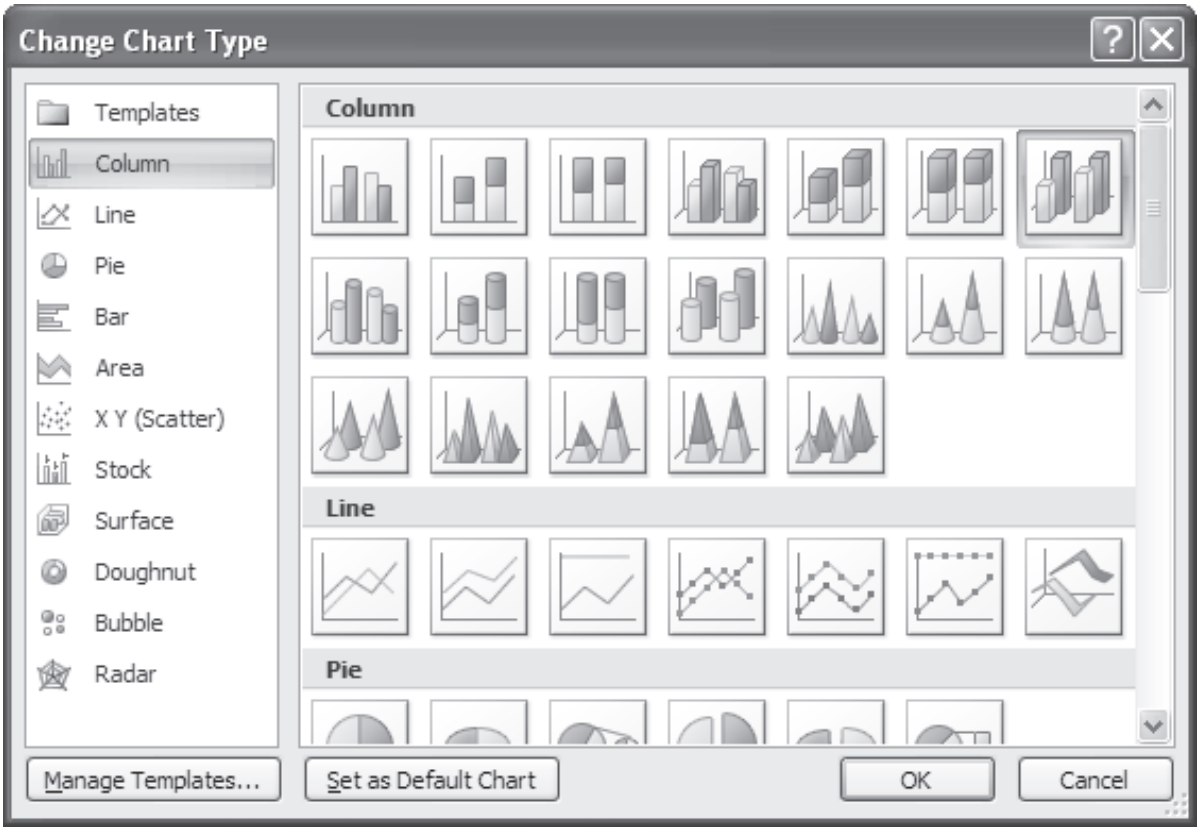


Figure 2.12: Selecting 3-D Column Chart

The PivotChart report with 3-D Column chart type is displayed in figure 2.13.

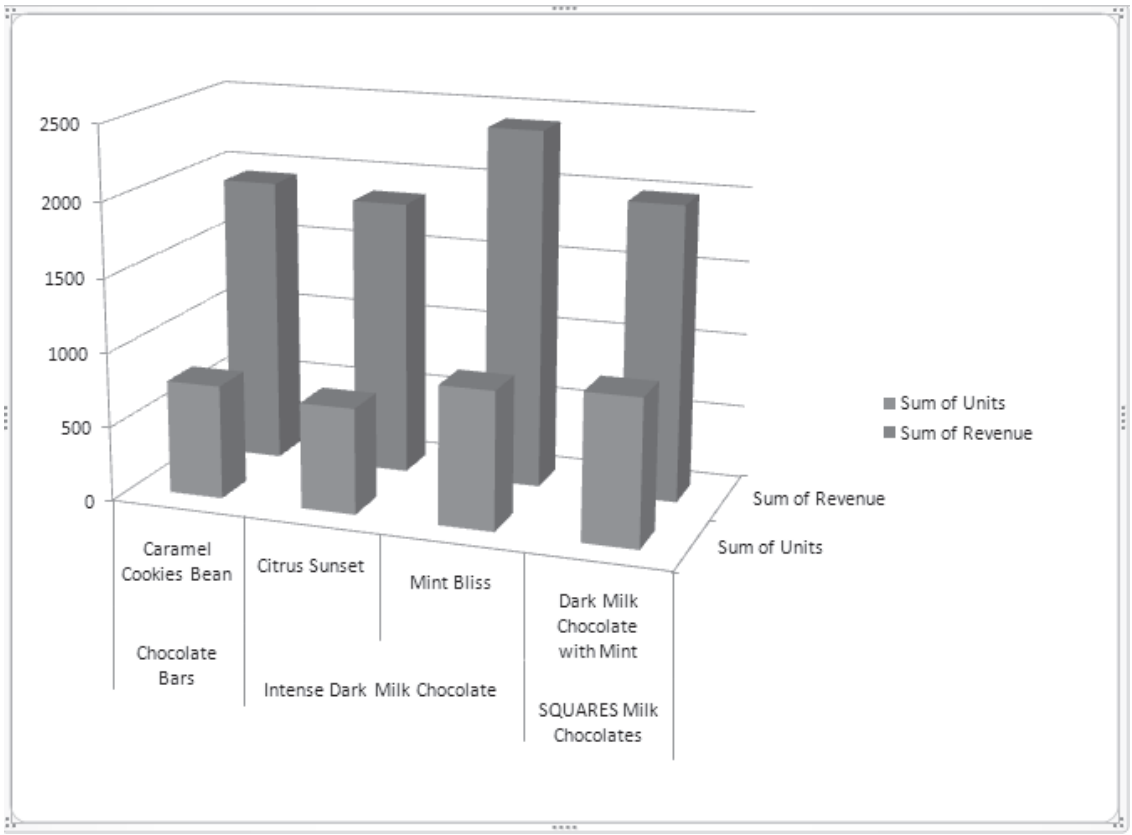


Figure 2.13: PivotChart Report with 3-D Column Chart Type

18. **To display the legends at the top, on the PivotTable Tools tab, click Layout → Legend → Show Legend at Top.**
19. **To display a chart title at the top, on the PivotTable Tools tab, click Layout → Chart Title → Above Chart.**
20. **In the Chart Title box, type San Diego – 2007 Sales.**

Figure 2.14 shows the PivotChart report with a chart title and legends at the top.

V 1.0 © 2008 Aptech Limited

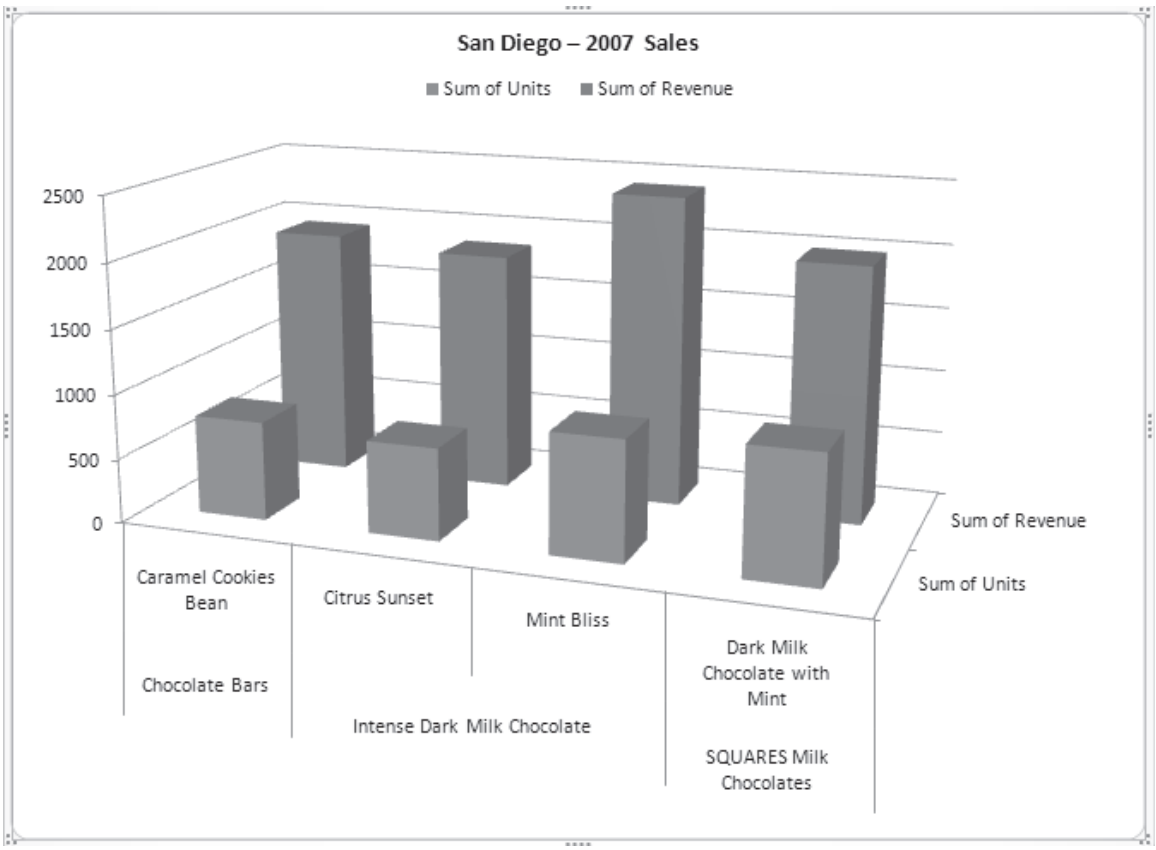


Figure 2.14: PivotChart Report – legends and Chart Title

21. To display the values of units and revenue generated by each chocolate, on the PivotTable Tools tab, click Layout → Data Labels → Show.

The final PivotChart report is displayed in figure 2.15.

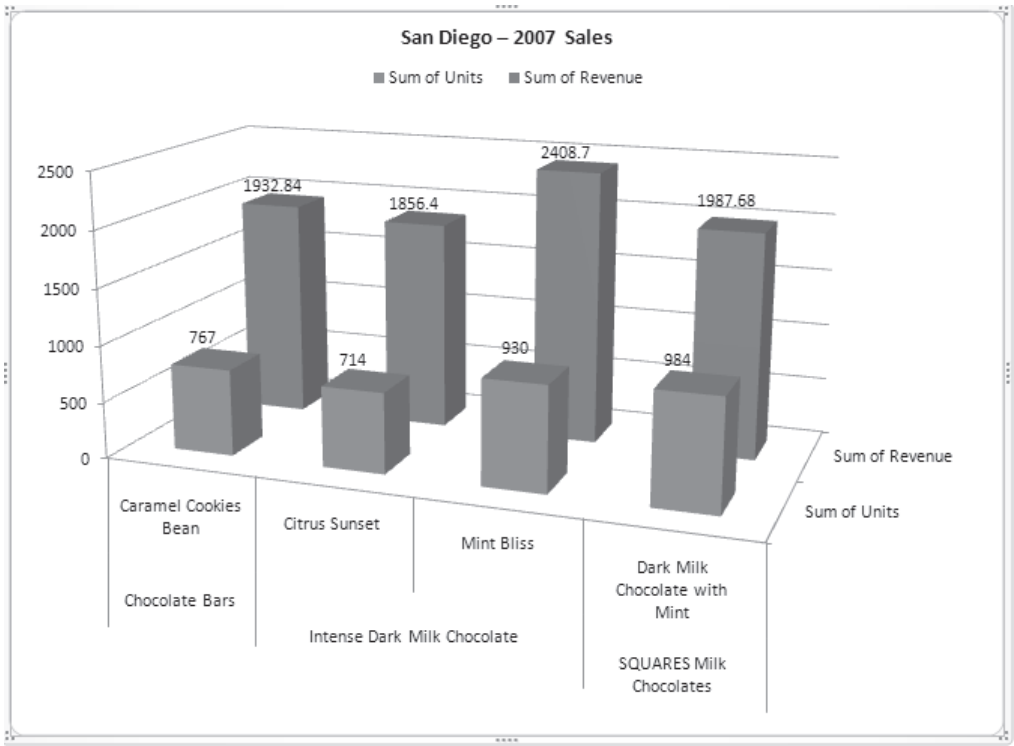


Figure 2.15: PivotChart Report – San Diego, 2007 Sales Data

Exercise 2: PivotTable Report - Citizen International

Problem

Citizen International sells beverages. The organization has branches in the East, West, North, South, and Northwest regions of California. Every sales executive is supposed to provide the half-yearly sales data to the Sales Manager. The Sales Manager wants all the sales executives to provide the following information in their reports:

- List of all the order IDs
- Date of each order
- Order amount for each region
- Order amounts of orders between the period 01/01/2007 and 30/06/2007

The sales data of each executive is maintained in an Excel sheet as shown in figure 2.16.

	A	B	C	D	E	F
1	Order ID	Executive	Order Date	Region	Order Amount	
2	1101	Roman	1/10/2007	North	$1,000.00	
3	1102	Cathy	4/13/2007	South	$777.00	
4	1103	Simon	8/17/2007	North	$2,023.00	
5	1104	Glen	12/10/2007	West	$1,234.56	
6	1105	Glen	2/14/2007	East	$14,456.00	
7	1106	Cathy	6/14/2007	South	$3,123.00	
8	1107	Simon	7/16/2007	West	$8,564.00	
9	1108	Roman	10/8/2007	South	$3,454.80	
10	1109	Cathy	3/20/2007	East	$449.00	
11	1110	Roman	5/10/2007	North	$1,500.00	
12	1111	Cathy	9/15/2007	East	$1,926.00	
13	1112	Glen	11/13/2007	NorthWest	$50.00	
14	1113	Simon	2/10/2007	West	$987.00	
15	1114	Ronald	5/21/2007	North	$1,750.00	
16	1115	Roman	10/13/2007	NorthWest	$555.00	
17	1116	Ronald	12/17/2007	NorthWest	$290.00	
18	1117	Glen	9/10/2007	South	$567.80	
19	1118	Simon	1/15/2007	East	$1,111.80	
20	1119	Roman	4/20/2007	North	$867.00	
21	1120	Ronald	6/13/2007	West	$500.00	
22						

Sheet1 Sheet2 Sheet3

Figure 2.16: Citizen International's Sales Data

One of the sales executives Cathy, wants to prepare her sales data.

Analysis

The PivotTable report is to be prepared for Cathy. Therefore, the **`Executive`** field should be used as a report filter. In other words, the **`Executive`** field should be added to the Report Filter area. The Sales Manager wants the order ID and date of all the orders. Therefore, the fields **`Order ID`** and **`Order Date`** should be added to the Row Labels area. The order amounts of each order can be displayed in the report by adding the **`Order Amount`** field to the

V 1.0 © 2008 Aptech Limited

Values area. Finally, the order amounts should be displayed for each region. Hence, the **Region** field should be added to the Row Labels area. However, two labels, namely **Order ID** and **Order Date**, are already added to the report. Therefore, to keep the report's layout simple and easy to understand, the **Region** field can be added to the Column Labels area.

Solution

1. **To create PivotTable report for the source data in figure 2.16, select a cell in the source data.**
2. **On the Ribbon, click Insert → PivotTable → PivotTable.**

The **Create PivotTable** dialog box is displayed, as shown in figure 2.17.

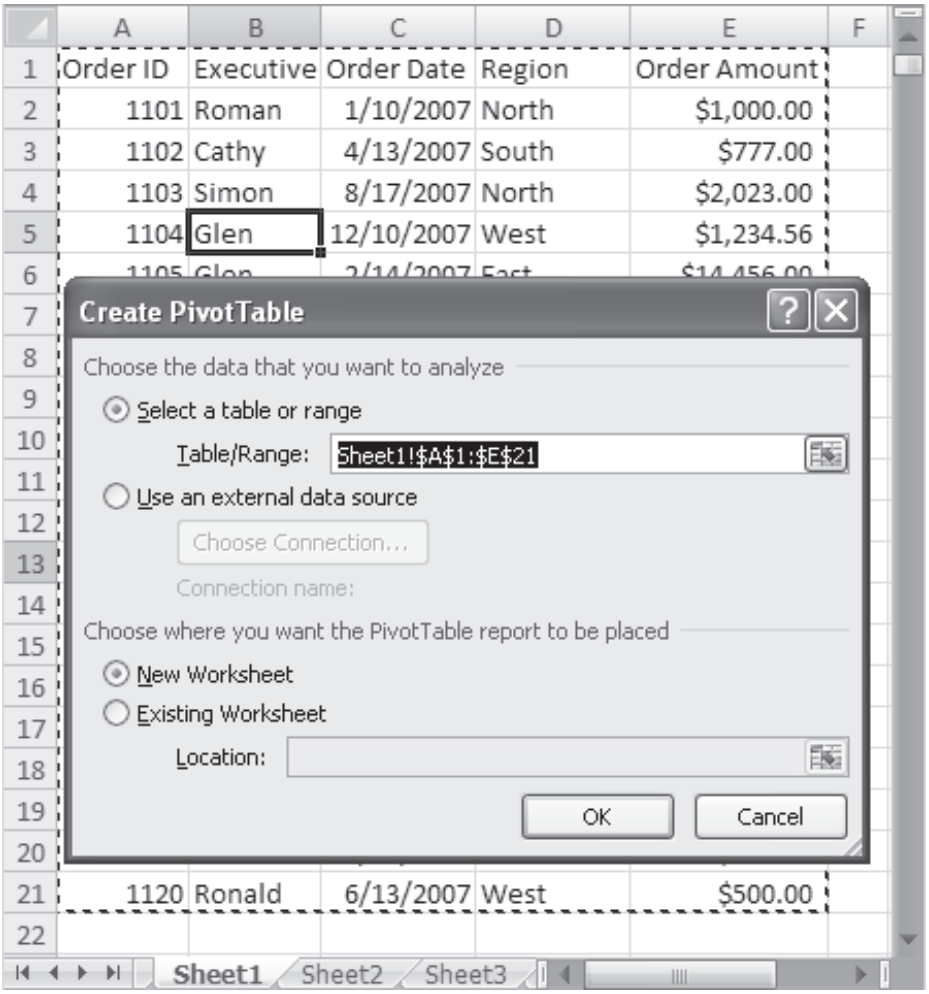


Figure 2.17: Create PivotTable dialog box

3. **Click OK on the Create PivotTable dialog box.**

A new worksheet named **Sheet4** is automatically created and an empty PivotTable report is added to the worksheet. Figure 2.18 shows the empty PivotTable report created.

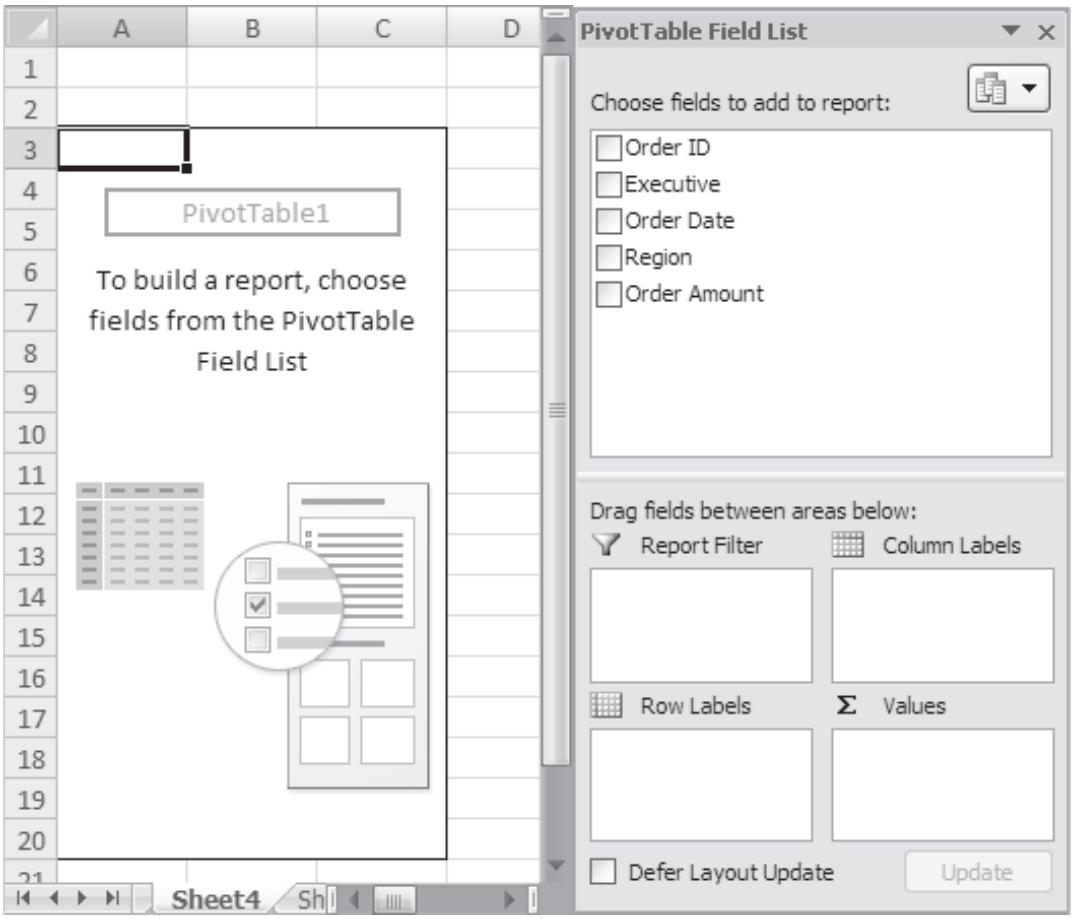


Figure 2.18: Empty PivotTable Report

4. To display sales data about specific executives, in the PivotTable Field List pane, drag the Executive field to the Report Filter area.

Figure 2.19 shows the **`Executive`** field added to the Report Filter area in the PivotTable report.

V 1.0 © 2008 Aptech Limited

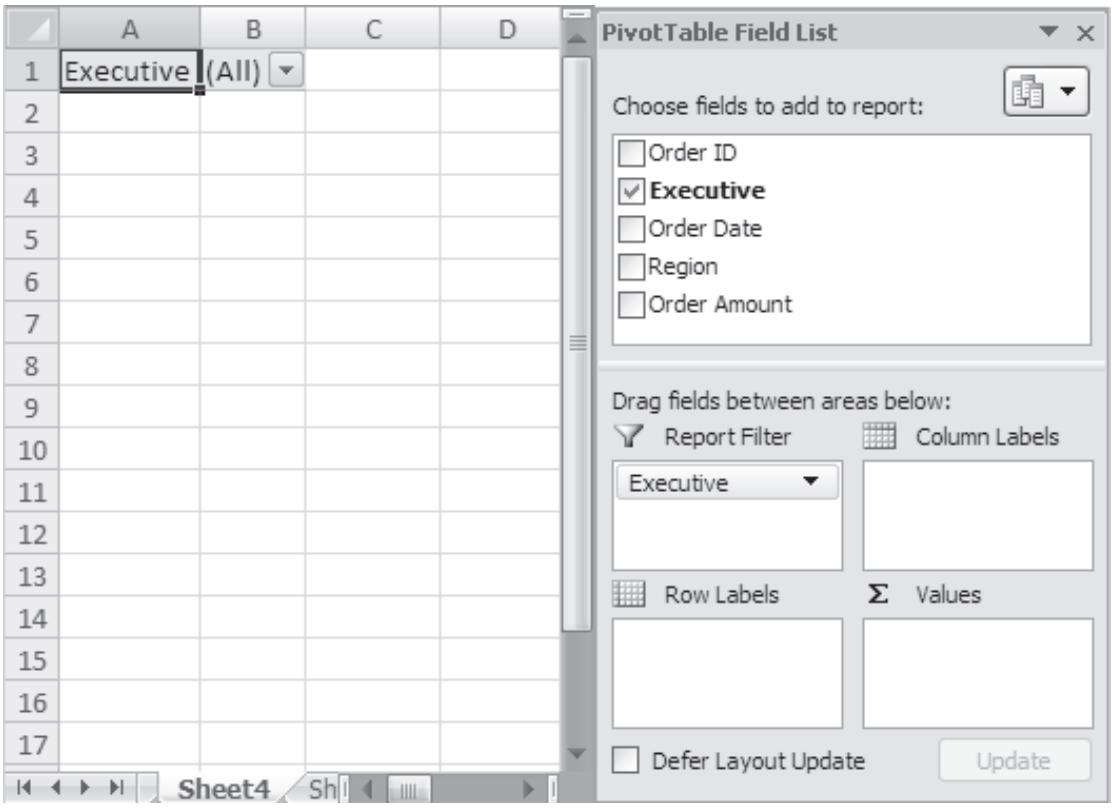


Figure 2.19: Executive Field added to the Report Filter area

5. To display sales data only for Cathy, click the arrow in cell B1, select Cathy from the list and then click OK.

Figure 2.20 shows the name **`Cathy`** selected in the filter **`Executive`**.

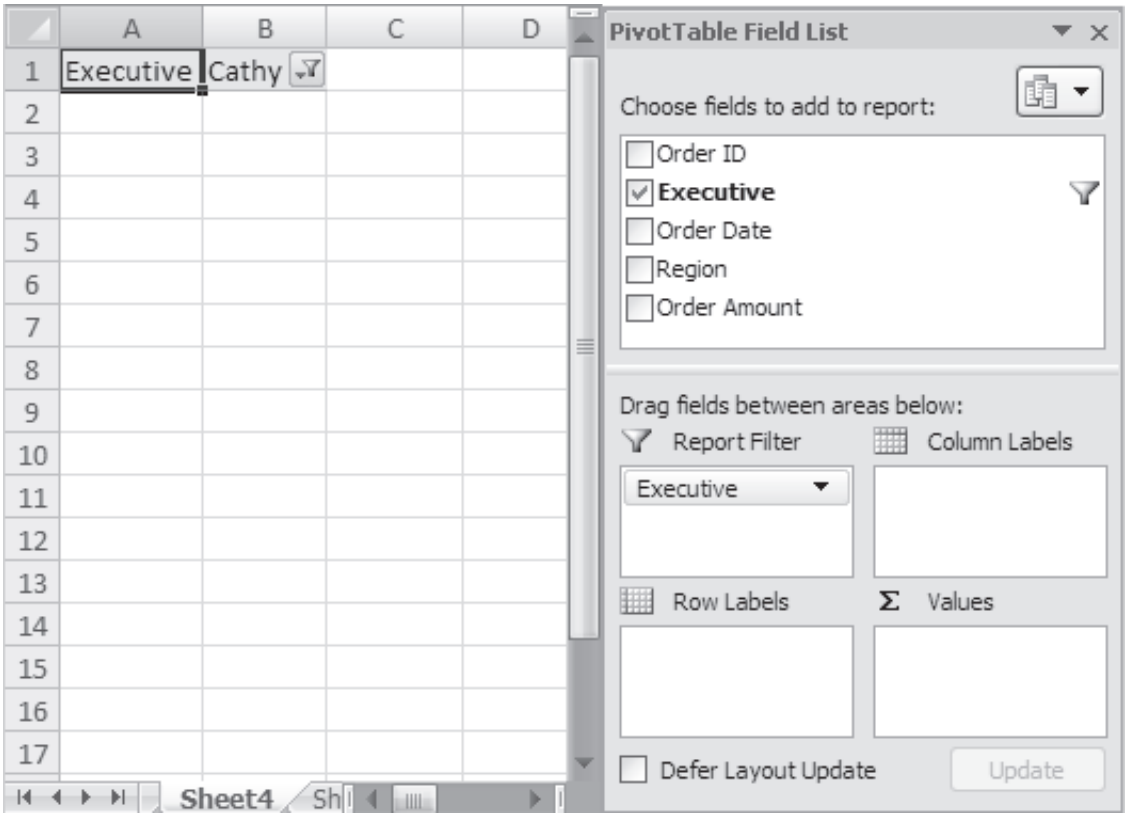


Figure 2.20: Cathy selected in the Filter

6. **To display each order ID, in the PivotTable Field List pane, drag the field Order ID to the Row Labels area.**
7. **To display the dates on which the orders were placed, in the PivotTable Field List pane, click the Order Date check box.**

Figure 2.21 shows the PivotTable report with the order dates and the orders obtained by Cathy.

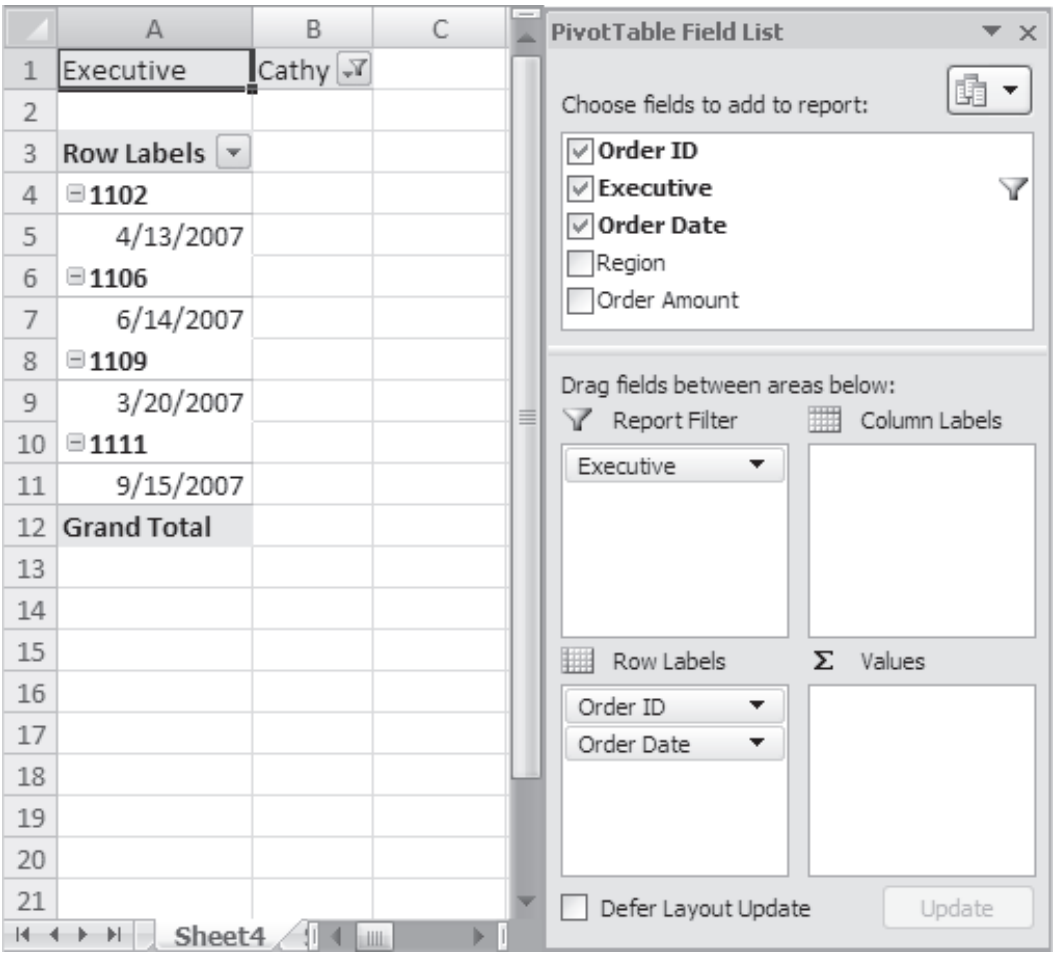


Figure 2.21: Cathy's Order details

8. **To display the order amount of each order, in the PivotTable Field List pane, click the Order Amount check box.**
9. **To display the region in which the orders were obtained, in the PivotTable Field List pane, drag the field Region to the Column Labels area.**

Figure 2.22 shows the report with the order amounts and the regions in which Cathy obtained the orders.

V 1.0 © 2008 Aptech Limited

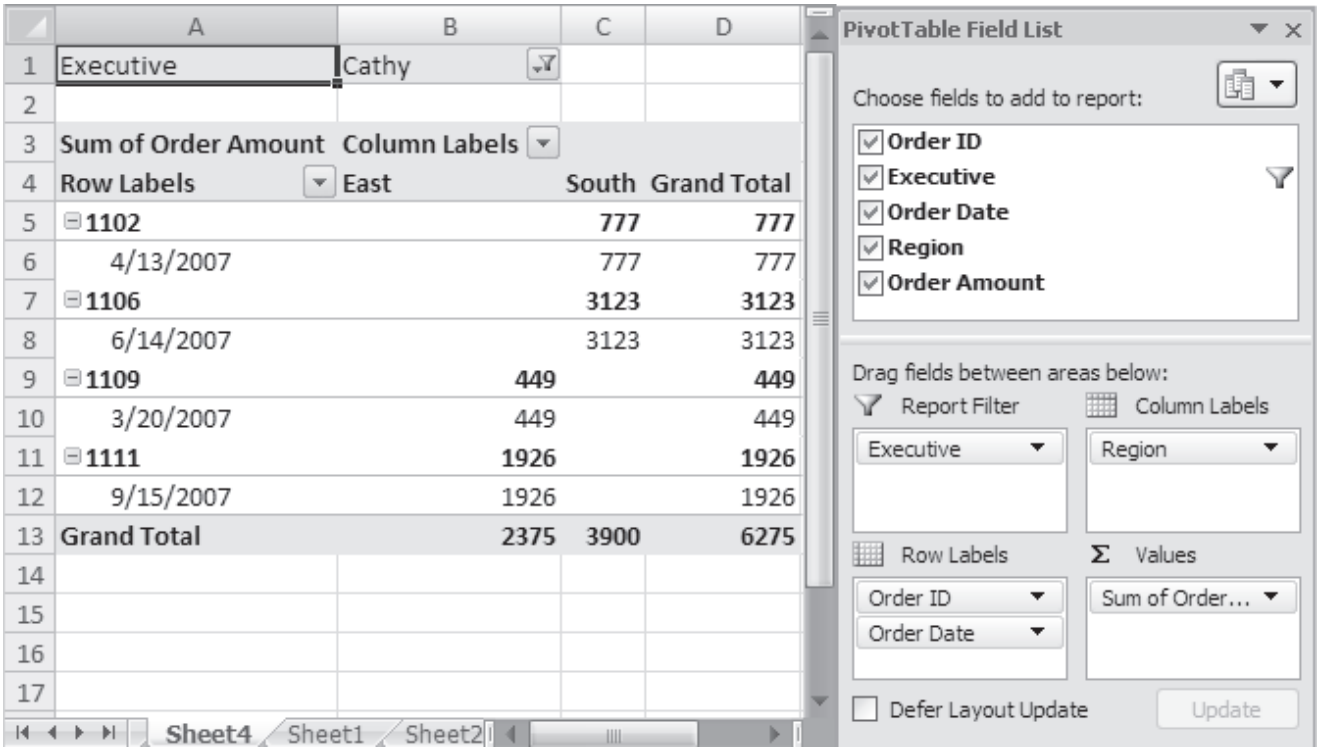


Figure 2.22: PivotTable Report with Regions and Order Amounts

10. To display orders only for the period 01/01/2007 to 30/06/2007, select a date label, for example cell A6.

11. To display the shortcut menu, click the arrow in Row Labels (cell A4).

The shortcut menu is displayed, as shown in figure 2.23.

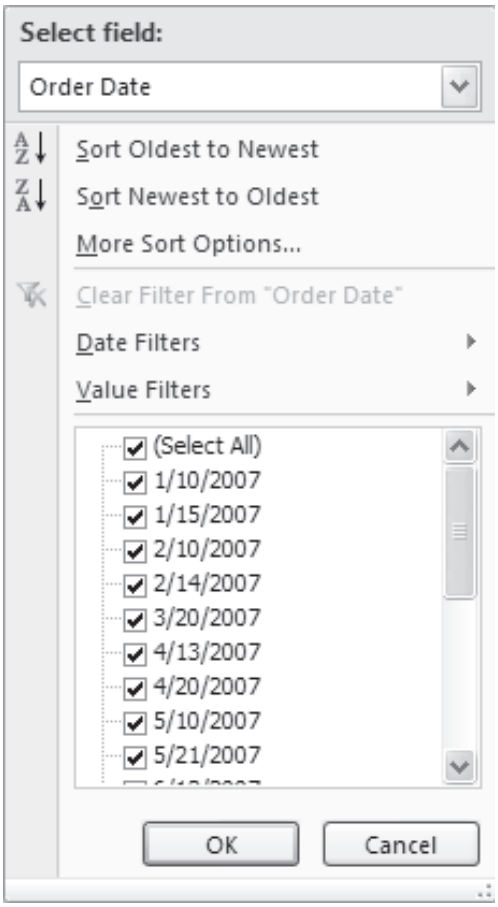


Figure 2.23: Shortcut Menu for Date Label

Note that the field, `Order Date`, is selected in the **Select field** drop-down list.

12. To apply a date filter for a specific period, point to Date Filters and then click Between.

The **Date Filter (Order Date)** dialog box is displayed, as shown in figure 2.24.

Figure 2.24: Date Filter (Order Date) Dialog Box

13. To display orders only for the period 01/01/2007 to 30/06/2007, in the first empty box, type 1/1/2007.

14. In the second empty box, type 30/06/2007 and then click OK.

Figure 2.25 shows that Cathy obtained only three orders in the first half of the year 2007.

V 1.0 © 2008 Aptech Limited

	A	B	C	D
1	Executive	Cathy		
2				
3	Sum of Order Amount	Column Labels		
4	Row Labels	East	South	Grand Total
5	1102		777	777
6	4/13/2007		777	777
7	1106		3123	3123
8	6/14/2007		3123	3123
9	1109	449		449
10	3/20/2007	449		449
11	Grand Total	449	3900	4349
12				
13				

Sheet4 Sheet1 Sheet2

Figure 2.25: Cathy's Sales Data for the period 01/01/2007 to 30/06/2007

Part II - 60 Minutes

1. Reuse the PivotTable report created in Exercise 2 of Part I and create a PivotChart report showing the following information:
   - ✧ Sales data for Simon for the East region.
   - ✧ Order amounts greater than 1000.

   Hints:
   - ✧ Use the value Simon as the filter for **`Executive`** field.
   - ✧ Create a Label Filter using the **`Region`** field.
   - ✧ Create a Value Filter using the **`Order ID`** field.
2. Basien Corporation is into retail business. It sells various products such as beverages, meat and so on. The sales data is maintained in an Excel sheet, as shown in figure 2.26.

	A	B	C	D	E	F	G
1	Executive	Sales	Units	Region	Product	Month	Year
2	John	345	565	East	Beverages	Jan	2007
3	Anita	5656	3535	West	Dairy	Jan	2007
4	Suzanne	234	565	South	Meat	Jan	2007
5	Brian	6774	345	North	Produce	Mar	2007
6	Shane	23234	789	North	Beverages	Jan	2007
7	Jenna	4564	2332	West	Meat	Feb	2007
8	Heather	8785	6767	East	Produce	Feb	2007
9	Derrick	2345	989	West	Produce	Feb	2007
10	Peter	7658	543	East	Meat	Feb	2007
11	Angel	7733	2344	South	Beverages	Feb	2007
12	Roy	7782	5451	East	Dairy	Mar	2007
13	Joby	3454	233	South	Beverages	Mar	2007
14							

Sheet1 Sheet2 Sheet3

Figure 2.26: Source Data

The Sales Manager wants a PivotChart report showing the following:

- Sales trend of meat in the months of January and February
- Executive's name and region in which the sales were done
- Sales amount and the number of units values
- Sales amount with the highest values should be displayed on the left

Hint: Use the fields, **`Month`** and **`Executive`** as filters in the PivotTable report

V 1.0 © 2008 Aptech Limited

Do It Yourself

1. Consider the source data shown in figure 2.27.

	A	B	C	D	E
1	State	Model No.	Business Type	Revenue	
2	Georgia	4055T	Parts	$3,454	
3	South Dakota	2500P	Printer Sale	$15,682	
4	South Dakota	2500C	Copier Sale	$12,968	
5	Maryland	4055T	Parts	$2,789	
6	Maryland	4500C	Service Plan	$32,605	
7	Maryland	3002P	Service Plan	$52,437	
8	Maryland	2500P	Printer Sale	$17,391	
9	Maryland	4055T	Parts	$2,468	
10	Maryland	3002C	Parts	$375	
11	Mississippi	4500C	Service Plan	$67,592	
12	Mississippi	3002C	Service Plan	$21,454	
13	Washington	3002C	Parts	$979	
14	Washington	4055T	Parts	$4,539	
15	Washington	3002P	Parts	$3,209	
16	Washington	3002P	Parts	$34,612	
17	California	3002C	Service Plan	$14,051	
18	Utah	5001P	Parts	$249	
19	Utah	5001P	Service Plan	$3,951	
20	Iowa	3002P	Service Plan	$157,434	
21	Iowa	3002P	Parts	$19,087	

Sheet4 Sheet1 Shee

Figure 2.27: Source Data

Create a Column PivotChart report providing the following information for the models 2500C, 2500P, and 3002C:

- ✧ Revenue generated by each state
- ✧ Revenue for each business type

Additionally, the PivotChart report should display a title above the report and data labels above each column.

This page has been intentionally left blank

Session 3

Data Tables and Scenario Manager

Session Objectives:

At the end of this session, you will be able to -

- Describe What-if Analysis
- Explain the working of Data Tables
- Explain the working of Scenarios

3.1 What-if Analysis

One of the most interesting features of Excel is its ability to create a dynamic model. In dynamic model, a formula is used which gets recalculated whenever the values in the cell changes. What-if analysis is the process of changing the values in worksheet cells to see how these changes affect the result of formulas on the formula cells. You can change the values in one or more cells and observe the effects on specific formula cells. For example, you can modify the interest rate for a loan amount taken in order to determine the payment that has to be made over a 15-year or 30-year term.

What-if analysis is a way of asking questions such as, “What will be the profit earned if the interest rate on the loan amount reduces from 9.5 percent to 9.0 percent?” or “What will happen if the share price of the company reduces by more than 5 percent?”.

By using What-if analysis tools in Microsoft Office Excel, you can use several different sets of values in one or more formulas to view different results. Excel provides a number of tools to help you find answers to "What-if" type of questions. The various What-if analysis tools present in Excel are:

- **Goal Seek**: Goal Seek is used when the desired result is known but one of the input values for the desired result is unknown. For example, if the desired result is 72, and the first number is 9, then what is the second number? Will it be 9 multiplied by 8 or 7? In such cases, Goal seek can be used to find the value. It is easy to use but limited in power and flexibility.
- **Data Table**: Data Table is used for viewing different results by altering an input cell in a formula. It is easy to use, but limited in power and flexibility.
- **Scenario Manager**: A scenario is a set of values and formulas that are saved and can be substituted automatically in worksheet cells based on the requirement. You can create and save different groups of values on a worksheet and then change to any defined scenarios to view the different results. It is easy to use but limited in power and flexibility.
- **Solver**: Solver is used to find the best solution to complex problems that involves manipulation of multiple variables and constraints. It is difficult to use but very powerful and extremely flexible.

3.2 Using Data Tables

Data Tables are used to view different results by altering the value in an input cell that is used in the formula. For example, if the interest rate is altered, the data table shows how much a $20,000 loan amount would cost each month. Thus, data tables are used to see how different input values affect the results of a formula. In data tables, the different combinations of input values and results are presented in a table format for easy comparison.

Consider the following example that converts the decimal numbers into binary values.

	A	B
1	Decimal	Binary
2	1	1
3	2	10
4	3	11
5	4	100
6	5	101
7	6	110
8	7	111
9	8	1000
10	9	1001
11	10	1010

Figure 3.1: Decimal-to-Binary Conversion Table

Figure 3.1 displays the conversion table, which shows the decimal value in one column, along with its corresponding binary value in the right column. The binary equivalent value of cell A2 is calculated by typing the formula, =DEC2BIN(A2) in cell B2. Similarly, cells A3 through A11 list the decimal numbers from 2 to 10, and cells B3 through B11 calculate and list the corresponding binary values.

You can create either one-input or two-input data tables. The difference between the two types of data tables is in the number of input values contained in the table. In one-input data table, the input values consist of one input cell whereas, in two-input data table, input values consist of two input cells. These input cells contain the values that are replaced in the formula. For one-input data table the input values are present either in row or column whereas, for two-input data table, the input values are present in row and column. A data table also contains result values, which represent the desired result after changing the input values in the formula.

3.2.1 Creating One-Input Data Table

A one-input data table is used to see how different values of one variable in one or more formulas will change the results of those formulas in the worksheet. For example, one-input data table can be used to calculate the effect of different dividend rates on the total Stock Payment amount. While creating a one-input data table, the input values should be listed either down a column or across a row.

To create a one-input data table, first type the list of values that are to be substituted in the input cell, either down a column or across a row. If the values are listed down a column, type the formula to be used in your data table in the right column of the input values. You can also enter any additional formulas in the cells to the right of the first formula. On the other hand, if the input values are present in a row, type the formula in the column one cell below the row. You can enter any additional formulas in the cells below the first formula.

Select the range of cells containing the formulas and the values that need to be substituted. Select the **Data** tab from the Ribbon. Under the **Data Tools group**, click the drop-down arrow of **What-if Analysis**. From the drop-down list, select **Data Table** option, as shown in figure 3.2. This will display the **Data Table** dialog box.

V 1.0 © 2008 Aptech Limited

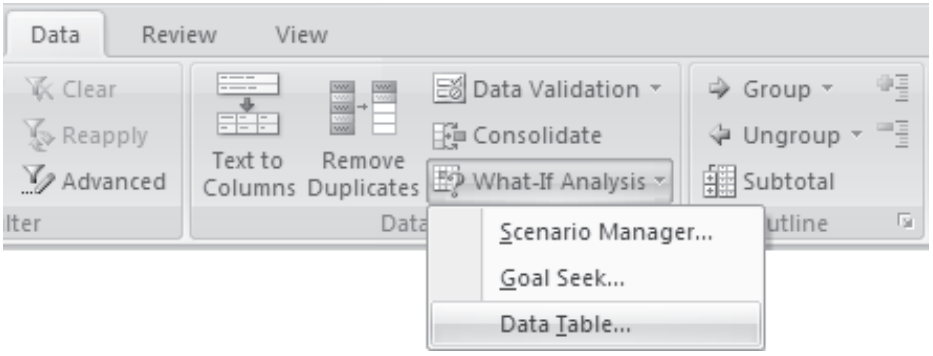


Figure 3.2: Selecting Data Table from What-If Analysis Menu

To enter the cell reference of the input cell, you have to take into account the way the input values have been entered. If the input values have been entered across a row, the input cell reference is provided in the Row Input box. If the input values have been entered down a column, the input cell reference is provided in the Column Input box. Excel will calculate the result for the formula(s) using the different values of the input variable, and display each result in the data table.

For example, figure 3.3 shows how to set up a one-input data table. Here, the square root of ten numbers are calculated using one-input data table. The input values are listed from cells A3:A12.

B2 =SQRT(A2)

	A	B	C	D
1	Number	Square Root		
2		0		
3	1	1		
4	2	1.414213562		
5	3	1.732050808		
6	4	2		
7	5	2.236067977		
8	6	2.449489743		
9	7	2.645751311		
10	8	2.828427125		
11	9	3		
12	10	3.16227766		

Figure 3.3: Setting up One-Input Data Table

To calculate the square root of the given numbers, first type the formula, =SQRT(A2), in cell B2. Since there is no input value specified in cell A2, the square root of cell A2 is displayed as 0 in cell B2. So when generating the data table, select the cell that contains the formula and the input value, which is A2:B12. Also, specify the input cell reference, which is B2, in the Column input cell box. On clicking **OK**, the data table automatically calculates the square root from cells B3 through B12.

Figure 3.4 shows how to set up a one-input data table with input values across a row.

B2 fx =SQRT(B1)

	A	B	C	D	E	J	K	L
1	Number		1	2	3	8	9	10
2	Square Root	0	1	1.414213562	1.732050808	2.828427125	3	3.16227766
3								

Figure 3.4: Setting up a one-input data table with input values across a row

To calculate the square root of cell B1, type the formula, =SQRT(B1), in cell B2. This calculates the square root of cell B1 and displays the result in cell B2. So when generating the data table, select the cells that contains the formula and the input value, which is B1:L2. Also, specify the input cell reference, which is B2, in the Row input cell box. On clicking **OK**, the data table automatically calculates the square root from cells C2 through L2.

3.2.2 Creating Two-Input Data Table

A two-input data table allows you to see the effects of changing the values of two variables (input values) used in a formula. You can create a two-input data table to accept different values for two input variables and display the varying results for a single formula. Thus, data table avoids the need for reentering the same formula or function repeatedly for each different input value or combination of input values. Data table shows the effect of changing the input values by displaying the different results in a table.

For example, you could use a two-input data table to calculate the effect of different interest rates and number of repayments on the repayment amount for a loan. A two-input data table can only use one formula, which must refer to both the input cells.

To create a data table to accept two input variables, first enter the formula that refers to both the input cells in any worksheet cell. In the column below the formula, enter one set of input values to be used in the data table. In the same row, to the right of the formula, enter the other set of input values. Select the group of cells that contains the formula and the column and row of input values that you want to substitute. Select the **Data** tab from the Ribbon and, in the **Data Tools** group, click the drop-down arrow of **What-if Analysis**. From the drop-down list, select **Data Table** to open the **Data Table** dialog box. In the Row Input box, enter the cell reference of the input cell that corresponds to the values in the row. Similarly, in the Column Input box, enter the cell reference of the input cell that corresponds to the values in the column. Excel will calculate, using the formula repeatedly using each combination of the different values for the input variables, and display the respective result for each calculation in the data table.

Consider an example of two-input data tables to calculate stock dividend payments. Cells B1 through B4 contain the initial input data. Cell B1 shows the price per share of stock. Cell B2 displays the number of shares held and cell B3 displays the dividend rate per share of stock (in percentage). Cell B4 calculates the total dividends paid, using the formula B1*B2*B3. Refer figure 3.5.

	A	B
1	Price Per Share	10
2	Shares Hold	1200
3	Dividend Rate	0.005
4	Total Dividends Paid	=B1*B2*B3

Figure 3.5: Initial Data Before Creating Two-Input Data Table

To calculate the stock dividend payments for the shares of stock held from 30,000 to 360,000 with stock prices from $30.00 per share to $60.00 per share, a data table needs to be created.

V 1.0 © 2008 Aptech Limited

To create the data table and format the results, first type 30000 and 60000 in cells B5 and B6 respectively. Select cells B5 and B6 together, and then drag using the mouse through B16. Click Fill Series option from the Smart Tag option. Similarly, from cells C4 through I4, display the values from 30 to 60, with an increment of 5 in each cell respectively, using the Fill Series option. Select cells B4 through I16. Select the **Data** tab from the Ribbon, and click the drop-down arrow of **What-if Analysis** option. From the drop-down list, select **Data Table** to open the **Data Table** dialog box. In the Row Input Cell box, type or click cell B1. In the Column Input Cell box, type or click cell B2. Change the format of cells C4 through I16 to Currency.

The output generated is shown in figure 3.6.

	A	B	C	D	E	F	G	H	I
1	Value Per Share	$10.00							
2	Shares Hold	1200							
3	Dividend Rate	0.005							
4	Total Dividends Paid	$60.00	$30.00	$35.00	$40.00	$45.00	$50.00	$55.00	$60.00
5		30000	$4,500.00	$5,250.00	$6,000.00	$6,750.00	$7,500.00	$8,250.00	$9,000.00
6		60000	$9,000.00	$10,500.00	$12,000.00	$13,500.00	$15,000.00	$16,500.00	$18,000.00
7		90000	$13,500.00	$15,750.00	$18,000.00	$20,250.00	$22,500.00	$24,750.00	$27,000.00
8		120000	$18,000.00	$21,000.00	$24,000.00	$27,000.00	$30,000.00	$33,000.00	$36,000.00
9		150000	$22,500.00	$26,250.00	$30,000.00	$33,750.00	$37,500.00	$41,250.00	$45,000.00
10		180000	$27,000.00	$31,500.00	$36,000.00	$40,500.00	$45,000.00	$49,500.00	$54,000.00
11		210000	$31,500.00	$36,750.00	$42,000.00	$47,250.00	$52,500.00	$57,750.00	$63,000.00
12		240000	$36,000.00	$42,000.00	$48,000.00	$54,000.00	$60,000.00	$66,000.00	$72,000.00
13		270000	$40,500.00	$47,250.00	$54,000.00	$60,750.00	$67,500.00	$74,250.00	$81,000.00
14		300000	$45,000.00	$52,500.00	$60,000.00	$67,500.00	$75,000.00	$82,500.00	$90,000.00
15		330000	$49,500.00	$57,750.00	$66,000.00	$74,250.00	$82,500.00	$90,750.00	$99,000.00
16		360000	$54,000.00	$63,000.00	$72,000.00	$81,000.00	$90,000.00	$99,000.00	$108,000.00

Figure 3.6: Two-Input Data Table Calculating Stock Dividend Payments

Data tables are useful, but they have a few limitations. For instance, only one or two input cells can be changed at a time.

3.3 Scenarios

A Scenario is a collection of cell values and formulas that can be saved as a group and then swapped automatically for another group of cell values in a worksheet. However, scenarios can accept more than two input variables. You can later switch to any of these scenarios to view the results. Scenarios are helpful as they help to analyze situations by comparing sets of cell values.

For example, the interest rate forecast for a house loan may depend on several factors, such as best case, worst case, and most likely case. You can view any of these scenarios by selecting the scenario from a list. Based on the scenario selected, Excel substitutes the appropriate input values in your worksheet and recalculates the formulas.

3.3.1 Using Scenario

Before the creation of scenario, construct a model on the worksheet to display the Total Revenue, Total Costs, and Profit, based on the different values of Units Sold, Selling Price per Unit, Variable Cost per Unit, and Fixed Costs.

	A	B	C	D	E
1		Input/Changing Cells			Result Cells
2	Units Sold	20000		Total Revenue	$2,000,000.00
3	Selling Price Per Unit	$100.00		Total Costs	$1,040,000.00
4	Variable Cost Per Unit	$50.00		Profit	$960,000.00
5	Fixed Cost	$40,000.00			

Figure 3.7: Constructing a Scenario Model

As shown in figure 3.7, all the changing cells and results cells are named using the labels present on their left, to make the scenarios more meaningful. The input or changing cells are B2:B5 and the results cells are E2:E4.

The advantages of using Scenarios over Goal Seeking and data table are as follows:

- Using goal seeking, you can change only one of the cell values in cell B2, B3, B4 or B5 at a time. Scenarios can change any combination of cell values in these cells at a time.
- A data table can show the calculated result of cells E2 and E3 in cell E4, but it cannot show the updated values of both the cells E2 and E3, along with the calculated result in cell E4. With scenarios, you can see values of all these cells.

At a time, you can substitute up to 32 input cell values. The number of scenarios stored in a worksheet depends on the computer's available memory.

Before the creation of a new scenario, you can define the names for the input/changing cells and results cells by selecting the respective cells and their corresponding labels. To do this, select the **Formulas** tab and click the Create from Selection command. In the **Create Names from Selection** dialog box, select the Left column checkbox for the location of the labels, as shown in figure 3.8.

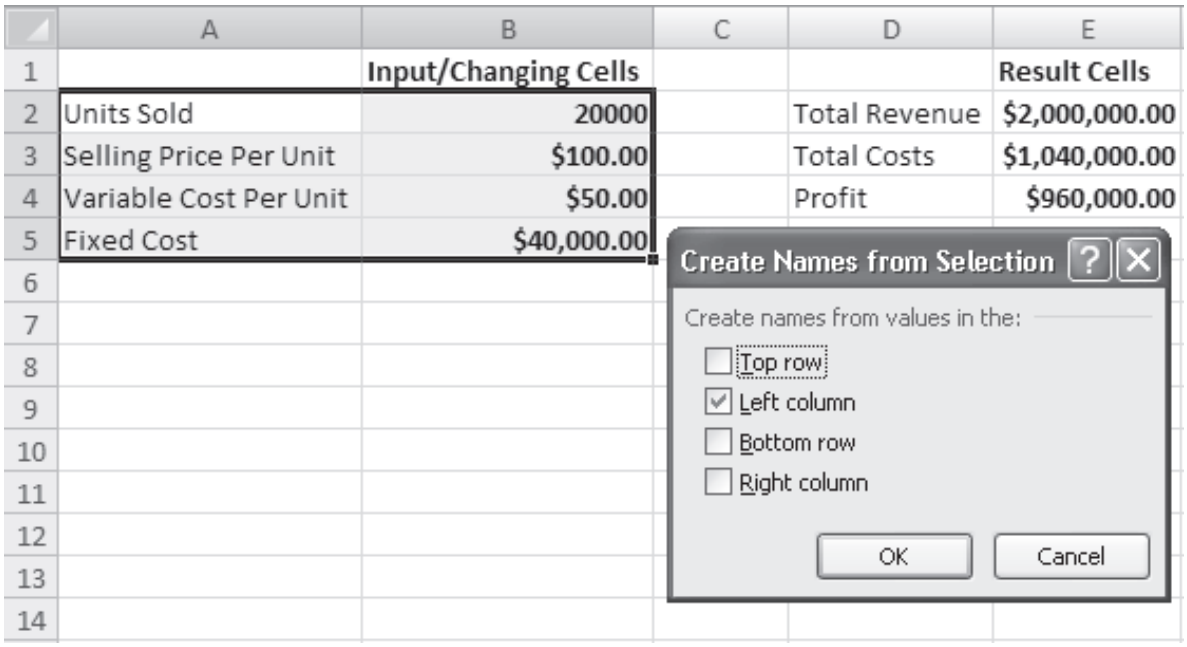


Figure 3.8: Create Names from Selection Dialog Box

Excel will name each cell according to its label, inserting underscore characters to replace any spaces in the label.

From the **Data Tools group**, present in the **Data** tab, click the drop-down arrow of **What-if Analysis** option, and select the **Scenario Manager** option, as shown in figure 3.9.

V 1.0 © 2008 Aptech Limited

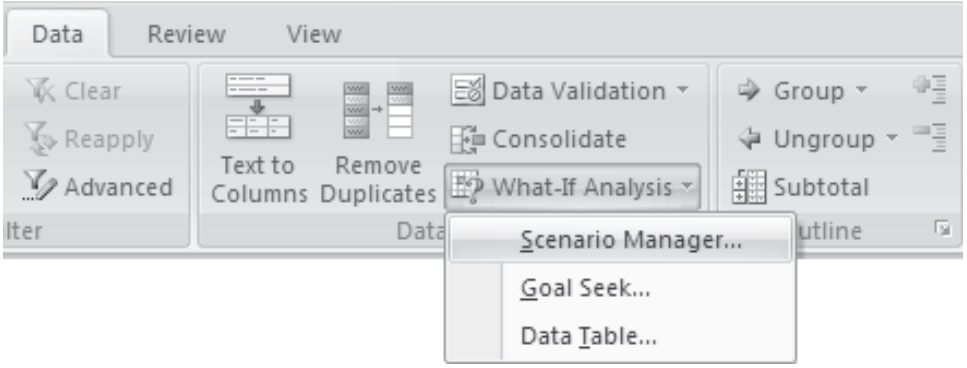


Figure 3.9: Selecting Scenario Manager

On selecting the Scenario Manager option, the **Scenario Manager** dialog box is displayed, as shown in figure 3.10.

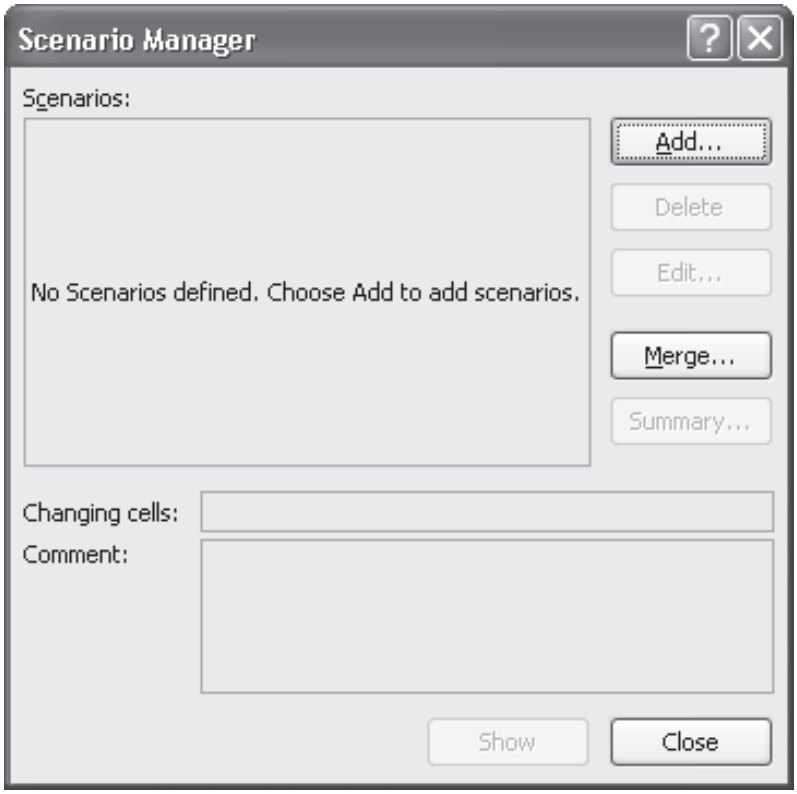


Figure 3.10: Scenario Manager Dialog Box

The **Scenario Manager** dialog box contains the following controls:

- The **Scenarios** section displays all the available scenarios of the selected worksheet.
- The **Add** button displays the **Add Scenario** dialog box, so that you can add a scenario.
- The **Delete** button deletes the selected scenario from the worksheet. This button is disabled if there are no scenarios in the worksheet.
- The **Edit** button displays the **Edit Scenario** dialog box, so that you can modify an existing scenario. This button is disabled if there are no scenarios in the worksheet.
- The **Merge** button displays the **Merge Scenarios** dialog box, so that you can add scenarios from one worksheet to another worksheet.

- The **Summary** button displays the **Scenario Summary** dialog box, so that you can view a summary or a PivotTable report. The **Summary** button is disabled if there are no scenarios in the worksheet.
- The **Changing cells** box displays the cell reference for the selected scenario's changing cell values. This box is empty if there are no scenarios in the worksheet.
- The **Comment** box displays any comments associated with the selected scenario. This box is empty if there are no scenarios in the worksheet.
- The **Show** button changes the cell values associated with the selected scenario. This button is disabled if there are no scenarios in the worksheet.
- The **Close** button exits the **Scenario Manager** dialog box.

3.3.2 Creating a New Scenario

To create a new scenario, click **Add** in the **Scenario Manager** dialog box. The **Add Scenario** dialog box is displayed, as shown in figure 3.11.

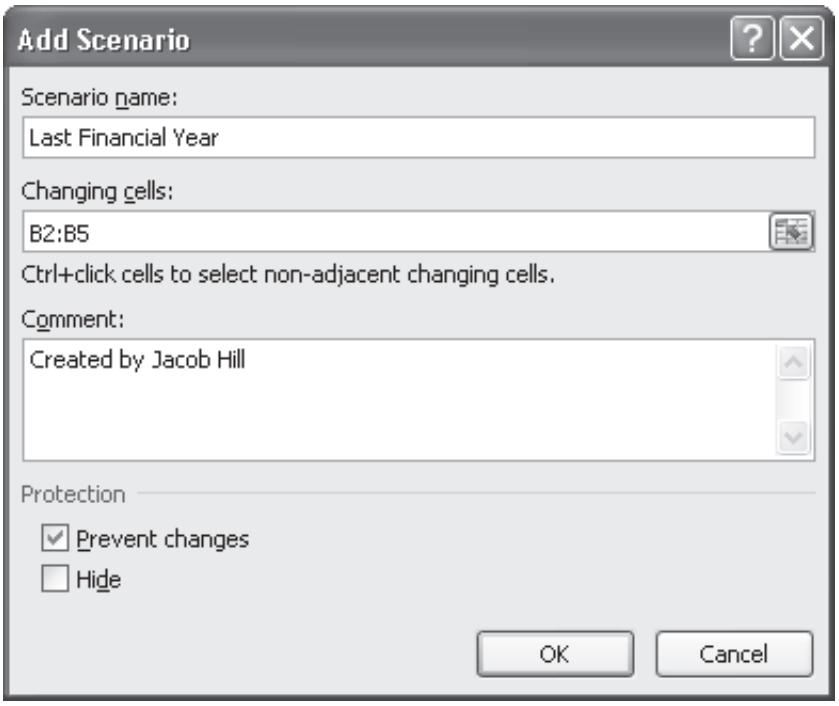


Figure 3.11: Add Scenario Dialog Box

In the **Add Scenario** dialog box, type a name for the scenario in the Scenario name box. In the Changing cells box, enter the cell references whose values need to be changed. For example, B4:B7. You can use commas to separate noncontiguous cells. Type a comment associated with the scenario, such as the scenario creator's name and creation date in the Comment box. The Prevent changes and Hide options under Protection should be selected only when the worksheet needs to be protected. Note that the Hide checkbox, when checked, removes the scenario from the Scenario list present in the Scenario Manager dialog when the worksheet is protected.

On clicking **OK** in the **Add Scenario** dialog box, the **Scenario Values** dialog box is displayed, as shown in figure 3.12.

V 1.0 © 2008 Aptech Limited

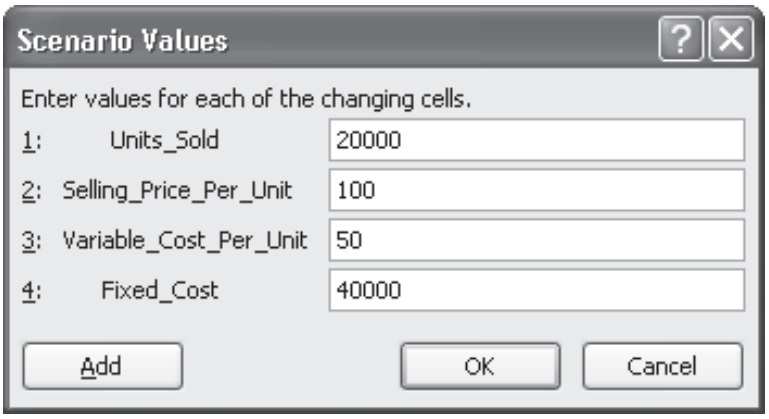


Figure 3.12: Scenario Values Dialog Box

This dialog box contains one field for each changing cell reference, along with the changing cell's value.

In order to change the current values, enter the appropriate values for this scenario in the respective boxes. Click **Add** to return to the **Add Scenario** dialog box to define another scenario.

You can create multiple scenarios by changing one or more values for the changing cells. After the creation of the last scenario, click **OK** instead of **Add** to return to the **Scenario Manager** dialog box. After all the scenarios have been defined, the Scenario Manager is used to add new scenarios or to display, edit, and delete the existing scenarios.

3.3.3 Displaying a Scenario

After defining all the scenarios and returning to the **Scenario Manager** dialog box, the dialog box displays the names of all the defined scenarios, as shown in figure 3.13.

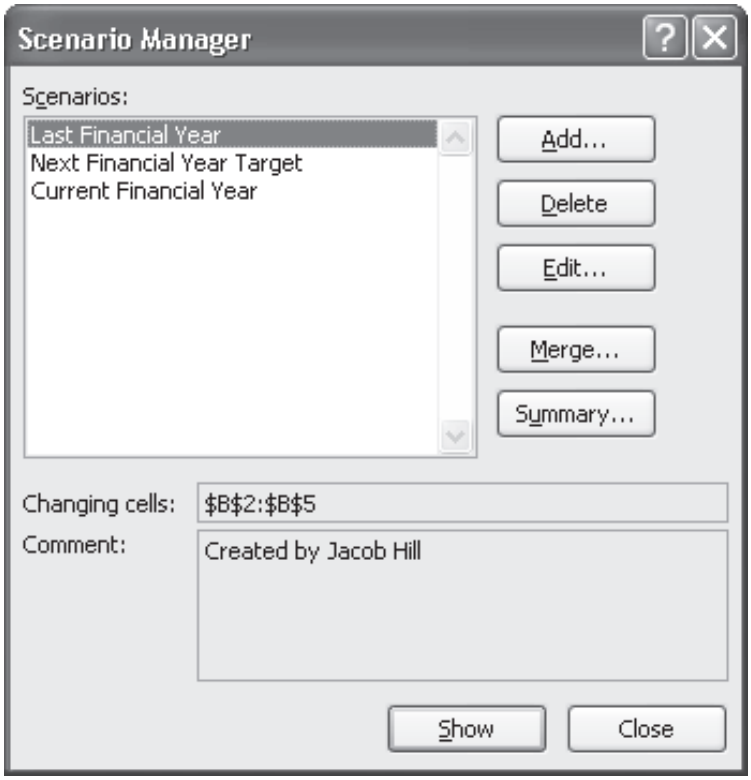


Figure 3.13: List of Defined Scenarios

To display a particular scenario, select one of the scenarios from the list and click **Show**. Excel automatically inserts the corresponding values into the input cells. It calculates the input based on the formula entered in the cells E2, E3, and E4 and displays the result for that scenario, as shown in figure 3.14.

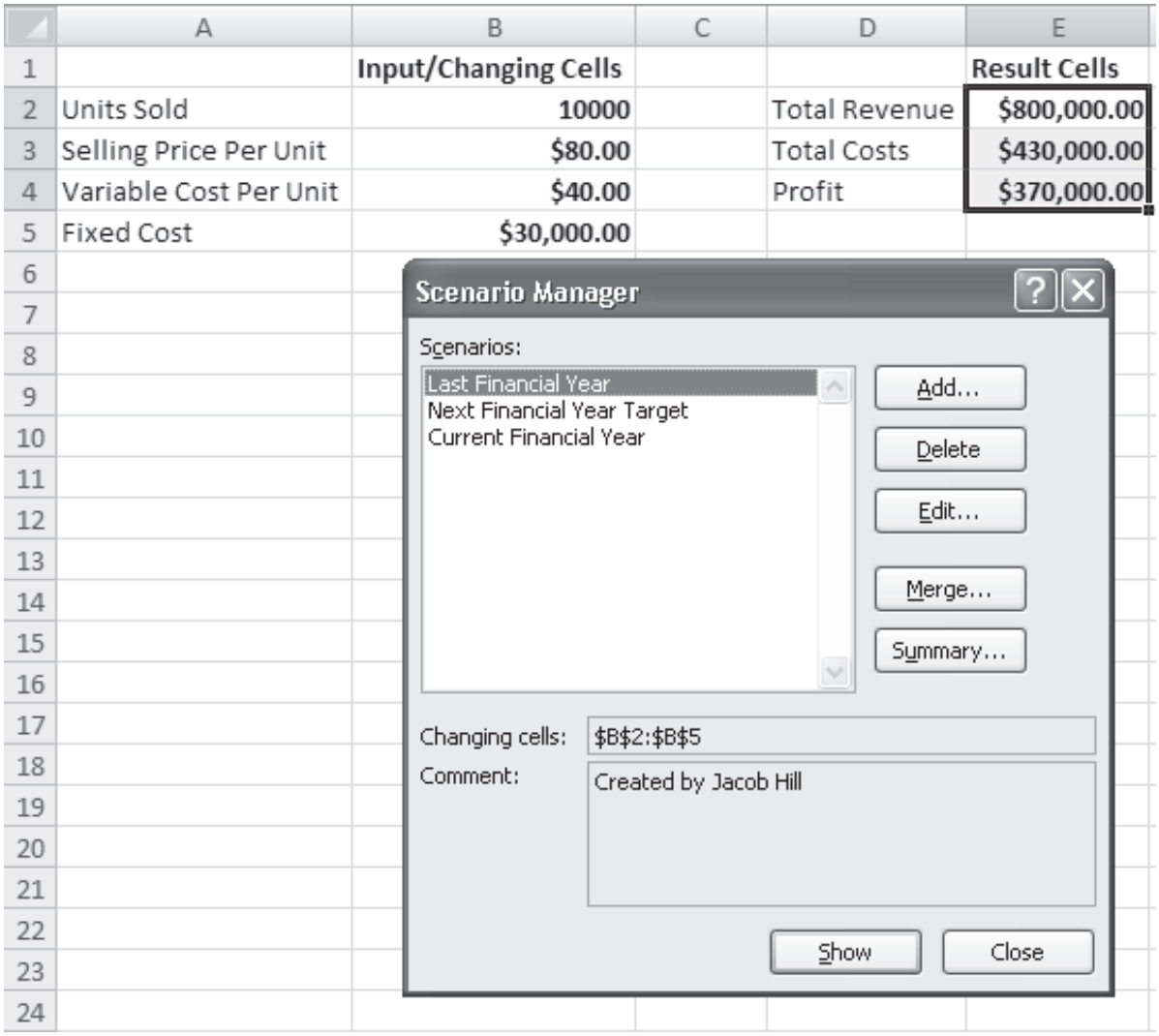


Figure 3.14: Calculated Result for Last Financial Year Scenario

Figure 3.14 displays the new value in the input/changing cells and its corresponding value in the result cells. To view each scenario result any time on the worksheet, select the scenario and click **Show**.

3.3.4 Editing an Existing Scenario

To edit the details of an existing scenario on the current worksheet, select the scenario you want to change from the scenario-list, and then click **Edit** in the **Scenario Manager** dialog box. Excel displays the **Edit Scenario** dialog box, as shown in figure 3.15.

V 1.0 © 2008 Aptech Limited

Figure 3.15: Edit Scenario Dialog box

In the **Edit Scenario** dialog box, you can change the scenario name in the Scenario Name box, type or select the cell reference for the cells that will change in the Changing Cells box, or update a comment in the Comment box. Modify the Scenario Name as "Previous Financial Year", and click **OK**. The modified scenario name is reflected in the **Scenario Manager** dialog box, as shown in figure 3.16.

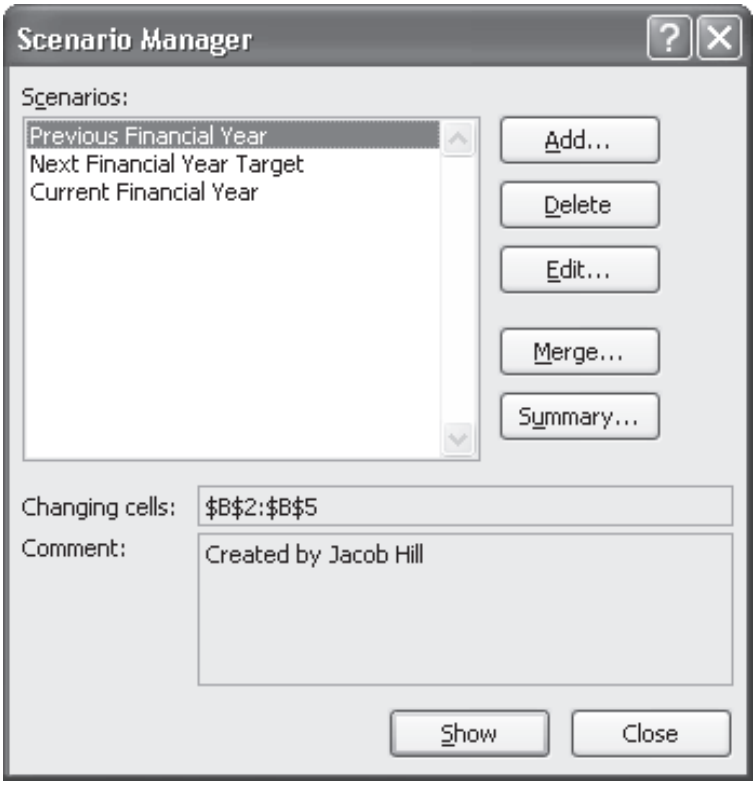


Figure 3.16: Updated Name of the Scenario

3.3.5 Deleting a Scenario

To remove a scenario from the current list of defined scenarios, select the scenario name you want to remove from the list, and click **Delete**. The selected scenario gets deleted from the Scenario list in the **Scenario Manager** dialog box, as shown in figure 3.17.

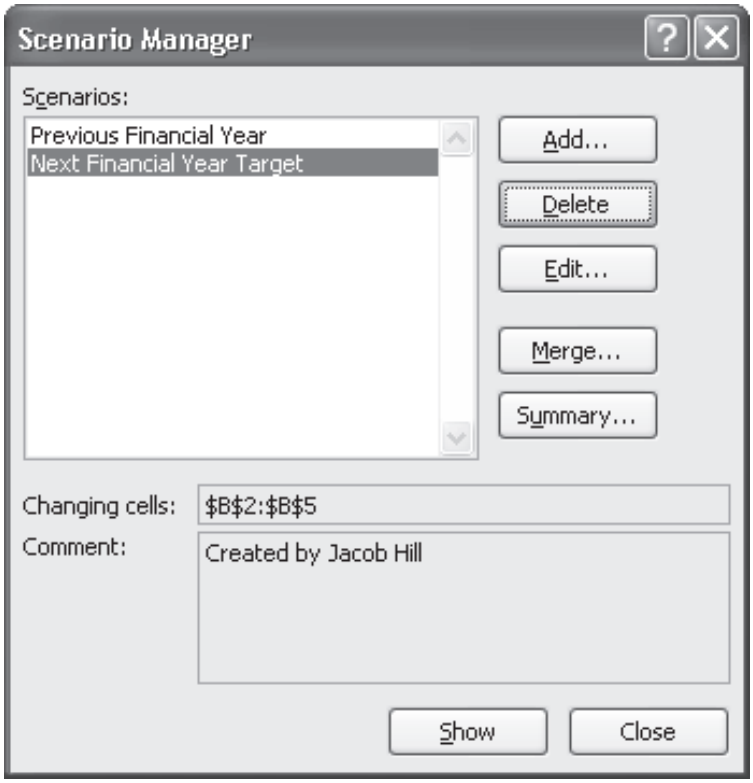


Figure 3.17: Current Financial Year Deleted from the Scenario Manager

Note that the scenario will be removed from the Scenario Manager without any warning. Hence, be careful before using this feature.

3.3.6 Creating a Scenario Summary Report

A summary report or PivotTable can be created to view a new worksheet outlining all the scenarios. Generally, a Scenario Summary is best for viewing all the scenarios. A PivotTable report may be used to display a large number of scenarios with multiple result cells.

To create a summary of all the scenarios, click **Summary** from the **Scenario Manager** dialog box. The **Scenario Summary** dialog box shows the option for the type of report to be generated, as shown in figure 3.18.

V 1.0 © 2008 Aptech Limited

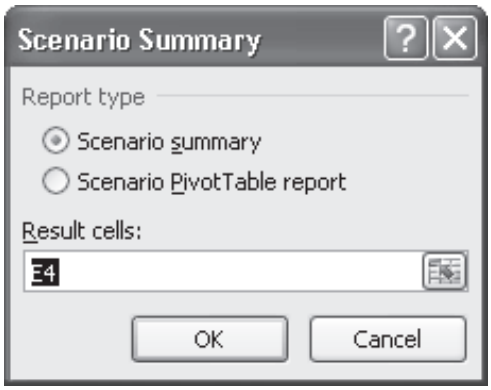
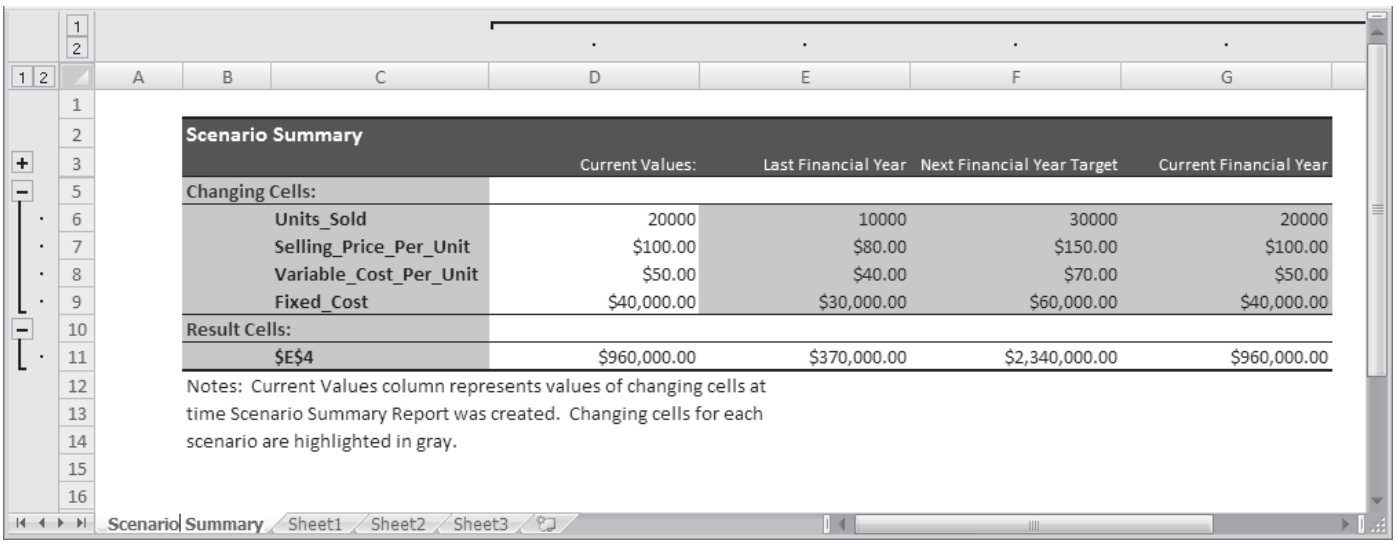


Figure 3.18: Scenario Summary Dialog Box for the Report Type

The Summary Report can be displayed in two forms: Scenario Summary and Scenario PivotTable. In Scenario summary report type, the summary report appears in the form of a worksheet outline whereas, in Scenario PivotTable report type, the summary report appears in the form of a PivotTable. The Result cells box displays the cell reference for the scenarios' result cells on the summary report or PivotTable. In the Results cells box, enter the references for the results cells in the worksheet, and click **OK**. There should be at least one results cell to generate a Scenario PivotTable report. The Scenario Summary or Scenario PivotTable report will be displayed on a new worksheet.

The Scenario Summary report gets displayed, as shown in figure 3.19.



Scenario Summary	Current Values:	Last Financial Year	Next Financial Year Target	Current Financial Year
Changing Cells:				
Units_Sold	20000	10000	30000	20000
Selling_Price_Per_Unit	$100.00	$80.00	$150.00	$100.00
Variable_Cost_Per_Unit	$50.00	$40.00	$70.00	$50.00
Fixed_Cost	$40,000.00	$30,000.00	$60,000.00	$40,000.00
Result Cells:				
$E$4	$960,000.00	$370,000.00	$2,340,000.00	$960,000.00

Notes: Current Values column represents values of changing cells at time Scenario Summary Report was created. Changing cells for each scenario are highlighted in gray.

Figure 3.19: Scenario Summary Report

The Scenario PivotTable gets displayed, as shown in figure 3.20.

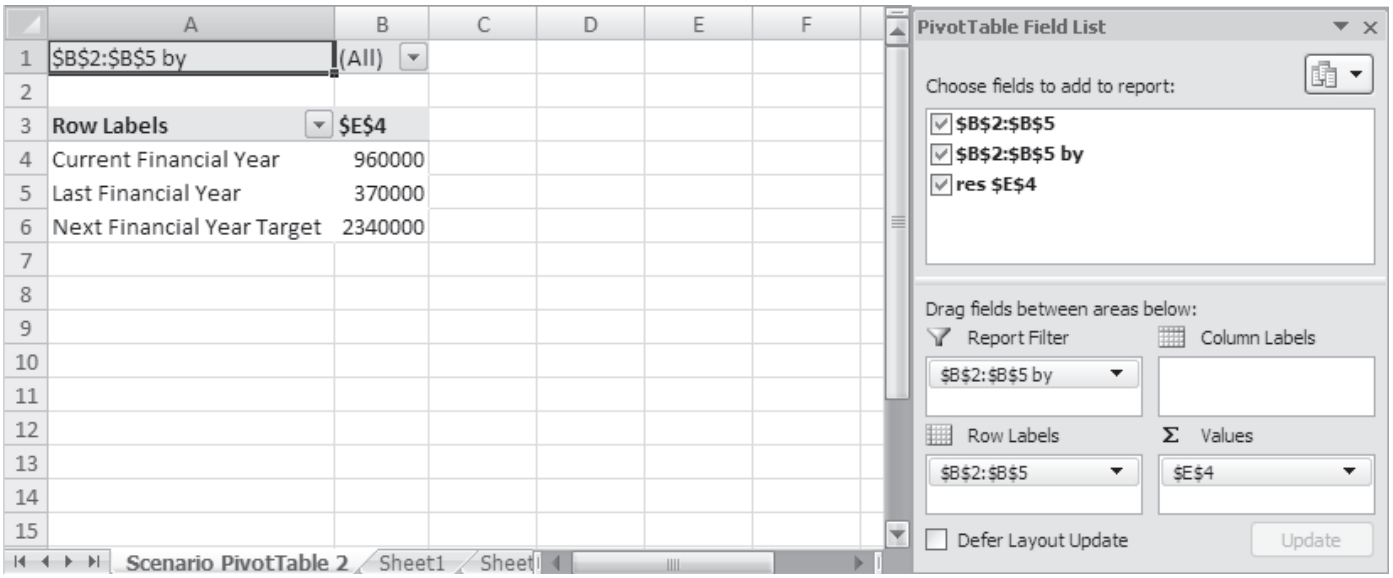


Figure 3.20: Scenario PivotTable

3.3.7 Merging Scenarios from another Worksheet

To merge scenarios from another worksheet into the current worksheet, click **Merge** in the **Scenario Manager** dialog box. This will allow you to create a summary of all the scenarios. To merge scenarios from other worksheets into the active worksheet, first open the **Scenario Manager** dialog box, and then click **Merge** to open the **Merge Scenarios** dialog box.

The **Merge Scenarios** dialog box contains the following controls:

- The **Book** list contains the list of all open workbooks.
- The **Sheet** list contains the list of all worksheets for the workbook selected in the Book list.

From the **Book** dropdown list, select the **Dividend.xlsx** workbook, which contains the scenario worksheets. Select the worksheet that contains the scenario(s) to be merged from the dropdown list, as shown in figure 3.21.

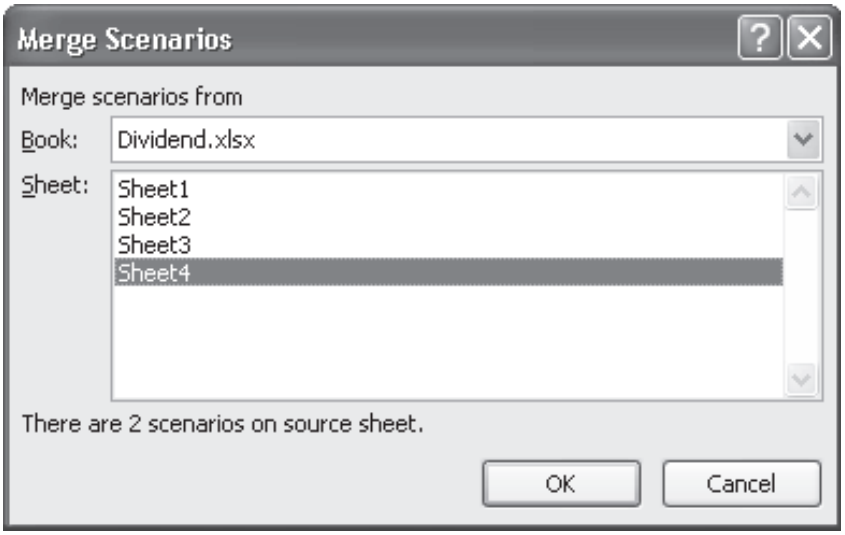


Figure 3.21: Merge Scenarios Dialog Box

On clicking **OK** in the **Merge Scenario** dialog box, the two scenarios created in **Sheet4** of **Dividend.xlsx** file are added to the existing scenario as shown in figure 3.22.

V 1.0 © 2008 Aptech Limited

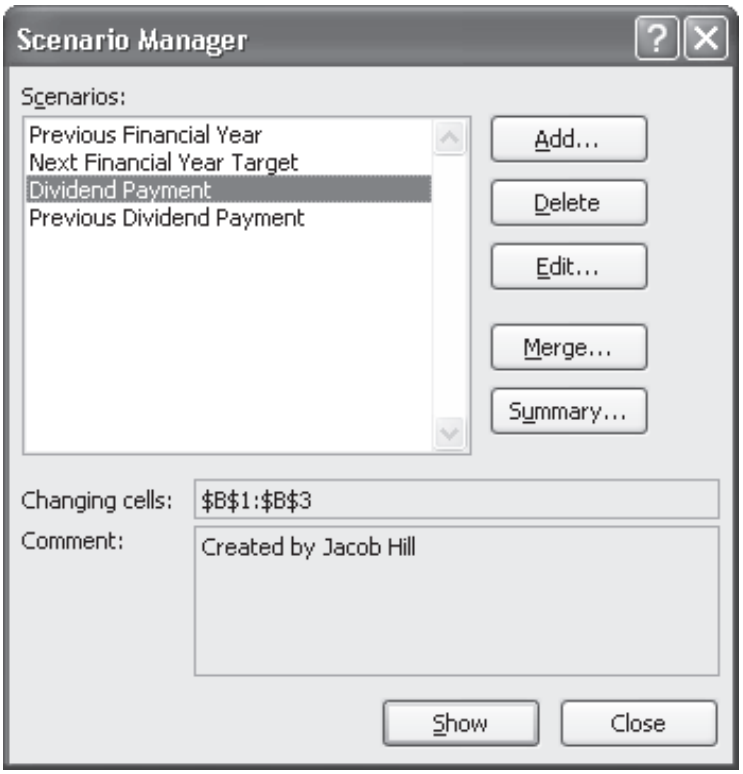


Figure 3.22: Two Scenarios from Different Worksheet Added to the Existing Scenario

3.3.8 Preventing Changes to a Scenario

To protect the scenarios on the current worksheet from editing, in the **Scenario Manager** dialog box, click the scenario you want to protect, and then click **Edit**. In the **Edit Scenario** dialog box, select the **Prevent changes** checkbox to avoid changes to the selected scenario, or select the **Hide** checkbox to prevent the scenario from appearing in the Scenarios list, as shown in figure 3.23.

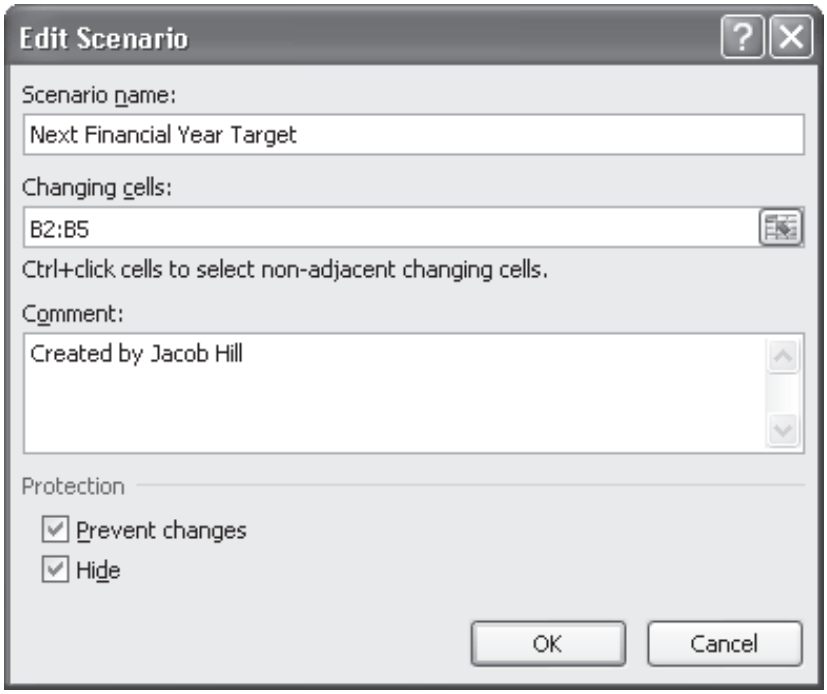


Figure 3.23: Protecting the Scenario from Editing

The Session In Brief

- Using What-if analysis tools in Microsoft Office Excel, you can use several different sets of values in one or more formulas to view different results.
- Data Table is used for viewing different results by altering an input cell in a formula.
- A one-input data table is used to see how different values of one variable in one or more formulas will change the results of those formulas in the worksheet.
- A two-input data table allows you to see the effects of changing the values of two variables (input values) used in a formula.
- A Scenario is a collection of cell values and formulas that can be saved as a group and then swapped automatically for another group of cell values in a worksheet.

V 1.0 © 2008 Aptech Limited

Check Your Progress

1. In _____, the different combinations of input values and results are presented in a table format for easy comparison.

   a. Data Tables

   b. Scenario

   c. Goal Seek

   d. Solver

2. To open the Data Table dialog box, click the drop-down arrow of _____, and from the drop-down list, select Data Table.

   a. Scenario

   b. Data Table

   c. What-if Analysis

   d. Goal Seek

3. To create a data table to accept two input variables, first enter the _____ that refers to both the input cells.

   a. values

   b. name

   c. formula

   d. label

4. To create a new scenario, click _____ in the Scenario Manager dialog box.

   a. New

   b. Add

   c. Open

   d. Create

5. To display a particular scenario, select one of the scenarios from the list and click _____.

   a. Open

   b. Summary

   c. Create

   d. Show

6. A _____ report may be used to display a large number of scenarios with multiple result cells.

   a. PivotTable

   b. Summary

   c. Result

   d. Scenario

Check Your Progress

7. In the Edit Scenario dialog box, select the _____ checkbox to avoid changes to the selected scenario.

   a. Prevent Changes

   b. Hide

   c. Security

   d. Protection

V 1.0 © 2008 Aptech Limited

Session 4

Data Tables and Scenario Manager (Lab)

Session Objectives:

At the end of this session, you will be able to -

- Use Data Tables
- Use Scenarios

Part I - 90 Minutes

Exercise 1: Working with Data Table

Problem

Heaven Constructions have built a colony of multi-storey buildings. Jack Smith wants to buy a flat in that colony. He needs $50,000 in five years for the down payment on the flat. If the simple interest rate were 7.5%, how much would he have to invest today to accumulate a maturity amount of $50,000 in five years?

Analysis

The goal here is to achieve an amount of $50,000 in five years. The present value refers to today's value of a future amount. This can be achieved by calculating the Principal amount.

Maturity Amount = $50000

Years = 5

Rate of Interest = 7.5% = 7.5/100 = 0.075

The formula to calculate the Principal Amount is:

Principal Amount = Maturity Amount/(1 + Years*Rate of Interest)

Thus, if you invest $36363.64 today at 7.5% simple interest, you will have $50,000 in five years.

The explanation for the calculation is given below:

Interest per year = $36363.64 x 0.075 = $2727.27

Interest for five years = $2727.27 x 5 = $13636.36

Therefore, the amount available for down payment at the end of five years is $36363.64 + $13636.36 = $50000.

Solution

1. **Enter the data and formula (Cell B4) shown in figure 4.1 to calculate the Principal Amount.**

B4 | fx | =B1/(1+B2*B3)

	A	B	C	D
1	Maturity Amount	$50,000		
2	Number of Years	5		
3	Rate of Interest	0.075		
4	Principal Amount	$36,363.64		
5				

Figure 4.1: Initial Data before creating Two-Input Data Table

2. **Type the values 0.080 and 0.085 in cells C4 and D4 respectively.**
3. **Select cells C4 and D4 together, and then drag the Fill Handle to H4.**

The rate of interest is incremented by 0.005 to calculate the amount to be invested over a period.

4. **Type the values 6 and 7 in cells B5 and B6 respectively.**
5. **Select cells B5 and B6 together, and then drag the Fill handle to B14.**

The worksheet gets displayed, as shown in figure 4.2.

	A	B	C	D	E	F	G	H
1	Maturity Amount	$50,000						
2	Number of Years	5						
3	Rate of Interest	0.075						
4	Principal Amount	$36,363.64	0.080	0.085	0.090	0.095	0.100	0.105
5		6						
6		7						
7		8						
8		9						
9		10						
10		11						
11		12						
12		13						
13		14						
14		15						

Figure 4.2: Substituting the Values for Number of Years and Rate of Interest

6. **Select cells B4 through H14.**
7. **On the Data tab, click What-If Analysis arrow → Data Table, as shown in figure 4.3.**

V 1.0 © 2008 Aptech Limited

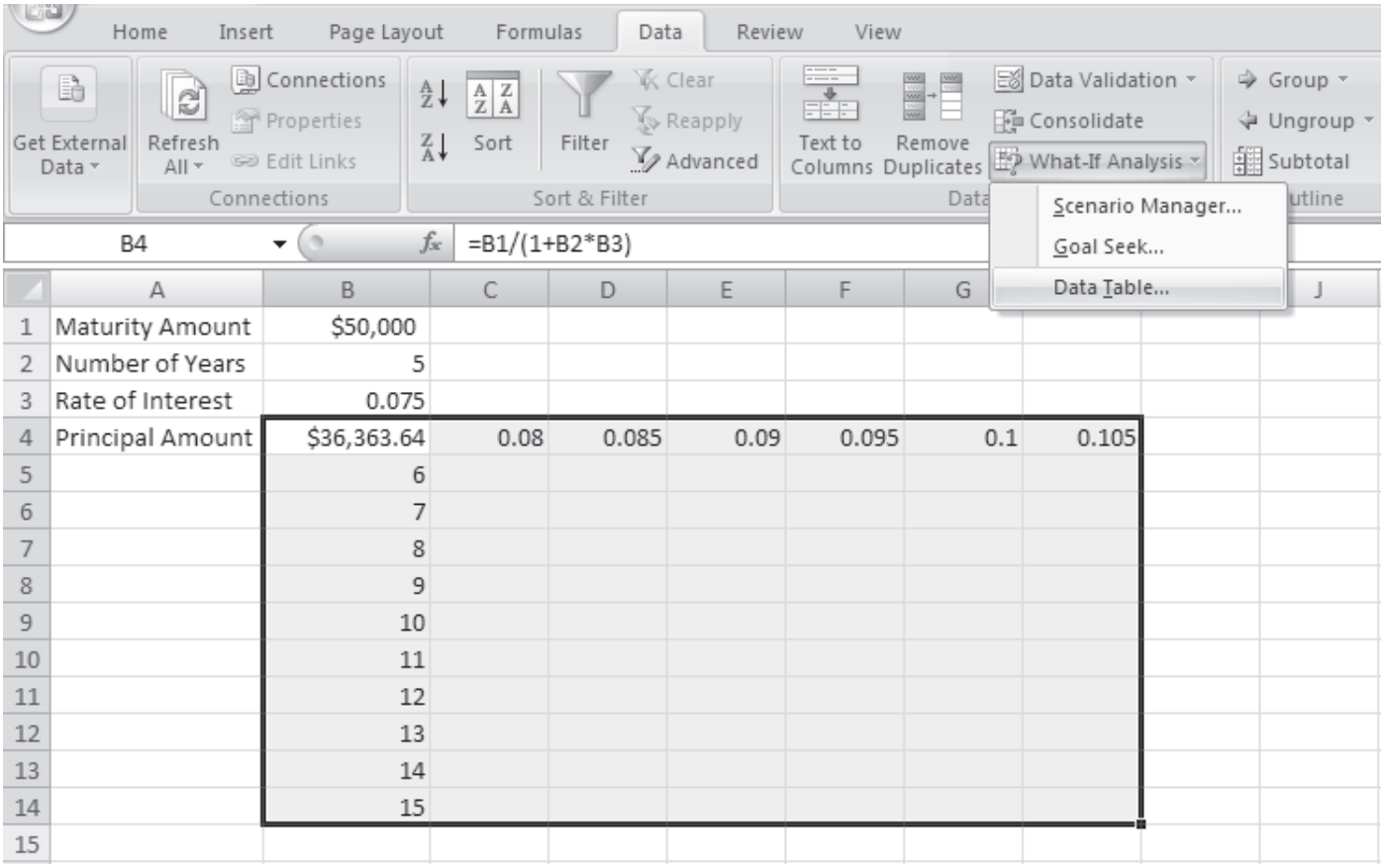


Figure 4.3: Selecting Data Table option from What-If Analysis

The Data Table dialog box gets displayed.

8. In the Row input cell box, select cell B3, and in the Column input cell box, select cell B2 as shown in figure 4.4.

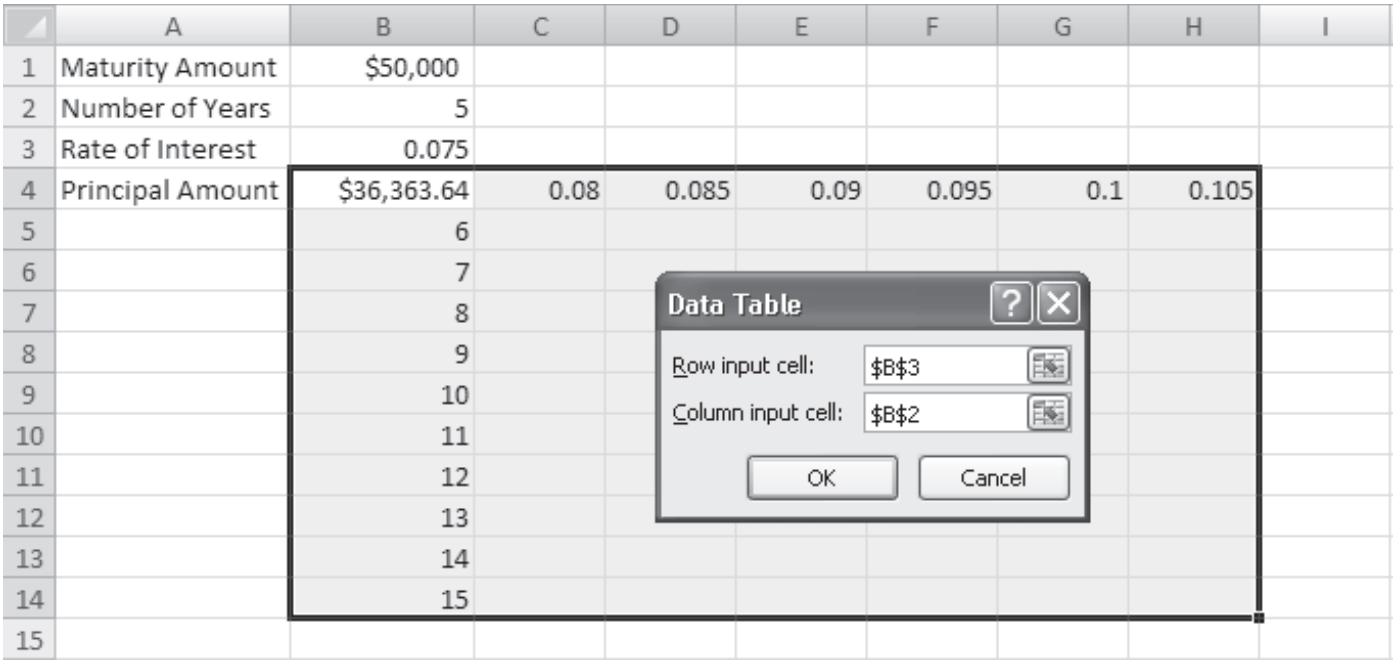


Figure 4.4: Data Table Dialog Box

9. Click OK.

The Principal Amount gets generated, as shown in figure 4.5.

	A	B	C	D	E	F	G	H	I
1	Maturity Amount	$50,000							
2	Number of Years	5							
3	Rate of Interest	0.075							
4	Principal Amount	$36,363.64	0.080	0.085	0.090	0.095	0.100	0.105	
5		6	33783.78	33112.58	32467.53	31847.13	31250.00	30674.85	
6		7	32051.28	31347.96	30674.85	30030.03	29411.76	28818.44	
7		8	30487.80	29761.90	29069.77	28409.09	27777.78	27173.91	
8		9	29069.77	28328.61	27624.31	26954.18	26315.79	25706.94	
9		10	27777.78	27027.03	26315.79	25641.03	25000.00	24390.24	
10		11	26595.74	25839.79	25125.63	24449.88	23809.52	23201.86	
11		12	25510.20	24752.48	24038.46	23364.49	22727.27	22123.89	
12		13	24509.80	23752.97	23041.47	22371.36	21739.13	21141.65	
13		14	23584.91	22831.05	22123.89	21459.23	20833.33	20242.91	
14		15	22727.27	21978.02	21276.60	20618.56	20000.00	19417.48	
15									

Figure 4.5: Two-Input Data Table Calculating Investment Amount

Figure 4.5 displays the amount to be invested for various rates of interest and number of years. For instance, the value in cell F5 states that the amount to be invested is 31847.13 for a period of 6 years at a rate of interest of 9.5%.

10. Select cells C5:H14.

11. On the Home tab, click Number Format arrow → Currency.

The currency format is applied to the cells, as shown in figure 4.6.

	A	B	C	D	E	F	G	H	I
1	Maturity Amount	$50,000							
2	Number of Years	5							
3	Rate of Interest	0.075							
4	Principal Amount	$36,363.64	0.080	0.085	0.090	0.095	0.100	0.105	
5		6	$33,783.78	$33,112.58	$32,467.53	$31,847.13	$31,250.00	$30,674.85	
6		7	$32,051.28	$31,347.96	$30,674.85	$30,030.03	$29,411.76	$28,818.44	
7		8	$30,487.80	$29,761.90	$29,069.77	$28,409.09	$27,777.78	$27,173.91	
8		9	$29,069.77	$28,328.61	$27,624.31	$26,954.18	$26,315.79	$25,706.94	
9		10	$27,777.78	$27,027.03	$26,315.79	$25,641.03	$25,000.00	$24,390.24	
10		11	$26,595.74	$25,839.79	$25,125.63	$24,449.88	$23,809.52	$23,201.86	
11		12	$25,510.20	$24,752.48	$24,038.46	$23,364.49	$22,727.27	$22,123.89	
12		13	$24,509.80	$23,752.97	$23,041.47	$22,371.36	$21,739.13	$21,141.65	
13		14	$23,584.91	$22,831.05	$22,123.89	$21,459.23	$20,833.33	$20,242.91	
14		15	$22,727.27	$21,978.02	$21,276.60	$20,618.56	$20,000.00	$19,417.48	
15									

Figure 4.6: Cells in Currency Format

V 1.0 © 2008 Aptech Limited

Exercise 2: Working with Scenarios

Problem

William Smith and his friends decided to take a trip around the world on their vehicle by road. Create a best-case scenario and a worst-case scenario to forecast the total cost of the trip depending on the average fuel consumption of the vehicle. Also, find the number of days that will be required to complete the trip based on the data provided as shown in figure 4.7.

	A	B	C	D	E
1		**Changing Cells**			**Result Cells**
2	Distance to be travelled	15000		Cost of trip	
3	Average fuel consumption	75		Number of days travelled	
4	Cost of fuel	$20.00			
5	Estimated distance to be travelled per day	400			

Figure 4.7: Initial Data for Constructing the Scenario Model

Analysis

William would like to estimate the cost and duration of the journey before he starts his journey. William needs to determine the cost and duration for travelling in both, the best as well as worst-case scenarios. The best-case scenario would be achieving the optimal fuel consumption rate and covering the maximum distance possible. The worst-case scenario would be achieving maximum fuel consumption rates and covering minimum distance.

Total cost of the trip = (Total distance / Average fuel consumption) * Cost of fuel

Total number of days = Total distance to be travelled / Estimated distance travelled per day

Here, the total distance to be travelled is mentioned in miles, and the average fuel consumption of the vehicle is in miles per gallon.

Solution

1. **To calculate the cost of the trip, type "= (B2/B3)*B4", in cell E2.**
2. **To calculate the number of days required to complete the trip, type "= B2/B5", in cell E3.**

	A	B	C	D	E
1		**Changing Cells**			**Result Cells**
2	Distance to be travelled	15000		Cost of trip	$4,000.00
3	Average fuel consumption	75		Number of days travelled	37.5
4	Cost of fuel	$20.00			
5	Estimated distance to be travelled per day	400			

Figure 4.8: Constructing Scenario Model for the Trip Cost

3. **To name the changing cells, select cells A2:B5.**
4. **To display Create Names from Selection dialog box, on the Formulas tab, click Create from Selection.**
5. **In the Create Names from Selection dialog box, select Left column checkbox, as shown in figure 4.9.**

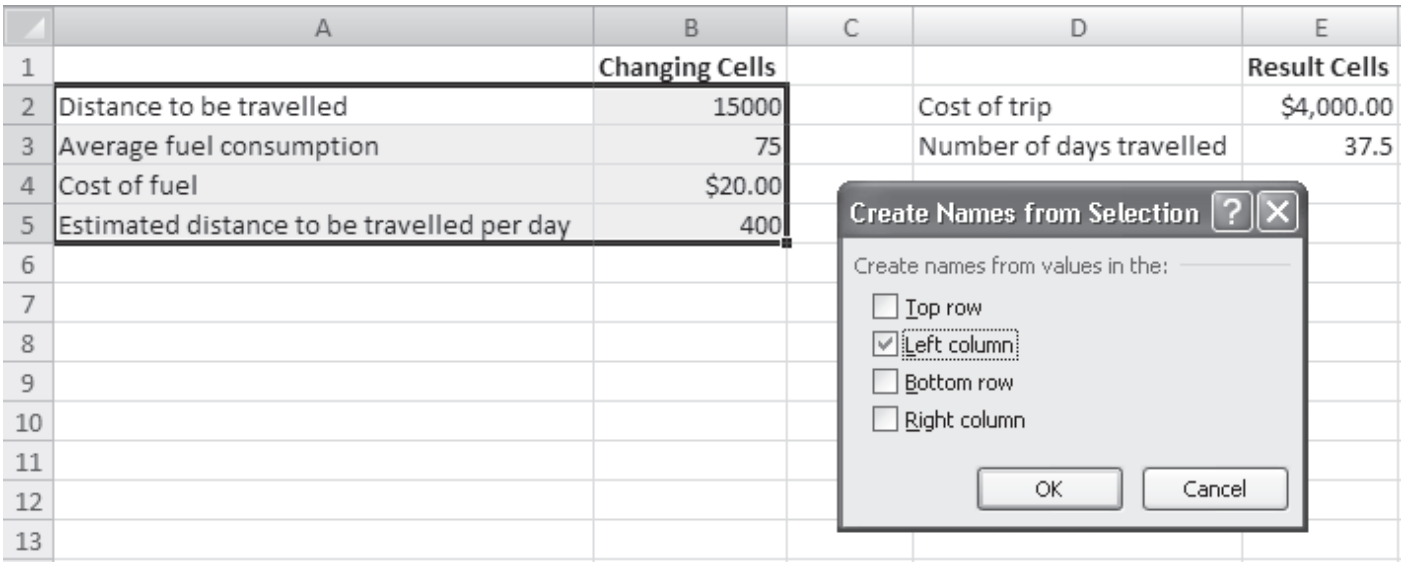


Figure 4.9: Create Names from Selection

6. Click OK.

Excel will name each cell according to its label by replacing spaces in the label with underscores. For example, the label 'Distance to be travelled' is converted to 'Distance_to_be_travelled'.

7. Similarly, name the Result Cells (E2:E3) using labels from cells D2:D3.

8. To display the Scenario Manager dialog box, from the Data Tools group, present in the Data tab, click What-If Analysis arrow → Scenario Manager, as shown in figure 4.10.

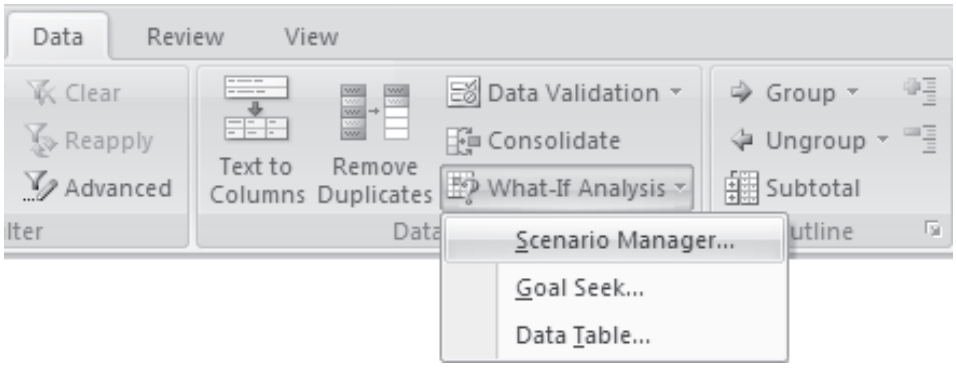


Figure 4.10: Selecting Scenario Manager

The Scenario Manager dialog box is displayed, as shown in figure 4.11.

V 1.0 © 2008 Aptech Limited

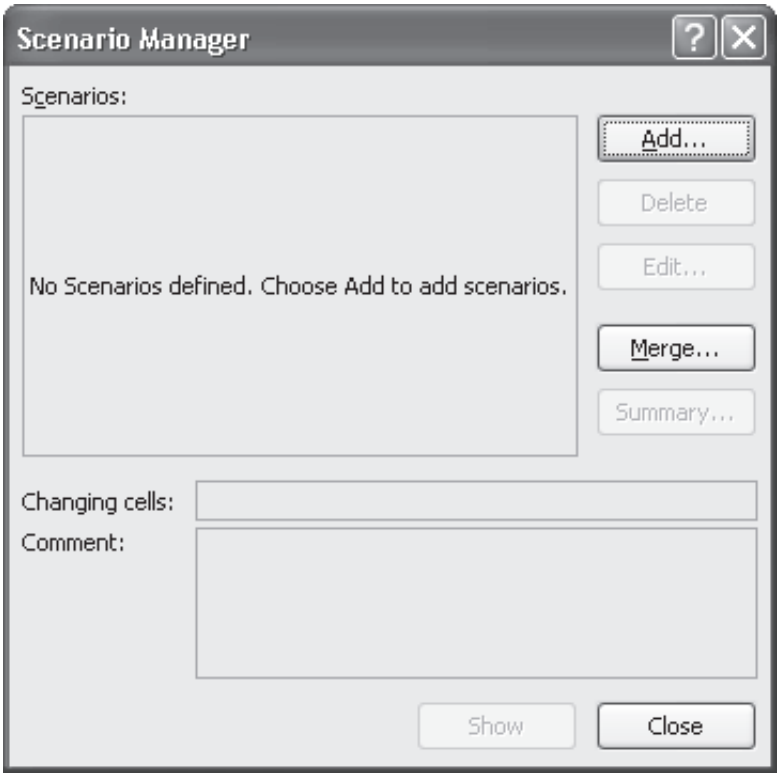


Figure 4.11: Scenario Manager Dialog Box

9. **To create a new scenario, click Add.**
10. **In the Scenario name box, type "Best Case".**
11. **In the Changing cells box, type "B2:B5".**

The Add Scenario dialog box for the first scenario is displayed, as shown in figure 4.12.

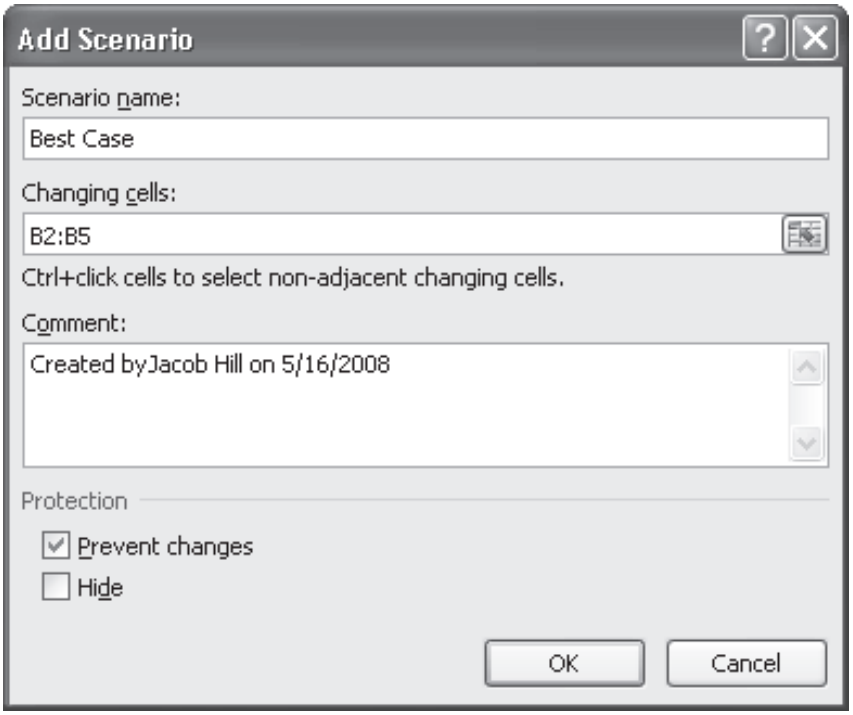


Figure 4.12: Add Scenario Dialog Box – Best Case Scenario

12. Click OK.

The Scenario Values dialog box is displayed, as shown in figure 4.13.

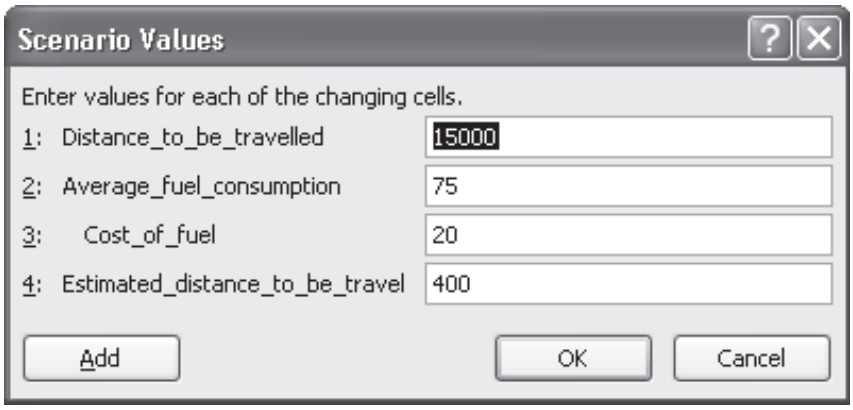


Figure 4.13: Scenario Values Dialog Box – Best Case Scenario

13. To create another Scenario, click Add.

The Add Scenario dialog box gets displayed.

14. In the Scenario name box, type "Worst Case", as shown in figure 4.14.

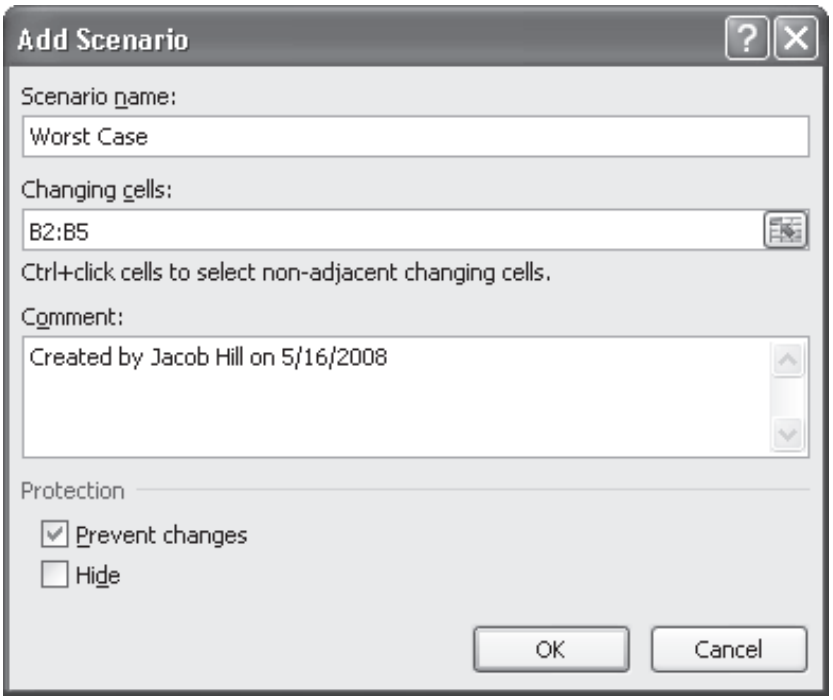


Figure 4.14: Add Scenario Dialog Box – Worst Case Scenario

15. Click OK.

The Scenario Values dialog box gets displayed.

16. In the Average_fuel_consumption box, type "65".

17. Similarly, in the Estimated_distance_to_be_travelled_per_day box, type "300" as shown in figure 4.15.

V 1.0 © 2008 Aptech Limited

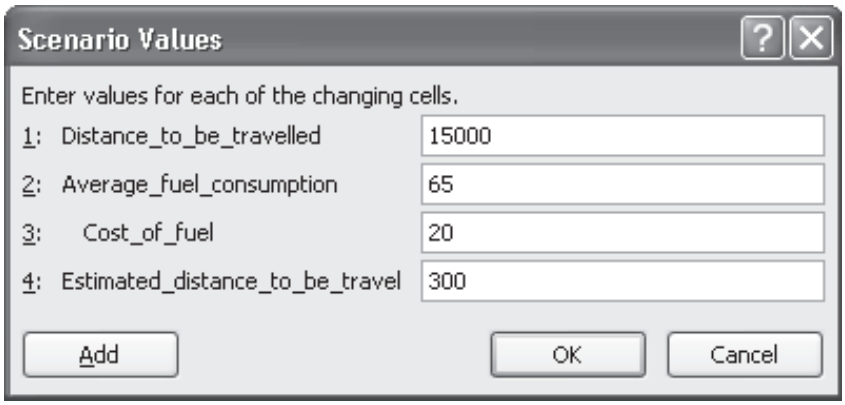


Figure 4.15: Scenario Values Dialog Box – Scenario 2

18. To return to the Scenario Manager dialog box, click OK.

The dialog box displays the names of the two scenarios created, as shown in figure 4.16.

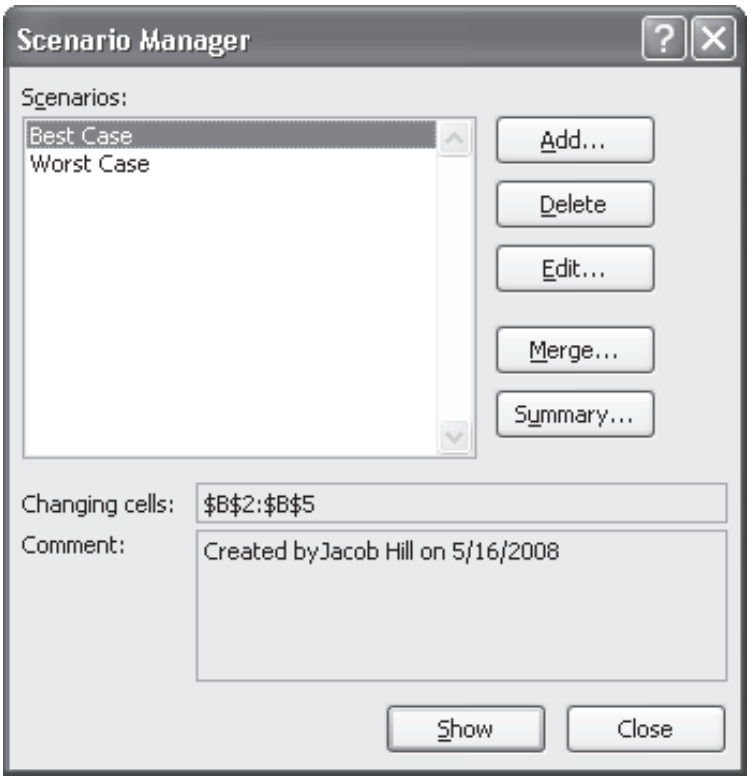


Figure 4.16: List of Defined Scenarios

19. To display the result for Worst case scenario, select Worst Case → Show.

Excel automatically inserts the corresponding values into the input cells. It calculates the input based on the formula entered in the cells E2 and E3, and displays the result for Worst Case scenario, as shown in figure 4.17.

	A	B	C	D	E	F
1		Changing Cells			Result Cells	
2	Distance to be travelled	15000		Cost of trip	$4,615.38	
3	Average fuel consumption	65		Number of days travelled	50	
4	Cost of fuel	$20.00				
5	Estimated distance to be travelled per day	300				

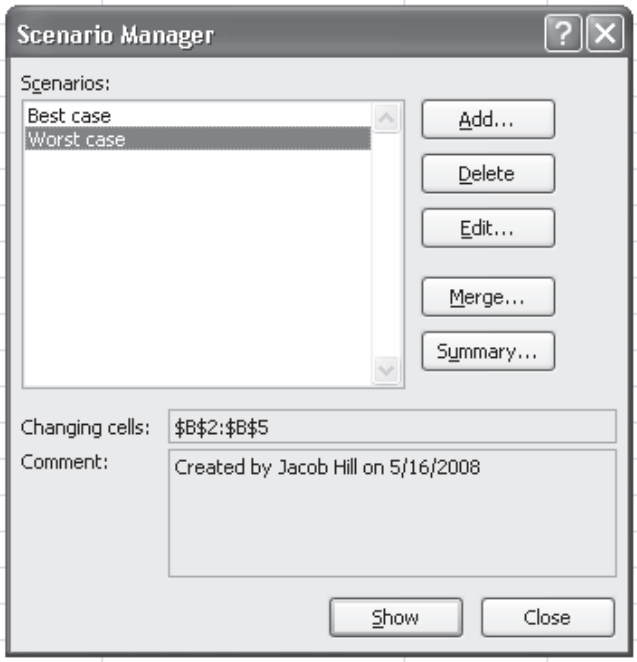


Figure 4.17: Calculated Result for Worst Case Scenario

Figure 4.17 displays the updated value in the changing cells and its corresponding value in the result cells for the Worst Case scenario.

Similarly, you can also view the result of the Best Case Scenario.

As shown in figure 4.17, after decreasing the average fuel consumption, the cost of trip increased from $4000 to $4615.38. Similarly, after decreasing the estimated distance to be travelled per day, the numbers of days required to complete the trip increased from 37.5 days to 50 days. Hence, this is a Worst Case scenario.

20. To create a summary of both the scenarios, click Summary.

The Scenario Summary dialog box gets displayed, as shown in figure 4.18.

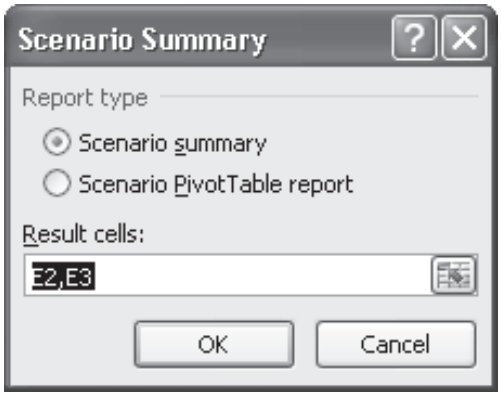


Figure 4.18: Scenario Summary Dialog Box for the Report Type

By default, cells E2 and E3 are displayed in the Result cells box.

V 1.0 © 2008 Aptech Limited

21. To create a report of Scenario Summary type, click OK.

The Scenario Summary report gets displayed, as shown in figure 4.19.

Scenario Summary	Current Values:	Best case	Worst case
Changing Cells:			
Distance_to_be_travelled	15000	15000	15000
Average_fuel_consumption	75	75	65
Cost_of_fuel	$20.00	$20.00	$20.00
Estimated_distance_to_be_travel	400	400	300
Result Cells:			
Cost_of_trip	$4,000.00	$4,000.00	$4,615.38
Number_of_days_travelled	37.5	37.5	50

Notes: Current Values column represents values of changing cells at time Scenario Summary Report was created. Changing cells for each scenario are highlighted in gray.

Figure 4.19: Scenario Summary Report

22. To create a report of Scenario PivotTable report type, click Scenario PivotTable report from Scenario Summary dialog box.

23. Click OK.

The Scenario PivotTable report type gets displayed, as shown in figure 4.20.

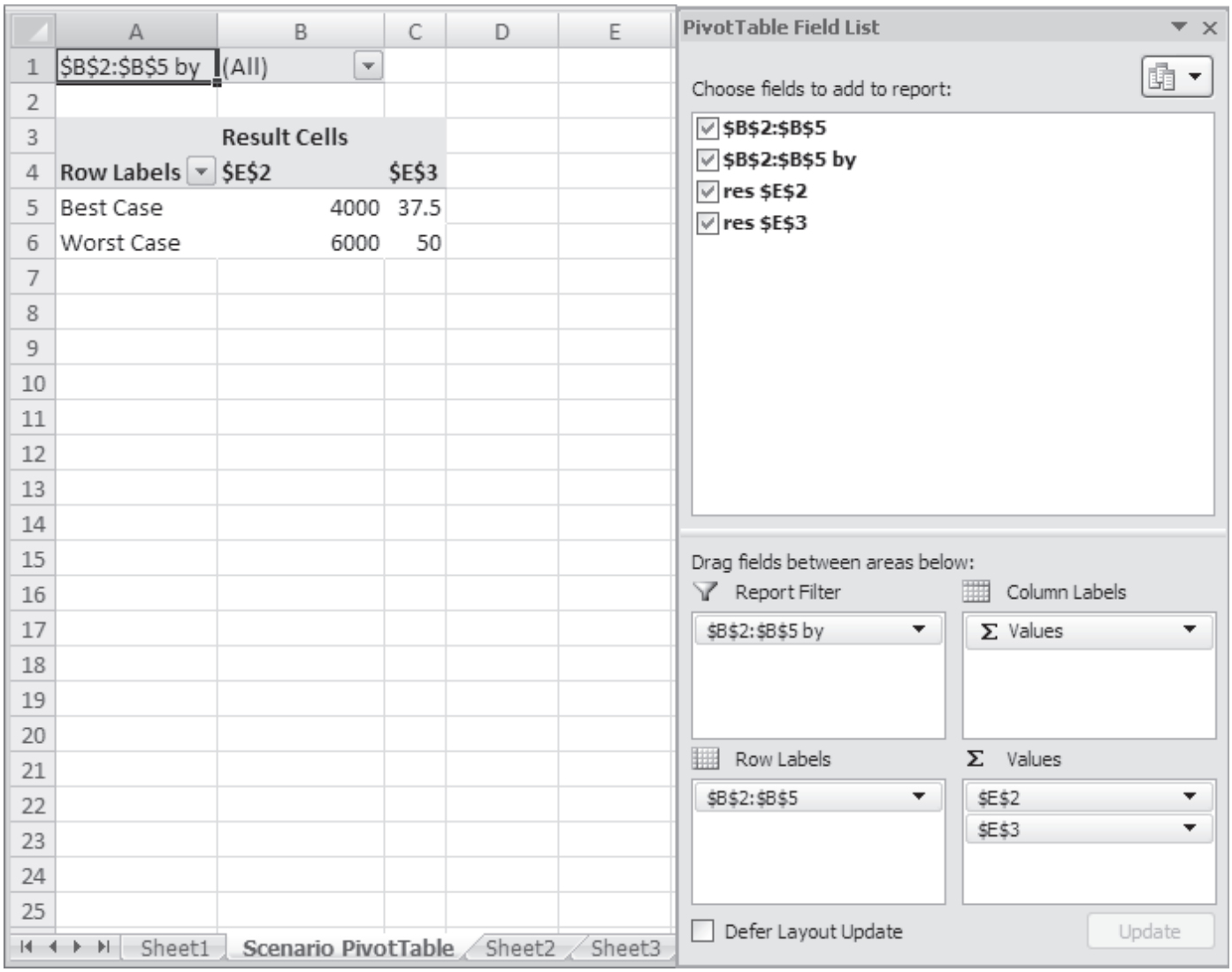


Figure 4.20: Scenario PivotTable Report

Part II - 30 Minutes

1. InfraSoft Solutions have been in the business of manufacturing chips in Europe. The new manager of the company wants to forecast the profit that the company should make in the future. Create a scenario to forecast the growth rate as well as the profit of the company for the next five years. Consider the following factors while creating the scenario: Gross Sales, Cost of Production, Sales Tax, and Profit in the current year.

V 1.0 © 2008 Aptech Limited

Do It Yourself

1. Hillwater Corporation wants to make a loan analysis for the company by making a two-variable data table. Create a data table to determine the total payment amount for different loan amounts at various interest rates. The interest rates to be considered are 6%, 6.25%, 6.50%, 6.75%, 7%, 7.25%, and 7.50%. The loan amounts should range from $50,000 to $60,000 with $1,000 increments, and the term should be from 5 years to 10 years with an increment of 1.

   Hint: The formula to be used is: **=PMT(Annual Interest/12, Term [Months], Loan Amount)**.

This page has been intentionally left blank

Session 5

Goal Seek and Solver

Session Objectives:

At the end of this session, you will be able to -

- Explain What-If analysis using Goal Seek tool
- Describe the Solver tool
- Describe the steps involved in using the Solver tool
- Describe the various Solver tool settings

5.1 Goal Seek

Many Excel users input data into a worksheet, use simple functions or formulas to calculate results, and then use this result for taking decisions. Although this is a perfectly legal use of Excel, it basically turns Excel into a calculator.

What-If analysis involves asking questions of the type "What if a value changes?". Excel provides a number of tools to find answers to "What-If" type of questions. What-If analysis is the procedure of changing the cell values to see how those changes affect the result of formulas. Excel provides features to analyze data and solve complex problems by employing variables and constraints. Sometimes while analyzing a problem, an end user is clear about the end result to be achieved but not about the data that should be entered in the Excel sheet to achieve the desired result. The Goal Seek option of Excel allows accomplishing this easily. Thus, it can be said that, using Goal Seek, Excel calculates the input value required by varying the value in the formula until the desired result is achieved.

Goal seeking feature says, "This is the final value that is to be achieved, what is the required input value ?"

Consider a situation where temperature need to be converted from Celsius to Fahrenheit, as shown in figure 5.1. So, in one cell of the Excel sheet, A1, a value representing temperature in Celsius has been entered and in cell A2, the conversion formula `=CONVERT(A1, "C", "F")` has been entered. The `C` in the `CONVERT` formula represents temperature in Celsius and `F` represents temperature in Fahrenheit. If the temperature has been entered as 37° C then the temperature is converted to its Fahrenheit value and returned. But, sometimes the user may want the result to be 100° F and be unaware of the input value. In such a case the user could type one value after another in cell A1 in a trial-and-error fashion (10, 11, 12, 12.5, and so on), until cell A2 displays 100. But this is a much time consuming method. A much quicker and more accurate method of solving this problem is to use the Goal Seek feature of Excel.

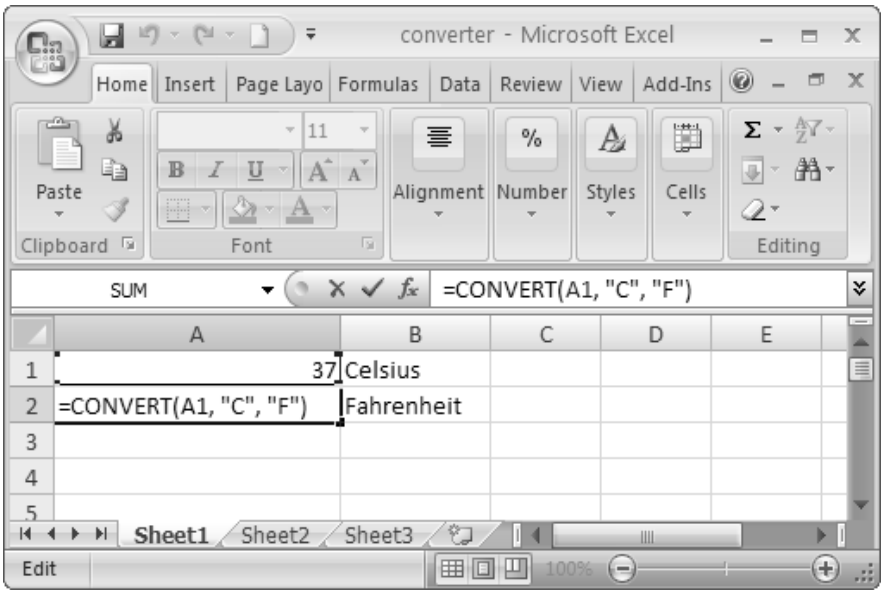


Figure 5.1: Converting Celsius to Fahrenheit

5.1.1 Goal Seeking

To use Goal Seek, click the **Data** tab, click **What-If Analysis**, and then click **Goal Seek** option, as shown in figure 5.2.

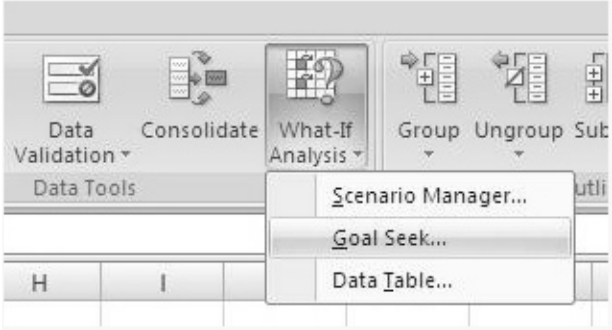


Figure 5.2: Locating Goal Seek

The **Goal Seek,** dialog box appears, as shown in figure 5.3.

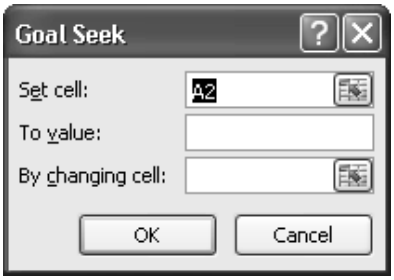


Figure 5.3: Goal Seek Dialog Box

To perform **Goal Seek** the values of three options need to be specified. These three options are:

- **Set cell** box: Specifies the address of the cell containing the formula to be set.
- **To value** box: Specifies the value that is to be displayed in the cell that was provided in the **Set cell** box.
- **By changing cell** box: Specifies the address of the cell whose value needs to be changed.

V 1.0 © 2008 Aptech Limited

Note: The cell where the new value is displayed must contain the formula and the cell whose value needs to be changed should be specified in the formula.

After entering the details in the **Goal Seek** dialog box, click **OK** or press **Enter** to display the **Goal Seek Status** dialog box. The **Goal Seek Status** dialog box displays the required target value in the **To value** box and the current value of the cell in the **By changing cell** box. Sometimes these two values may not match, each other when Excel does not find a solution.

In the Goal Seek example of Celsius to Fahrenheit conversion, the goal to seek **100** Fahrenheit is specified in cell **A2**. Cell **A1** needs to be changed to achieve this. The values in the **Goal Seek** dialog box for the example are entered as shown in figure 5.4.

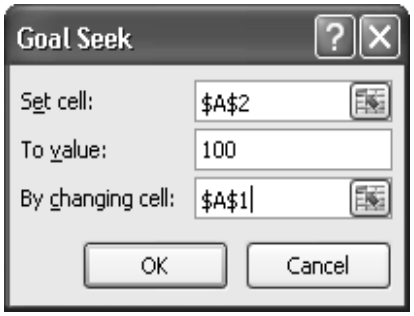


Figure 5.4: Goal Seek for Celsius to Fahrenheit

Click **OK**. The result of the conversion example is displayed in figure 5.5.

Figure 5.5: Goal Seek Result

In this case, the Celsius temperature in **A1** has been adjusted to 37.77777778. If the result is correct, click **OK** to retain the adjustments, else click **Cancel** to reject the changes.

Consider an example where a person wants to calculate how the term of the loan will be affected if the amount of the monthly loan repayments is increased to $5,000. The worksheet for the Proposed Housing loan is shown in figure 5.6.

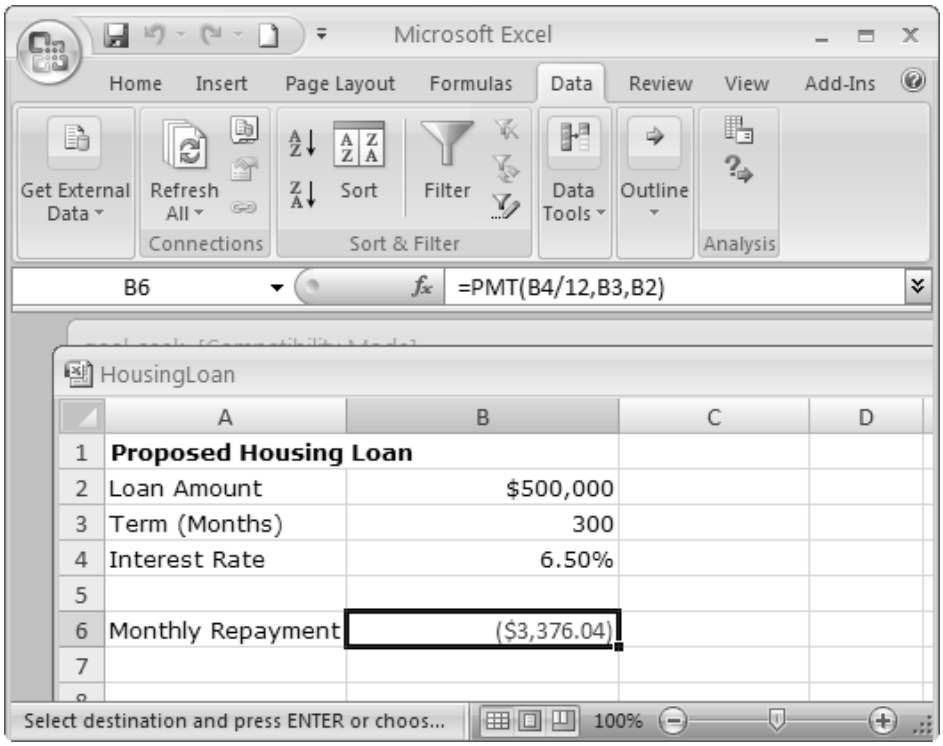


Figure 5.6: Worksheet for Proposed Housing Loan

To perform Goal Seek operation, the following steps needs to be followed:

1. Click on **Goal Seek** option from the **What-If Analysis** drop-down list to display the **Goal Seek** dialog box.
2. Type B6 in the **Set cell** box.
3. Type **-5000** in the **To value** box.
4. To achieve the desired result in **B6**, the cell that needs to be modified is **B3**. Type **B3** in the **By changing cell** box.
5. Click **OK** to start goal seeking.

The new term (144 months) of the loan is visible, as shown in figure 5.7. The user can choose to either accept (click **OK**) the new values or reject (click **Cancel**) the changes and return the worksheet to its original state.

V 1.0 © 2008 Aptech Limited

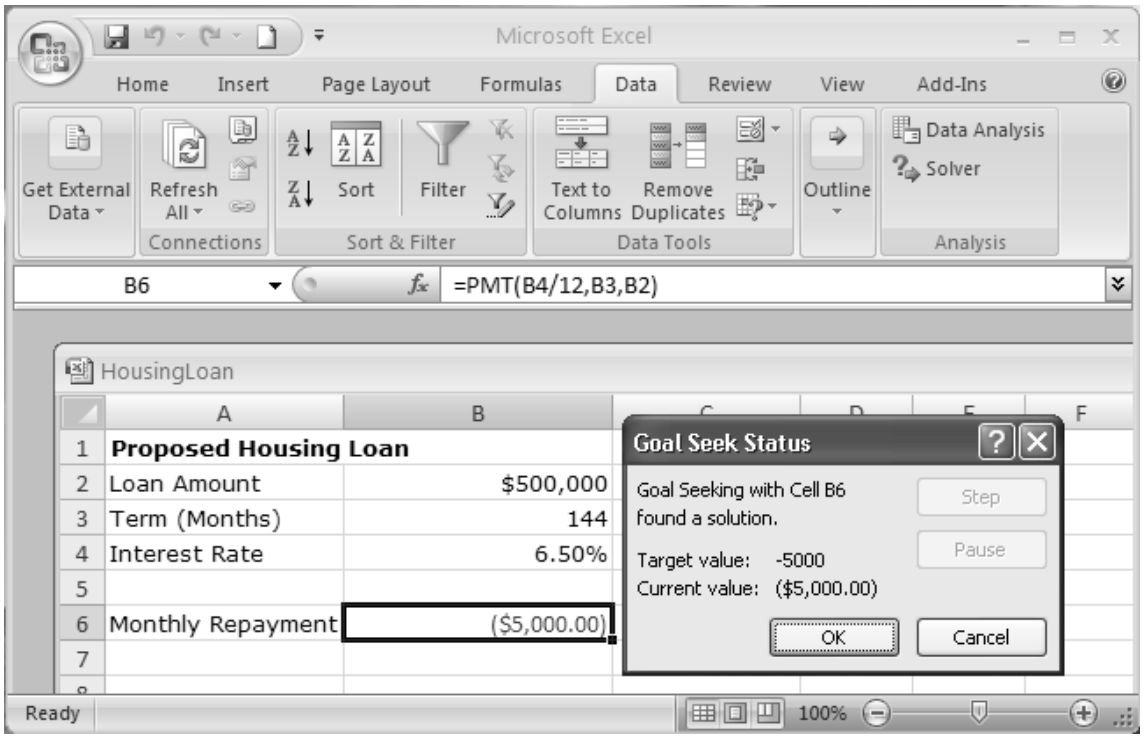


Figure 5.7: Goal Seek Status of Proposed Housing Loan

Some important points to be considered when troubleshooting a Goal Seek are:

- The **Set cell** reference is set to a cell that contains a formula, function, or cell reference.
- The value in the **To value** cell is a reasonable value.
- The **By changing cell** reference is a number or blank, and not a formula, function, or cell reference.
- There should be a link by formulas between the two cells entered in the Goal Seek option.

5.2 Solver Tool

Goal Seek is used for simple problems where there is a direct relationship between the input value and the results in the formula. The disadvantage of using Goal Seek is that it can be used for one adjustable cells and can return only one value. For complex problems, Solver needs to be used. More than one changing cell can be specified in order to increase the number of possibilities. Constraints can be built in to restrict the analysis to operate only under specific conditions.

The basis for using the Solver tool is usually to alter many values to produce the optimum result for a single formula. This could mean, for example, altering price values to maximize profits. It could mean adjusting expenditure to minimize costs and so on. The variable values to be adjusted must have an influence, either directly or indirectly, on the overall result in any case. It means, the changing cells must affect the formula to be optimized. Up to 200 changing cells can be included in the solving process, and up to 100 constraints can be built in to limit the Solver results.

5.2.1 What is Solver?

Solver is a tool that changes the cell value in a worksheet to a specified value, highest value or lowest value for a worksheet cell formula. It finds a set of decision variables that achieve the best possible value of an output. Solvers are software tools available in Excel, which help the users find the best way to allocate scarce resources such as raw materials, machine time, people time, and so on, in limited supply. The best solution is maximizing profits, minimizing costs or achieving the best possible quality. Solver tackles much more complicated problems, including scenarios that have multiple changing

cells, additional rules, and subtle relationships. It helps answer questions such as:

- How should we allocate the budget to maximize profit?
- How much inventory should be stocked of each type of product, with the given constraints of shelf size and budget?
- Which suppliers will give you the optimal combination of price and delivery?

Some of the terminologies that are used in Solver are:

- **Target cell**: Represents a cell whose value will be set to a specified, maximum and minimum value by the Solver.
- **Objective:** Represents the desired result of the problem that the Solver should achieve.
- **Adjustable cells**: Represents the cells whose values are changed by the Solver to achieve the desired results.
- **Constraints**: Represents the limitations that are imposed on the adjustable cells. Solver must adhere to these constraints while trying to achieve the desired results.
- **Model**: Represents the target cell, the adjustable cells and the constraints for the problem that the Solver is solving.

Solver should build a model that comprises constraints, objectives and adjustable cells. Applications of Solver include financial, budgeting, and investment problems. Solver also helps solve mathematical models, which have been entered into an Excel worksheet. These mathematical models can either be:

- **Linear Programming problem**: Specifies a linear relationship among all constraints and the objective function.
- **Integer Programming problem**: Specifies that the decision variables can only take integer values in a given range (these integer values can also be boolean = 0 or 1 only).

Solver is designed to use an iterative process. It starts with an initial solution provided as a guess by the user, and then solves the problem iterating over it, using different values in the changing variables. When a change to a variable results in a better solution, Solver makes another change to the same variable in the same direction. When a change results in a worse solution, it does not make any more changes to that variable in that direction. It continues to make changes and re-run the problem, until it arrives at a solution that is not significantly better than the previous one. At this point, it reports success.

The differences between Solver and Goal Seek, when both are used together with a single target cell, are listed in table 5.1.

Goal Seek	Solver
It can work with only a single changing cell	It can work with several changing cells
Finds out the exact possible values	Finds out the highest, lowest and exact possible values
Does not allow specifying restrictions, or constraints for cell values	Allows specifying restrictions, or constraints for cell values
Settings in goal seek are not saved when the worksheet is closed	The last solver settings for the worksheet are saved

Table 5.1: Differences between Goal Seek and Solver

5.2.2 Setting Solver Parameters

Solver is not enabled by default. To add Solver, click **Office**, and click **Excel Options** to display the **Excel Options** dialog box. In the **Excel Options** dialog box, click **Add-Ins** from the left pane, select **Add Solver add-in**, and finally click **Go** to display the **Add-Ins** dialog box. Select the **Solver Add-in**, as shown in figure 5.8.

V 1.0 © 2008 Aptech Limited

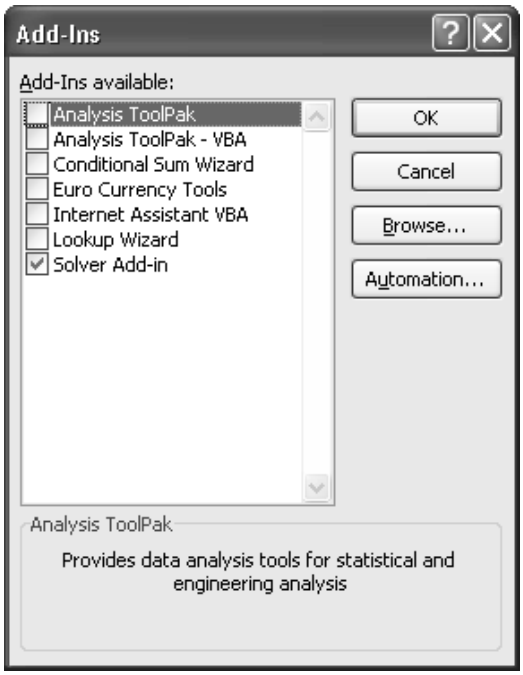


Figure 5.8: Add-Ins Dialog Box

To use Solver, click **Data tab** and then click **Solver** from the **What-If** drop-down list. The **Solver Parameters** dialog box is displayed, as shown in figure 5.9.

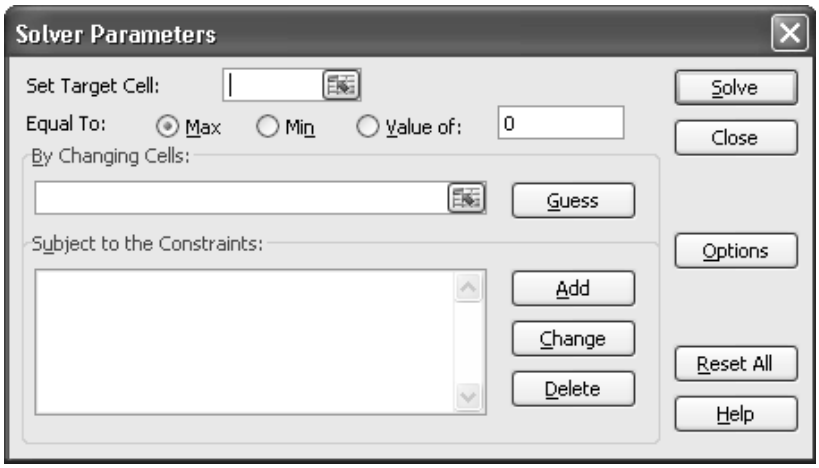


Figure 5.9: Solver Parameters Dialog Box

The Solver tool needs quite a lot of information in order for it to be able to come up with a realistic solution. This information is known as Solver parameters. The **Solver Parameters** dialog box consists of the following fields:

- **Set Target Cell**: Represents a target cell in the worksheet that needs to be set to a specified, maximum or minimum value. This cell must contain a formula.
- **Equal To**: Gives the option of treating the target cell in three alternative ways. **Max** (the default) option tells Excel to maximize the target cell value with the given constraints. **Min** option attempts to reach the minimum value with the specified constraints. **Value of** option attempts to reach the specified value of the target.
- **By Changing Cells:** Represents the cells that can be adjusted until the constraints in the problem are satisfied and the cell in the **Set Target Cell** box reaches its target. The adjustable cells must be related, directly or indirectly, to the target cell.

- **Guess:** Guesses all non-formula cells referred to by the formula in the **Set Target Cell** box, and places their references in the **By Changing Cells** box.
- **Subject to the Constraints:** Lists the current restrictions on the problem.
- **Add:** Displays the **Add Constraint** dialog box to add constraints.
- **Change:** Displays the **Change Constraint** dialog box to change constraints.
- **Delete**: Removes the selected constraint from the **Subject to the Constraints** list.
- **Solve**: Starts the solution process for the defined problem.
- **Close:** Closes the **Solver Parameters** dialog box without solving the problem. Retains any changes made by using the **Options**, **Add**, **Change**, or **Delete** buttons.
- **Options:** Displays the **Solver Options** dialog box, to load and save problem models and control advanced features of the solution process.
- **Reset All** - Clears the current problem settings, and resets all settings to their original values.

Each Solver option has a default setting that is appropriate for most problems, as shown in figure 5.10.

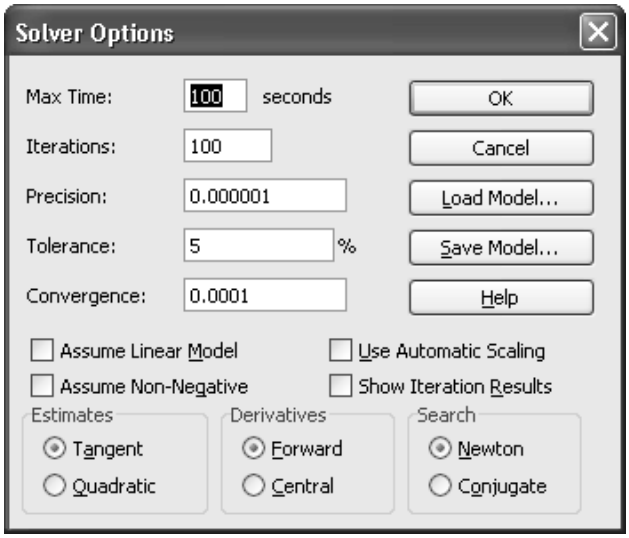


Figure 5.10: Solver Options Dialog Box

The **Solver Options** dialog box consists of the following fields:

- **Max Time** - Specifies the maximum time in seconds that the Solver should try to reach the objective. The maximum value that can be entered is 32,767 seconds and by default it displays 100 seconds.
- **Iterations** - Specifies the maximum number of times that the Solver will calculate to reach the objective or give up.
- **Precision** - Specifies the precision level that the constraint cell should achieve to meet the target value. For example, if the value in this box is 0.000001, and the constraint says J3 must equal 14, Solver will stop when the value in cell J3 is within 0.000001 of 14, that is, between 13.999999 and 14.000001.
- **Tolerance** - Similar to **Precision**, but used for integer constraints. Solver stops when the answer is within this percentage of the constraint value.
- **Convergence** - Specifies the amount of change that the Solver is allowed to make in its last five calculations before the Solver stops.
- **Assume Linear Model** - If the problem can be solved with linear functions, the Solver can be speeded up by checking this box.

V 1.0 © 2008 Aptech Limited

- **Show Iteration Results** - Instructs Solver to pause to show the results of each iteration.
- **Use Automatic Scaling** - If the problem involves large differences in magnitude — for example, when maximizing the percentage of profit based on million-dollar investments, this option is used to scale the magnitude automatically.
- **Assume Non-Negative:** Causes Solver to assume a lower limit of 0 (zero) for all adjustable cells for which a lower limit has not been set in the **Constraint** box in the **Add Constraint** dialog box.
- **Estimates:** Specifies two approaches used to obtain initial values of the basic variables. They are:
  - **Tangent:** Instructs Solver to use an algorithm that performs a linear extrapolation from a tangent vector.
  - **Quadratic:** Instructs Solver to use an algorithm that performs a quadratic extrapolation, which can improve the results on highly nonlinear problems.
- **Derivatives:** Specifies two algorithms that the Solver uses for calculations of solutions by estimating partial derivatives of the objective and constraint functions. They are:
  - **Forward:** Specifies use for most of the problems, in which the constraint values change relatively slowly.
  - **Central:** Specifies use for problems in which the constraint values change rapidly, especially near the limits. Although this option requires more calculations, it can help when Solver returns a message saying that it could not improve the solution.
- **Search:** Specifies two algorithms that the Solver will use for each iteration to determine the direction to search. They are:
  - **Newton:** Uses a quasi-Newton method that typically requires more memory but fewer iterations than the Conjugate gradient method. It results in a highly accurate search for possible solutions.
  - **Conjugate:** Requires less memory than the Newton method but typically needs more iterations to reach a particular level of accuracy. This option is used when there is a large problem and memory usage is a concern, or when stepping through iterations reveals slow progress.
- **Load Model:** Displays the **Load Model** dialog box, where the reference for the model to be loaded is specified.
- **Save Model**: Displays the **Save Model** dialog box, to specify where to save the model. This is used only when more than one model with a worksheet is required to be saved — the first model is automatically saved.

Note: After applying the "Keep Solver Solution" option, the undo operation cannot be performed on the changes to the values in the cells.

5.2.3 Adding Constraints and Creating Solver Reports

Business decisions are often subject to constraints or business rules to solve business problems. Adding constraints restrict the changes that Excel can make to the cell values. The **Add Constraint** dialog box is shown in figure 5.11.

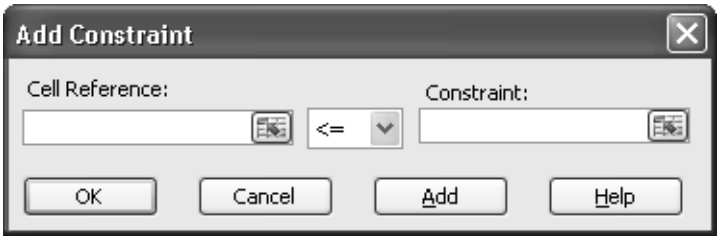


Figure 5.11: Add Constraint Dialog Box

In most models, constraints play a key role in determining what values can be assumed by the decision variables. Constraints

reflect real-world limits on variables. The **Add Constraint** dialog box consists of the following fields:

- **Cell Reference** box: Specifies the cell references for which to restrict the values.
- **Operator list between the Cell Reference and Constraint** boxes: Specifies the relationship between the referenced cells and the constraint. The different relationship choices are listed in table 5.2.

Operator List	Description
<=	Specifies that the referenced cell values must be less than or equal to the constraint
=	Specifies that the referenced cell values must be equal to the constraint
>=	Specifies that the referenced cell values must be greater than or equal to the constraint
int	Specifies that the referenced cell values must be integers. If this operator is selected, Integer is displayed in the Constraint box
bin	Specifies that the referenced cell values must be only one of two values, such as Yes/No, True/False, or 0/1. If this operator is selected, Binary is displayed in the Constraint box

Table 5.2: Operator List in the Add Constraints Dialog Box

- **Constraint** box: Contains the cell reference, a number, or a formula.

When maximizing or minimizing a formula value, it is important to include constraints, which sets upper or lower limits on the changing values. For instance, when maximizing Profits by changing Income figures, the Solver could conceivably increase these figures to infinity. If the income figures are not limited by an upper constraint, the Solver will return an error message. Similarly, minimizing total costs could be achieved by making one of the contributing costs, that is, salaries and production costs, infinitely less than zero. A constraint should be included, therefore, to set a minimum level on these values.

After setting the solver parameters and adding constraints, to solve, click **Solve.** When the Solver finds a solution, the **Solver Results** dialog box is displayed, as shown in figure 5.12.

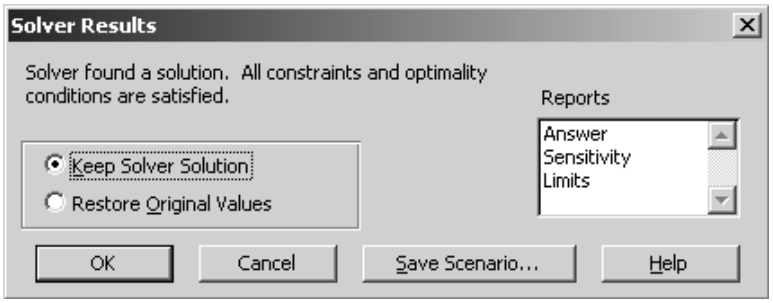


Figure 5.12: Solver Results Dialog Box

The **Solver Results** dialog box provides the following features:

- The **Keep Solver Solution** option replaces the values in the adjustable cells in the worksheet with the suggested Solver values.
- The **Restore Original Values** option retains the original values in the adjustable cells.
- The **Save Scenario** option saves the solution as a scenario that can be used in the **Scenario Manager**.
- The report options from the list creates an **Answer**, **Sensitivity** and/or a **Limits** report to view how Solver reached the solution.

V 1.0 © 2008 Aptech Limited

If the maximum time or iterations are reached before Solver finds a solution, or if the **Show Iteration Results** option is selected, the **Show Trial Solution** dialog box appears, as shown in figure 5.13.

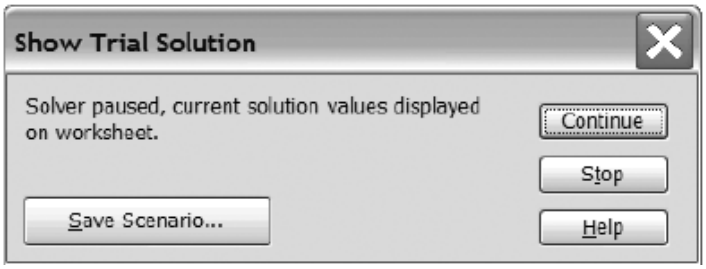


Figure 5.13: Show Trial Solution Dialog Box

The different messages displayed in the **Show Trial Solution** dialog box are:

- **Solver paused, current solution values displayed on worksheet**: This message is displayed when the S**how Iteration** box is selected in the **Solver Options** dialog box or when the Solver is interrupted by pressing of the **Esc** key.
- **The maximum iteration limit was reached; continue anyway?**: This message is displayed when the maximum number of allowed **Iterations** has been completed by the Solver.
- **The maximum time limit was reached; continue anyway?**: This message is displayed when Solver has finished running for the maximum number of seconds specified in the **Max Time** box in the **Solver Options** dialog box.

The solution process of the Solver can be stopped in the **Show Trial Solution** dialog box using **Stop**. The solution process of the Solver can be continued to the next trial solution in the **Show Trial Solution** dialog box using **Continue**. The current trial solution's results can be saved as a scenario, using **Save Scenario**.

If a solution is found by the Solver, it allows creation of one or a combination of reports. The different types of reports are:

- **Answer Report**: It lists the target cell and the changing cells with their original and final values, constraints, and information about the constraints. The Answer Report consists of the following details:

  The address of the target and adjustable cells, the names assigned to them, the original values of the cells before the Solver was run and the final values after the Solver was run.

  The address of the constraints cell, the names assigned to them, the values of the cells, the formulas of the constraints, the status of the constraints (Binding, means a slack value of 0, or Not Binding, means a nonzero slack value), and the slack values of the constraints. A slack value is the absolute difference in values between the left side and the right side of the constraint.
- **Sensitivity Report**: The Sensitive report specifies how sensitive the solution is to minor changes made in the target cell's formula or constraints. Excel can generate a Sensitivity Report in two parts - a changing cells report and a constraints report. For the Changing Cells report, the allowable increase and decrease refers to how much the objective function decision variable coefficient can change without changing the value of any of the decision variables. For the constraints report it includes shadow prices, objective coefficient (with allowable increase and decrease), and constraint right-hand side ranges. The shadow price is the amount that the objective function value would change if the named constraint is changed by one unit. The shadow price is valid up to the allowable increase or decrease in the constraint.
- **Limits Report**: It lists the target cell and the changing cells with their respective values, lower and upper limits, and target values. The lower limit is the smallest value that the changing cell can take while holding all other changing cells fixed and still satisfying the constraints. The upper limit is the greatest value.

5.2.4 Working with Solver

Consider a factory that manufactures three electronic products: TV sets, stereos, and speakers. Each product is assembled using different parts present in the inventory. The five different parts are: chassis, picture tubes, speaker cones, power supplies, and electronics units. The aim is to produce a mix of products which will maximize profits, given the inventory of products on hand.

In this example, the decision variables are the number of products to build. The objective function is (gross) profit. Assume selling the TV sets for a gross profit of $75 each, stereos for a profit of $50 each, and speaker cones for $35 each.

To assemble a TV set, 1 chassis, 1 picture tube, 2 speaker cones, 1 power supply, and 2 sets of electronics units are required. To make a stereo, 1 chassis, 2 speaker cones, 1 power supply, and 1 set of electronics units are required. To build a speaker, 1 speaker cone, and 1 set of electronics unit is required. The different parts available in the inventory are 450 chassis, 250 picture tubes, 800 speaker cones, 450 power supplies, and 600 sets of electronics units.

Assume the following symbols: x for the number of TV sets to be assembled, y for the number of stereos, and z for the number of speakers. The objective function is to find a total profit and can be represented as:

Maximize 75 x + 50 y + 35 z (Profit).

Building each product requires a certain number of parts of each type. The number of parts used depends on the mix of products built (constraint left hand side), and the total number of parts of each type available (constraint right hand side):

1 x + 1 y + 0 z <= 450 (Chassis)

1 x + 0 y + 0 z <= 250 (Picture tubes)

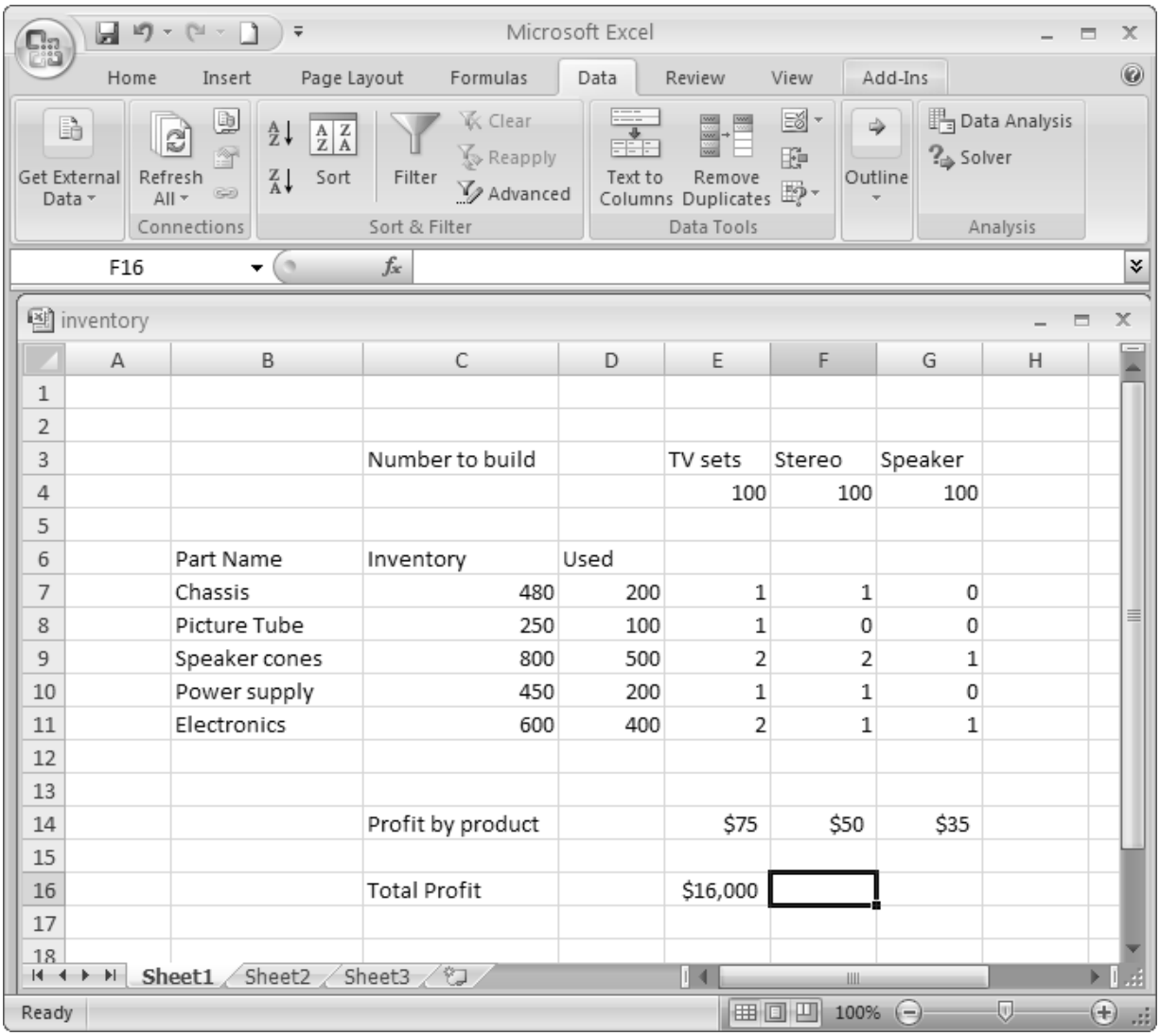
2 x + 2 y + 1 z <= 800 (Speaker cones)

1 x + 1 y + 0 z <= 450 (Power supplies)

2 x + 1 y + 1 z <= 600 (Electronics)

The next step is to create a worksheet where the formulas for the objective function and the constraints are calculated, as shown in figure 5.14.

V 1.0 © 2008 Aptech Limited



	A	B	C	D	E	F	G	H
1								
2								
3			Number to build		TV sets	Stereo	Speaker	
4					100	100	100	
5								
6		Part Name	Inventory	Used				
7		Chassis	480	200	1	1	0	
8		Picture Tube	250	100	1	0	0	
9		Speaker cones	800	500	2	2	1	
10		Power supply	450	200	1	1	0	
11		Electronics	600	400	2	1	1	
12								
13								
14			Profit by product		$75	$50	$35	
15								
16			Total Profit		$16,000			
17								
18								

Figure 5.14: Inventory Worksheet

Click **Solver** in the **Data** tab. The **Solver** dialog box appears. Set the target cell, which is the value of the objective function, the Total Profit, to cell E16. Set the **By Changing Cells** or the value of the decision variables, in this case, the number of TVs, stereos, and speakers to E4:G4.

Specify the constraints: Number of used parts must be less than inventory, or D7:D11 <= C7:C11, and number of TV sets, stereos, and speakers must be non-negative, or E4:G4 >= 0, as shown in figure 5.15.

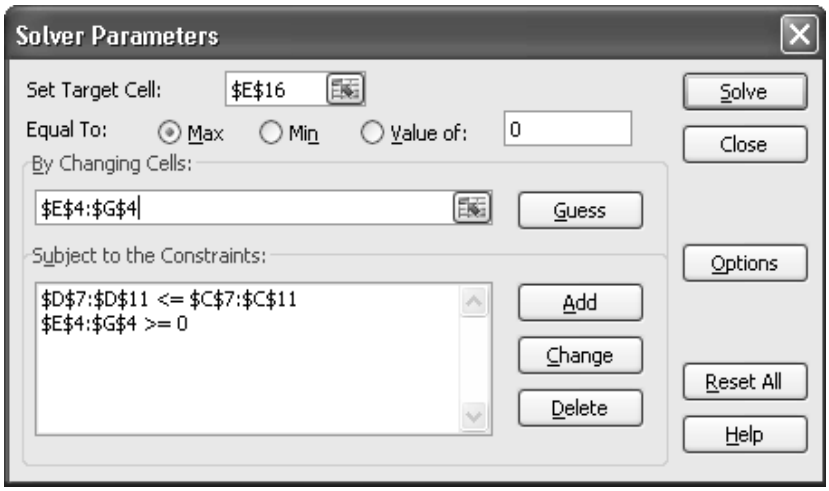


Figure 5.15: Setting the Solver Parameters for Inventory

Excel Solver is ready to run. Click the **Solve** button. The **Solver Results** dialog box appears, as shown in figure 5.16.

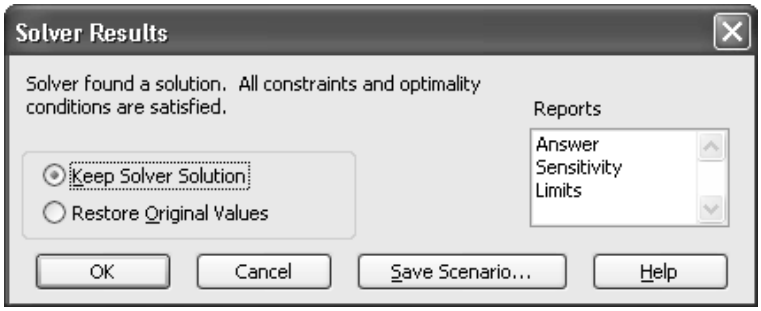
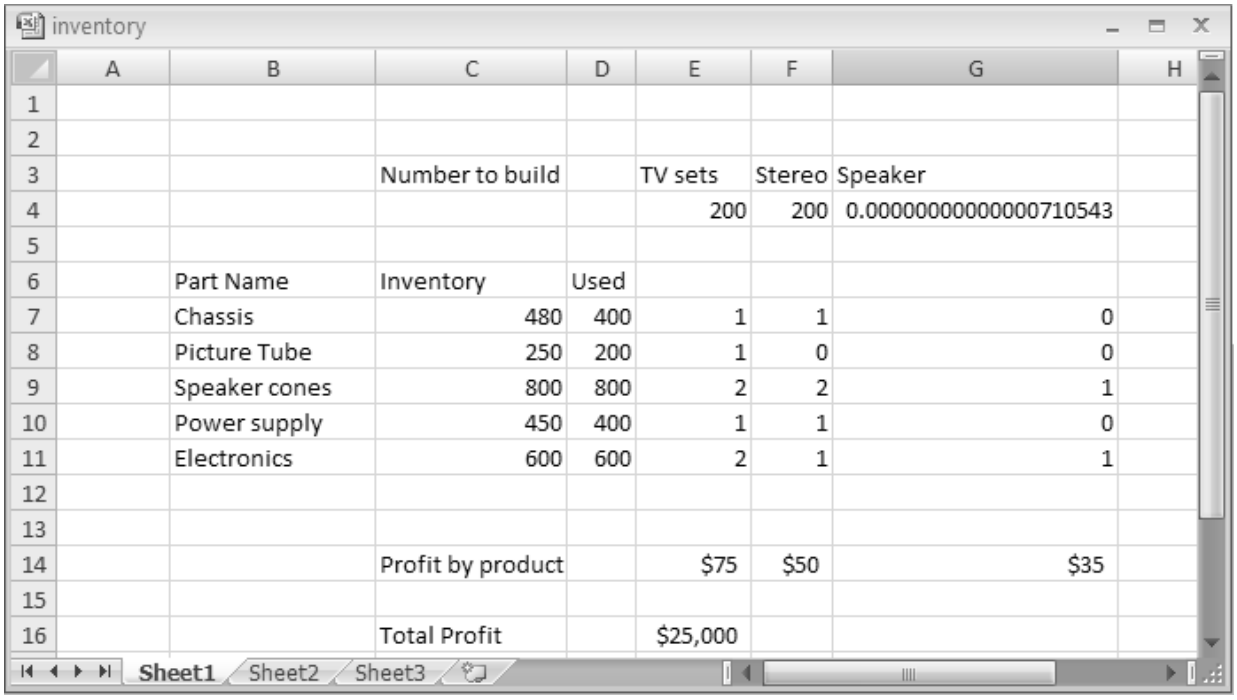


Figure 5.16: Solver Results for Inventory

The solution generated by the solver for the example is shown in figure 5.17.



	A	B	C	D	E	F	G	H
1								
2								
3			Number to build		TV sets	Stereo	Speaker	
4					200	200	0.00000000000000710543	
5								
6		Part Name	Inventory	Used				
7		Chassis	480	400	1	1	0	
8		Picture Tube	250	200	1	0	0	
9		Speaker cones	800	800	2	2	1	
10		Power supply	450	400	1	1	0	
11		Electronics	600	600	2	1	1	
12								
13								
14			Profit by product		$75	$50	$35	
15								
16			Total Profit		$25,000			

Figure 5.17: Solution

V 1.0 © 2008 Aptech Limited

The Session In Brief

- Goal Seek is an easy-to-use, timesaving tool. It is used for simple problems where there is a direct relationship between the input value and the results in the formula.
- Solver helps find the specified, minimum, and maximum value for a single worksheet cell formula.
- Unlike Goal Seek, Solver allows specifying restrictions, or constraints for cell values.
- Unlike Goal Seek, the last settings in Solver are saved when the worksheet is closed.

Check Your Progress

1. _____ report shows how sensitive the solution is to minor changes in the objective function formula or constraints.

   a. Answer

   b. Sensitivity

   c. Analysis

   d. Limits

2. The _____ box in the Goal Seek dialog box specifies the location of the formula used to get the end result.

   a. Set Cell

   b. To value

   c. By Changing Cell

   d. Location

3. The _____ in the Solver options dialog box specifies the percentage by which the target cell of a solution satisfying the integer constraints can differ from the true optimal value and still be considered acceptable.

   a. Tolerance

   b. Iterations

   c. Precision

   d. Max Time

4. Which of the following finds out the highest, lowest and exact possible values?

   a. Solver

   b. Goal Seek

   c. Pivot Tables

   d. Data Tables

5. Settings in ____ are not saved when the worksheet is closed.

   a. Goal Seek

   b. Solver

   c. Solver Parameters

   d. Solver Results

V 1.0 © 2008 Aptech Limited

Session 6

Goal Seek and Solver (Lab)

Session Objectives:

At the end of this session, you will be able to -

- Use the Goal Seek tool
- Use the Solver tool

Part I - 60 Minutes

Exercise 1: Working with Goal Seek

Problem

The Accounts manager of the ABC company, has been provided with a new budget, where the total annual allowance of his department has been raised to to $430,000.00. This new budget was authorized so that he can add one new employee to the department. Before the authorization of the new budget, the total allowance of the department was $377,825.00. Use Goal Seek to determine the new employee's salary.

Analysis

The new budget of $430,000.00 is the goal to be reached. The goal has to be achieved on the total annual allowance. The formula used here for the Totals is `=SUM(B3:B11)`, that is, the total annual salary for all the members of his department.

Solution

1. **Create a worksheet with the data, as shown in figure 6.1.**
2. **Type the function, =SUM(B3:B11), in cell B13.**

The `=SUM(B3:B11)` calculates the total salary of the cell range, B3 to B11.

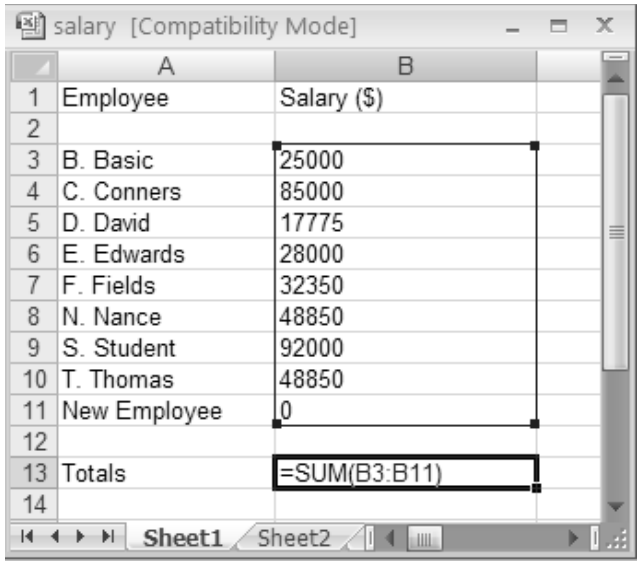


Figure 6.1: Worksheet

3. **Click the Data tab.**
4. **Click the What-If Analysis option and then select the Goal Seek option.**

The **Goal Seek** dialog box is displayed.

5. **Click on cell B13 to set the value in the Set cell box to $B$13.**
6. **Type the value '430,000' in the To value box.**
7. **Click on the By changing cell box and select the result cell, B11.**
8. **Click OK.**

The **Goal Seek** dialog box will display the values, as shown in figure 6.2.

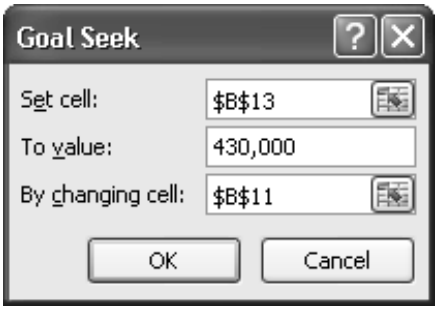


Figure 6.2: Goal Seek Dialog Box

The **Set cell** box represents the the cell address containing the formula. The value in the **To value** box is the target value that needs to be set, that is 430,000.The **By changing cell** $B$11 specifies the cell address whose value needs to be changed. Here, it specifies the salary of the new employee. On clicking **OK**, the **Goal Seek Status** dialog box appears, as shown in figure 6.3.

V 1.0 © 2008 Aptech Limited

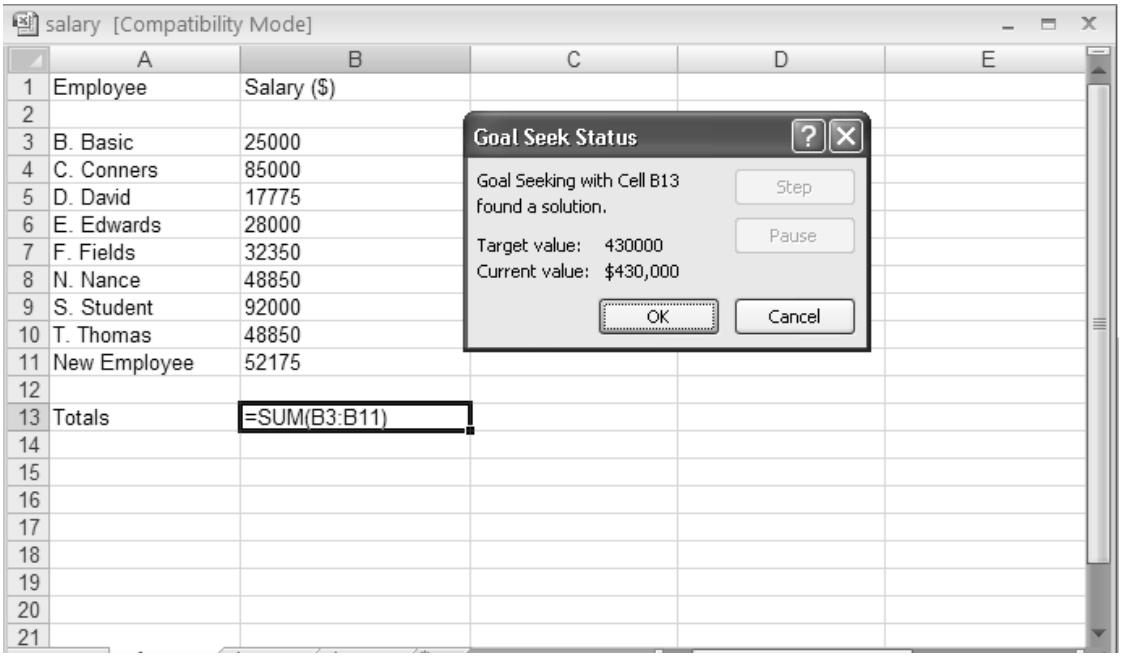


Figure 6.3: Goal Seek Status Dialog Box

The result is shown in figure 6.3. Notice that in cell B11, the cell corresponding to the new employee's salary, has now changed to the amount that can be allocated, as it is within the budget restrictions.

9. Click OK to accept the result.

Thus, Goal Seek modifies the contents of one cell to achieve a new result in another cell.

Exercise 2: Working with Solver

Problem

A farmer has 2400 feet of land and wants to fence off a rectangular field that borders a straight river. Find the dimensions of the fence to maximize the area enclosed by the fence.

Analysis

The fence land represents a rectangular area. To maximize the area, A, of the rectangle having width and height of y and x respectively, the formula will be A=xy. Since, 2400 ft of fence is already available, the linear equation is 2x + y = 2400. Using this constraint for the amount of fence represented by y, the equation will be y = 2400 – 2x. Now, substituting the value of y in the original equation, which calculated the area (A=xy), results in a function where x is maximized. The function is:

$A = x(2400 - 2x) = 2400x - 2x^2$.

Solution

By default, all the add-ins are not loaded when MS Excel is installed.

1. **To load the Solver add-in, click Office → Excel Options.**
2. **Select Add-Ins from the left pane.**
3. **Select Solver Add-in from the View and manage Microsoft Office add-ins, as shown in figure 6.4.**

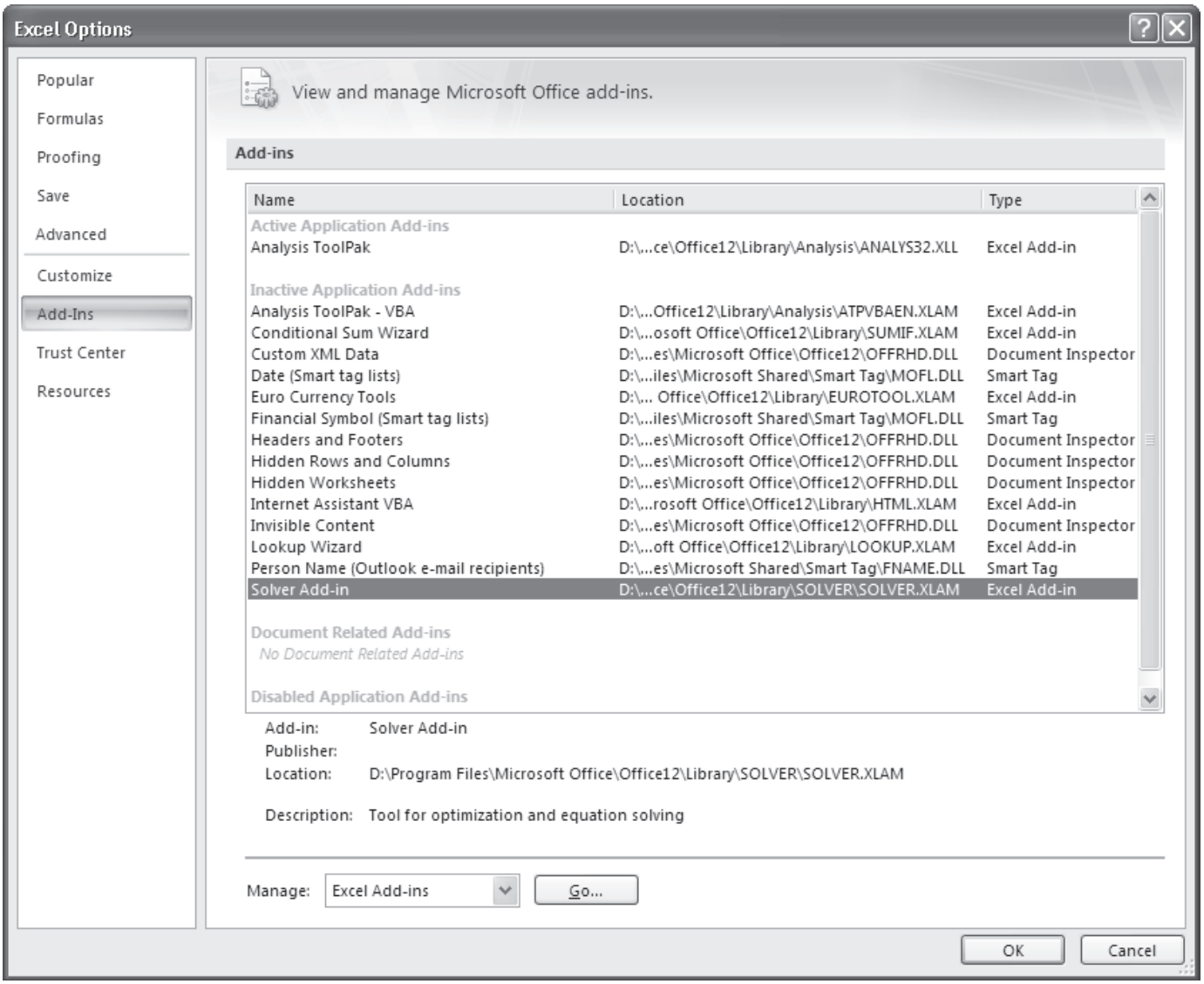


Figure 6.4: Excel Options

4. **Click Go to display the Add-Ins dialog box.**
5. **Select the Solver Add-in from the Add-Ins dialog box, as shown in figure 6.5.**

V 1.0 © 2008 Aptech Limited

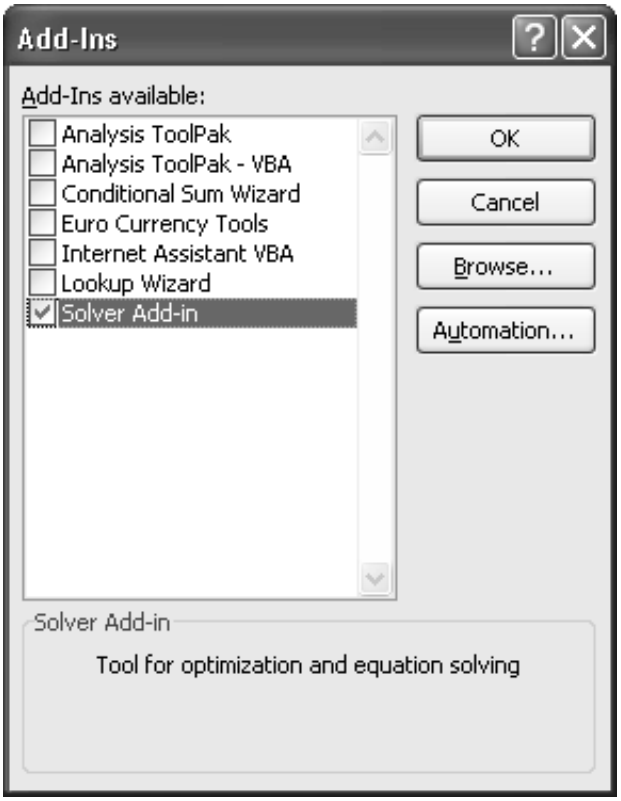


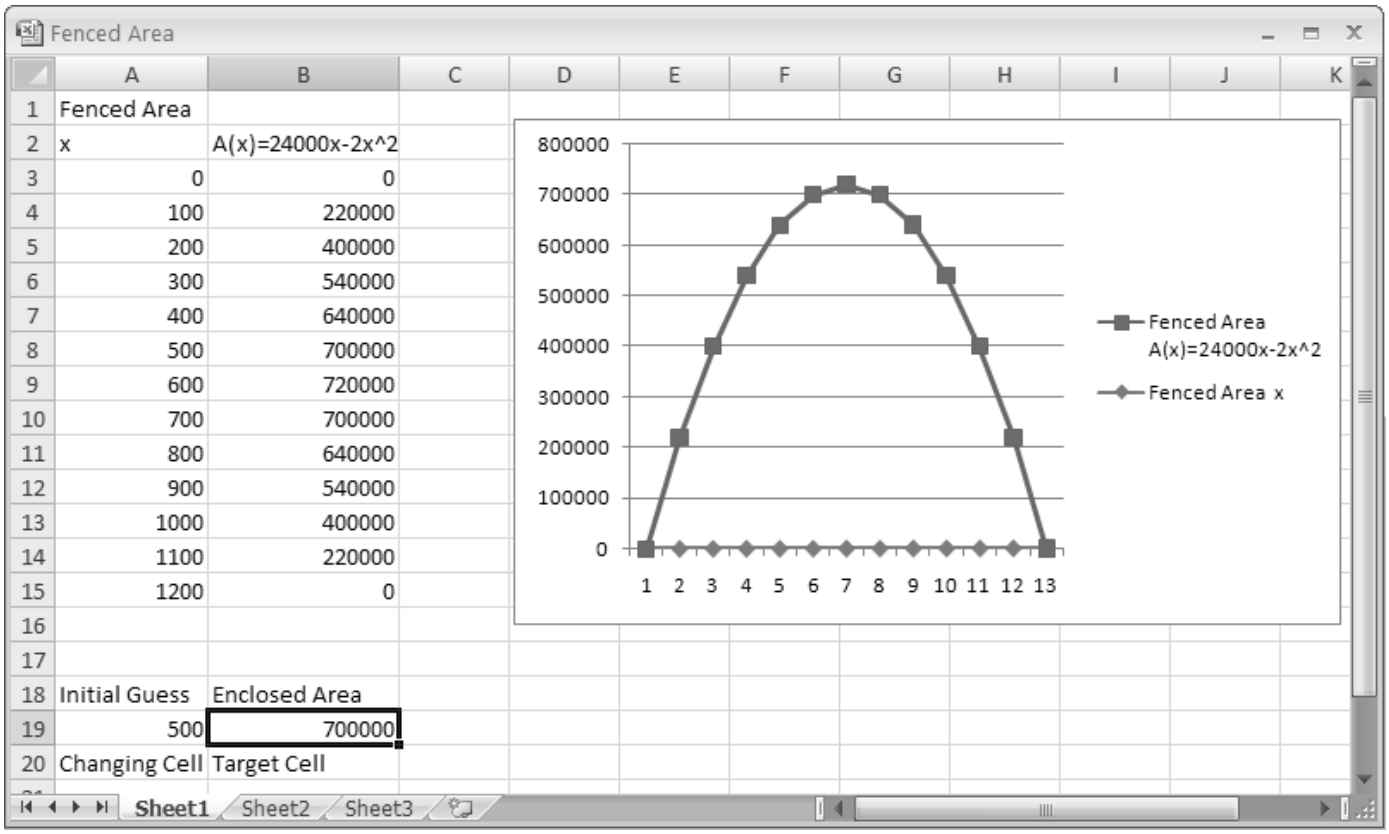
Figure 6.5: Add-Ins Dialog Box

6. Click OK.

The **Solver** option is now available in the **Data** tab. When installing the first time, it will ask for confirmation.

7. Create a worksheet named Fenced Area with the data and line chart, as shown in figure 6.6.

The B3-B15 and cell B19 use the formula A(x)=24000x-2x^2 to calculate the fenced area.



	A	B
1	Fenced Area	
2	x	A(x)=24000x-2x^2
3	0	0
4	100	220000
5	200	400000
6	300	540000
7	400	640000
8	500	700000
9	600	720000
10	700	700000
11	800	640000
12	900	540000
13	1000	400000
14	1100	220000
15	1200	0
16		
17		
18	Initial Guess	Enclosed Area
19	500	700000
20	Changing Cell	Target Cell

Figure 6.6: Worksheet

8. **Click B19 cell, representing the Enclosed Area.**
9. **Click Solver from the Data tab.**

The **Solver Parameters** dialog box is displayed, as shown in figure 6.7.

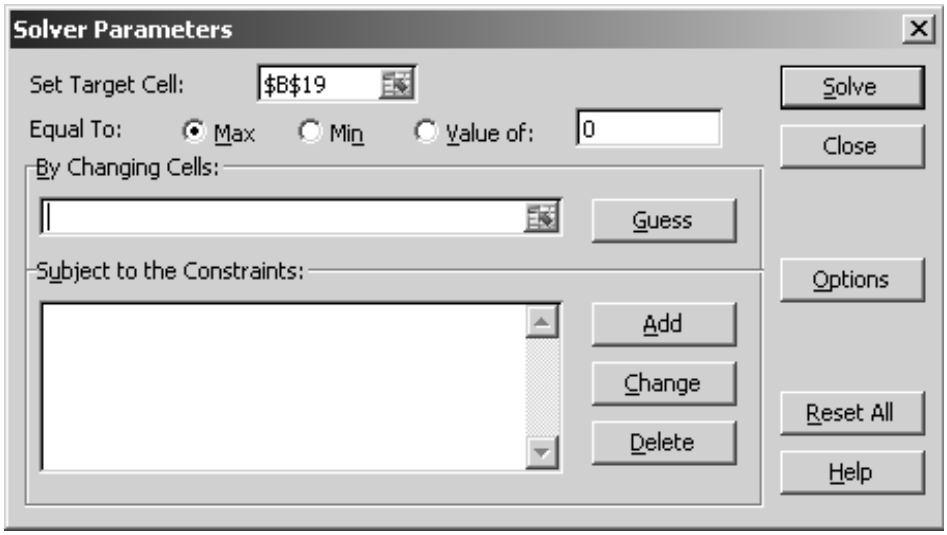


Figure 6.7: Solver Parameters Dialog Box

10. **Select Max from the Equal To choice group.**
11. **In the By Changing Cells box, select A19.**
12. **Click Solve.**

The **Solver Results** dialog box appears, as shown in figure 6.8.

V 1.0 © 2008 Aptech Limited

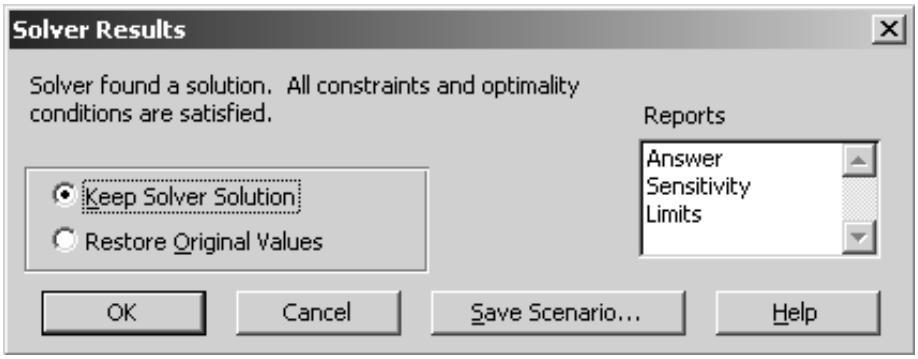


Figure 6.8: Solver Results Dialog Box

The solution found by the Solver is displayed in figure 6.9. The value in the cell, A19, changes to 599.9999984.

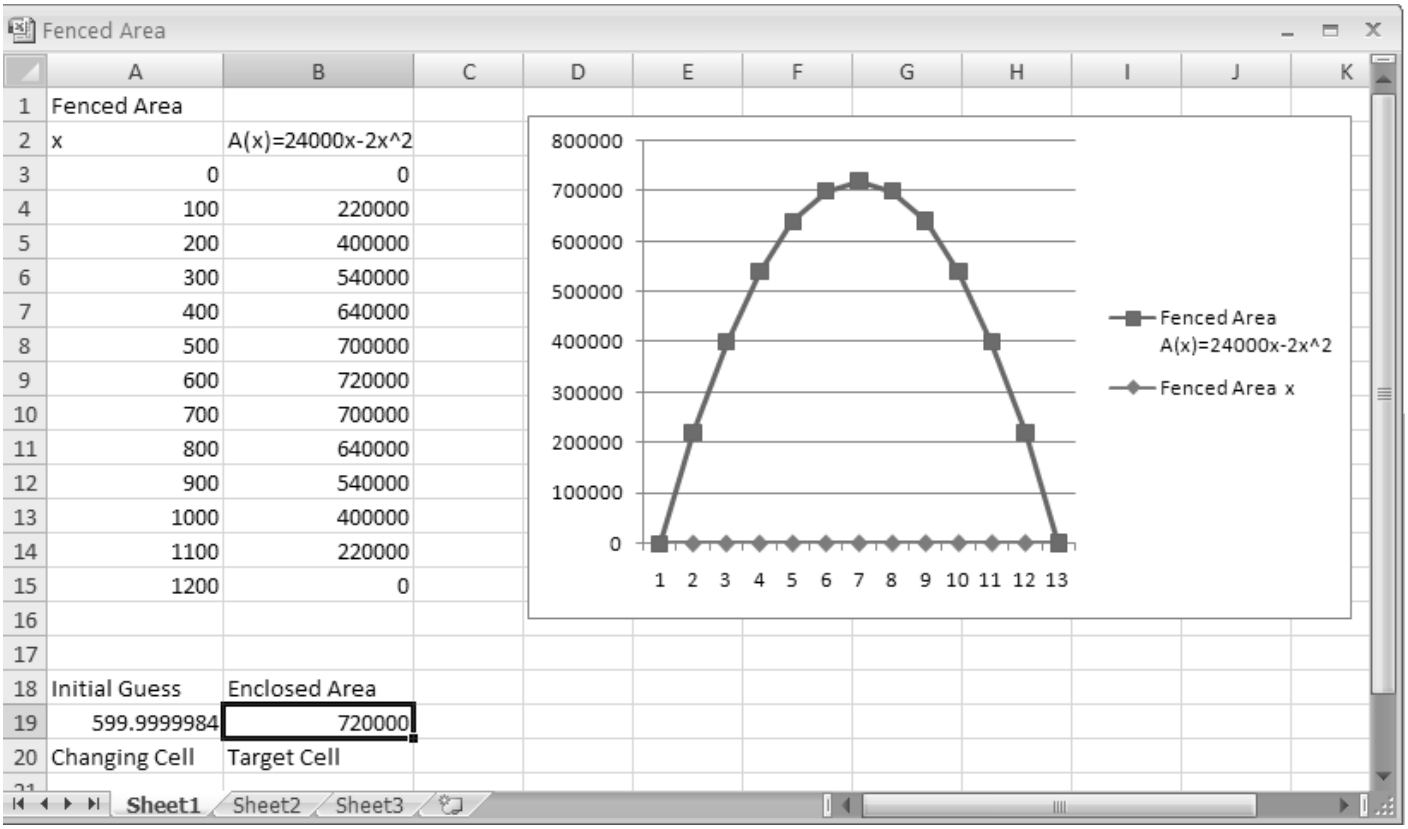
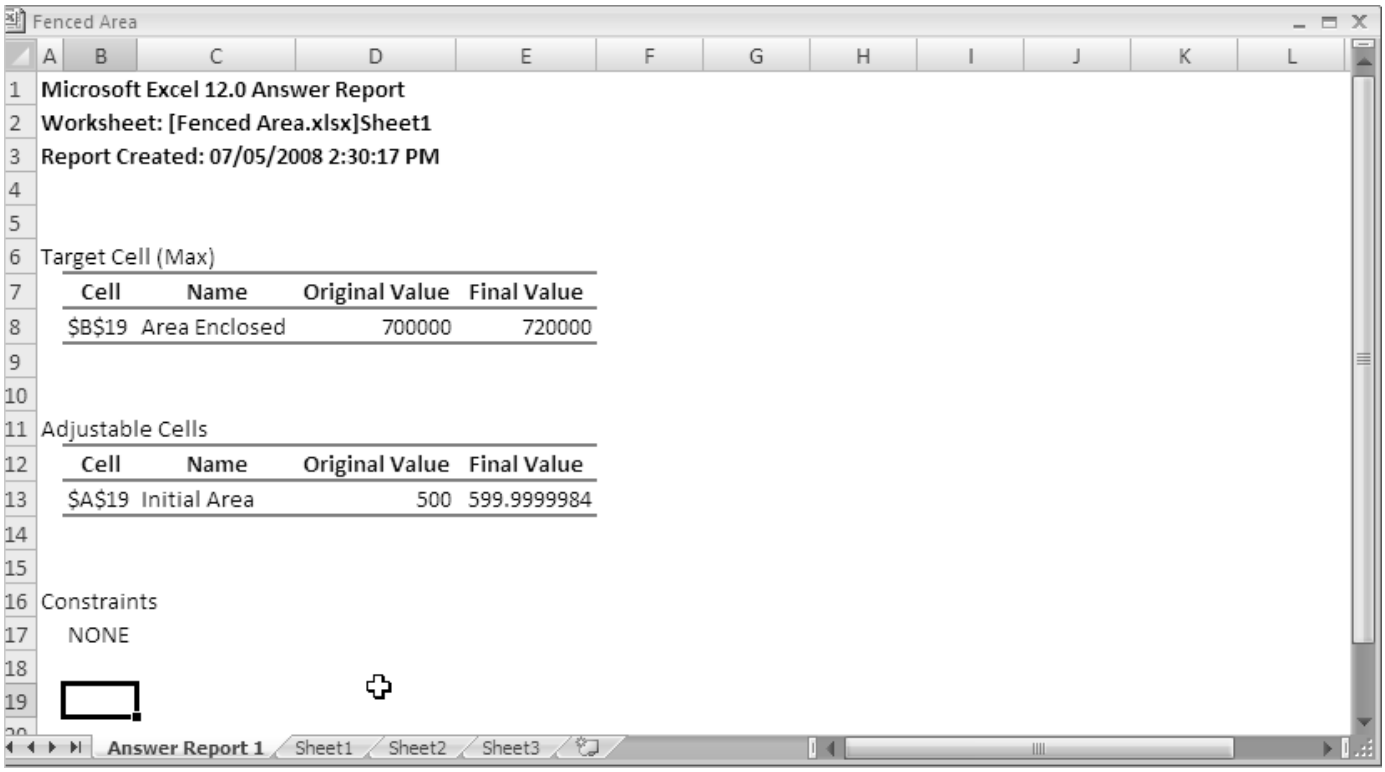


Figure 6.9: Solution

13. Select the Answer option from the Report pane, in the Solver Results dialog box, to create the Answer report.

Solver creates the answer report and stores it in a separate worksheet called “Answer Report 1”. The Answer report generated is shown in figure 6.10.



Fenced Area

Microsoft Excel 12.0 Answer Report
Worksheet: [Fenced Area.xlsx]Sheet1
Report Created: 07/05/2008 2:30:17 PM

Target Cell (Max)

Cell	Name	Original Value	Final Value
$B$19	Area Enclosed	700000	720000

Adjustable Cells

Cell	Name	Original Value	Final Value
$A$19	Initial Area	500	599.9999984

Constraints

NONE

Answer Report 1 | Sheet1 | Sheet2 | Sheet3

Figure 6.10: Answer Report

Exercise 3: Adding Constraints in Solver

Problem

The Colorado Cattle Company (CCC) can purchase three types of raw feed ingredients from a wholesale distributor. The company has certain nutritional requirement standards with respect to fat, protein, calcium, and iron content for the cow. Each cow requires at least 10 units of calcium, not more than 7.5 units of fat, at least 12 units of iron, and at least 15 units of protein per day. Table 6.1 shows the amount of fat, protein, calcium, and iron content available in each pound of the three types of feed ingredients. Grade 1, Grade 2, and Grade 3 feeds costs $.25, $.10, and $.08 per pounds respectively. The cow can be fed a mixture of the three types of raw feed. CCC would like to feed its herd as cheaply as possible.

	Grade 1	Grade 2	Grade 3
Calcium	0.7	0.8	0.0
Iron	0.9	0.8	0.8
Protein	0.8	1.5	0.9
Fat	0.5	0.6	0.4

Table 6.1: Feed Ingredients (Units Per Pound)

Analysis

A linear programming (LP) model of this problem follows:

Assume,

grade1 = total amount of grade 1 feed (in lbs.) to be used in feeding a cow per day

grade2 = total amount of grade 2 feed (in lbs.) to be used in feeding a cow per day

grade3 = total amount of grade 3 feed (in lbs.) to be used in feeding a cow per day

V 1.0 © 2008 Aptech Limited

Minimize

0.25 grade1 + 0.1 grade2 + 0.08 grade3 (objective function)

subject to (constraints)

0.7 grade1 + 0.8 grade2 + 0.0 grade3 ≥ 10 (Calcium)

0.9 grade1 + 0.8 grade2 + 0.8 grade3 ≥ 12 (Iron)

0.8 grade1 + 1.5 grade2 + 0.9 grade3 ≥ 15 (Protein)

0.5 grade1 + 0.6 grade2 + 0.4 grade3 ≤ 7.5 (Fat)

grade1, grade2, grade3 ≥ 0

Solution

1. Create a worksheet named CCC with the data, as shown in figure 6.11.

CCC [Compatibility Mode]

	A	B	C	D	E	F	G
4							
5	Grade	1	2	3			
6	Cost /lb	0.25	0.1	0.08		Minimum	Maximum
7	Calcium /lb	0.7	0.8	0	Calcium	10	
8	Iron /lb	0.9	0.8	0.8	Iron	12	
9	Protein /lb	0.8	1.5	0.9	Protein	15	
10	Fat /lb	0.5	0.6	0.4	Fat		7.5
11							
12	Decision Variables						
13							
14		Grade 1	Grade 2	Grade 3			
15	Quantity	0	0	0			
16							
17							
18	Model Outputs (i.e Constraints)						
19		Amount					
20	Calcium	=SUMPRODUCT(B7:D7,B15:D15)	>=	10			
21	Iron	=SUMPRODUCT(B8:D8,B15:D15)	>=	12			
22	Protein	=SUMPRODUCT(B9:D9,B15:D15)	>=	15			
23							
24	Fat	=SUMPRODUCT(B10:D10,B15:D15)	<=	7.5			
25							
26	Total Cost	=SUMPRODUCT(B6:D6,B15:D15)	Objective Function				

Sheet1 Sheet2 Sheet3

Figure 6.11: Worksheet

The `SUMPRODUCT` function multiplies corresponding components in the arrays, and returns the sum of those products, which is a dot product. For example, the amount of calcium used is SUMPRODUCT (B7:D7, B15:D15) = (B7*B15 + C7*C15 + D7*D15).

2. Click Solver from the Data tab.

3. Type B26 in the Set Target Cell box.

4. Select Min from the Equal To choice group (for minimization).

5. Click B15:D15 in the By Changing Cells box.

The **Solver Parameters** dialog box, with the specified settings, is displayed in figure 6.12.

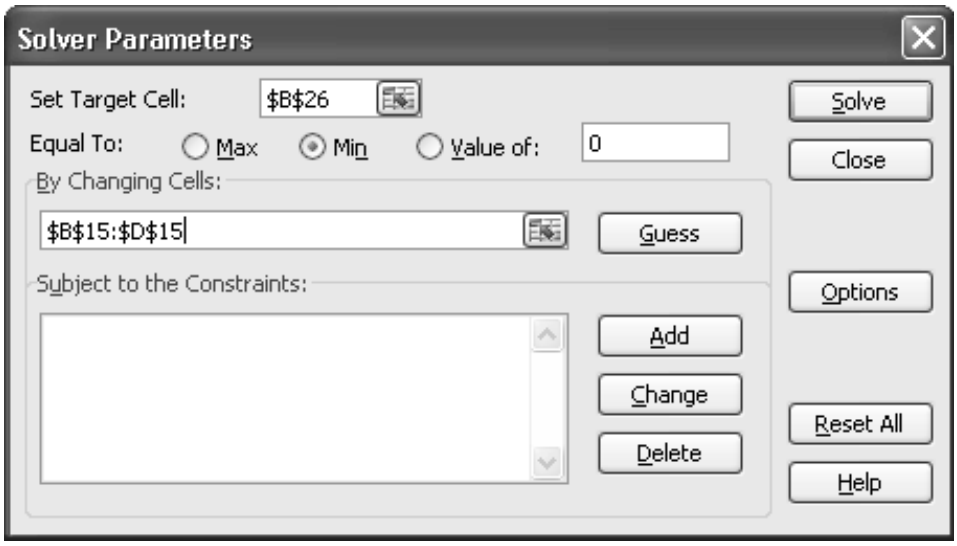


Figure 6.12: Solver Parameters Dialog Box

6. **To add constraints, click Add.**

The **Add Constraints** dialog box is displayed. In this dialog box, three values need to be entered. They are entered in **Cell Reference** box, which contains the reference of the cell to which the constraint will be added, the symbol box, which denotes whether the Cell Reference value should be less than, or greater than, or equal to the constraint value, and the Constraint box, which contains the cell address where the constraint is present.

7. **To enter the minimum requirement constraints, in the Cell Reference box click, and select the cells from B20 to B22, select >= type and then click on the Constraint box and type F7:F9.**
8. **Click Add to keep adding constraints.**
9. **To enter the maximum allowed constraints, in the Cell Reference box, click and select cell B24, select <= type and then click on the Constraint box and type G10.**

The **Add Constraint** dialog box, after entering the last constraint, is shown in figure 6.13.

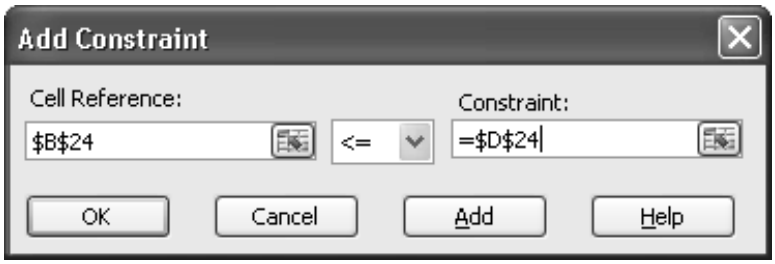


Figure 6.13: Add Constraint Dialog Box

10. **Click OK.**
11. **Click Options to display the Solver Options dialog box.**
12. **Select Assume Linear Model and Assume Non-Negative from the Solver Options dialog box, since this is an LP, as shown in figure 6.14.**

V 1.0 © 2008 Aptech Limited

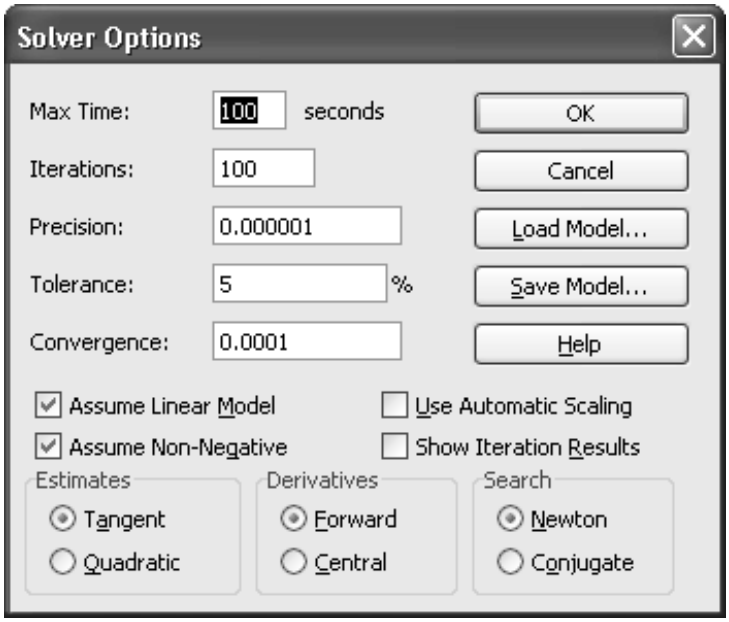


Figure 6.14: Solver Options Dialog Box

13. Click OK to return to the Solver Parameters dialog box.

The **Solver Parameters** dialog box, after setting the parameters and constraints required for the LP problem, appears, as shown in figure 6.15.

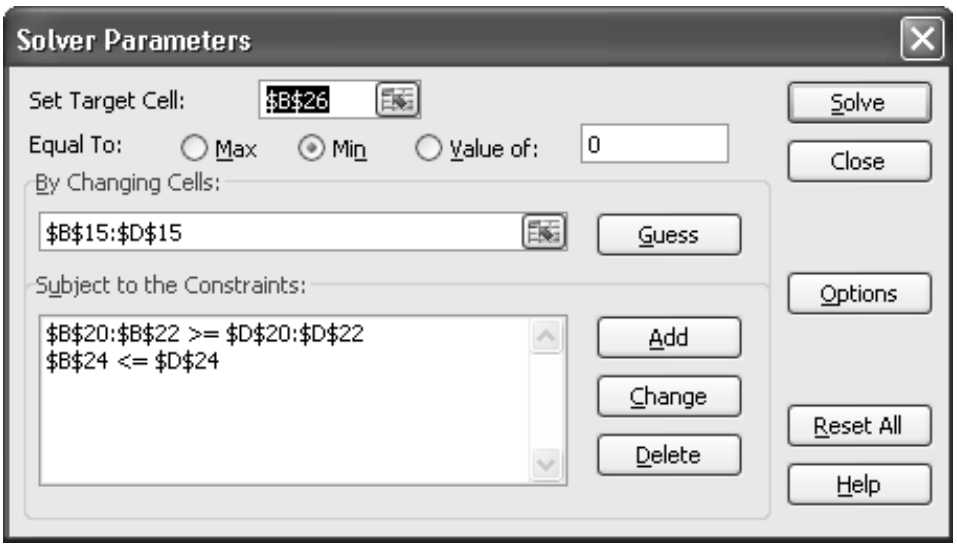


Figure 6.15: Solver Parameters Dialog Box

14. Click Solve.

The **Solver Results** dialog box is displayed, as shown in figure 6.16. You have the option of keeping the solution found by Solver in the worksheet or restoring the original values. In the **Reports** pane of the **Solver Results** dialog box, three types of reports, which can be automatically generated, are displayed.

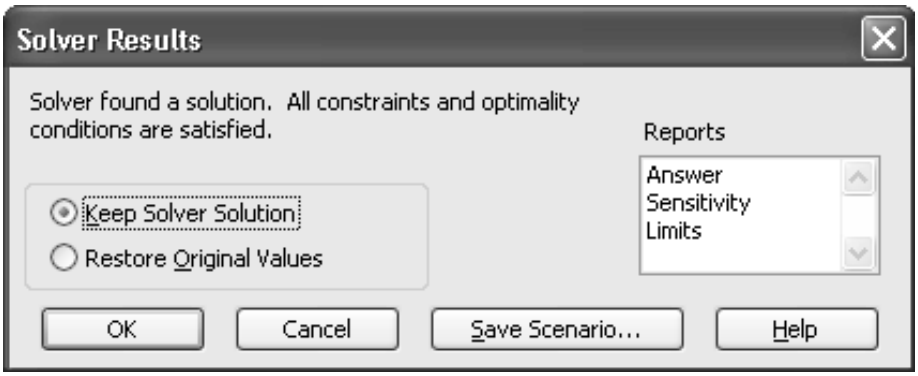


Figure 6.16: Solver Results Dialog Box

15. Select Keep Solver Solution.

16. Select all three options in the Reports section to generate reports, as shown in figure 6.17.

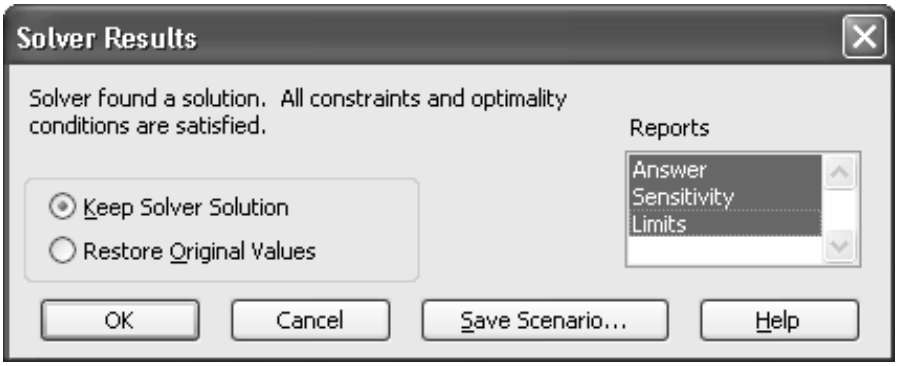


Figure 6.17: Solver Parameters Dialog Box

The reports are automatically placed in separate Excel Sheets titled “Answer Report 1”, “Sensitivity Report 1”, and “Limits Report 1”.

The Answer Report gives the original and final value of the Target Cell and all Adjustable Cells as well as a listing of each constraint and its status. The Answer Report for the Colorado Cattle Company model is shown in figure 6.18.

V 1.0 © 2008 Aptech Limited

CCC [Compatibility Mode]

Microsoft Excel 12.0 Answer Report
Worksheet: [CCC.xls]Sheet1
Report Created: 5/8/2008 10:35:06 AM

Target Cell (Min)

Cell	Name	Original Value	Final Value
$B$26	Total Cost Amount	0	2.59

Adjustable Cells

Cell	Name	Original Value	Final Value
$B$15	Quantity Grade 1	0	8
$C$15	Quantity Grade 2	0	5.5
$D$15	Quantity Grade 3	0	0.5

Constraints

Cell	Name	Cell Value	Formula	Status	Slack
$B$20	Calcium Amount	10	$B$20>=$D$20	Binding	0
$B$21	Iron Amount	12	$B$21>=$D$21	Binding	0
$B$22	Protein Amount	15.1	$B$22>=$D$22	Not Binding	0.1
$B$24	Fat Amount	7.5	$B$24<=$D$24	Binding	0

Answer Report 1 | Sensitivity Report 1

Figure 6.18: Answer Report

Excel can generate a Sensitivity Report in two parts - a changing cells report and a constraints report. It displays the following: the optimal value of each adjustable cell, its reduced cost, the objective function co-efficient and the increase and decrease of the objective function coefficient for which the current solution will remain optimal. In the Colorado Cattle Company Sensitivity Report shown in figure 6.19, the cost of grade 1 may decrease to 0.1425 and the current solution will remain optimal. A slack value is the absolute difference in values between the left side and the right side of the constraint. Constraints with a 'Slack' value of zero are said to be tight or binding. Constraints which are not tight are called loose or not binding. The shadow price for each constraint (the change in the objective function per unit of increase in the right-hand side) is provided, along with the increase and decrease of the right-hand side value for which this price is valid. If the units of fat amount are increased by 0.5 (which is within the allowable increase on that constraint), it will result in a change in the objective function of (0.5) x (-1.14) = -0.57.

CCC [Compatibility Mode]

Microsoft Excel 12.0 Sensitivity Report
Worksheet: [CCC.xls]Sheet1
Report Created: 5/8/2008 10:35:06 AM

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$B$15	Quantity Grade 1	8	0	0.25	1E+30	0.1425
$C$15	Quantity Grade 2	5.5	0	0.1	0.162857143	1E+30
$D$15	Quantity Grade 3	0.5	0	0.08	0.177142857	2.68

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$B$20	Calcium Amount	10	0.31	10	0.137931034	4
$B$21	Iron Amount	12	0.67	12	0.054794521	2
$B$22	Protein Amount	15.1	0	15	0.1	1E+30
$B$24	Fat Amount	7.5	-1.14	7.5	1	0.016949153

Sensitivity Report 1 | Limits Report 1

Figure 6.19: Sensitivity Report

The Limits Report, as shown in figure 6.20, gives the lower and upper limits of each adjustable cell holding all other adjustable cells at their current value, after satisfying the constraints.

CCC [Compatibility Mode]

Microsoft Excel 12.0 Limits Report
Worksheet: [CCC.xls]Limits Report 1
Report Created: 5/8/2008 10:35:07 AM

Cell	Target Name	Value
$B$26	Total Cost Amount	2.59

Cell	Adjustable Name	Value	Lower Limit	Target Result	Upper Limit	Target Result
$B$15	Quantity Grade 1	8	8	2.59	8	2.59
$C$15	Quantity Grade 2	5.5	5.5	2.59	5.5	2.59
$D$15	Quantity Grade 3	0.5	0.5	2.59	0.5	2.59

Limits Report 1 | Sheet1 | Shee

Figure 6.20: Limits Report

V 1.0 © 2008 Aptech Limited

16. Click on Sheet1.

The optimal solution is displayed in "Sheet1" in cells B15 - D15, as shown in figure 6.21.

CCC [Compatibility Mode]

	A	B	C	D	E	F	G
1	Colorado Cattle Company						
2							
3							
4							
5	Grade	1	2	3			
6	Cost /lb	$0.25	$0.10	$0.08		Minimum	Maximum
7	Calcium /lb	0.7	0.8	0	Calcium	10	
8	Iron /lb	0.9	0.8	0.8	Iron	12	
9	Protein /lb	0.8	1.5	0.9	Protein	15	
10	Fat /lb	0.5	0.6	0.4	Fat		7.5
11							
12	Decision Variables						
13							
14		Grade 1	Grade 2	Grade 3			
15	Quantity	8	5.5	0.5			
16							
17							
18	Model Outputs (i.e Constraints)						
19		Amount					
20	Calcium	10	>=	10			
21	Iron	12	>=	12			
22	Protein	15.1	>=	15			
23							
24	Fat	7.5	<=	7.5			
25							
26	Total Cost	2.59	Objective Function				

Limits Report 1 | Sheet1 | Sheet2

Figure 6.21: Worksheet after Keeping Solver Solution

Part II - 60 Minutes

1. The Globus Company wants to build 11 new stores in Luxembourgh city. These stores will be categorized as convenience stores, standard stores, or expanded service stores based on their sizes. A convenience store requires $4.125 million to build and 30 employees to operate. A standard store requires $8.25 million to build and 15 employees to operate. An expanded-services store requires $12.375 million to build and 45 employees to operate. The company can spend $82.5 million in construction and 300 employees to staff the 11 stores. On an average, a convenience store nets a profit of $1.2 million annually, a standard store nets a profit of $2 million annually, and an expanded-services store nets a profit of $2.6 million annually. How many of each types of stores should they build to maximize the annual revenue?

   Hint:

   Assign the variables:

   x1 = number of convenience stores

   x2 = number of standard stores

x3 = number of expanded services stores

The constraints:

$x1 + x2 + x3 \leq 11$

$4.125 x1 + 8.25x2 + 12.375x3 \leq 82.5$

$30x1 + 15x2 + 45x3 \leq 300$

$x1 \geq 0$, $x2 \geq 0$, and $x3 \geq 0$

The objective function is: $1.2x1 + 2x2 + 2.6x3$ (in millions)

2. A company named Technet has a sales representative for its packaging business. The sales representative must achieve the sales target of $100,000. Table 6.2 displays the sale situation in which he has sold only 2000 units of the product at a price of $3.46 per unit. How many units must he sell to achieve the $100,000 target?

Current Sales Status		
Boxes Sold	**Unit Price**	**Total Sales to Date**
2000	$3.46	$6,920

Table 6.2: Current Sales Status

Hint: Use Goal Seek to achieve the target of $100,000.

V 1.0 © 2008 Aptech Limited

Do It Yourself

1. Jason is a jeweller in Larks Jewellery Company where he makes diamond bracelets, necklaces, and earrings. He wants to work 160 hours per month, at the most. He has 800 ounces of diamonds. The profit, labor time, and ounces of diamonds required to produce each product are given in table 6.3. If demand for each product is unlimited, how can Jason maximize his profit?

Product	Unit Profit	Labor Hours Per Unit	Ounces of Diamonds Per Unit
Bracelet	$300	0.35	1.2
Necklace	$200	0.15	0.75
Earrings	$100	0.05	0.5

Table 6.3: Product Details

2. An office manager needs to purchase new filing cabinets. At the local superstore, Office Min, Ace cabinets costs $40 each, requires 6 square feet of floor space, and hold 24 cubic feet of files. On the other hand, each Excello cabinet costs $80, require 8 square feet of floor space, and holds 36 cubic feet of files. The maximum amount that can be spent on cabinets is $560. The office has a maximum space of 72 square feet for the cabinets. The manager wants the greatest storage capacity within the limitations imposed by funds and space. How many cabinets of each type should be purchased?

3. A salesperson of Innovative Services earns $200 a week with a 15% commission on his sales. Use Goal Seek to determine how much he will have to sell to make $1,200 next week.

This page has been intentionally left blank

Session 7

Analysis ToolPak

Session Objectives:

At the end of this session, you will be able to -

- Describe the Analysis ToolPak
- List and describe the various Analysis ToolPak Tools
- List the steps involved in using the various Analysis ToolPak tools

7.1 Analysis ToolPak Tools

MS Excel provides some medium scale statistical analysis tools in the form of the **Analysis ToolPak** add-in. To display the analysis tools, on the **Data** tab, click **Data Analysis**. Figure 7.1 shows the **Data Analysis** dialog box.

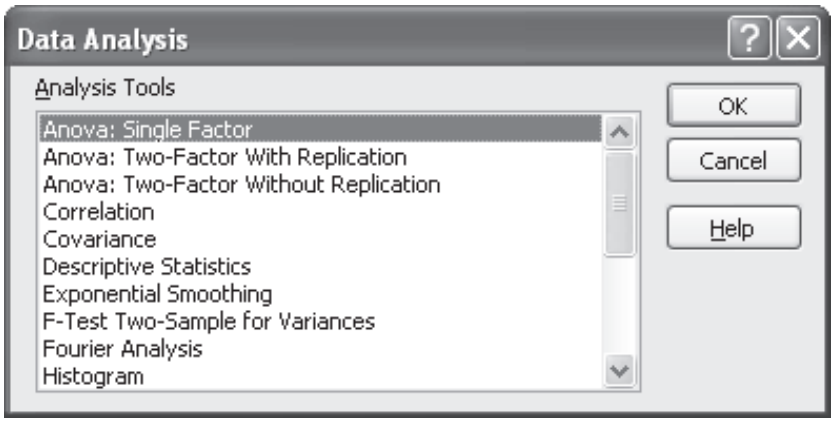


Figure 7.1: Data Analysis Dialog Box

The Analysis ToolPak consists of 19 tools. With the help of these medium scale statistical analysis tools, you can perform various statistical operations such as analysis of variances, exponential smoothing, fourier analysis and so on. These tools can save a lot of effort in the form of tedious mathematical calculations. Table 7.1 briefly introduces these tools.

Tool	Function
Anova: Single Factor Anova: Two-Factor With Replication Anova: Two-Factor Without Replication	Determines if there is a statistically significant difference between two or more given sets of data by analyzing their variances
Correlation	Computes correlation coefficients for two or more sets of data
Covariance	Computes pair-wise covariance coefficients for two or more sets of data
Descriptive Statistics	Generates a report of univariate statistics for a given set of data

Tool	Function
Exponential Smoothing	Computes corrected values for a given set of data after smoothing out errors present in the given set of data
F-Test Two-Sample for Variances	Tests whether variances of the two given sets of data differ or not
Fourier Analysis	Computes the Discrete Fourier Transform and its inverse for a given set of data
Histogram	Creates a histogram for a given set of data
Moving Average	Computes moving averages for a given set of data
Random Number Generation	Generates random numbers
Rank and Percentile	Computes rank and percentile for the values of a given set of data
Regression	Performs regression analysis on two sets of data
Sampling	Generates samples from a given set of data
t-Test: Paired Two Sample for Means t-Test: Two-Sample Assuming Equal variances t-Test: Two-Sample Assuming Unequal Variances	Performs Student's t-Test on two sets of data
z-Test: Two Sample for Means	Performs z-Test on two sets of data with known variances

Table 7.1: Analysis ToolPak Tools

7.1.1 Anova: Single Factor Tool

The **Anova: Single Factor** tool compares two or more data sets by analyzing their variances. It generates a single factor variance analysis report for the given sets of data.

Consider the sales data shown in figure 7.2.

	A	B	C	D
1	Year	Product 'A'	Product 'B'	Product 'C'
2	2003	10	12	7
3	2004	11	13	9
4	2005	12	15	8

Figure 7.2: Sales Data

To display the **Anova: Single Factor** dialog box, on the **Data** tab, click **Data Analysis**, select **Anova: Single Factor** from the **Analysis Tools** list, and then click **OK**. Figure 7.3 shows the **Anova: Single Factor** dialog box.

V 1.0 © 2008 Aptech Limited

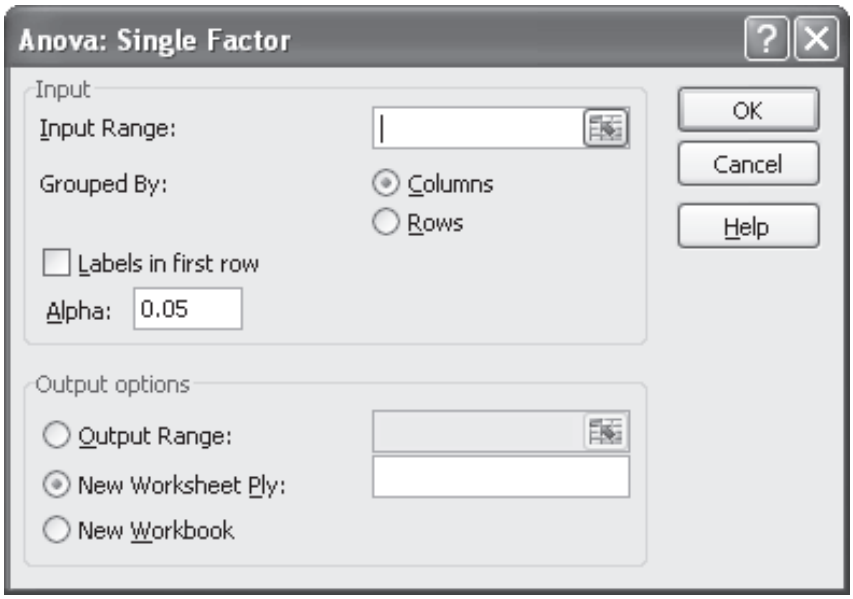


Figure 7.3: Anova: Single Factor Dialog Box

The **Input Range** box specifies the range of cells on which the analysis is to be done. The **Grouped By** option specifies whether the data sets are to be treated as columns or rows. The **Labels in first row** option specifies whether the data sets include labels in their first rows or columns. The **Output Range** box specifies the cell address from where the report is to be displayed.

To analyze the sales data shown in figure 7.2, fill the **Anova: Single Factor** dialog box, as shown in figure 7.4, and then click **OK**.

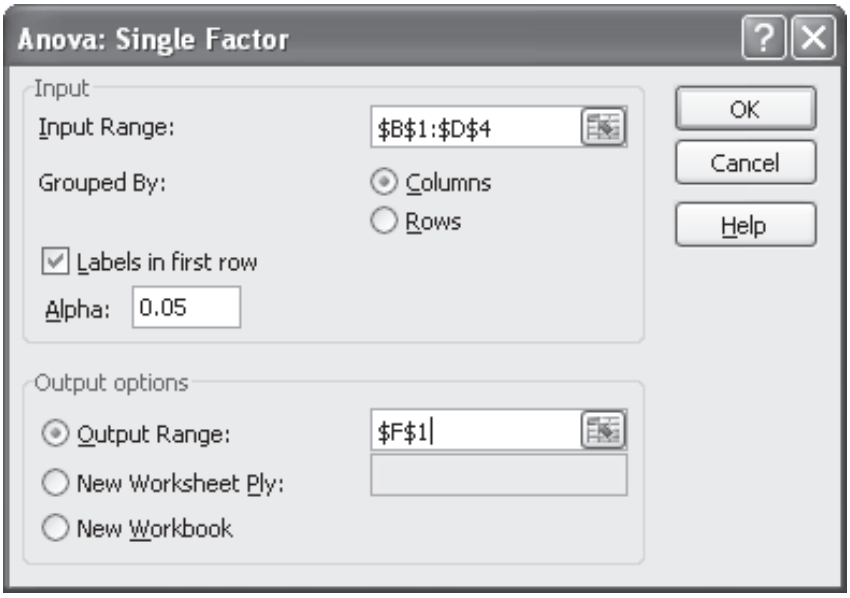


Figure 7.4: Filled Anova: Single Factor Dialog Box

Figure 7.5 shows the results of the Anova: Single Factor analysis.

	F	G	H	I	J	K	L
1	Anova: Single Factor						
2							
3	SUMMARY						
4	*Groups*	*Count*	*Sum*	*Average*	*Variance*		
5	Product 'A'	3	33	11	1		
6	Product 'B'	3	40	13.33333	2.333333333		
7	Product 'C'	3	24	8	1		
8							
9							
10	ANOVA						
11	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
12	Between Groups	42.88888889	2	21.44444	14.84615385	0.004750397	5.14325285
13	Within Groups	8.666666667	6	1.444444			
14							
15	Total	51.55555556	8				

Figure 7.5: Anova: Single Factor Results

The `F` value (cell J12) is greater than `F crit` (cell L12). This means that there is a statistically significant difference in the sales of the three products. The report concludes that `Product 'B'` is the best selling product and there is a "significant" statistical difference between it and the other 2 products.

7.1.2 Anova: Two-Factor Without Replication

The **Anova: Two-Factor Without Replication** tool generates a two factor variance analysis report for the given sets of data.

Consider the revised sales data figures, as shown in figure 7.6. The **Anova: Single Factor** tool analyzes the input data for one factor while the **Anova: Two-Factor Without Replication** tool analyzes the data for two factors.

	A	B	C	D
1	Year	Product 'A'	Product 'B'	Product 'C'
2	2003	3.74	4.58	4.58
3	2004	4.47	6.78	5.19
4	2005	5.65	7	6.08

Figure 7.6: Revised Sales Data

To display the **Anova: Two-Factor Without Replication** dialog box, on the **Data** tab, click **Data Analysis**, select **Anova: Two-Factor Without Replication** from the **Analysis Tools** list, and then click **OK**. Figure 7.7 shows the **Anova: Two-Factor Without Replication** dialog box.

V 1.0 © 2008 Aptech Limited

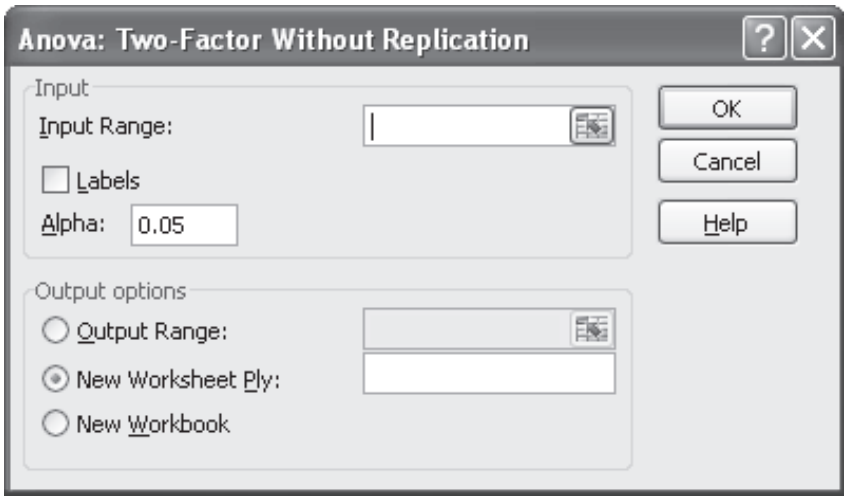


Figure 7.7: Anova: Two-Factor Without Replication Dialog Box

The **Input Range** box specifies the range of cells on which analysis is to be done. The **Labels** option specifies whether the data sets include labels in their first rows or columns. The **Output Range** box specifies the cell address from where the report is to be displayed.

To analyze the revised sales data shown in figure 7.6, fill the **Anova: Two-Factor Without Replication** dialog box, as shown in figure 7.8, and then click **OK**.

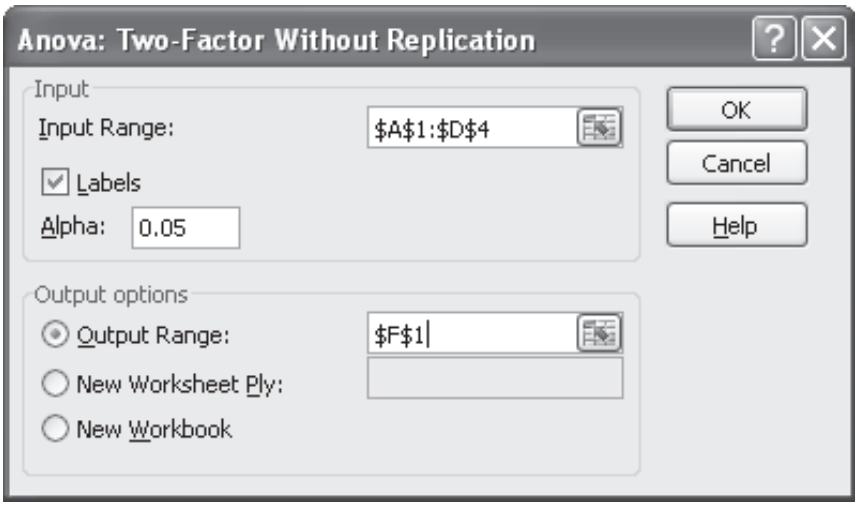


Figure 7.8: Filled Anova: Two-Factor Without Replication Dialog Box

Figure 7.9 shows the results of the **Anova: Two-Factor Without Replication** analysis.

	F	G	H	I	J	K	L
1	Anova: Two-Factor Without Replication						
2							
3	*SUMMARY*	*Count*	*Sum*	*Average*	*Variance*		
4	2003	3	12.9	4.3	0.2352		
5	2004	3	16.44	5.48	1.3971		
6	2005	3	18.73	6.243333333	0.475633333		
7							
8	Product 'A'	3	13.86	4.62	0.9289		
9	Product 'B'	3	18.36	6.12	1.7908		
10	Product 'C'	3	15.85	5.283333333	0.569033333		
11							
12							
13	ANOVA						
14	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
15	Rows	5.751622222	2	2.875811111	13.92906924	0.015764463	6.94427191
16	Columns	3.390022222	2	1.695011111	8.20983236	0.038372736	6.94427191
17	Error	0.825844444	4	0.206461111			
18							
19	Total	9.967488889	8				

Figure 7.9: Anova: Two-Factor Without Replication Results

From figure 7.9, it can be concluded that sales for the year 2004 were better than the other two years and there is a "significant" statistical difference between the sales in 2004 and the other two years. Similarly, `Product 'B'` is the best selling product and there is a "significant" statistical difference between it and the other two products.

7.1.3 Anova: Two-Factor With Replication

The **Anova: Two-Factor With Replication** tool works exactly like the **Anova: Two-Factor Without Replication**. The **Anova: Two-Factor Without Replication** tool deals with data sets that have one sample for each combination, while the **Anova: Two-Factor With Replication** tool deals with data sets that have more than one sample for each combination. Consider the revised sales data figures, as shown in figure 7.10.

	A	B	C	D
1	Year	Product 'A'	Product 'B'	Product 'C'
2	2003	3	4	0
3		4	7	7
4		5	2	4
5		5	1	4
6	2004	0	3	3
7		7	4	8
8		4	5	1
9		4	5	2
10	2005	4	7	0
11		7	2	7
12		2	6	4
13		1	1	4

Figure 7.10: Revised Sales Data

V 1.0 © 2008 Aptech Limited

To display the **Anova: Two-Factor With Replication** dialog box, on the **Data** tab, click **Data Analysis**, select **Anova: Two-Factor With Replication** from the **Analysis Tools** list, and then click **OK**. Figure 7.11 shows the **Anova: Two-Factor With Replication** dialog box.

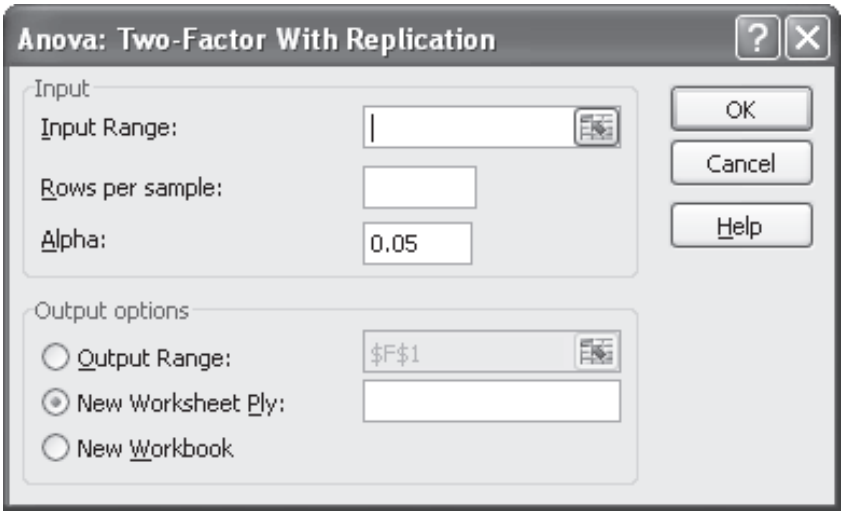


Figure 7.11: Anova: Two-Factor With Replication Dialog Box

The **Input Range** box specifies the range of cells on which the analysis is to be done. The **Rows per sample** box specifies the number of samples in each combination. The **New Worksheet Ply** box creates a new worksheet by the name specified in it. The analysis report is displayed in this new worksheet.

To analyze the revised sales data shown in figure 7.10, fill the **Anova: Two-Factor With Replication** dialog box, as shown in figure 7.12, and then click **OK**.

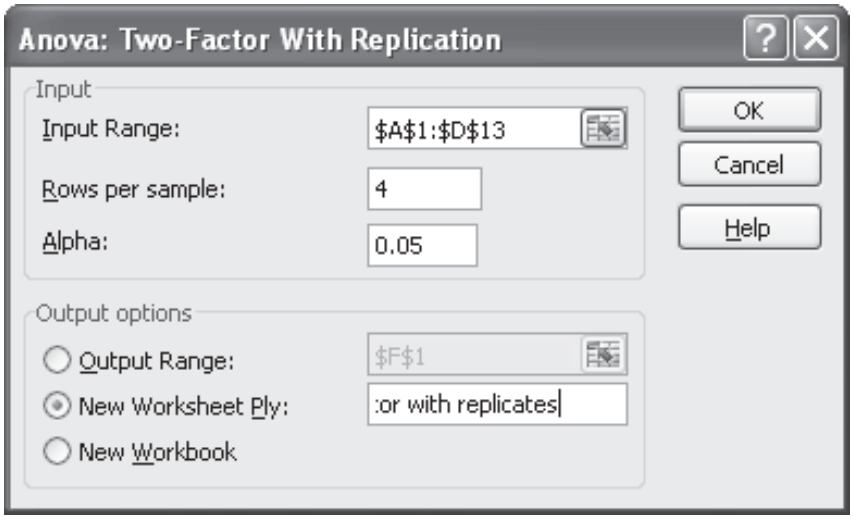


Figure 7.12: Filled Anova: Two-Factor With Replication Dialog Box

Figure 7.13 shows the results of the **Anova: Two-Factor With Replication** analysis.

	A	B	C	D	E	F	G
1	Anova: Two-Factor With Replication						
3	SUMMARY	Product 'A'	Product 'B'	Product 'C'	Total		
4	*2003*						
5	Count	4	4	4	12		
6	Sum	17	14	15	46		
7	Average	4.25	3.5	3.75	3.833333333		
8	Variance	0.916666667	7	8.25	4.515151515		
10	*2004*						
11	Count	4	4	4	12		
12	Sum	15	17	14	46		
13	Average	3.75	4.25	3.5	3.833333333		
14	Variance	8.25	0.916666667	9.666666667	5.242424242		
16	*2005*						
17	Count	4	4	4	12		
18	Sum	14	16	15	45		
19	Average	3.5	4	3.75	3.75		
20	Variance	7	8.666666667	8.25	6.568181818		
22	*Total*						
23	Count	12	12	12			
24	Sum	46	47	44			
25	Average	3.833333333	3.916666667	3.666666667			
26	Variance	4.515151515	4.628787879	7.151515152			
27							
28	ANOVA						
29	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
30	Sample	0.055555556	2	0.027777778	0.004243281	0.995766	3.354131
31	Columns	0.388888889	2	0.194444444	0.02970297	0.970766	3.354131
32	Interaction	2.444444444	4	0.611111111	0.093352192	0.98374	2.727765
33	Within	176.75	27	6.546296296			
34							
35	Total	179.6388889	35				

Figure 7.13: Anova: Two-Factor With Replication Results

The initial part of the Anova: Two-Factor With Replication analysis comprises the count, sum, average and variance for the three products for each individual year. The latter part consists of the analysis to determine if the sales data for the three products are statistically significantly different from each other or not. From figure 7.13, it can be concluded that the sales data are not statistically different (values of `F` are lesser than `F crit`).

7.1.4 Correlation

The **Correlation** tool computes the correlation coefficients for two or more sets of data. Consider the shopping mall data shown in figure 7.14.

V 1.0 © 2008 Aptech Limited

	A	B	C
1	Shopper's Point		
2	Month	Gift Vouchers	Customers
3	1	51	320
4	2	42	362
5	3	70	485
6	4	65	436
7	5	49	386
8	6	52	414

Figure 7.14: Shopping Mall Data

To display the **Correlation** dialog box, on the **Data** tab, click **Data Analysis**, select **Correlation** from the **Analysis Tools** list, and then click **OK**. Figure 7.15 shows the **Correlation** dialog box.

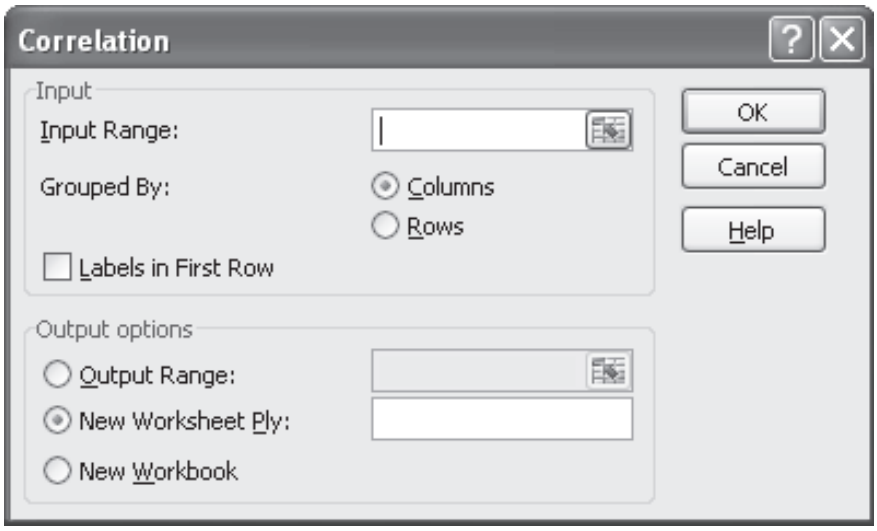


Figure 7.15: Correlation Dialog Box

The **Input Range** box specifies the range of cells on which the analysis is to be done. The **Grouped By** option specifies whether the data sets are to be treated as columns or rows. The **Labels in First Row** option specifies whether the data sets include labels in their first rows or columns. The **Output Range** box specifies the cell address from where the report is to be displayed.

To analyze the shopping mall data shown in figure 7.14, fill the **Correlation** dialog box, as shown in figure 7.16, and then click **OK**.

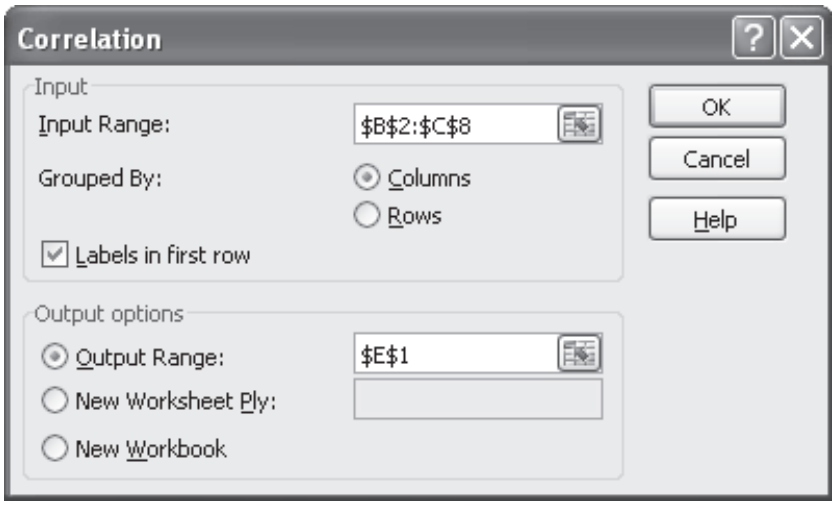


Figure 7.16: Filled Correlation Dialog Box

Figure 7.17 shows the results of the **Correlation** analysis.

	A	B	C	D	E	F	G
1		Shopper's Point				*Gift Vouchers*	*Customers*
2	Month	Gift Vouchers	Customers		Gift Vouchers	1	
3	1	51	320		Customers	0.817148378	1
4	2	42	362				
5	3	70	485				
6	4	65	436				
7	5	49	386				
8	6	52	414				

Figure 7.17: Correlation Results

Cells F2:G3 display the correlation coefficients for the `Gift Vouchers` and `Customers` data sets.

7.1.5 Covariance

The **Covariance** tool computes pair-wise covariance coefficients for two or more sets of data. Covariance can be positive or negative. If the covariance is positive, the values in data sets are directly proportional, that is, if values in one increase then values in the second also increase. If the covariance is negative, then the values in the data sets are inversely proportional, that is, if the values in the first decrease then the values in the second increase. If the covariance is zero, the data sets are not correlated.

Consider the shopping mall data shown in figure 7.14. To display the **Correlation** dialog box, on the **Data** tab, click **Data Analysis**, select **Correlation** from the **Analysis Tools** list, and then click **OK**. Figure 7.18 shows the **Covariance** dialog box.

V 1.0 © 2008 Aptech Limited

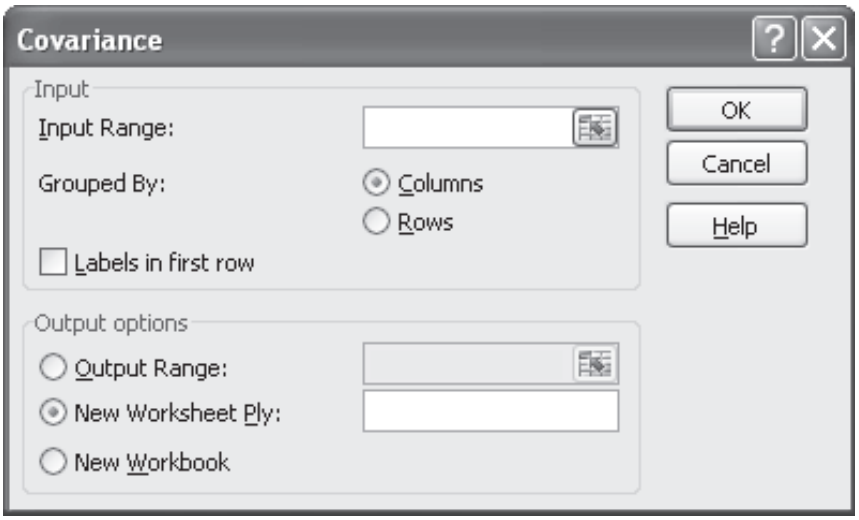


Figure 7.18: Covariance Dialog Box

The **Input Range** box specifies the range of cells on which the analysis is to be done. The **Grouped By** option specifies whether the data sets are to be treated as columns or rows. The **Labels in first row** option specifies whether the data sets include labels in their first rows or columns. The **Output Range** box specifies the cell address from where the report is to be displayed.

To analyze the shopping mall data shown in figure 7.14, fill the **Covariance** dialog box as shown in figure 7.19 and then click **OK**.

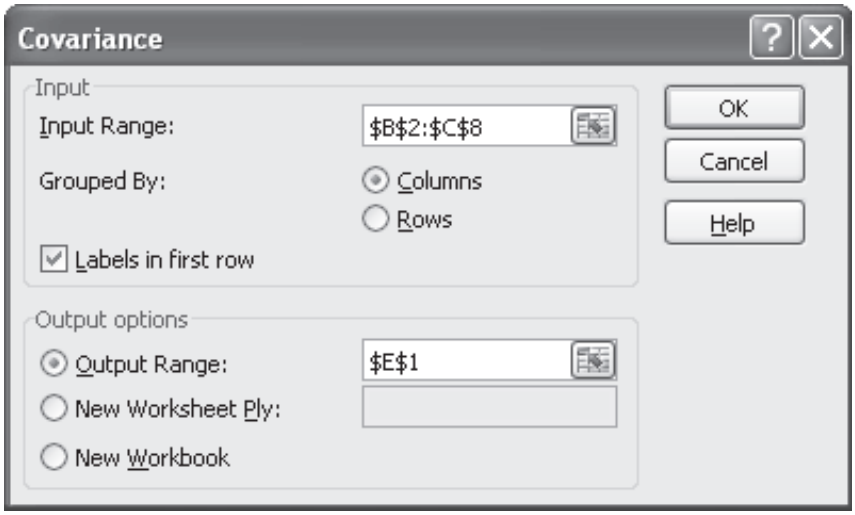


Figure 7.19: Filled Covariance Dialog Box

Figure 7.20 shows the results of the **Covariance** analysis.

	A	B	C	D	E	F	G
1	Shopper's Point					*Gift Vouchers*	*Customers*
2	Month	Gift Vouchers	Customers		Gift Vouchers	92.47222222	
3	1	51	320		Customers	415.25	2792.583333
4	2	42	362				
5	3	70	485				
6	4	65	436				
7	5	49	386				
8	6	52	414				

Figure 7.20: Covariance Results

Cell F3 displays the covariance. The covariance is positive (415.25), so as values of the `Gift Vouchers` data set increase, values of the `Customers` data set also increase.

7.1.6 Descriptive Statistics

The **Descriptive Statistics** tool generates a report of univariate statistics for a given set of data.

Consider the shopping mall data shown in figure 7.14. To display the **Descriptive Statistics** dialog box, on the **Data** tab, click **Data Analysis**, select **Descriptive Statistics** from the **Analysis Tools** list, and then click **OK**. Figure 7.21 shows the **Descriptive Statistics** dialog box.

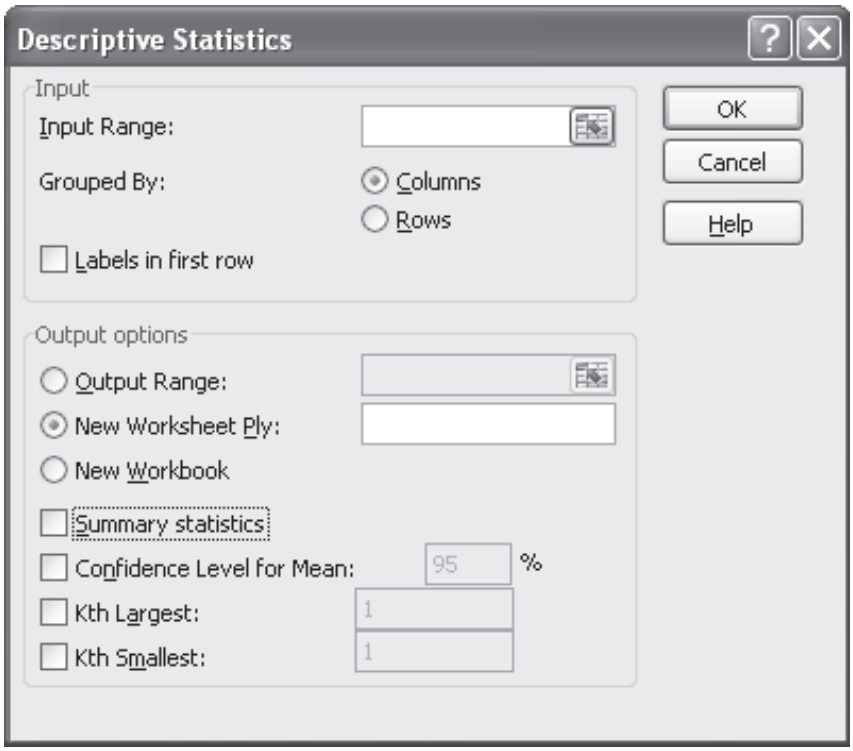


Figure 7.21: Descriptive Statistics Dialog Box

The **Input Range** box specifies the range of cells on which the analysis is to be done. The **Grouped By** option specifies whether the data sets are to be treated as columns or rows. The **Labels in first row** option specifies whether the data sets include labels in their first rows or columns. The **Output Range** box specifies the cell address from where the report is to be displayed. The **Summary statistics** option generates a complete summary of the univariate statistics.

V 1.0 © 2008 Aptech Limited

To analyze the shopping mall data shown in figure 7.14, fill the **Descriptive Statistics** dialog box, as shown in figure 7.22, and then click **OK**.

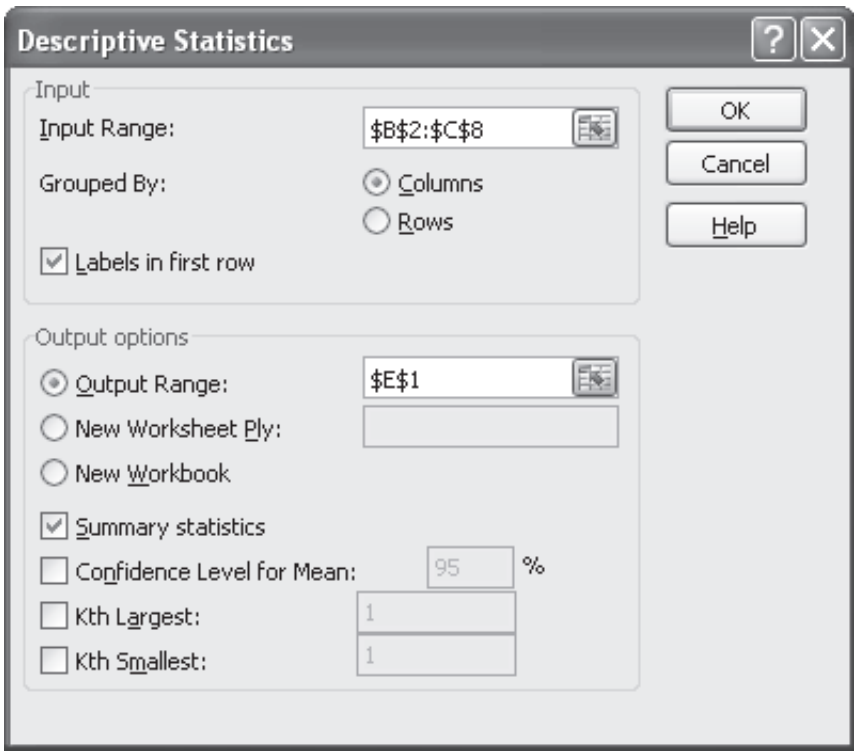


Figure 7.22: Filled Descriptive Statistics Dialog Box

Figure 7.23 shows the results of the **Descriptive Statistics** analysis.

	A	B	C	D	E	F	G	H
1	Shopper's Point				*Gift Vouchers*		*Customers*	
2	Month	Gift Vouchers	Customers					
3	1	51	320		Mean	54.83333333	Mean	400.5
4	2	42	362		Standard Error	4.300516765	Standard Error	23.63295721
5	3	70	485		Median	51.5	Median	400
6	4	65	436		Mode	#N/A	Mode	#N/A
7	5	49	386		Standard Deviation	10.5340717	Standard Deviation	57.88868629
8	6	52	414		Sample Variance	110.9666667	Sample Variance	3351.1
9					Kurtosis	-1.010982346	Kurtosis	-0.213629102
10					Skewness	0.551387063	Skewness	0.106358149
11					Range	28	Range	165
12					Minimum	42	Minimum	320
13					Maximum	70	Maximum	485
14					Sum	329	Sum	2403
15					Count	6	Count	6

Figure 7.23: Descriptive Statistics Results

Cells E1:H15 show the various univariate statistics generated for the `Gift Vouchers` and `Customers` data sets.

7.1.7 Exponential Smoothing

The **Exponential Smoothing** tool computes corrected values for a given set of data after smoothing out errors present in the given set of data. The **Exponential Smoothing** tool generates the corrected values using a **Damping factor**. The Damping factor is an assumption. It represents the percentage to which the existing value needs to be adjusted. This percentage represents the error that might have occurred in the existing value's previous forecast. 0.2 to 0.3 are reasonable damping factor values.

Consider the shopping mall data shown in figure 7.14. To display the **Exponential Smoothing** dialog box, on the **Data** tab, click **Data Analysis**, select **Exponential Smoothing** from the **Analysis Tools** list, and then click **OK**. Figure 7.24 shows the **Exponential Smoothing** dialog box.

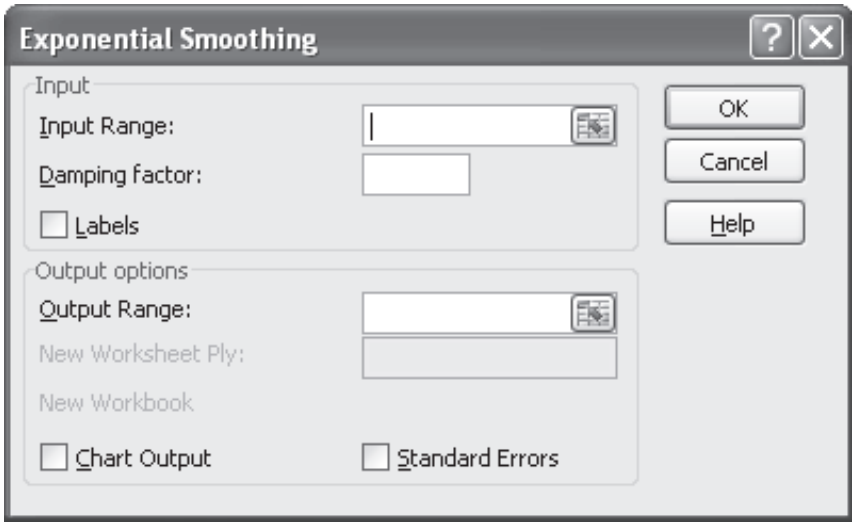


Figure 7.24: Exponential Smoothing Dialog Box

The **Input Range** box specifies the range of cells on which the analysis is to be done. Only a single row or column of data can be analyzed at any given point of time. The **Damping factor** box specifies the damping factor to be considered while predicting the new values. The **Labels** option specifies whether the data set includes labels. The **Output Range** box specifies the cell address from where the output is to be displayed.

To analyze the shopping mall data shown in figure 7.14, fill the **Exponential Smoothing** dialog box, as shown in figure 7.25, and then click **OK**.

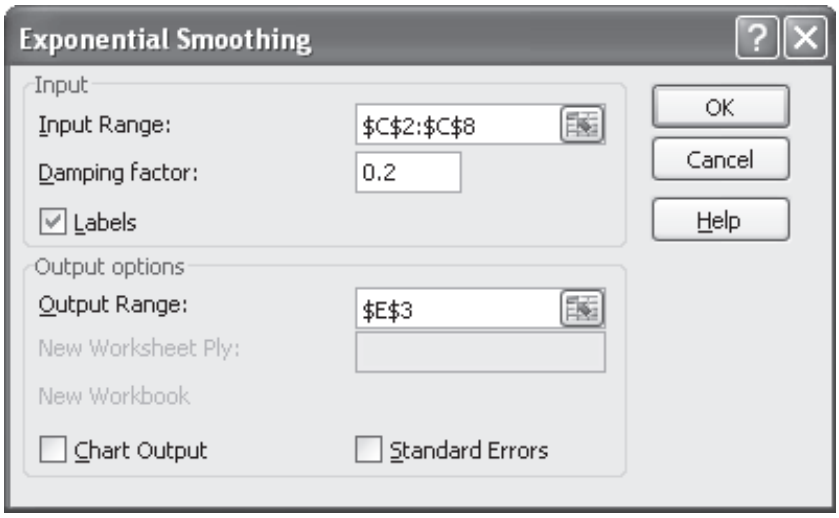


Figure 7.25: Filled Exponential Smoothing Dialog Box

V 1.0 © 2008 Aptech Limited

Figure 7.26 shows the results of the **Exponential Smoothing** analysis.

	A	B	C	D	E
1	Shopper's Point				
2	Month	Gift Vouchers	Customers		
3	1	51	320		#N/A
4	2	42	362		320
5	3	70	485		353.6
6	4	65	436		458.72
7	5	49	386		440.544
8	6	52	414		396.9088

Figure 7.26: Exponential Smoothing Results

The predicted or corrected values for the `Customers` data set are displayed in cells E3:E8.

7.1.8 F-Test Two-Sample for Variances

The **F-Test Two-Sample** tool compares variances between two data sets. By comparing variances, the F-Test compares the spread of the two data sets. It tests whether variances are different or not.

Consider the shopping mall data shown in figure 7.14. To display the **F-Test Two-Sample for Variances** dialog box, on the **Data** tab, click **Data Analysis**, select **F-Test Two-Sample for Variances** from the **Analysis Tools** list, and then click **OK**. Figure 7.27 shows the **F-Test Two-Sample for Variances** dialog box.

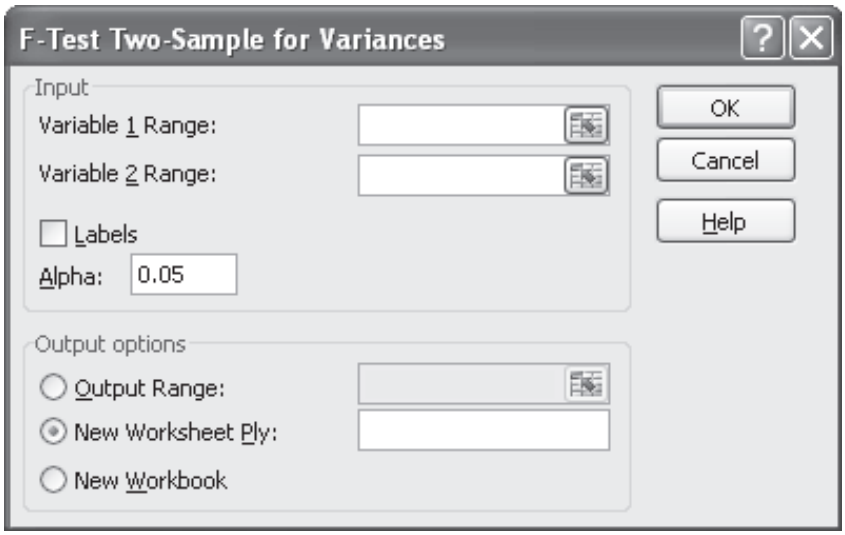


Figure 7.27: F-Test Two-Sample for Variances Dialog Box

The **Variable 1 Range** box specifies the range of cells to be considered as the first data set. The **Variable 2 Range** box specifies the range of cells to be considered as the second data set. The **Labels** option specifies whether the data sets include labels in their first rows or columns. The **Output Range** box specifies the cell address from where the report is to be displayed.

To analyze the shopping mall data shown in figure 7.14, fill the **F-Test Two-Sample for Variances** dialog box, as shown in figure 7.28, and then click **OK**.

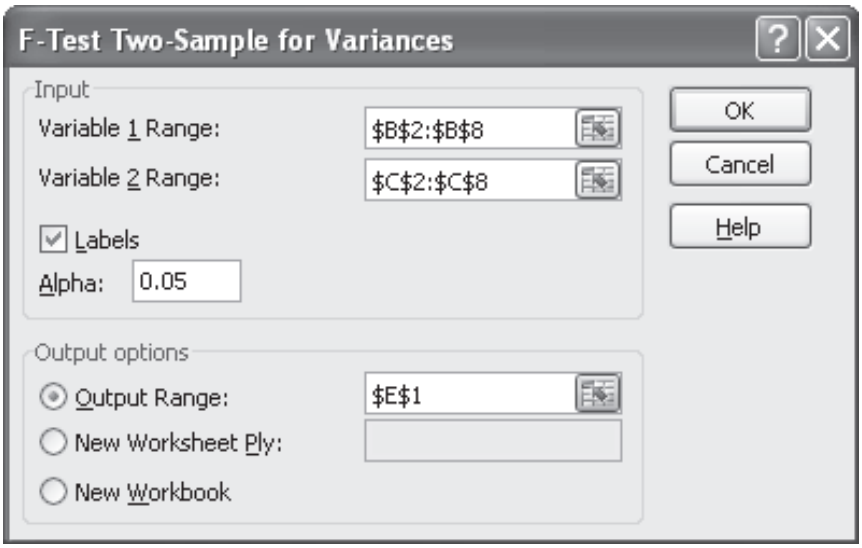


Figure 7.28: Filled F-Test Two-Sample for Variances Dialog Box

Figure 7.29 shows the results of the F-Test Two-Sample for Variances analysis.

	A	B	C	D	E	F	G
1	Shopper's Point				F-Test Two-Sample for Variances		
2	Month	Gift Vouchers	Customers				
3	1	51	320			*Gift Vouchers*	*Customers*
4	2	42	362		Mean	54.83333333	400.5
5	3	70	485		Variance	110.9666667	3351.1
6	4	65	436		Observations	6	6
7	5	49	386		df	5	5
8	6	52	414		F	0.033113505	
9					P(F<=f) one-tail	0.000965077	
10					F Critical one-tail	0.1980069	

Figure 7.29: F-Test Two-Sample for Variances Results

Since `F` (cell F8) is less than `F Critical` (cell F10), it can be concluded that the variances are not different and any change in them is due to random error.

7.1.9 Fourier Analysis

The **Fourier Analysis** tool calculates the Discrete Fourier Transform (DFT) for a range of discrete values by using the Fast Fourier Transform (FFT) algorithm. The **Fourier Analysis** tool also enables inverse transformations, that is, calculation of the original values from the transformed values. Fourier Analysis generally computes the frequency domain representation of an original function.

Consider a sinusoidal wave's time domain data shown in figure 7.30.

V 1.0 © 2008 Aptech Limited

	A
1	Time Domain Data
2	0
3	1
4	0
5	1
6	0
7	1
8	0
9	1

Figure 7.30: Frequency Data

To display the **Fourier Analysis** dialog box, on the **Data** tab, click **Data Analysis**, select **Fourier Analysis** from the **Analysis Tools** list, and then click **OK**. Figure 7.31 shows the **Fourier Analysis** dialog box.

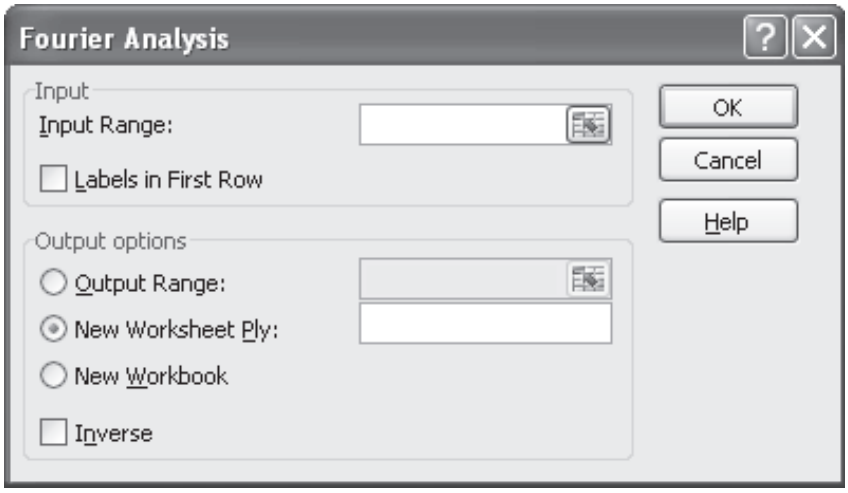


Figure 7.31: Fourier Analysis Dialog Box

The **Input Range** box specifies the range of cells on which the Fast Fourier Transform is to be done. The **Labels in First Row** option specifies whether the data sets include labels in their first rows or columns. The **Output Range** box specifies the cell address from where the report is to be displayed.

To analyze the sales data shown in figure 7.30, fill the **Fourier Analysis** dialog box, as shown in figure 7.32, and then click **OK**.

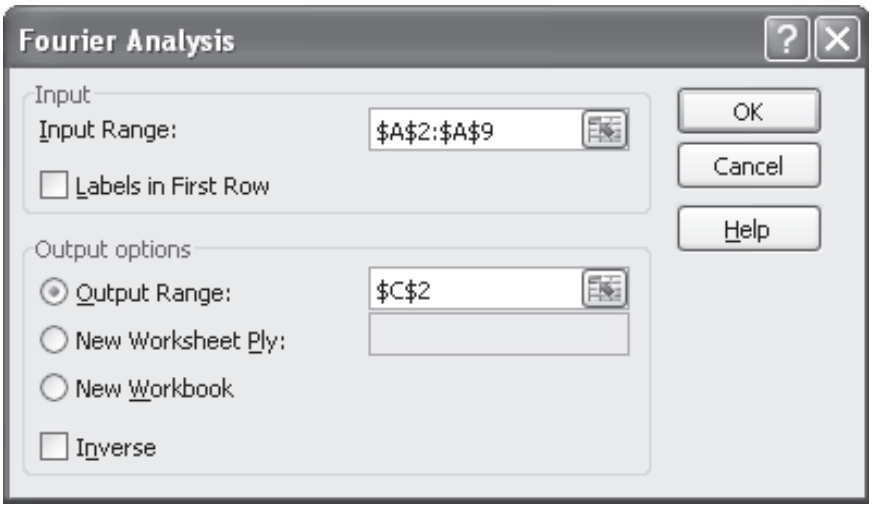


Figure 7.32: Filled Fourier Analysis Dialog Box

Figure 7.33 shows the results of the Fourier analysis.

	A	B	C
1	Time Domain Data		
2	0		4
3	1		0
4	0		0
5	1		0
6	0		-4
7	1		0
8	0		0
9	1		0

Figure 7.33: Fourier Analysis Results

The Fourier analysis generates the frequency domain representation for the given time domain values.

7.1.10 Histogram

The **Histogram** tool generates histograms for a given set of data. Histograms are normally used to summarize large sets of data. Histograms summarize data by first segregating data into the various ranges provided. A Histogram then calculates the total number of values in each range and represents each range along with the final count in it. Range values are also called bin values.

Consider the players' data shown in figure 7.34.

V 1.0 © 2008 Aptech Limited

	A	B	C	D
1	Match	Player 'A'	Player 'B'	Player 'C'
2	1	57	45	99
3	2	34	33	67
4	3	25	47	57
5	4	89	89	29
6	5	100	150	45
7	6	102	24	29
8	7	77	65	99
9	8	34	7	29
10	9	0	36	30
11	10	45	25	38
12				
13	Scores			
14	10			
15	50			
16	100			

Figure 7.34: Players' Scores Data

To display the **Histogram** dialog box, on the **Data** tab, click **Data Analysis**, select **Histogram** from the **Analysis Tools** list, and then click **OK**. Figure 7.35 shows the **Histogram** dialog box.

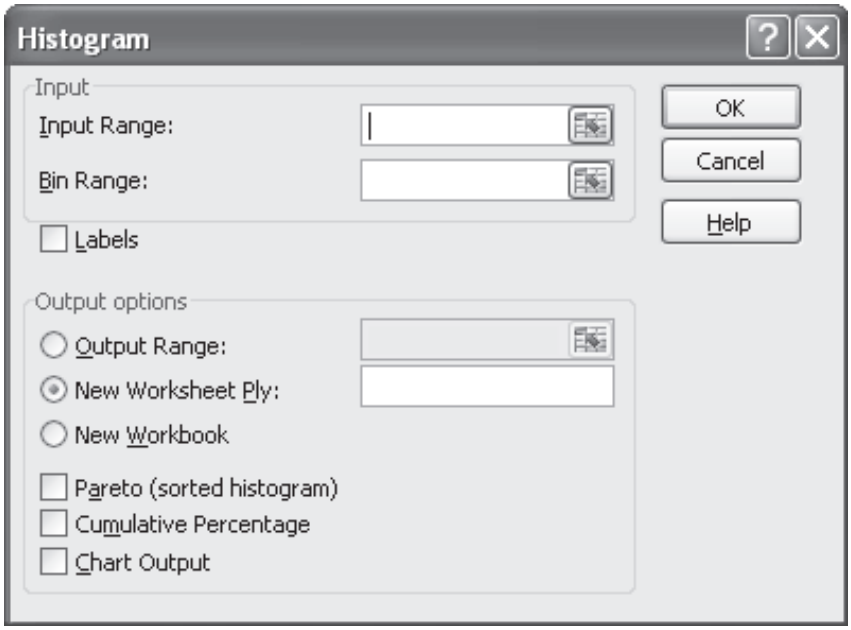


Figure 7.35: Histogram Dialog Box

The **Input Range** box specifies the range of cells on which the analysis is to be done. The **Bin Range** box specifies the cells whose values will be treated as periods while plotting the histogram. The **Output Range** box specifies the cell address from where the report is to be displayed. The **Chart Output** option creates a graph.

To analyze the players' data shown in figure 7.34, fill the **Histogram** dialog box, as shown in figure 7.36, and then click **OK**.

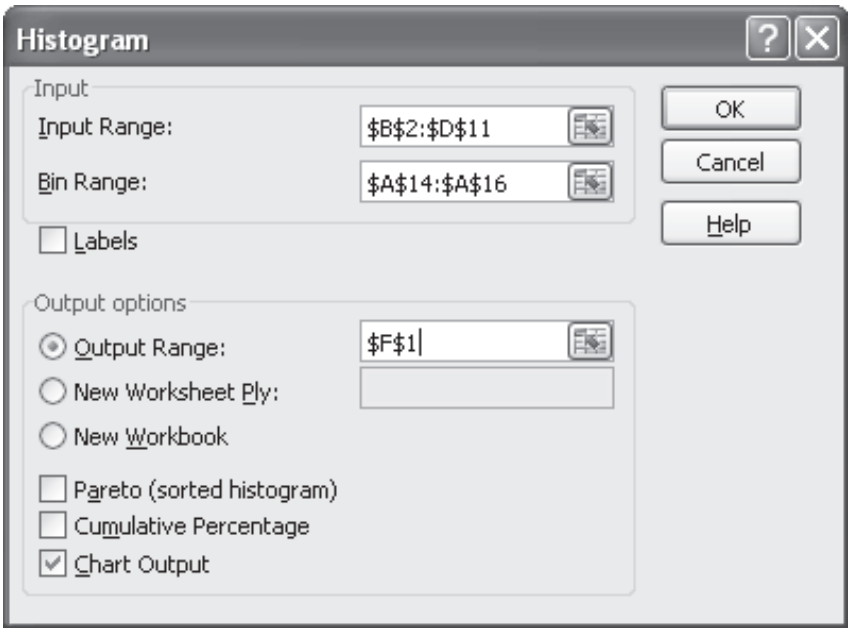
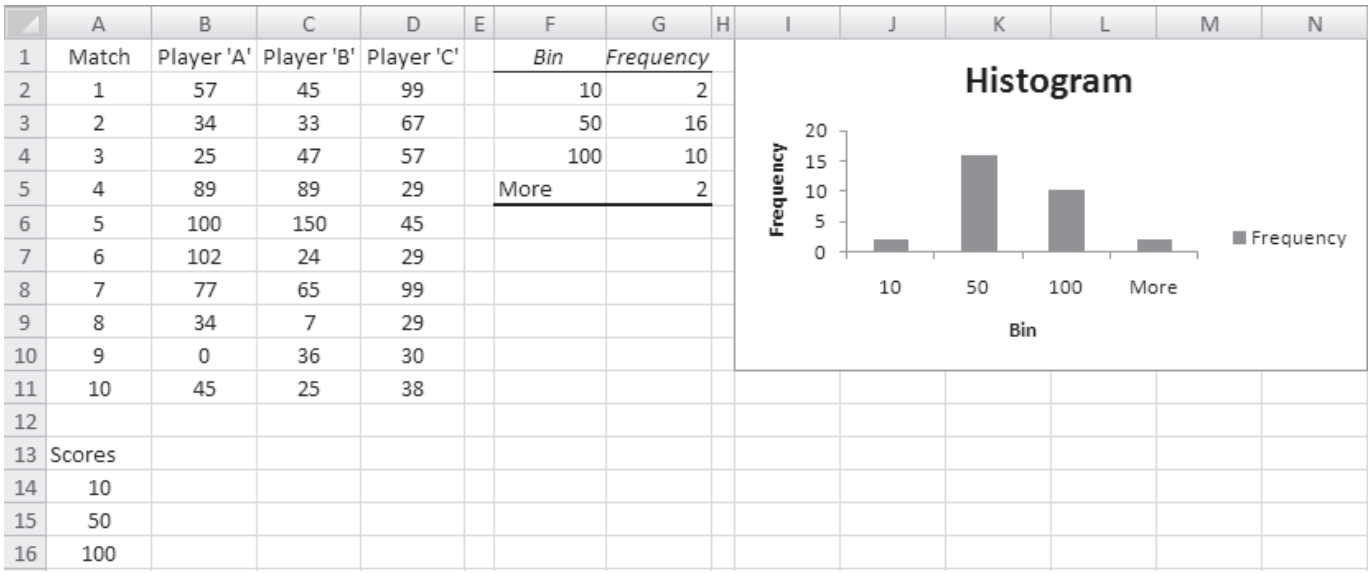


Figure 7.36: Filled Histogram Dialog Box

Figure 7.37 shows the results of the Histogram analysis.



	A	B	C	D	E	F	G
1	Match	Player 'A'	Player 'B'	Player 'C'		*Bin*	*Frequency*
2	1	57	45	99		10	2
3	2	34	33	67		50	16
4	3	25	47	57		100	10
5	4	89	89	29		More	2
6	5	100	150	45			
7	6	102	24	29			
8	7	77	65	99			
9	8	34	7	29			
10	9	0	36	30			
11	10	45	25	38			
12							
13	Scores						
14	10						
15	50						
16	100						

Figure 7.37: Histogram Results

The **Histogram** tool, first calculates the number of observations it finds for the ranges 0 - 10, 11 - 50, 51 - 100 and then above 100. This calculation can be seen in cells F1:G5. The entire data set, cells B2:D11, consists of only two values in the range 0 - 10, that is, `0` (cell B10) and `7` (cell C9). Similarly, the data set also consists of only two values above the 100 mark, `102` in cell B7 and `150` in cell C6. The histogram plots this calculation (cells B2:G5) in the form of a graph with the bin ranges on the X-axis and the number of observations found in a particular bin range on the Y-axis.

7.1.11 Moving Average

The **Moving Average** tool generates moving averages for a given data set depending on an interval. The **Moving Average** tool is handy in situations where data set values show a continuous trend but the varied values make it very hard to pinpoint the trend.

V 1.0 © 2008 Aptech Limited

Consider the five-year sales details for a particular brand of cold cream shown in figure 7.38. Sales figures in figure 7.38 clearly depict a rising trend for each year, that is, the sales increase over the first quarter to the last quarter, with a very sharp increase in the last quarter. However, the sharp difference in sales between the last and first quarters of two consecutive years makes it difficult to recognize or follow this trend.

	A	B	C
1	Year	Quarter	Sales (in million $)
2	1	1	573.9
3		2	577.9
4		3	637.9
5		4	972.4
6	2	1	701.0
7		2	668.0
8		3	639.0
9		4	1115.0
10	3	1	847.0
11		2	806.0
12		3	851.0
13		4	1429.0
14	4	1	1084.0
15		2	1100.0
16		3	1134.0
17		4	1946.0
18	5	1	1530.0
19		2	1387.0
20		3	1463.0
21		4	2541.0

Figure 7.38: Sales Details

Figure 7.39 shows the sales data being represented using a scatter chart.

Year	Quarter	Sales (in million $)
1	1	573.9
	2	577.9
	3	637.9
	4	972.4
2	1	701.0
	2	668.0
	3	639.0
	4	1115.0
3	1	847.0
	2	806.0
	3	851.0
	4	1429.0
4	1	1084.0
	2	1100.0
	3	1134.0
	4	1946.0
5	1	1530.0
	2	1387.0
	3	1463.0
	4	2541.0

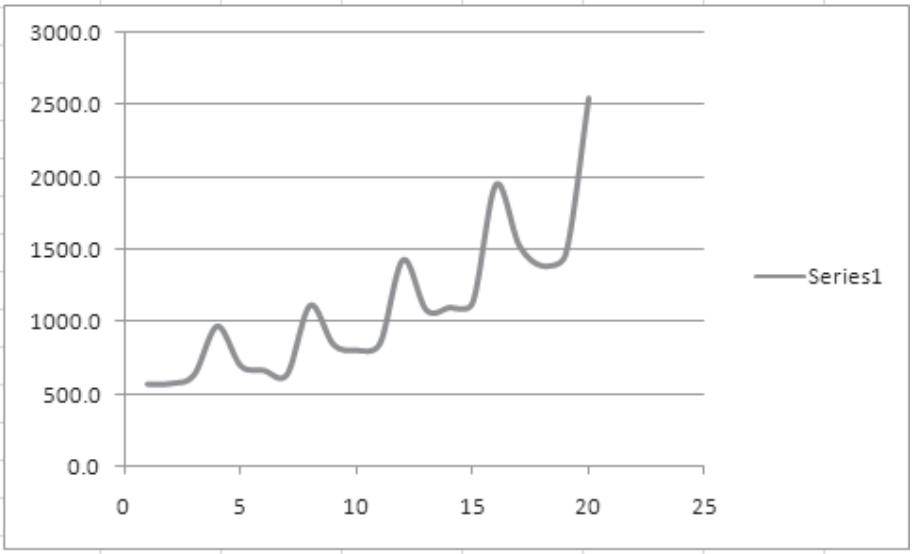


Figure 7.39: Sales Details Represented Using Scatter Chart

The scatter chart in the figure does represent the rising trend but not so effectively. The **Moving Average** tool clearly depicts the hidden rising trend.

To display the **Moving Average** dialog box, on the **Data** tab, click **Data Analysis**, select **Moving Average** from the **Analysis Tools** list, and then click **OK**. Figure 7.40 shows the **Moving Average** dialog box.

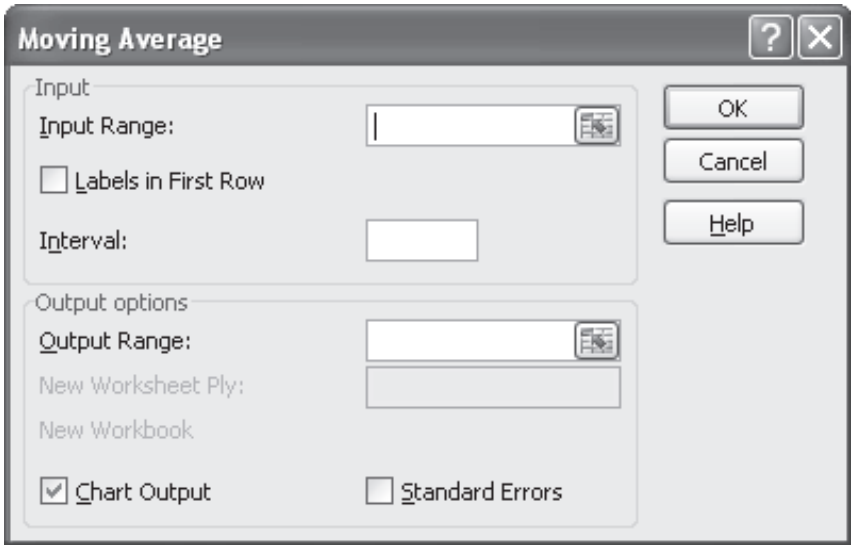


Figure 7.40: Moving Average Dialog Box

The **Input Range** box specifies the range of cells on which the analysis is to be done. The **Intervals** box specifies the interval at which the moving averages will be calculated. The **Output Range** box specifies the cell address from where the report is to be displayed. The **Chart Output** option creates a graph.

V 1.0 © 2008 Aptech Limited

To analyze the data shown in figure 7.38, fill the **Moving Average** dialog box as shown in figure 7.41 and then click **OK**.

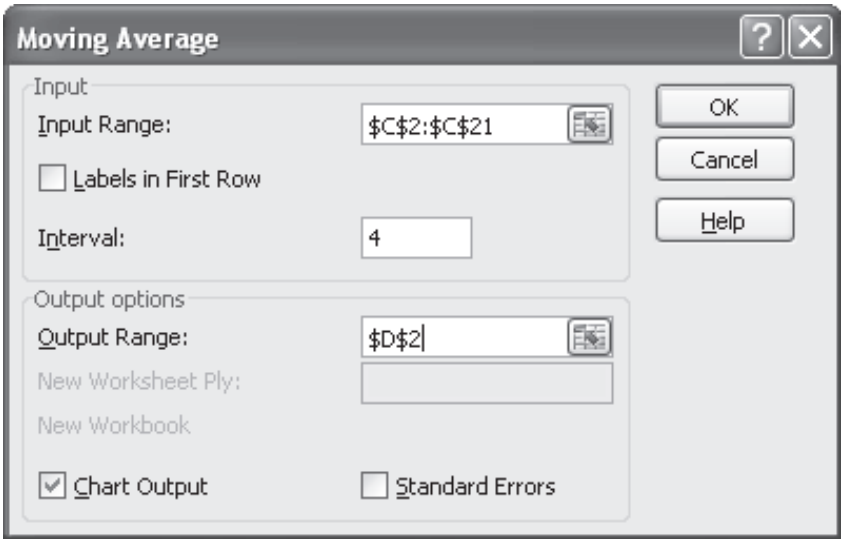


Figure 7.41: Filled Moving Average Dialog Box

Figure 7.42 shows the results of the Moving Average analysis.

	A	B	C	D	E	F	G	H	I	J	K
1	Year	Quarter	Sales (im million $)								
2	1	1	573.9	#N/A							
3		2	577.9	#N/A							
4		3	637.9	#N/A							
5		4	972.4	690.5							
6	2	1	701.0	722.3							
7		2	668.0	744.8							
8		3	639.0	745.1							
9		4	1115.0	780.8							
10	3	1	847.0	817.3							
11		2	806.0	851.8							
12		3	851.0	904.8							
13		4	1429.0	983.3							
14	4	1	1084.0	1042.5							
15		2	1100.0	1116.0							
16		3	1134.0	1186.8							
17		4	1946.0	1316.0							
18	5	1	1530.0	1427.5							
19		2	1387.0	1499.3							
20		3	1463.0	1581.5							
21		4	2541.0	1730.3							

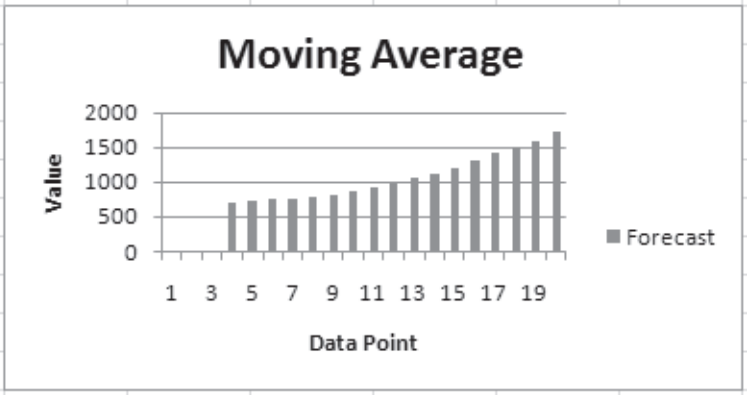


Figure 7.42: Moving Average Results

The **Moving Average** chart, as seen in figure 7.42, clearly depicts the rising trend in sales. The moving average values generated by Excel (cells D2:D21) can be seen just beside the input range values (cells C2:C21). Since the interval is set to 4, Excel calculates the moving average values from year 1's quarter 4 onwards. Values for year 1's quarters 1 to 3 (cells C2:C4) are replaced by the #N/A error in the output (cells D2:D4). From year 1's fourth quarter onwards, each value in the

output (cells D2:D21) is an average of values in the previous three quarters and itself. Thus, the moving average for the year 1's fourth quarter is nothing but the average sales value of the first four quarters.

```
(573.9 + 577.9 + 637.9 + 972.4) / 4 = 690.5
```

Similarly, the moving average for year 4's fourth quarter is the average sales value for year 4.

```
(1530 + 1387 + 1463 + 2541) / 4 = 1730.3
```

Since the process of calculating the average keeps moving from value to value, the process is called **Moving Average**.

7.1.12 Random Number Generation

The **Random Number Generation** tool generates random numbers. It can generate random numbers using seven different types of distribution methods. They are:

- **Uniform**: It generates random numbers by drawing numbers with equal probability between an upper and lower bound.
- **Normal**: It generates random numbers based on the mean and standard deviation values.
- **Bernoulli**: It generates random numbers in a two-step process. It initially generates values between 0 and 1. It then checks these values against a probability value. If the generated value is lesser than or equal to the probability value, a final value of 1 is generated. If the generated value is greater than the probability value, a final value of 0 is generated. Thus, at the end, the Random Number Generation tool generates a batch of 0 and 1 values.
- **Binomial**: It generates random numbers based on 'number of trials'. The 'number of trials' value represents the number of times an action, such as throwing a dice or pulling a card, is allowed. Depending on the probability value, the tool decides which of the nth trial is successful. This n^{th} trial is set as the output value. Thus, at the end, the Random Number Generation tool generates a batch of values between 0 and the 'number of trial' value entered by the user.
- **Poisson**: Poisson distributions are normally used to represent events that are accounted against units of time. For example, the number of customers visiting a mall every hour. The distribution generates random variables based on the probability that the event may occur 'n' number of times within a given interval of time.
- **Patterned**: It generates random values in a repetitive manner. It repeats generated values creating patterns. It generates values within a set range.
- **Discrete**: It generates random numbers as well as their probabilities. The sum of the probabilities is 1.

To display the **Random Number Generation** dialog box, on the **Data** tab, click **Data Analysis**, select **Random Number Generation** from the **Analysis Tools** list, and then click **OK**. Figure 7.43 shows the **Random Number Generation** dialog box.

V 1.0 © 2008 Aptech Limited

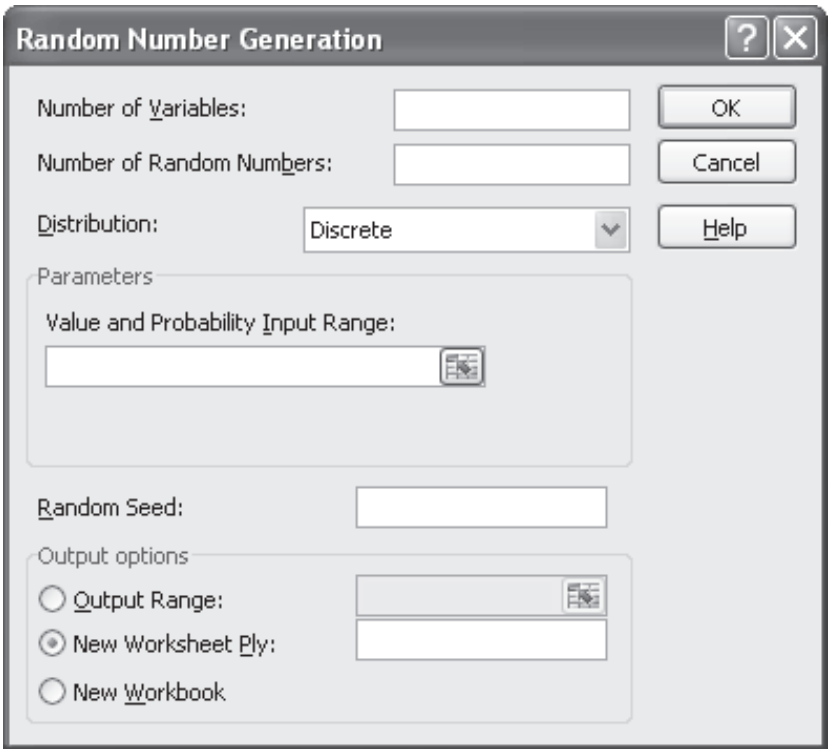


Figure 7.43: Random Number Generation Dialog Box

The **Number of Variables** box specifies the number of columns of random numbers that will be generated. The **Number of Random Numbers** box specifies the number of rows of random numbers that will be generated. The **Distribution** drop-down specifies the distribution method to be used for generating the random numbers. The **Parameters** option specifies the upper and lower boundaries within which the random numbers are to be generated. The **Parameters** option keeps changing with respect to the distribution method selected. The **Random Seed** box specifies a random number from which the random numbers will be generated. The **Output Range** box specifies the cell address from where the report is to be displayed.

To generate a few random numbers using the **Uniform** distribution method, fill the **Random Number Generation** dialog box, as shown in figure 7.44, and then click **OK**.

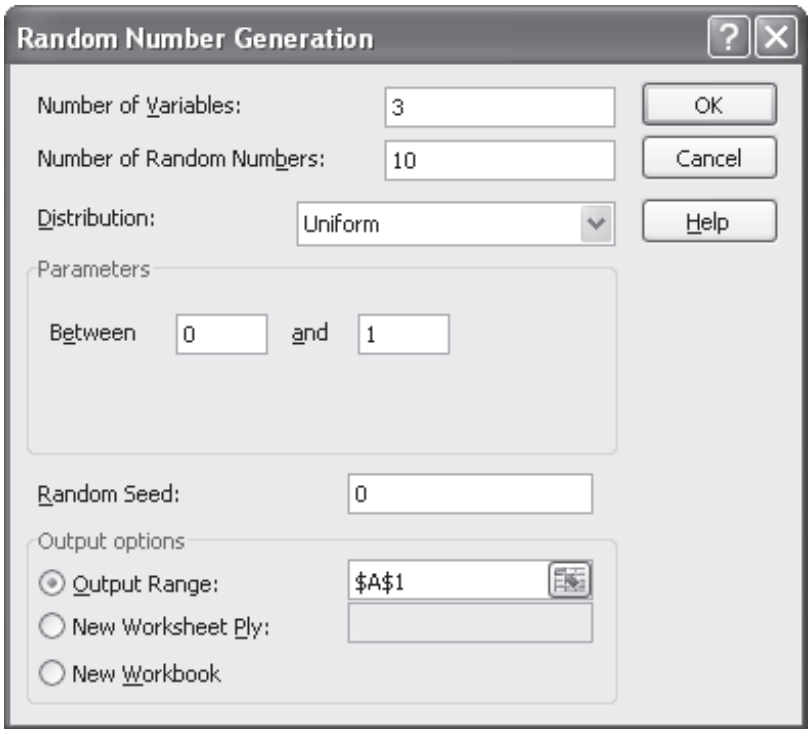


Figure 7.44: Filled Random Number Generation Dialog Box

Figure 7.45 shows the random numbers generated as a result of the Random Number Generation operation. The numbers are generated between 0 and 1 using uniform distribution.

	A	B	C
1	0.21396527	0.069795831	0.584063234
2	0.381603442	0.711783197	0.821802423
3	0.882961516	0.947141942	0.247627186
4	0.659566027	0.605883969	0.352916044
5	0.456862087	0.300698874	0.229346599
6	0.781212806	0.457472457	0.897640919
7	0.277138585	0.217902158	0.120212409
8	0.852137822	0.001373333	0.033875546
9	0.450209052	0.295937986	0.065981017
10	0.560258797	0.98586993	0.518570513

Figure 7.45: Random Number Generation Results

7.1.13 Rank and Percentile

The **Rank and Percentile** tool computes rank and percent rank for the values of a given set of data. Consider the score details of three different players shown in figure 7.46. A Rank and Percentile operation on one of the player's scores would rate each score in ascending order and even evaluate each score's percent.

V 1.0 © 2008 Aptech Limited

	A	B	C	D
1	Match	Player 'A'	Player 'B'	Player 'C'
2	1	57	45	99
3	2	34	33	67
4	3	25	47	57
5	4	89	89	29
6	5	100	150	45
7	6	102	24	29
8	7	77	65	99
9	8	34	7	29
10	9	0	36	30
11	10	45	25	38

Figure 7.46: Players' Scores Data

To display the **Rank and Percentile** dialog box, on the **Data** tab, click **Data Analysis**, select **Rank and Percentile** from the **Analysis Tools** list, and then click **OK**. Figure 7.47 shows the **Rank and Percentile** dialog box.

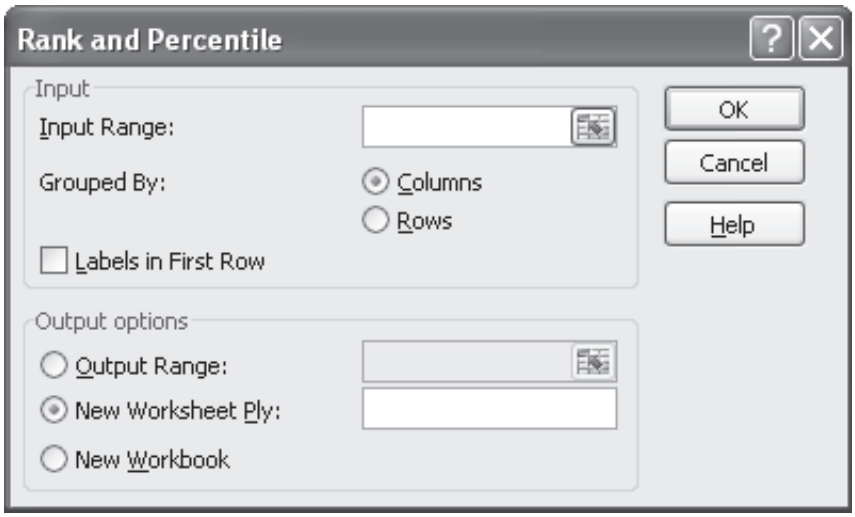


Figure 7.47: Rank and Percentile Dialog Box

The **Input Range** box specifies the range of cells on which the analysis is to be done. The **Grouped By** option specifies whether the data sets are to be treated as columns or rows. The **Labels in First Row** option specifies whether the data sets include labels in their first rows or columns. The **Output Range** box specifies the cell address from where the report is to be displayed.

To analyze the players' data shown in figure 7.46, fill the **Rank and Percentile** dialog box, as shown in figure 7.48, and then click **OK**.

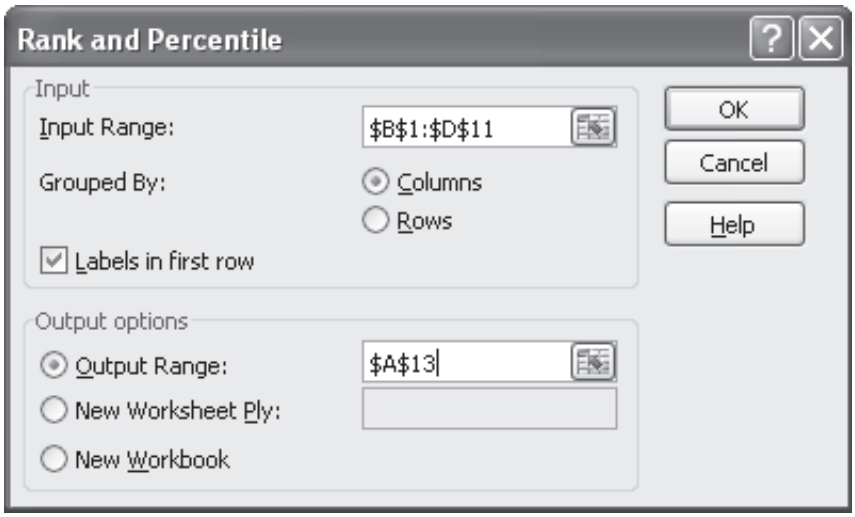


Figure 7.48: Filled Rank and Percentile Dialog Box

Figure 7.49 shows the results of the **Rank and Percentile** analysis.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Match	Player 'A'	Player 'B'	Player 'C'								
2	1	57	45	99								
3	2	34	33	67								
4	3	25	47	57								
5	4	89	89	29								
6	5	100	150	45								
7	6	102	24	29								
8	7	77	65	99								
9	8	34	7	29								
10	9	0	36	30								
11	10	45	25	38								
12												
13	*Point*	*Player 'A'*	*Rank*	*Percent*	*Point*	*Player 'B'*	*Rank*	*Percent*	*Point*	*Player 'C'*	*Rank*	*Percent*
14	6	102	1	100.00%	5	150	1	100.00%	1	99	1	88.80%
15	5	100	2	88.80%	4	89	2	88.80%	7	99	1	88.80%
16	4	89	3	77.70%	7	65	3	77.70%	2	67	3	77.70%
17	7	77	4	66.60%	3	47	4	66.60%	3	57	4	66.60%
18	1	57	5	55.50%	1	45	5	55.50%	5	45	5	55.50%
19	10	45	6	44.40%	9	36	6	44.40%	10	38	6	44.40%
20	2	34	7	22.20%	2	33	7	33.30%	9	30	7	33.30%
21	8	34	7	22.20%	10	25	8	22.20%	4	29	8	0.00%
22	3	25	9	11.10%	6	24	9	11.10%	6	29	8	0.00%
23	9	0	10	0.00%	8	7	10	0.00%	8	29	8	0.00%

Figure 7.49: Rank and Percentile Results

Cells A13:L23 show the rank and percent rank calculation for each player's scores. Cells A14:D23 represent the rank and percent rank evaluation for **`Player 'A'`**. The **`Point`** column represents the value's serial number, for example, the first value `102` (cell B14) appears sixth in the data set. The **`Rank`** column represents the rank of the scores, for example, on evaluating cells B2:B11, it can be observed that **`Player 'A'`** has the highest score of `102`. Hence, `102` is ranked at number 1. Similarly, `0` is the lowest score made by **`Player 'A'`**. Hence it is ranked 10. The **`Percent`** column represents the percent rank for each value in the data set.

V 1.0 © 2008 Aptech Limited

7.1.14 Regression

The **Regression** tool performs regression analysis on two or more sets of data.

Consider the shopping mall data shown in figure 7.14. To display the **Regression** dialog box, on the **Data** tab, click **Data Analysis**, select **Regression** from the **Analysis Tools** list, and then click **OK**. Figure 7.50 shows the **Regression** dialog box.

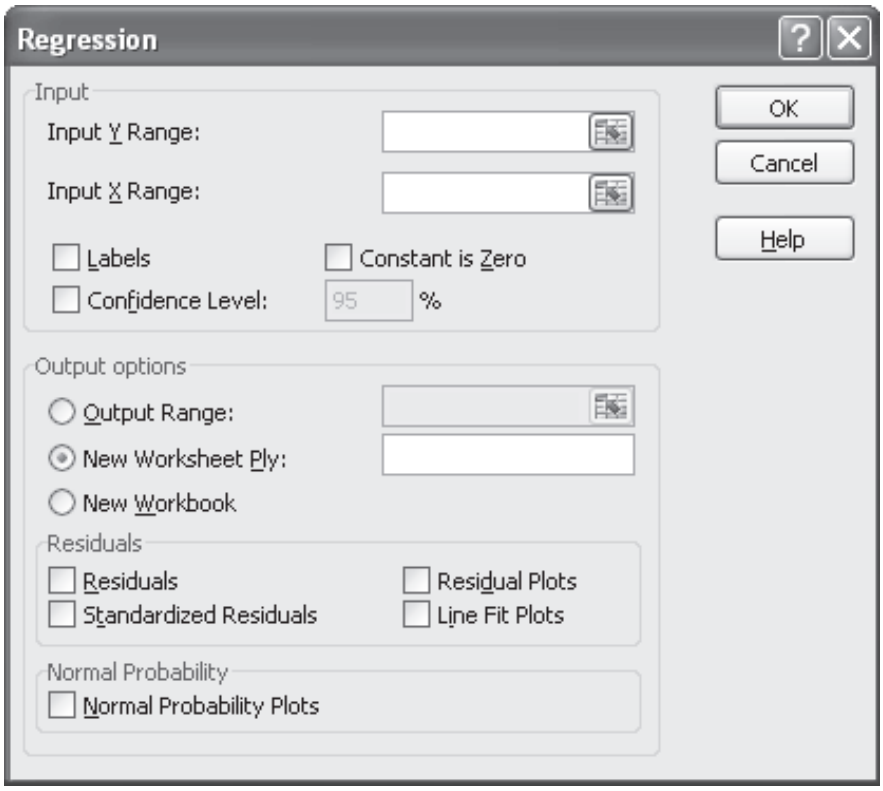


Figure 7.50: Regression Dialog Box

The **Input Y Range** box specifies the dependent range of cells. The **Input X Range** box specifies the independent range of cells. The **Labels** option specifies whether the data sets include labels in their first rows or columns. The **New Worksheet Ply** box specifies the name of a new worksheet in which the analysis report will be displayed. The **Line Fit Plots** option plots a graph depicting the regression analysis.

To analyze the shopping mall data shown in figure 7.14, fill the **Regression** dialog box, as shown in figure 7.51, and then click **OK**.

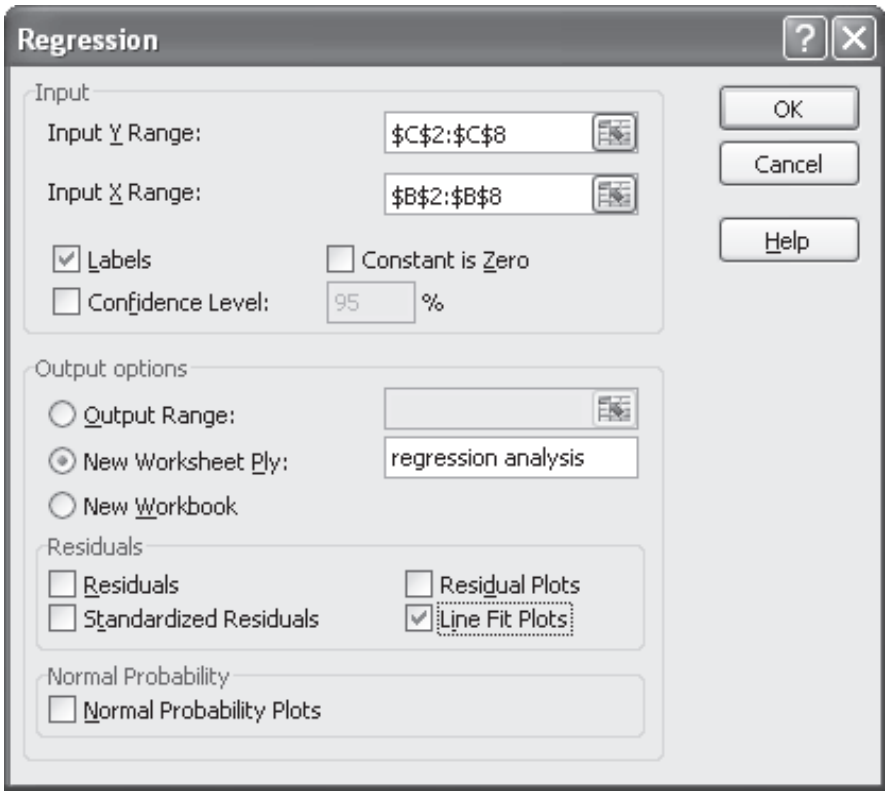


Figure 7.51: Filled Regression Dialog Box

Figure 7.52 shows the results of the **Regression** analysis and figure 7.53 shows the line plot created using the existing and predicted customer data set values.

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
3	*Regression Statistics*								
4	Multiple R	0.817148378							
5	R Square	0.667731472							
6	Adjusted R Square	0.58466434							
7	Standard Error	37.30725574							
8	Observations	6							
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*Significance F*			
12	Regression	1	11188.17468	11188.17468	8.038455832	0.047095277			
13	Residual	4	5567.325323	1391.831331					
14	Total	5	16755.5						
15									
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
17	Intercept	154.2688495	88.17277271	1.749620033	0.155085022	-90.53801369	399.0757127	-90.53801369	399.0757127
18	Gift Vouchers	4.490537699	1.583842644	2.83521707	0.047095277	0.093085544	8.887989854	0.093085544	8.887989854
22	RESIDUAL OUTPUT								
24	*Observation*	*Predicted Customers*	*Residuals*						
25	1	383.2862722	-63.28627215						
26	2	342.8714329	19.12856714						
27	3	468.6064884	16.39351157						
28	4	446.1537999	-10.15379994						
29	5	374.3051968	11.69480324						
30	6	387.7768099	26.22319015						

Figure 7.52: Regression Results

V 1.0 © 2008 Aptech Limited

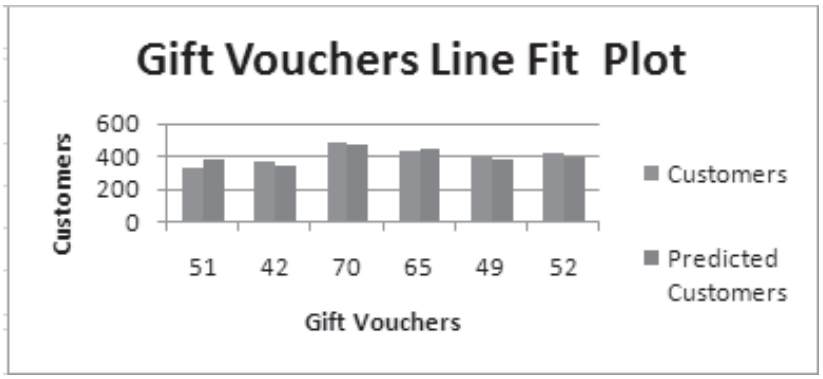


Figure 7.53: Regression Result Line Plot

The line fit plot displays the present and the predicted number of customers.

7.1.15 Sampling

The **Sampling** tool generates samples from data sets. The samples contain values from the data set itself.

Consider the sales details shown in Figure 7.38. To display the **Sampling** dialog box, on the **Data** tab, click **Data Analysis**, select **Sampling** from the **Analysis Tools** list, and then click **OK**. Figure 7.54 shows the **Sampling** dialog box.

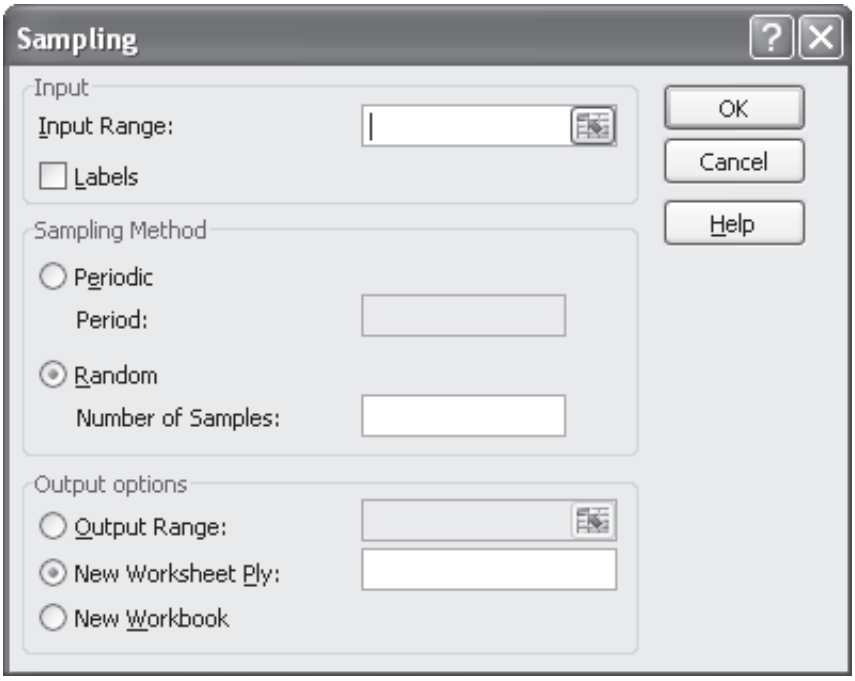


Figure 7.54: Sampling Dialog Box

The **Input Range** box specifies the range of cells on which the analysis is to be done. The **Labels** option specifies whether the data sets include labels in their first rows or columns. The **Number of Samples** box specifies the number of samples that will be created. The **Output Range** box specifies the cell address from where the report is to be displayed.

To analyze the sales data shown in figure 7.38, fill the **Sampling** dialog box, as shown in figure 7.55, and then click **OK**.

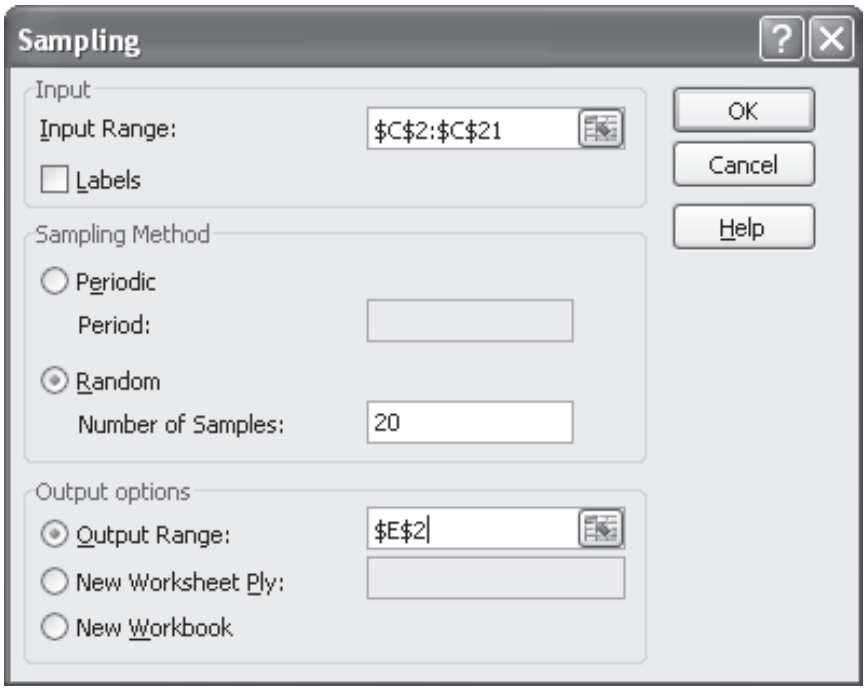


Figure 7.55: Filled Sampling Dialog Box

Figure 7.56 shows the results of the **Sampling** operation.

	A	B	C	D	E
1	Year	Quarter	Sales (im million $)		
2	1	1	573.9		847
3		2	577.9		1115
4		3	637.9		573.9
5		4	972.4		847
6	2	1	701.0		1946
7		2	668.0		573.9
8		3	639.0		577.9
9		4	1115.0		1100
10	3	1	847.0		637.9
11		2	806.0		573.9
12		3	851.0		1084
13		4	1429.0		639
14	4	1	1084.0		1463
15		2	1100.0		1100
16		3	1134.0		701
17		4	1946.0		639
18	5	1	1530.0		851
19		2	1387.0		1463
20		3	1463.0		1115
21		4	2541.0		577.9

Figure 7.56: Sampling Results

V 1.0 © 2008 Aptech Limited

The generated sample is shown in cells E2:E21.

7.1.16 t-Test

The t-Test tools tests two data sets to determine if their means are equal. The test can be conducted on three different assumptions. The assumptions are:

- The data set variances are equal
- The data set variances are not equal
- The data sets represent before and after treatment observations on the same object(s)

Consider the players' score data shown in figure 7.34. To display the **t-Test: Paired Two Sample for Means** dialog box, on the **Data** tab, click **Data Analysis**, select **t-Test: Paired Two Sample for Means** from the **Analysis Tools** list, and then click **OK**. Figure 7.57 shows the **t-Test: Paired Two Sample for Means** dialog box.

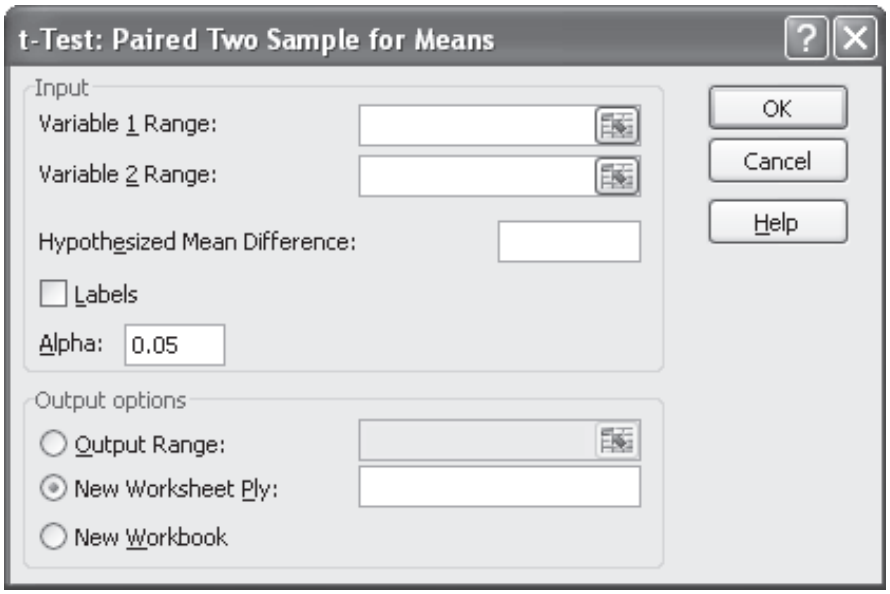


Figure 7.57: t-Test: Paired Two Sample for Means Dialog Box

The **Variable 1 Range** and **Variable 2 Range** boxes specify the first and second range of cells on which the analysis is to be done. The **Hypothesized Mean Difference** box specifies a hypothesized difference between the means of the two data sets. The **Labels** option specifies whether the data sets include labels in their first rows or columns. The **Output Range** box specifies the cell address from where the report is to be displayed.

To compare the means of players 'A' and 'B', as shown in figure 7.34, fill the **t-Test: Paired Two Sample for Means** dialog box, as shown in figure 7.58, and then click **OK**.

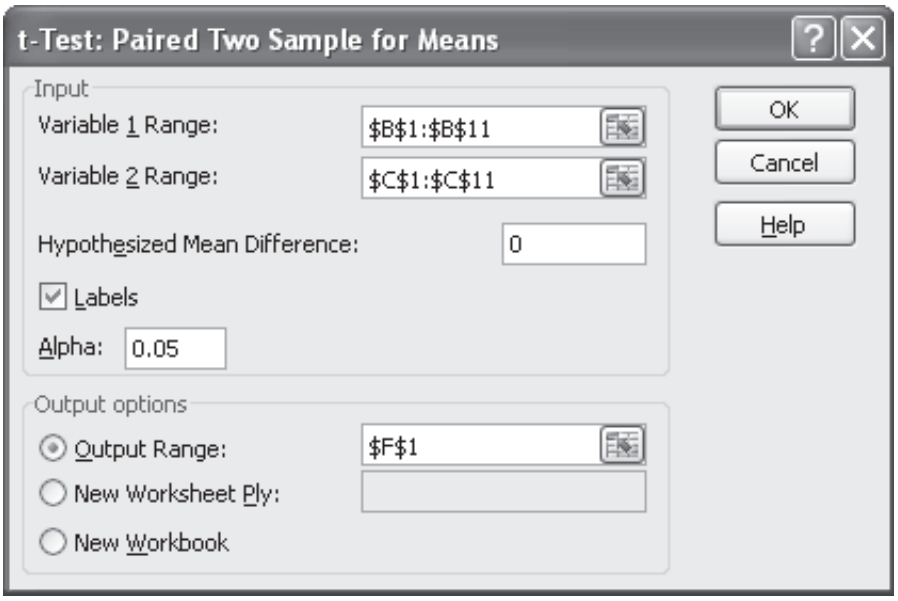


Figure 7.58: Filled t-Test: Paired Two Sample for Means Dialog Box

Figure 7.59 displays the result of the t-Test: Paired Two Sample for Means operation.

	A	B	C	D	E	F	G	H
1	Match	Player 'A'	Player 'B'	Player 'C'		t-Test: Paired Two Sample for Means		
2	1	57	45	99				
3	2	34	33	67			*Player 'A'*	*Player 'B'*
4	3	25	47	57		Mean	56.3	52.1
5	4	89	89	29		Variance	1196.455556	1707.877778
6	5	100	150	45		Observations	10	10
7	6	102	24	29		Pearson Correlation	0.564830373	
8	7	77	65	99		Hypothesized Mean Difference	0	
9	8	34	7	29		df	9	
10	9	0	36	30		t Stat	0.369859473	
11	10	45	25	38		P(T<=t) one-tail	0.360020299	
12						t Critical one-tail	1.833112923	
13						P(T<=t) two-tail	0.720040597	
14						t Critical two-tail	2.262157158	

Figure 7.59: t-Test: Paired Two Sample for Means Results

The **t-Test: Paired Two Sample for Means** tool generates the `t Stat` and `t Critical` (for one-tail and two-tail distributions) values as desired. Since, `t Stat` (cell 10) is lesser than `t Critical` (cell 14), there is no significant difference in the means of the scores of players `A` and `B`.

Figure 7.60 shows the results generated by the **t-Test: Two-Sample Assuming Equal Variances** tool and Figure 7.61 shows the results generated by the **t-Test: Two-Sample Assuming Unequal Variances** tool.

V 1.0 © 2008 Aptech Limited

	A	B	C	D	E	F	G	H
1	Match	Player 'A'	Player 'B'	Player 'C'		t-Test: Two-Sample Assuming Equal Variances		
2	1	57	45	99				
3	2	34	33	67			*Player 'A'*	*Player 'B'*
4	3	25	47	57		Mean	56.3	52.1
5	4	89	89	29		Variance	1196.455556	1707.877778
6	5	100	150	45		Observations	10	10
7	6	102	24	29		Pooled Variance	1452.166667	
8	7	77	65	99		Hypothesized Mean Difference	0	
9	8	34	7	29		df	18	
10	9	0	36	30		t Stat	0.246448433	
11	10	45	25	38		P(T<=t) one-tail	0.404061867	
12						t Critical one-tail	1.734063592	
13						P(T<=t) two-tail	0.808123735	
14						t Critical two-tail	2.100922037	

Figure 7.60: t-Test: Two-Sample Assuming Equal Variances Results

	A	B	C	D	E	F	G	H
1	Match	Player 'A'	Player 'B'	Player 'C'		t-Test: Two-Sample Assuming Unequal Variances		
2	1	57	45	99				
3	2	34	33	67			*Player 'A'*	*Player 'B'*
4	3	25	47	57		Mean	56.3	52.1
5	4	89	89	29		Variance	1196.455556	1707.877778
6	5	100	150	45		Observations	10	10
7	6	102	24	29		Hypothesized Mean Difference	0	
8	7	77	65	99		df	17	
9	8	34	7	29		t Stat	0.246448433	
10	9	0	36	30		P(T<=t) one-tail	0.40414317	
11	10	45	25	38		t Critical one-tail	1.739606716	
12						P(T<=t) two-tail	0.808286339	
13						t Critical two-tail	2.109815559	

Figure 7.61: t-Test: Two-Sample Assuming Unequal Variances Results

The **Two-Sample Assuming Equal Variances** and **Two-Sample Assuming Unequal Variances** evaluate to the same result. In both the cases, `t Stat` (cell F9) is lesser than `t Critical` (cell F13). Thus, both the tests confirm the result that there is no significant difference in the means of the scores of players `A` and `B`.

7.1.17 z-Test: Two Sample for Means

The **z-Test: Two Sample for Means** tool performs a z-test for samples whose variances are known. It tests the hypothesis that two samples, or the populations from which they have been drawn, have similar means. In case, the variances for the samples are not known, the **ZTEST** function should be used.

Consider the players' score data shown in figure 7.34. To display the **z-Test: Two Sample for Means** dialog box, on the **Data** tab, click **Data Analysis**, select **z-Test: Two Sample for Means** from the **Analysis Tools** list, and then click **OK**. Figure 7.62 shows the **z-Test: Two Sample for Means** dialog box.

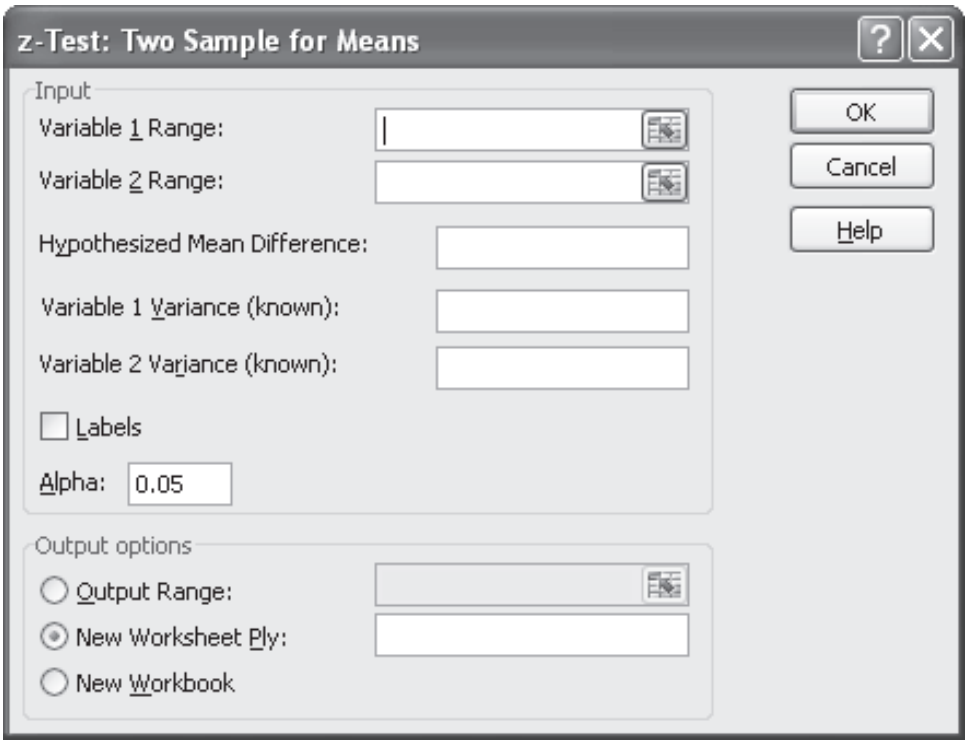


Figure 7.62: z-Test: Two-Sample for Means Dialog Box

The **Variable 1 Range** and **Variable 2 Range** boxes specify the range of cells to be analyzed. The **Hypothesized Mean Difference** box specifies an assumed difference between the means of the two data sets. The **Variable 1 Variance** and **Variable 2 Variance** boxes specify the variances of the two data sets respectively. The **Labels** option specifies whether the data sets include labels in their first rows or columns. The **Output Range** box specifies the cell address from where the report is to be displayed.

To analyze the players' data shown in figure 7.34, fill the **z-Test: Two Sample for Means** dialog box, as shown in figure 7.63, and then click **OK**.

V 1.0 © 2008 Aptech Limited

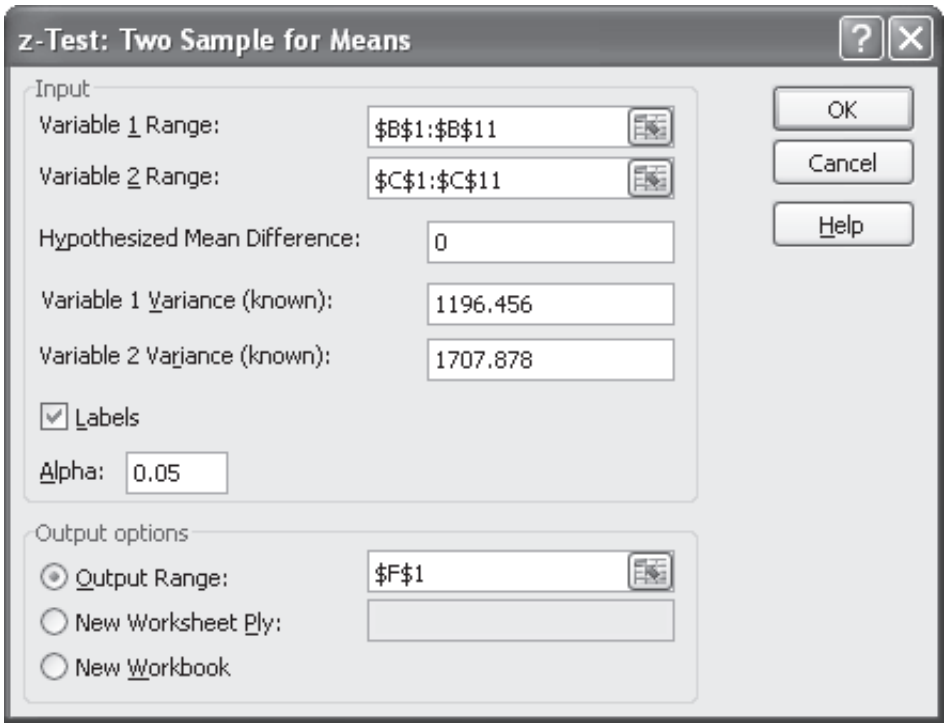


Figure 7.63: Filled z-Test: Two Sample for Means Dialog Box

Figure 7.64 shows the results of the z-Test: Two Sample for Means analysis.

	A	B	C	D	E	F	G	H
1	Match	Player 'A'	Player 'B'	Player 'C'		z-Test: Two Sample for Means		
2	1	57	45	99				
3	2	34	33	67			*Player 'A'*	*Player 'B'*
4	3	25	47	57		Mean	56.3	52.1
5	4	89	89	29		Known Variance	1196.456	1707.878
6	5	100	150	45		Observations	10	10
7	6	102	24	29		Hypothesized Mean Difference	0	
8	7	77	65	99		z	0.246448405	
9	8	34	7	29		P(Z<=z) one-tail	0.40266757	
10	9	0	36	30		z Critical one-tail	1.644853627	
11	10	45	25	38		P(Z<=z) two-tail	0.80533514	
12						z Critical two-tail	1.959963985	

Figure 7.64: z-Test: Two-Sample for Means Results

The **z-Test: Two Sample for Means** tool computes the `z` and `z Critical` (for one-tail and two-tail distributions) values as desired.

The Session In Brief

- MS Excel provides some medium scale statistical analysis tools in the form of the **Analysis ToolPak** add-in.
- The **Anova: Single Factor**, **Two-Factor Without Replication**, and **Two-Factor With Replication** tools compare two or more data sets by analyzing their variances.
- The **Correlation** tool computes the correlation coefficients for two or more sets of data, whereas the **Covariance** tool computes pair-wise covariance coefficients for two or more sets of data.
- The **Exponential Smoothing** tool computes corrected values for a given set of data after smoothing out errors present in the given set of data.
- The **F-Test Two-Sample** tool compares variances between two data sets. By comparing variances, the F-Test compares the spread of the two data sets.
- The **Fourier Analysis** tool calculates the Discrete Fourier Transform (DFT) for a range of discrete values using the Fast Fourier Transform (FFT) algorithm.
- The **t-Test** tool tests two data sets to determine if their means are equal, while the **z-Test: Two Sample for Means** tool performs a z-test for samples whose variances are known.

V 1.0 © 2008 Aptech Limited

Check Your Progress

1. Match the statistical tools with their correct functions.

	Tool		Function
1	Correlation	A	z-Test on two sets of data with known variances
2	F-Test Two-Sample for Variances	B	Performs regression analysis on two sets of data
3	Fourier Analysis	C	Computes rank and percentile for the values of a given set of data
4	Histogram	D	Generates random numbers
5	Moving Average	E	Computes correlation coefficients for two or more sets of data
6	Random Number Generation	F	Creates a histogram for a given set of data
7	Rank and Percentile	G	Performs Student's t-Test on two sets of data
8	Regression	H	Tests whether variances of the two given sets of data differ or not
9	t-Test: Paired Two-Sample for Means	I	Computes the Discrete Fourier Transform for a given set of data
10	z-Test: Two-Sample for Means	J	Computes moving averages for a given set of data

a. Correlation: ___

b. F-Test Two-Sample for Variances: ___

c. Fourier Analysis: ___

d. Histogram: ___

e. Moving Average: ___

f. Random Number Generation: ___

g. Rank and Percentile: ___

h. Regression: ___

i. t-Test Paired Two-Sample for Means: ___

j. z-Test Two-Sample for Means: ___

This page has been intentionally left blank

Session 8

Analysis ToolPak (Lab)

Session Objectives:

At the end of this session, you will be able to -

- Use the various Analysis ToolPak tools

Part I - 60 Minutes

Exercise 1: One Way Anova

Problem

A local coaching center has introduced three different levels of teaching: Beginners, Intermediate and Advanced. However, the coaching class claims that there is no difference in the quality of teaching material and the way classes are conducted. To prove its claim, it surveyed 5 students from each level of teaching. Each student was given a questionnaire, depending on which the student had to rate his/her level of teaching. Students were supposed to rate the teaching level on a scale of 0 to 60, with higher values representing higher satisfaction levels. Using the Anova: Single Factor test and the survey data determine if the claim is correct.

Beginners	Intermediate	Advanced
5	42	31
18	25	23
11	39	8
26	47	19
14	38	22

Table 8.1: Student's Survey Data

Analysis

The analysis of variance test involves the calculation of the F_{tab} and F_{cal} values.

Calculating F_{cal}

To begin with, first calculate the **Grand Mean** for the survey data.

	Beginners	Intermediate	Advanced
	5	42	31
	18	25	23
	11	39	8
	26	47	19
	14	38	22
Average	**14.8**	**38.2**	**20.6**

Grand Mean = (14.8 + 38.2 + 20.6)/3 = 24.53333

Now, calculate the Total Sum of Squares, Treatment Sum of Squares, and Unexplained Sum of Squares.

Total Sum of Squares

= (Grp 1 observation 1 – grand mean)^2 + … + (Grp 1 observation n – grand mean)^2 + … + (Grp n observation 1 – grand mean)^2 + … + (Grp n observation n – grand mean)^2

= (5 – 24.54)^2 + (18 – 24.54)^2 + (11 – 24.54)^2 + (26 – 24.54)^2 + (14 – 24.54)^2 + (42 – 24.54)^2 + (25 – 24.54)^2 + (39 – 24.54)^2 + (47 – 24.54)^2 + (38 – 24.54)^2 + (31 – 24.54)^2 + (23 – 24.54)^2 + (8 – 24.54)^2 + (19 – 24.54)^2 + (22 – 24.54)^2

= 2275.734

Treatment Sum of Squares

= (Grp 1 average – grand mean)^2 * number of observations + … + (Grp n average – grand mean)^2 * number of observations

= (14.8 – 24.54)^2 * 5 + (38.2 – 24.54)^2 * 5 + (20.6 – 24.54)^2 * 5

= 1484.933

Unexplained Sum of Squares

= (Grp 1 observation 1 – grp 1 average)^2 + … + (Grp 1 observation n – grp1 average)^2 + … + (Grp n observation 1 – grp n average)^2 + … + (Grp n observation n – grp n average)^2

= (5 – 14.8)2 + (18 – 14.8)^2 + (11 – 14.8)^2 + (26 – 14.8)^2 + (14 – 14.8)^2 + (42 – 38.2)^2 + (25 – 38.2)^2 + (39 – 38.2)^2 + (47 – 38.2)^2 + (38 – 38.2)^2 + (31 – 20.6)^2 + (23 – 20.6)^2 + (8 – 20.6)^2 + (19 – 20.6)^2 + (22 – 20.6)^2

= 790.8

SST = Treatment Sum of Squares = 1484.933

SSU = Unexplained Sum of Squares = 790.8

MST = SST / c – 1 = 1484.933 / 2 = 742.4667 … where c = number of data sets

MSU = SSU / n – c = 790.8 / 12 = 65.9 … where n = total number of values in all the data sets

F_{cal} = MST / MSU = 742.4667 / 65.9 = 11.26657

Calculating F_{tab}

Degree of Freedom = (c - 1, n - c) = (3 - 1, 15 - 3) = 2, 12

F_{tab} = F(2, 12) = 3.80

F_{cal} is much greater than **F_{tab}**. Thus, there is a significant difference between the three teaching levels. Thus, the claim made by the coaching center cannot be accepted.

Solution

Analysis of variance can be done in Excel 2007 using the Analysis ToolPak tools. To analyze the data using Excel 2007, do the following:

1. Replicate table 8.1 in a blank sheet, as shown in figure 8.1.

V 1.0 © 2008 Aptech Limited

	A	B	C
1	Beginners	Intermediate	Advanced
2	5	42	31
3	18	25	23
4	11	39	8
5	26	47	19
6	14	38	22

Figure 8.1: Teaching Level Data

2. On the Data tab, click Data Analysis → Anova: Single Factor → OK.

Figure 8.2 shows the **Anova: Single Factor** dialog box.

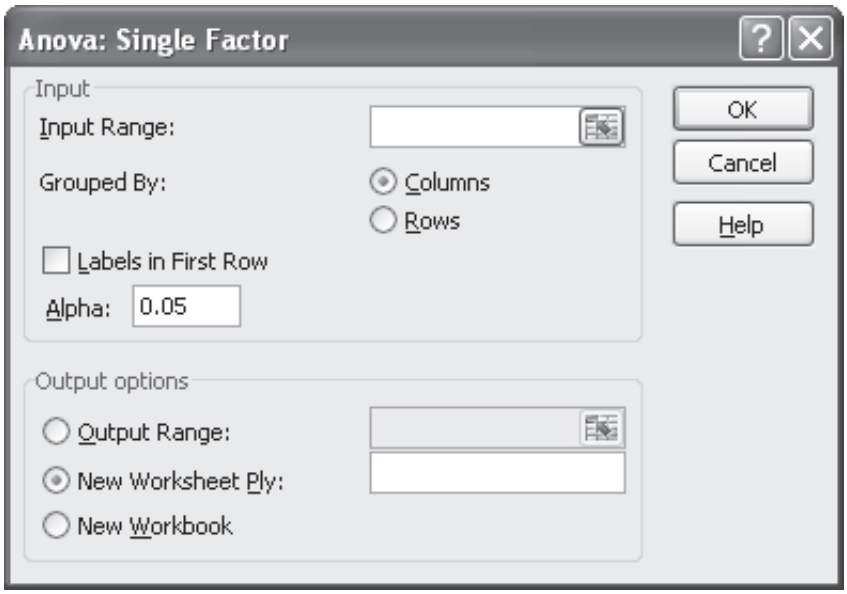


Figure 8.2: Anova: Single Factor Dialog Box

3. Set cells A1:C6 as the Input Range.

4. Select the Labels in First Row option.

5. Select the Output Range option and set it to cell E1.

Figure 8.3 shows the filled **Anova: Single Factor** dialog box.

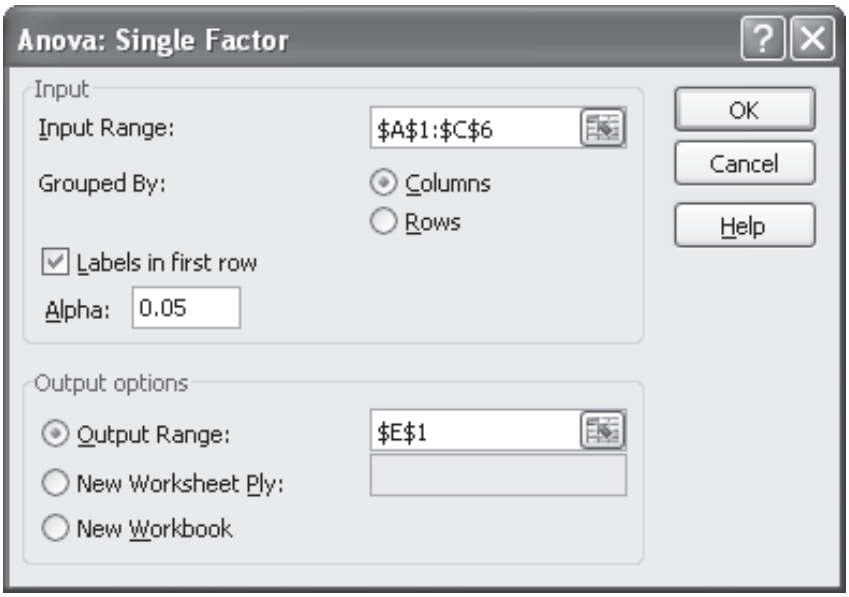


Figure 8.3: Filled Anova: Single Factor Dialog Box

6. Click OK.

	A	B	C	D	E	F	G	H	I	J	K
1	Beginners	Intermediate	Advanced		Anova: Single Factor						
2	5	42	31								
3	18	25	23		SUMMARY						
4	11	39	8		*Groups*	*Count*	*Sum*	*Average*	*Variance*		
5	26	47	19		Beginners	5	74	14.8	61.7		
6	14	38	22		Intermediate	5	191	38.2	66.7		
7					Advanced	5	103	20.6	69.3		
8											
9											
10					ANOVA						
11					*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
12					Between Groups	1484.933333	2	742.4666667	11.2665655	0.001760645	3.885293835
13					Within Groups	790.8	12	65.9			
14											
15					Total	2275.733333	14				

Figure 8.4: Anova: Single Factor Analysis Results

Figure 8.4 shows the analysis results. **F crit** represents the F_{tab} value and **F** represents the F_{cal} value. The values are almost the same. Again, since **F** is greater than **F crit**, it can be concluded that the teaching levels are significantly statistically different.

Exercise 2: t-Test

Problem

A coach at a local school wants to evaluate one of his student's performance over two seasons. He wants to know whether the student is performing consistently or if there is a significant difference in his performance. Table 8.2 shows the performance over two seasons. Using the t-Test: Two-Sample Assuming Unequal Variances test and the performance data, determine if there is a difference in the performance.

V 1.0 © 2008 Aptech Limited

Match	Season 1	Season 2
1	25	37
2	48	27
3	52	15
4	38	22
5	45	34
6	46	36
7	55	26
8	46	19
9	51	24
10	39	41

Table 8.2: Student Performance Data

Analysis

Table 8.2 shows the student's performance over two seasons. Consistency in the student's performance can be determined by evaluating whether the difference in the student's scores is accidental or due to other reasons. Since the number of samples is very small, a t-Test is appropriate to determine if the samples are statistically different.

Solution

1. Replicate table 8.2 in a blank sheet, as shown in figure 8.5.

	A	B	C
1	Match	Season 1	Season 2
2	1	25	37
3	2	48	27
4	3	52	15
5	4	38	22
6	5	45	34
7	6	46	36
8	7	55	26
9	8	46	19
10	9	51	24
11	10	39	41

Figure 8.5: Student Performance Data

2. On the Data tab, click Data Analysis → t-Test: Two-Sample Assuming Unequal Variances → OK.

Figure 8.6 shows the **t-Test: Two-Sample Assuming Unequal Variances** dialog box.

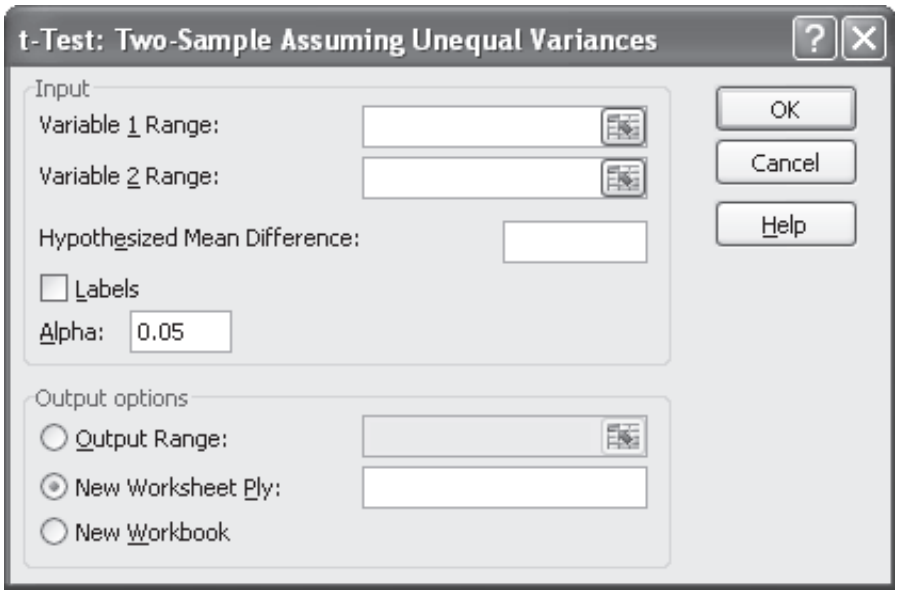


Figure 8.6: t-Test: Two-Sample Assuming Unequal Variances Dialog Box

3. **Set cells B1:B11 as Variable 1 Range.**
4. **Set cells C1:C11 as Variable 2 Range.**
5. **Set Hypothesized mean Difference as 0.**
6. **Select the Labels option.**
7. **Select the Output Range option and set it to cell E1.**

Figure 8.7 shows the filled **t-Test: Two-Sample Assuming Unequal Variances** dialog box.

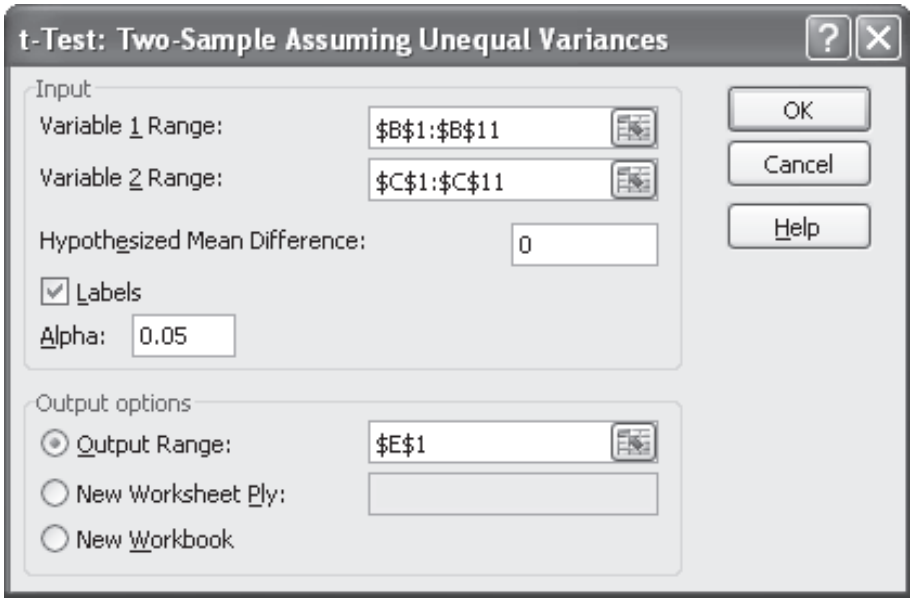


Figure 8.7: Filled t-Test: Two-Sample Assuming Unequal Variances Dialog Box

8. **Click OK.**

V 1.0 © 2008 Aptech Limited

	A	B	C	D	E	F	G
1	Match	Season 1	Season 2		t-Test: Two-Sample Assuming Unequal Variances		
2	1	25	37				
3	2	48	27			*Season 1*	*Season 2*
4	3	52	15		Mean	44.5	28.1
5	4	38	22		Variance	75.38888889	72.98888889
6	5	45	34		Observations	10	10
7	6	46	36		Hypothesized Mean Difference	0	
8	7	55	26		df	18	
9	8	46	19		t Stat	4.257546666	
10	9	51	24		P(T<=t) one-tail	0.000236832	
11	10	39	41		t Critical one-tail	1.734063592	
12					P(T<=t) two-tail	0.000473663	
13					t Critical two-tail	2.100922037	

Figure 8.8: t-Test: Two-Sample Assuming Unequal Variances Analysis Results

Figure 8.8 shows the results of the t-Test. From figure 8.8, it is seen that `t Stat` (cell F9) is greater than `t Critical` (F13). Thus, it can be concluded that there are significant differences in the performance and the change is surely not by chance.

Part II - 60 Minutes

1. Generate a sample of 30 random numbers using the data given below.

Input Variables	Probability
1	0.1
2	0.1
3	0.1
4	0.1
5	0.1
6	0.1
7	0.1
8	0.1
9	0.1
10	0.1

Hint: Use the Random Number Generation tool (Discrete distribution)

2. Compute the correlation and covariance coefficients for the following data.

Discount %	Increase in Sales (in '000s)
10	13
20	17
30	19

Discount %	Increase in Sales (in '000s)
40	20
50	25
60	29
70	35
80	39

Hint: Use the Correlation and Covariance tools

3. Generate histograms for the following data and bin ranges.

Sample Data		
1	57	20
81	59	49
36	90	82
75	18	86
71	52	31
2	10	37
16	17	99
45	13	1
2	38	54
58	61	61

Range 1	Range 2	Range 3
0.10	0.1	0.2
0.20	0.3	0.4
0.30	0.5	0.6
0.40	0.7	0.8
0.50	0.9	1
0.60		
0.70		
0.80		
0.90		
1.00		

V 1.0 © 2008 Aptech Limited

Do It Yourself

1. Using the students' marks list given below, calculate the rank and percentile for each subject individually.

Roll Number	Math	Physics	Chemistry
1	45	100	55
2	52	69	86
3	74	98	100
4	72	99	95
5	69	85	51
6	61	51	51
7	78	86	62
8	58	60	68
9	79	58	53
10	83	46	99
11	56	47	54
12	94	93	91
13	75	55	82
14	68	87	70
15	52	80	91
16	87	67	90
17	79	73	91
18	71	65	92
19	87	46	68
20	86	82	56

2. Identify the trend in the following demand data for a certain product over a span of three years.

Year	Quarters	Demand (in thousands)
1	1	13
	2	14
	3	18
	4	20
2	1	17
	2	19
	3	23
	4	24
3	1	19
	2	21
	3	25
	4	28

This page has been intentionally left blank

Session

9 Functions and Macros

Concepts

Objectives

At the end of this session, the student will be able to:

- *Describe lookup functions*
- *Describe financial functions*
- *Describe statistical functions*
- *Explain macros*

9.1 Introduction

Functions play a major role in MS Excel for making difficult tasks easy. MS Excel 2007 has introduced many new functions for performing statistical and financial analysis. Besides, macros help in automating many mundane and repetitive tasks. The session discusses lookup functions, financial and statistical functions, and macros in depth.

9.2 Lookup Functions

Lookup functions in MS Excel 2007 allows you to find a value within the defined range in a worksheet. MS Excel 2007 supports lookup functions, such as `LOOKUP()`, `MATCH()`, `HLOOKUP()`, and `VLOOKUP()`.

9.2.1 "HLOOKUP()" Function

The word HLOOKUP stands for horizontal lookup. The `HLOOKUP()` function searches for a value in the first row of a range of values. It finds a value from the same column in a specified row.

Syntax:

```
=HLOOKUP(key, range, row_index_num, range_lookup)
```

where,

`key`: A value to be searched in the first row of the range.

`range`: The range of data from which key will be searched.

`row_index_num`: The row number in the range that contains the key.

`range_lookup`: A logical value for `HLOOKUP()` to find for an exact or approximate match. If it is TRUE or omitted, an approximate value is returned. The next largest value that is less than lookup value is returned when an exact match is not found. Moreover, values in the first column of `range` must be placed in ascending sort order. If FALSE is specified, an exact match is returned. Values in the first column need not be sorted.

Consider a worksheet that has month wise sales data as shown in figure 9.1.

	A	B	C
1	**Month**	**Sales**	
2	January	1200.67	
3	February	2321.45	
4	March	4521.56	
5	April	2421.56	
6	May	3564.89	
7	June	3243.45	
8	July	1153.67	
9	August	4367.89	
10	September	4376.84	
11	October	3454.89	
12	November	3219.78	
13	December	5123.67	
14			

Figure 9.1: Month wise Sales Data

Table 9.1 shows how `HLOOKUP()` can be used to find values in a specified range of values.

Formula	Description
`=HLOOKUP("Sales",A1:B13,2,FALSE)`	The formula returns 1200.67. It searches for "Sales" in first row and returns the value from second row from the same column. Since the `range_lookup` is mentioned as FALSE, it searches for an exact match.
`=HLOOKUP("S",A1:B13,5,TRUE)`	The formula returns APRIL. It looks up for “S” in the first row and returns the value from fifth row. Because there is no exact match for “S”, it looks up for an approximate match. The next largest value that is less than “S” is “Month”, that value is used for search.

V 1.0 © 2008 Aptech Limited

Formula	Description
`=HLOOKUP("Sales",A1:B13,5)`	The formula returns 2421.56. The function looks up for "Sales" in first row within range A1:B13. It finds the word, and then returns corresponding value in fifth row.
`=HLOOKUP("d",A1:B13,2)`	The formula returns #N/A. The function cannot find an exact or nearest match of the lookup value. This is because there is no value which is less than "d" in the first row.

Table 9.1: Using `HLOOKUP()` Function

9.2.2 "VLOOKUP()" Function

The word VLOOKUP stands for vertical lookup. The `VLOOKUP()` function is used to search a value in the first column of a range of cells. It then returns a value from the same row in another column in the specified range.

Syntax:

```
=VLOOKUP(key, range, col_index_num, range_lookup)
```

where,

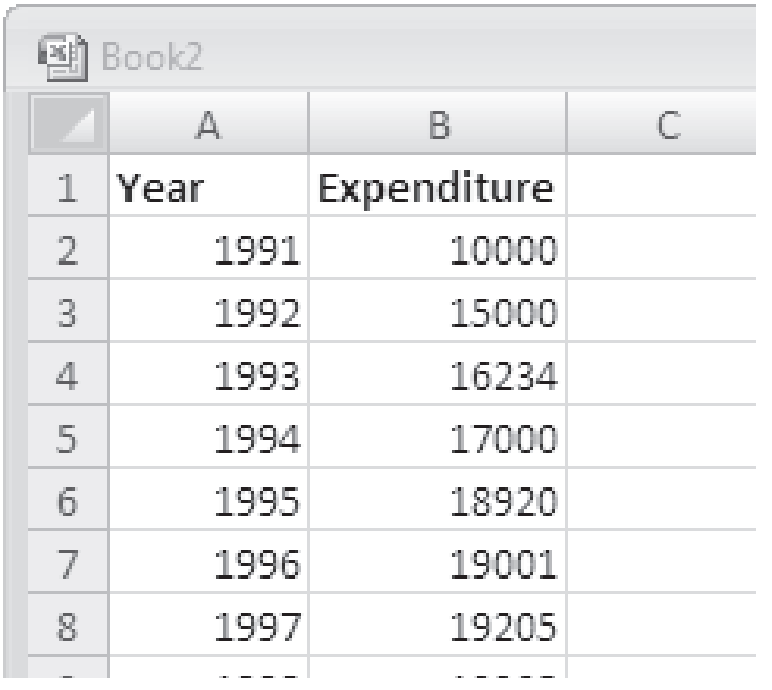
`key`: The value to be searched from the first column in the range. If the key value is smaller than the smallest value in the first column, it returns a #N/A error value.

`range`: The range of data from which key will be searched.

`col_index_num`: The index of column in the range that contains the key.

`range_lookup`: A logical value for `VLOOKUP()` to find for an exact or approximate match. If it is TRUE or omitted, an approximate value is returned. The next largest value that is less than lookup value is returned if an exact match is not found. Moreover, values in the first column of `range` must be placed in ascending sort order. If FALSE is specified, an exact match is returned. Values in the first column need not be sorted.

Figure 9.2 displays the worksheet that shows the commission and amount.



	A	B	C
1	Year	Expenditure	
2	1991	10000	
3	1992	15000	
4	1993	16234	
5	1994	17000	
6	1995	18920	
7	1996	19001	
8	1997	19205	

Figure 9.2: Commission and Amount

Table 9.2 shows how `VLOOKUP()` function can be used to find values in a specified range of values.

Formula	Description
`=VLOOKUP(0.05,B3:C14,2,FALSE)`	The formula returns 1234.8. The function looks up for 0.05 in column B and returns a value from column C in the same row.
`=VLOOKUP(0.12,B3:C14,2)`	The formula returns 2340. The function uses an approximate match. It searches for 0.12 in column B, finds the largest value that is less than or equal to 0.12, and returns a value from column C.
`=VLOOKUP(0.11,B3:C14,2,FALSE)`	The formula uses an exact match. It searches for the value, 0.11 in column B. Since there is no exact match for 0.11, it returns a #N/A error value.
`=VLOOKUP(0.04,B3:C14, 2, TRUE)`	The formula uses an approximate match. It searches for the value,0.04 in column B. Since the value, 0.04 is less than the smallest value in column B, it return the #N/A error value.

Table 9.2: Using `VLOOKUP()` Function

V 1.0 © 2008 Aptech Limited

9.2.3 "INDEX()" Function

The `INDEX()` function returns a value or reference to a value from a table or range. It operates both on an array and a reference form.

➢ **Array Form**

You can use the array form to return a value from a table or a range that is specified by a row and column indices.

Syntax:

```
=INDEX(array, row_index, column_index)
```

where,

`array`: Is a range of cells or an array constant. If the array has a single row or column, specifying a row index or a column index is optional. If the array has many rows and columns, and only a row index or a column index is used, then the function returns an entire row or column in the array.

`row_index`: Selects a particular row within `array` from which a value is returned. If `row_index` is omitted, `column_index` is required.

`column_index`: Selects a particular column within `array` from which a value is returned. It `column_index` is omitted, `row_index` is required.

Consider the following worksheet with item size, price, and count as shown in figure 9.3.

INDEX

	A	B	C	D
1	Item Size	Item Price	Count	
2	L	6.78	32	
3	M	4.56	42	
4	S	2.73	10	
5	XL	7.89	56	
6				

Figure 9.3: Item Details

Table 9.3 demonstrates the use of `INDEX()` function.

Formula	Description
`=INDEX (A2:C5,1,2)`	The function returns a value from the intersection of first row and the second column from the range. The result is 6.78.
`=INDEX(A2:C5,2,3)`	The function returns a value from the intersection of second row and the third column from the range. The result is 42.
`=INDEX({21,34,56;57,11,43},0,2)`	The function returns a value from the first row and the second column from the array constant. The result is 34.
`=INDEX({21,34,56;57,11,43},1,2)`	The function returns a value from the first row and the second column from the array constant. The result is 34.

Table 9.3: Using `INDEX()` Function

➢ **Reference Form**

The reference form returns the reference of a cell specified by a row and a column.

Syntax:

```
=INDEX(reference, row_index, column_index, area_index)
```

where,

`reference`: Is a reference to one or more cell ranges.

`row_index`: Is the number of the row in reference.

`column_index`: Is the number of the column in reference.

`area_index`: Selects a range in reference from which the specified `row_index` and `column_index` are returned. If there are more than one area, the index for the first one becomes 1, the second area becomes 2, and so on.

For example, if a reference describes cell ranges as (A1:C4, E1:F4, J1:K4), then A1:C4 is area index number 1, E1:F4 is area index number 2, and J1:K4 is area index number 3.

Consider the following worksheet with item size, price, and count as shown in figure 9.4.

INDEX

	A	B	C	D
1	Item Size	Item Price	Count	
2	L	6.78	32	
3	M	4.56	42	
4	S	2.73	10	
5	XL	7.89	56	
6				

Figure 9.4: Item Details

Table 9.4 demonstrates the use of `INDEX()` function with area number.

Formula	Description
`=INDEX((A2:C3,A4:C5),2,2,2)`	The formula returns the value of the intersection of the second row and second column from the second area number, that is, A4:C5. The result will be displayed from cell, B5, that is, 7.89.
`=INDEX((A2:C3,A4:C5),2,3,2)`	The formula returns 56 as the count value of XL item size. The count value for XL is located in second row, third column and second area number.
`=SUM(INDEX(A2:C5,0,3,1))`	This formula returns the sum of the third column in the first area in the range. Here 0 refers the first row and 3 refers the third column and 1 refers the first area of the range. The result will be the sum of C2:C5, that is, 140.

Table 9.4: Using `INDEX()` Function

9.2.4 "OFFSET()" Function

The `OFFSET()` function returns a reference to a single cell or a range of cells that is a specified number of rows and columns from a cell or a range of cells.

Syntax:

```
=OFFSET(ref, rows, cols, height, width)
```

where,

`ref`: Reference cell or a range of cells that is used as the initial reference point for the `OFFSET()` function.

`rows`: Represents number of rows away from the `ref` value. It is positive if you are moving down one or more rows to reach the offset value, otherwise it is negative.

`cols`: Represents number of columns away from the `ref` value. It is positive if you are moving in right direction to reach the offset value, otherwise it is negative.

`height`: Represents number of rows to be included. If it is omitted, it defaults to 1. It is optional.

`width`: Represents number of columns to be included. If it is omitted, it defaults to 1. It is optional.

Figure 9.5 shows the monthly sales of a company.

offset

	A	B	C
1	Month	Sales	
2	January	12321	
3	February	21987	
4	March	23543	
5	April	15256	
6	May	17890	
7	June	32897	
8	July	43562	
9	August	21678	
10	September	16786	
11	October	18324	
12	November	15478	
13	December	13876	
14			

Figure 9.5: Monthly Sales

Table 9.5 demonstrates the use of `OFFSET()` function.

Formula	Description
`=OFFSET(A1,7,1)`	Here the base cell specified is A1. Then it moves 7 rows below the base cell A1, and 1 column right of A1. The result will be 43562.

V 1.0 © 2008 Aptech Limited

Formula	Description
=OFFSET(B8,2,-1,1,1)	In this formula, the base cell specified is B8. Then it moves 2 rows down, 1 column left, with 1 row tall and 1 column width, that is, a single cell. This returns the value "September".
=SUM(OFFSET(B8,3,0,2,1))	Here the referenced cell is B8. It finds values 3 rows down and in the same column with 2 rows height and 1 column width. That means the sum of column values of B11 and B12. The result displayed is 33802.

Table 9.5: Using `OFFSET()` Function

9.2.5 "MATCH()" Function

The `MATCH()` function returns the position of an item that matches a user-specified value. It does not return the item. You can use `MATCH()` function to get the position of an item. Also, you can use lookup functions to find an item instead of its position.

Syntax:

```
=MATCH(lookup_val, lookup_range, match_type)
```

where,

`lookup_val`: Value that is searched within the specified range. It can be a value or a cell reference.

`lookup_range`: Range of cells in which the search will be performed.

`match_type`: Can be -1, 0, or 1. It specifies how `lookup_val` is matched with values within `lookup_range`. If `match_type` is 1, the largest value is found that is less than or equal to the lookup value. The lookup range must be sorted in the ascending order. If `match_type` is 0, then first value is found that exactly matches with the lookup value. The range can be unsorted also. If `match_type` is -1, the smallest value is found that is greater than or equal to the lookup value. The range must be sorted in descending order. If `match_type` is omitted, it is considered to be 1.

Figure 9.6 shows the ProductId, Price, and Count of some products. Here, the Price column is sorted in ascending order.

	A	B	C	D
1	ProductId	Price	Count	
2	P101	34.67	102	
3	P102	56.78	60	
4	P103	67.56	56	
5	P104	88.9	46	
6	P105	90.67	34	
7				

Figure 9.6: ProductId, Price, and Count

Table 9.6 demonstrates the use of `MATCH()` function.

Formula	Description
`=MATCH(67.56, B2:B6, 0)`	The formula returns 3, the position of 67.56 within the specified range, B2:B6. It searches for the exact match, because 0 is specified.
`=MATCH(67.28, B2:B6, 1)`	The formula returns 2, the position of 56.78. This is because 1 is specified, and the formula will return the position of the largest value that is less than or equal to specified value.
`=MATCH(50, C2:C6, -1)`	The formula returns 3, the position of 56. This is because -1 is specified, and the formula will find the smallest value that is greater than or equal to specified value.

Table 9.6: Using `MATCH()` Function

9.2.6 "LOOKUP()" Function

The `LOOKUP()` function returns a value from a range of values. It has two syntax forms: the vector form and the array form.

➢ **Vector Form**

A vector consists of a range of a single row or a single column. This type of `LOOKUP()` function searches within a one-row or one-column range (known as vector), and it returns a value from the same position within a second vector.

Syntax:

```
=LOOKUP(lookup_val, lookup_vector, output_vector)
```

where,

`lookup_val`: Value that is searched within first vector. It can be a number, reference, or a text.

`lookup_vector`: Range consisting of a single row or a column.

`output_vector`: Vector that consists of result. Both `output_vector` and `lookup_vector` must be of same size.

Figure 9.7 shows the sales data of a company. You want to find out the sales data for the month of June.

Lookup

	A	B	C
1	Month	Sales	
2	January	12321	
3	February	21987	
4	March	23543	
5	April	15256	
6	May	17890	
7	June	32897	
8	July	43562	
9	August	21678	
10	September	16786	
11	October	18324	
12	November	15478	
13	December	13876	
14			

Figure 9.7: Monthly Sales

Table 9.7 demonstrates the use of `LOOKUP()` function.

Formula	Description
`=LOOKUP("June",A2:A13,B2:B13)`	This formula looks up for "June" in column A and returns a value from column B within the same row. The result displayed is 32897.
`=LOOKUP("J", A2:A13,B2:B13)`	Before applying the formula you need to sort the data on Month column. Since there is no exact match, it finds the next smallest value, ie, February, and returns the result from column B that is in the same range. The result displayed is 21987.
`=LOOKUP(12320,B2:B13,A2:A13)`	Before applying the formula you need to sort the data on Sales column. It searches for the value, 12320 in column B. Since the smallest value in the column is 12321, it returns an error, #N/A.

Table 9.7: Using `LOOKUP()` Function with Vector Form

- **Array Form**

The array form of `LOOKUP()` searches for a value in the first row or the first column within an array, and returns a corresponding value from the same position from the last row or column of the array.

Syntax:

```
=LOOKUP(lookup_val, array)
```

where,

`lookup_val`: Value that is searched within an array. It can be a number, reference, or a text.

`array`: Range within which the lookup value is searched.

V 1.0 © 2008 Aptech Limited

Figure 9.8 shows the monthly sales data of a company.

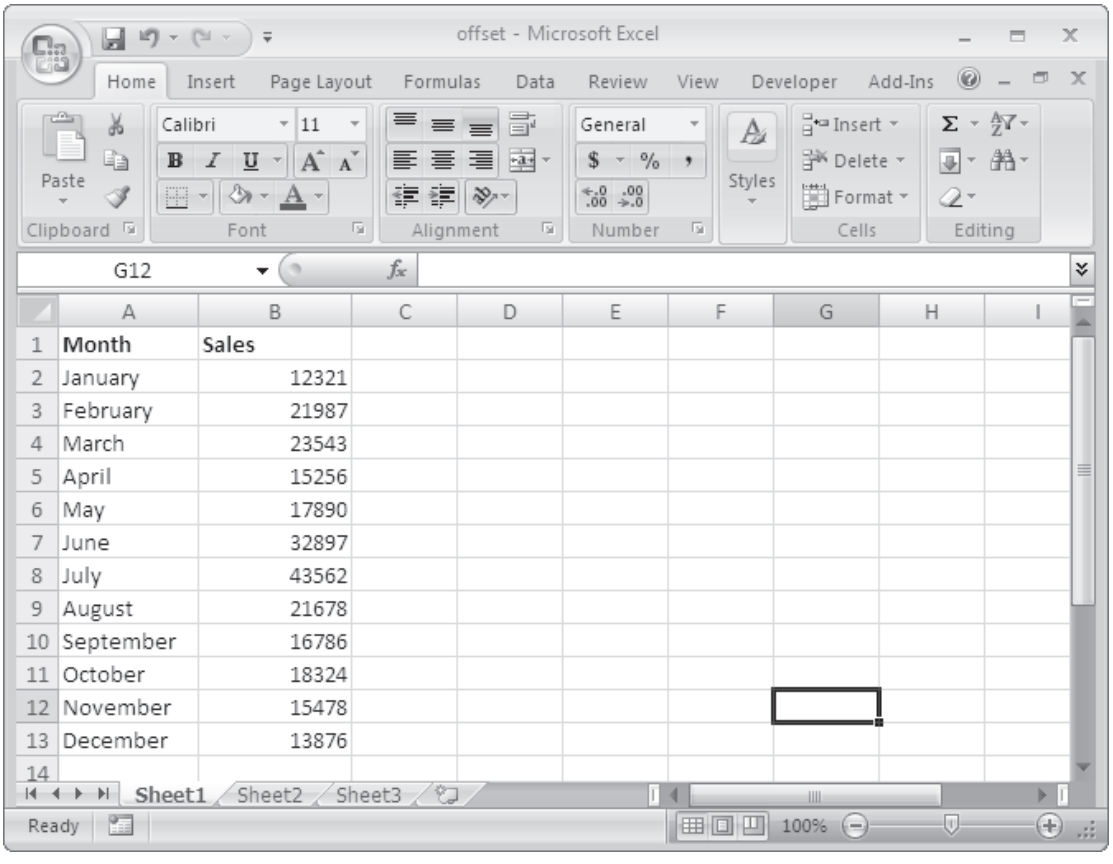


Figure 9.8: Monthly Sales

Table 9.8 demonstrates the use of `LOOKUP()` function with array form.

Formula	Description
`=LOOKUP(21987,{12321,21987,23543,15256,17890},{"January","February","March","April","May"})`	This formula looks up the value 21987 in the first row of the array, and then returns the value in the last row of the array in the same column. The result displayed is “February”.
`=LOOKUP(13870,{12321,13876,15256,15478,16786},{"January","December","April","November","September-"})`	This formula searches the value 13870 in the first row of the array, finds the largest value which is less than or equal to 13876, and then returns the value in the last row of the array in the same column. The result displayed is “January”.

Table 9.8: Using `LOOKUP()` Function with Array Form

Concepts

9.3 Financial Functions

Excel 2007 provides many useful financial functions. These functions can be used to find the present and future value of an investment, rate of interest of a loan, and so on.

9.3.1 "PMT()" Function

The Periodic Payment, the `PMT()` function calculates the loan payment based on a fixed interest rate and fixed payments.

Syntax:

```
=PMT(rate, num_installments, pv, fv, mode)
```

where,

`rate`: Is the interest rate for the loan.

`num_installments`: Is the number of payments for the loan.

`pv`: Is the present value of an investment, also known as principal amount.

`fv`: Is the future value for the investment.

`mode`: Indicates when payments are due. It can be either 0 or 1 as shown in table 9.9.

Value	Description
0	Payment is required at the end of the period. It is default value.
1	Payment is required at the beginning of the period.

Table 9.9: Values of Mode

Suppose, you want to find out how much you have to pay at the beginning of every month for a loan of principal amount $1,00,000 as shown in figure 9.9. The duration of the loan is 4 years or 48 months. The rate of interest is 10% that will be charged on the annual basis.

V 1.0 © 2008 Aptech Limited

PMT

	A	B	C
1	Rate	10%	
2	No. Of Payments	48	
3	Present Value	100000	
4	Future Value	0	
5			
6	Payment		
7			

Figure 9.9: Details of Loan

Table 9.10 demonstrates the use of `PMT()` function.

Formula	Description
`=PMT(B1/12,B2,B3,B4,1)`	This formula calculates the monthly payment for a loan at the beginning of every month. The principal value is 1,00,000 and future value is 0 for the loan. The result displayed is (-2515.30).
`=PMT(B1/12,B2,B3,B4,0)`	This formula calculates the monthly payment for a loan at the end of every month. The result displayed is (-2536.26).

Table 9.10: Using PMT() Function

9.3.2 "PV()" Function

The `PV()` function returns the present value of an investment. It is based on a rate of interest and a payment schedule.

Syntax:

```
=PV(rate, num_installments, pmt, fv, mode)
```

where,

`rate`: Is the annual rate of interest.

`num_installments`: Is the number of installments to be paid.

`pmt`: Is the amount of payment.

`fv`: Is the future value of payment. It is optional.

`mode`: It indicates when the payments are due. It is optional. It can have one of the following values as shown in table 9.11.

Value	Description
0	Payment is required at the end of the period. It is default value.
1	Payment is required at the beginning of the period.

Table 9.11: Values of Mode

Suppose, you want to find out the present value of an investment that requires a payment of $250 at the end of every month for 4 years as shown in figure 9.10. The invested money will earn 8% interest annually.

	A	B
1	Monthly Payment	250
2	No. of Months	48
3	Rate of Interest	8%

Figure 9.10: Present Value of an Investment

Table 9.12 demonstrates the use of `PV()` function.

Formula	Description
`=PV(B3/12, B2, B1,, 0)`	The formula calculates the present value for an investment. The result displayed is (-10,240.48).
`=PV(B3/12, B2, B1,, 1)`	The formula calculates the present value for an investment when the payment is due at the beginning of the period. The result displayed is (-10,308.75).

Table 9.12: Using `PV()` Function

V 1.0 © 2008 Aptech Limited

9.3.3 "FV()" Function

The `FV()` function calculates the future amount of an investment. It is based on periodic, constant interest rate and constant payments.

Syntax:

```
=FV(rate, num_installments, pmt, pv, mode)
```

where,

`rate`: Rate of interest.

`num_installments`: Number of installments.

`pmt`: Amount of the payment to be made for each period.

`pv`: The present value. It is optional.

`mode`: Indicates when the payments are due. It is optional. It can be one of the following values as shown in table 9.13.

Value	Description
0	Payment is required at the end of the period. It is default value.
1	Payment is required at the beginning of the period.

Table 9.13: Values of Mode

Suppose, you have made an initial deposit of $5000 in an investment. You want to deposit $350 at the beginning of every month as shown in figure 9.11. The investment will earn interest at the rate of 8% annually for the period of 4 years. You want to find out the future value of the investment.

	A	B	C
1	Deposit	5000	
2	Payment	350	
3	Earning	8%	
4	Tenure	4	
5			

Figure 9.11: Future Value of an Investment

Table 9.14 demonstrates the use of `FV()` function.

Formula	Description
`=FV(B3/12, B4*12, -B2, -B1, 1)`	The formula calculates the future value of an investment. The result displayed is 26,732.28
`=FV(B3/12, B4*12, -B2, -B1, 0)`	The formula calculates the future value of an investment when the payment is due at the end of the period. The result displayed is 26,600.80.

Table 9.14: Using `FV()` function

9.3.4 "RATE()" Function

The `RATE()` function returns the rate of interest for an investment.

Syntax:

```
=RATE(num_installments, pmt, pv, fv, mode, estimation)
```

where,

`num_installments`: Number of installments.

`pmt`: Amount of the payment to be made.

`pv`: Present value of an investment.

`fv`: Future value. It is optional.

`mode`: Indicates when the payments are due. It is optional. It can be one of the following values as shown in table 9.15.

Value	Description
0	Payment is required at the end of the period. It is default value.
1	Payments is required at the beginning of the period.

Table 9.15: Values of Mode

`estimation`: Approximate rate of interest. If it is omitted, the `RATE()` function assumes 10 percent.

V 1.0 © 2008 Aptech Limited

Suppose, you want to find out the interest rate for a loan amount of $5000 as shown in figure 9.12. The monthly payments for the loan is $200, to be made at the beginning of every month for the duration of 4 years.

	A	B	C
1	Loan Amount	5000	
2	Payment	200	
3	Tenure(Years)	4	
4			

Figure 9.12: Rate of Interest

Table 9.16 demonstrates the use of `RATE()` function.

Formula	Description
`=RATE(B3*12, -B2, B1, , 1)`	The formula returns 3% as the rate of interest for payments to be made at the beginning of every month.
`=RATE(B3*12, -B2, B1, , 0)`	The formula returns 3% as the rate of interest for payments to be made at the end of every month.

Table 9.16: Using `RATE()` Function

9.4 Statistical Functions

MS Excel 2007 provides a set of useful functions for statistical analysis on the data that are present in a spreadsheet.

9.4.1 "COUNT()" Function

The `COUNT()` function calculates the total number of cells of a selected range. It counts only cells that contain numbers. It does not count empty cells or cells with text. However, if a number is later added to an empty cell, the function updates the total.

Note: Cells that contain dates and time functions, and formulas are formatted as numbers in Excel. These are considered by `COUNT()` function.

Syntax:

```
=COUNT( value1, value2, value3,... )
```

where,

`value1, value2, value3`: Cells that contain numbers.

Figure 9.13 shows the spreadsheet that contains the marks of four students.

	A	B	C	D
1	Name	Subject	Marks	
2	Martha	Physics	50	
3	Martha	Computer Science	48	
4	Martha	Chemistry	49	
5	Charles	Physics	45	
6	Charles	Computer Science	47	
7	Charles	Chemistry	48	
8	Mary	Physics	50	
9	Mary	Computer Science	48	
10	Mary	Chemistry	47	
11	David	Physics	45	
12	David	Computer Science	48	
13	David	Chemistry	49	
14				

Figure 9.13: Marks of Students

Table 9.17 demonstrates the use of `COUNT()` function.

Formula	Description
`=COUNT(C2:C13)`	This formula counts the number of cells that contain numbers in the range. The result displayed is 12.
`=COUNT(C2:C13, 23)`	This formula counts the number of cells that contain numbers in the range and the value 23. The result displayed is 13.

Table 9.17: Using COUNT() Function

9.4.2 "COUNTIF()" Function

The `COUNTIF()` function counts the number of cells based on a given criteria within a selected range.

Syntax:

```
=COUNTIF(range, criteria)
```

where,

`range`: Is the range of cells for counting.

`criteria`: Is the condition in the form of a number, expression, or a text string. For example, the criteria can be 7, “7”, “>7”, “Vehicles”, or B5.

Figure 9.13 shows the spreadsheet that contains the marks of four students.

	A	B	C	D
1	Name	Subject	Marks	
2	Martha	Physics	50	
3	Martha	Computer Science	48	
4	Martha	Chemistry	49	
5	Charles	Physics	45	
6	Charles	Computer Science	47	
7	Charles	Chemistry	48	
8	Mary	Physics	50	
9	Mary	Computer Science	48	
10	Mary	Chemistry	47	
11	David	Physics	45	
12	David	Computer Science	48	
13	David	Chemistry	49	
14				

Figure 9.14: Marks of Students

Table 9.18 demonstrates the use of `COUNTIF()` function.

Formula	Description
`=COUNTIF(C2:C13, ">48")`	The formula counts the number of cells with a value greater than 48. The result displayed is 4.
`=COUNTIF(C2:C13, "<>"&C3)`	The formula counts the number of cells with a value not equal to 48. The result displayed is 8.
`=COUNTIF(B2:B13,"P*")`	The formula counts the number of cells starting with the letter “P”. The result displayed is 4.

Table 9.18: Using `COUNTIF()` Function

Concepts

9.4.3 "SUMIF()" Function

The `SUMIF()` function adds the cells based on a specified condition.

Syntax:

```
=SUMIF(range, criteria, sum_range)
```

where,

`range`: Is the range of cells for criteria evaluation.

`criteria`: Is the condition in which the cells are added or ignored.

`sum_range`: Data in this range will be added only if data in `range` meets `criteria`. If it is omitted, only `range` is summed up.

Figure 9.15 shows the data of student marks in various subjects.

	A	B	C	D
1	**Name**	**Subject**	**Marks**	
2	Martha	Physics	50	
3	Martha	Computer Science	48	
4	Martha	Chemistry	49	
5	Charles	Physics	45	
6	Charles	Computer Science	47	
7	Charles	Chemistry	48	
8	Mary	Physics	50	
9	Mary	Computer Science	48	
10	Mary	Chemistry	47	
11	David	Physics	45	
12	David	Computer Science	48	
13	David	Chemistry	49	
14				

Figure 9.15: Marks of Students

V 1.0 © 2008 Aptech Limited

Table 9.19 demonstrates the use of `SUMIF()` function.

Formula	Description
`=SUMIF(B2:B13, "Physics", C2:C13)`	The formula searches for only “Physics”, the criterion within the range B2:B13, and then returns total marks obtained in Physics in `sum_range`, C2:C13 by all students, that is, 190.
`=SUMIF(B2:B13, "Computer Science", C2:C13)`	The formula searches for only “Computer Science”, the criterion within the range B2:B13, and then returns total marks in range obtained in Computer Science in `sum_range` C2:C13 by all students, that is, 191.
`=SUMIF(B2:B13, "Chemistry", C2:C13)`	The formula searches for only “Chemistry”, the criterion within the range B2:B13, and then returns total marks in range obtained in Chemistry within `sum_range` C2:C13 by all students, that is, 193.

Table 9.19: Using `SUMIF()` Function

9.4.4 “STDEV()” Function

The word STDEV stands for Standard Deviation. The `STDEV()` function returns the standard deviation of data that are present in a worksheet.

Syntax:

```
=STDEV(num1, num2, num3, ...)
```

where,

`num1`, `num2`, `num3`: Are numeric values. They can be numbers, arrays, named ranges, or references to numbers. You can enter up to 30 values.

Suppose you want to find out the standard deviation of month wise sales figure of 2009 as shown in figure 9.16.

	A	B	C
1	Sales Figure for 2009		
2			
3	Month	Sales	
4	January	10000	
5	February	12000	
6	March	23123.67	
7	April	6754.89	
8	May	12769.56	
9	June	34562.78	
10	July	45671.34	
11	August	23421.67	
12	September	56734.12	
13	October	123216.78	
14	November	6745.9	
15	December	98769.23	
16			

Figure 9.16: Sales Figure in 2009

Table 9.20 demonstrates the use of `STDEV()` function.

Formula	Description
`=STDEV(B4:B15)`	The formula calculates the standard deviation for the entire range, B4:B15, and returns 37970.68168.
`=STDEV(10000, 12000, 6754.67)`	The formula calculates the standard deviation for specified numbers, and returns 2647.188874.
`=STDEV(B4:B15, 10210, 20234)`	The formula calculates the standard deviation for range, `B4:B15` and two more numbers, and returns 35932.33057.

Table 9.20: Using `STDEV()` Function

V 1.0 © 2008 Aptech Limited

9.4.5 "STDEVP()" Function

The `STDEVP()` function is used for calculating the standard deviation for a whole population of data in a worksheet.

Syntax:

```
=STDEVP( num1, num2, num3, ... )
```

where,

`num1`, `num2`, `num3`: Are the numbers for population. You can enter up to 30 numbers.

Figure 9.17 shows the sales data for the last three years.

	A	B	C	D	E	F	G
1			Annual Sales Figure				
2							
3	Sales in 2008		Sales in 2009		Sales in 2010		
4	Month	Sales	Month	Sales	Month	Sales	
5	January	23000	January	56908	January	57900	
6	February	345700	February	54670	February	76300	
7	March	56000	March	23190	March	88900	
8	April	12678	April	78456	April	75680	
9	May	45678	May	81500	May		
10	June	35780	June	87600	June		
11	July	67800	July	34580	July		
12	August	56890	August	96540	August		
13	September	34670	September	76500	September		
14	October	66079	October	56880	October		
15	November	45600	November	74908	November		
16	December	68450	December	86450	December		
17							

Figure 9.17: Using `STDEVP()` Function

Table 9.21 demonstrates the use of `STDEVP()` function.

Formula	Description
`=STDEVP(B5:B16, D5:D16, F5:F8)`	The formula returns 57213.76719 as the standard deviation for the entire population of data.
`=STDEVP(B5:B16,D5:D16,57900,76300,88900,75680)`	The formula returns 57213.76719 as the standard deviation for the entire population of data. The formula can consist of range as well numbers.

Table 9.21: Using `STDEVP()` Function

9.4.6 "AVERAGEIF()" Function

The `AVERAGEIF()` function calculates the average of values in cells based on specified condition.

Syntax:

```
=AVERAGEIF(range, criteria, avg_range)
```

where,

`range`: Is the range of cells for criteria evaluation.

`criteria`: Is the condition in which the cells are added or ignored.

`avg_range`: Data in this range will be averaged only if data in `range` meets `criteria`. If it is omitted, only `range` is averaged up.

V 1.0 © 2008 Aptech Limited

Figure 9.18 displays the sales and commission values.

	A	B	C
1	Sales	Commission	
2	100000	5000	
3	200000	8000	
4	300000	12000	
5	400000	15000	
6	500000	18000	
7			

Figure 9.18: Annual Sales Data

Table 9.22 demonstrates the use of `AVERAGEIF()` function.

Formula	Description
`=AVERAGEIF(A2:A6, ">300000")`	The formula calculates average of all sales data that are > 3,00,000 within the range, A2:A6, and returns 4,50,000.
`=AVERAGEIF(A2:A6, ">250000", B2:B6)`	The formula calculates average of commission data within range, B2:B6 based on the condition, sales data > 2,50,000. It returns 15,000.

Table 9.22: Using `AVERAGEIF()` Function

9.4.7 "MEDIAN()" Function

The `MEDIAN()` function calculates and returns a single middle value in an odd set of numbers. It calculates and returns the average of middle two values, if there is an even set of numbers.

Note: You can enter up to 255 numbers in the function.

Figure 9.19 shows the yearly household expenditure data.

Book2

	A	B	C
1	Year	Expenditure	
2	1991	10000	
3	1992	15000	
4	1993	16234	
5	1994	17000	
6	1995	18920	
7	1996	19001	
8	1997	19205	
9	1998	19305	
10	1999	19567	
11			

Figure 9.19: Annual Household Expenditure Data

Table 9.23 demonstrates the use of `MEDIAN()` function.

Formula	Description
`=MEDIAN(B2:B10)`	The formula calculates the median value of the expenditure data within the range, B2:B10. It returns 18920.
`=MEDIAN(1000, 2000, 3000)`	The formula calculates the median, that is, middle value of 1000, 2000, and 3000. It returns 2000.

Table 9.23: Using `MEDIAN()` Function

9.5 Macros

You perform many tasks in Excel worksheets that are repetitive in nature, such as the application of the same formatting styles to a group of data. Macros can combine the steps of a repetitive task into a single action, such as a mouse click or a key combination. You can assign a macro to a toolbar button, a graphic, or a control. You can also assign a shortcut key combination to a macro.

You can create simple macros in MS Excel 2007 environment. For more complex macros, you can write code in VBA (Visual Basic for Applications) programming language. The following sections demonstrate how to create a macro that saves the selected range of a worksheet to a print file, and how to run macros.

V 1.0 © 2008 Aptech Limited

9.5.1 Creating Macros

1. In the worksheet that contain your data, click **Developer** menu, and then select **Record Macro** option from the ribbon. The **Record Macro** dialog box is displayed as shown in figure 9.20.

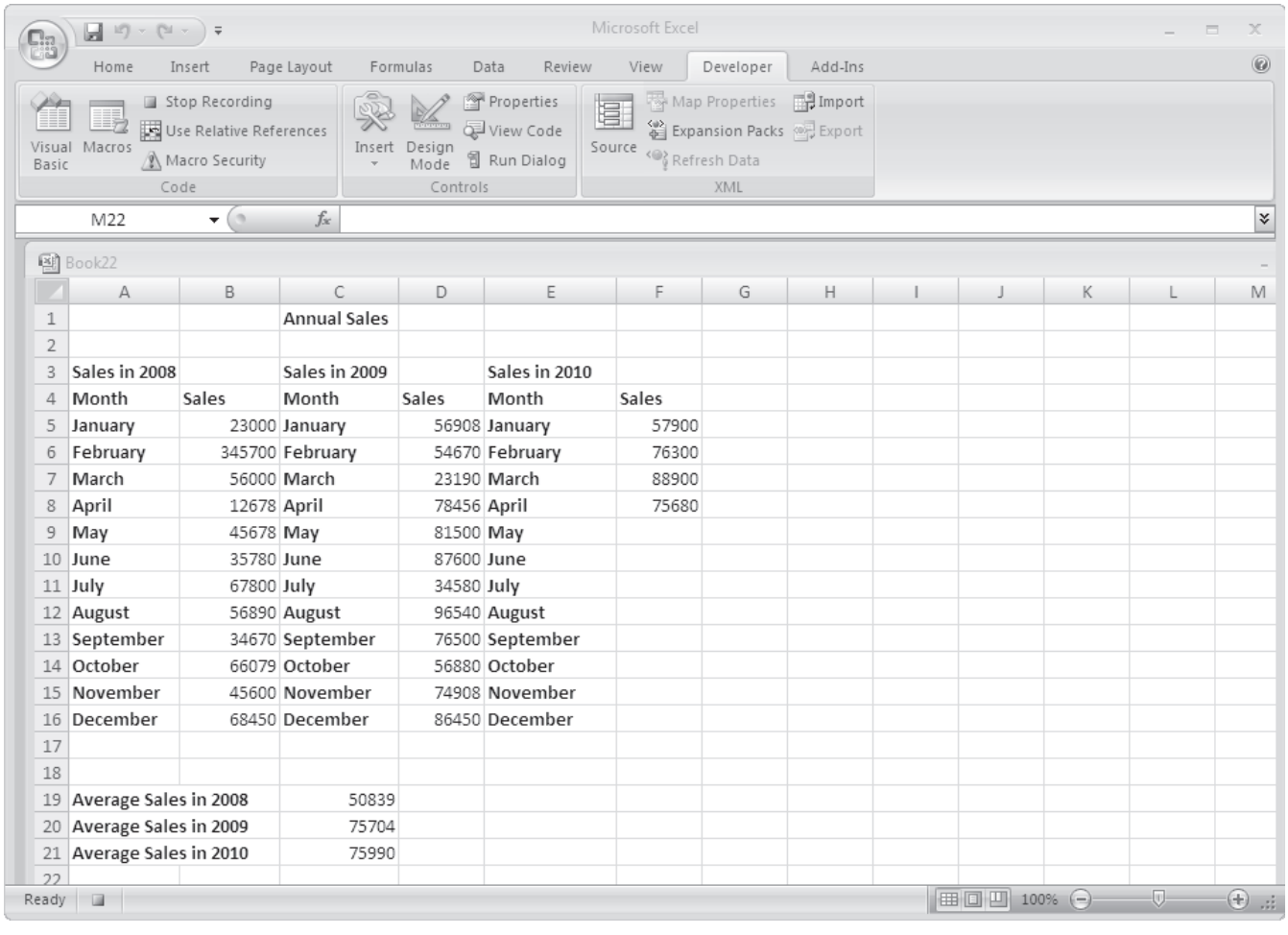


Figure 9.20: Record Macro Dialog Box

2. Type “Print_Selected_Range” in **Macro name:** text box.

3. Specify **Shortcut Key** as “Ctrl + Shift + P”.

4. Select the default “This Workbook” from **Store macro in:** drop-down list.

5. Type “Format cells and Print into a file” in **Description** text area, and click OK.

Now Excel will record all the actions that you perform. Figure 9.21 shows that recording of macros has been started.

Concepts

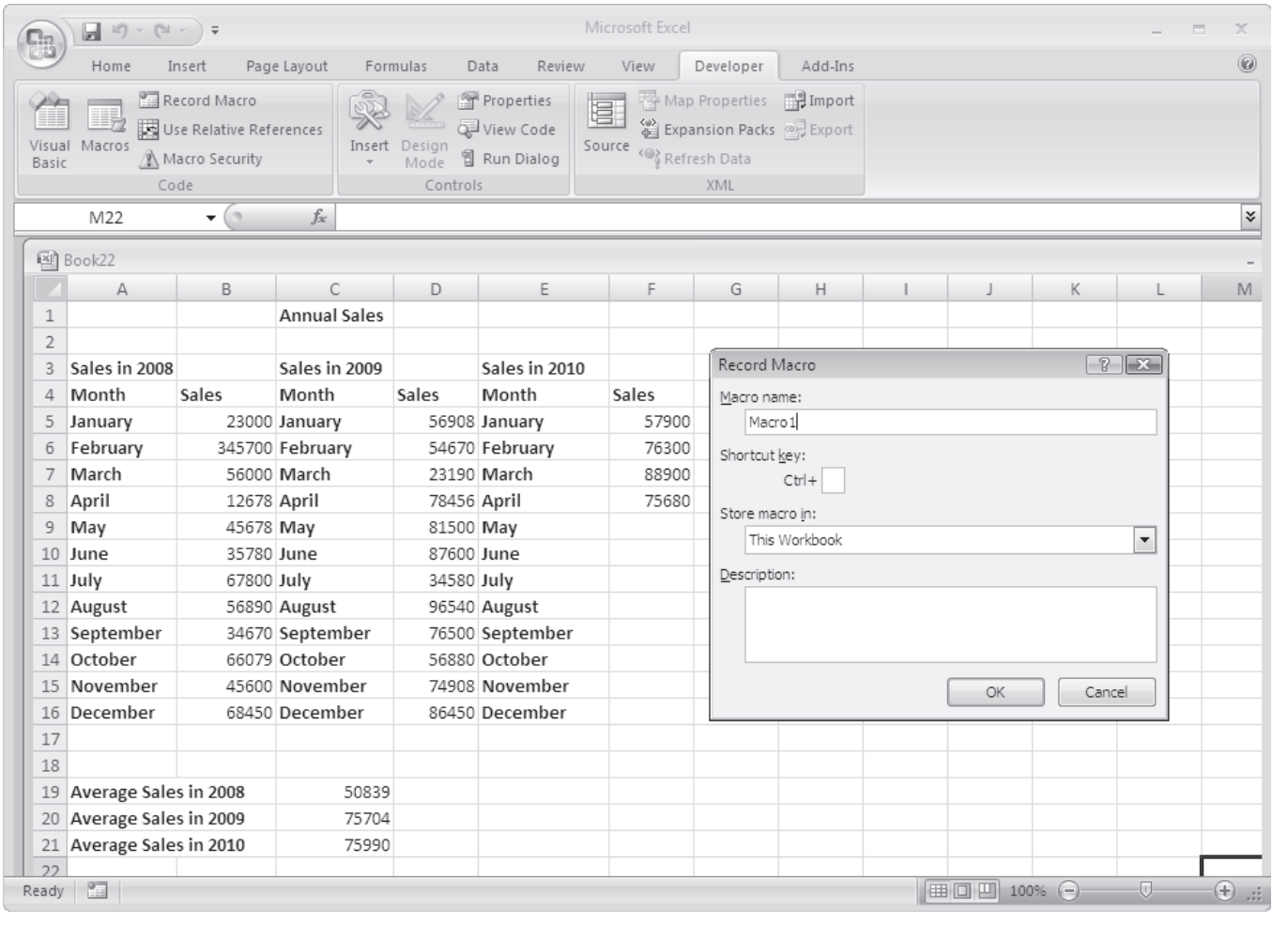


Figure 9.21: Recording of Macros

6. Select the range A19:C21 first.

7. Click **Office Button**, select **Print -> Print** option from the list.

8. Click **Selection** option under **Print What** area, and then select **Print to File** check box. The **Print to File Dialog Box** is displayed.

9. In **Print to File** dialog box, type a file name, and then click **OK**.

10. Click **Stop Recording** option to stop recording after selecting the range for printing to a file.

V 1.0 © 2008 Aptech Limited

9.5.2 Running Macros

1. Load your original worksheet, and in **Developer** ribbon select **Macros** option as shown in figure 9.22. The **Macro** dialog box is displayed.

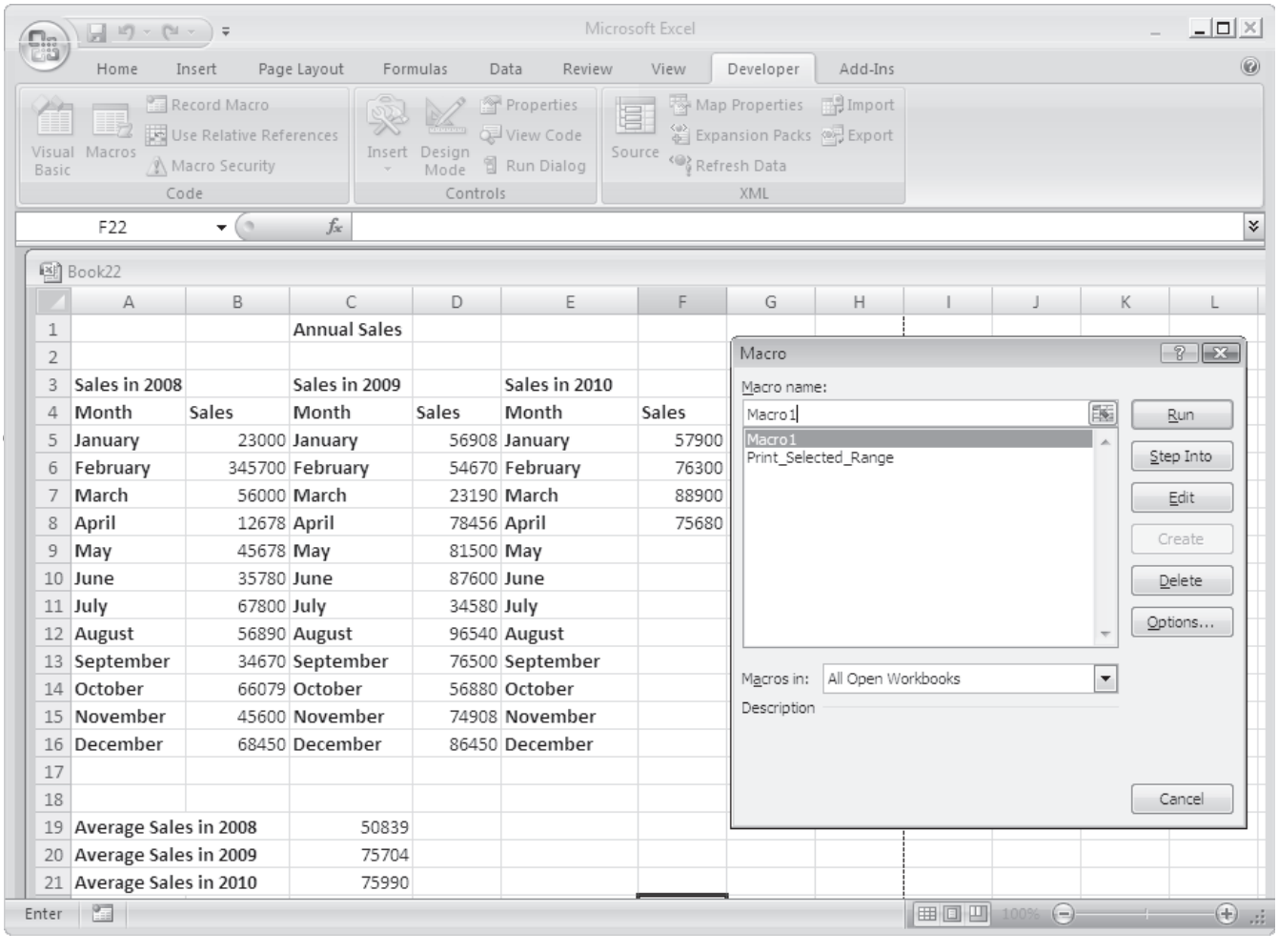


Figure 9.22: Running Macros

2. Select **Print_Selected_Range** from the list of macros, and click **Run** button. The macro is immediately executed, and the Average Sales data in range A19:C21 is selected. After some time, **Print to File** dialog box is displayed as shown in figure 9.23.

3. Type a file name in **Output File Name** text box.

Figure 9.23: Print to File Dialog Box

9.5.3 Macros in VBA

You can create macros using VBA. The following steps demonstrate how to create and run macros in VBA environment.

➢ **Creating macros**

1. Open a worksheet with data in MS Excel 2007 environment as shown in figure 9.24. Here the **Developer** ribbon is selected.

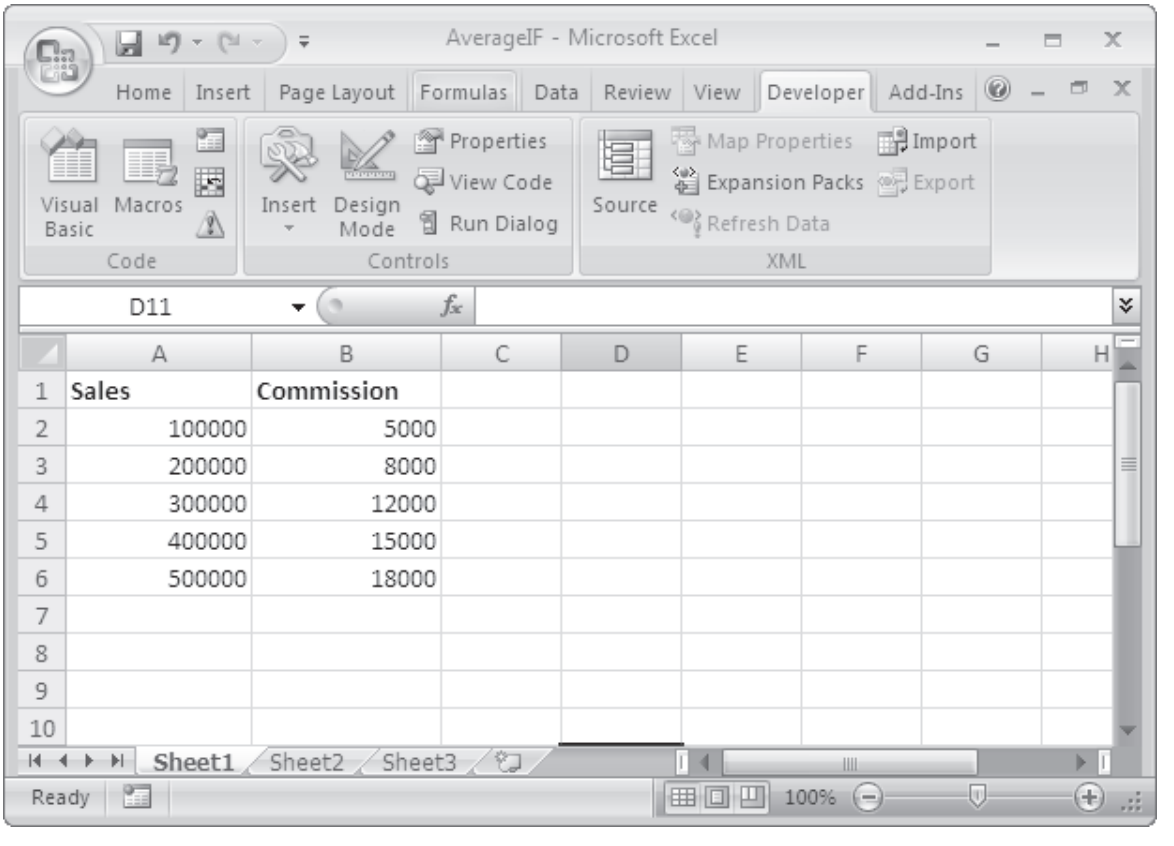


Figure 9.24: Sales Data

2. Click **Visual Basic** toolbar button from **Developer** ribbon. This displays the VBA Editor as shown in figure 9.25.

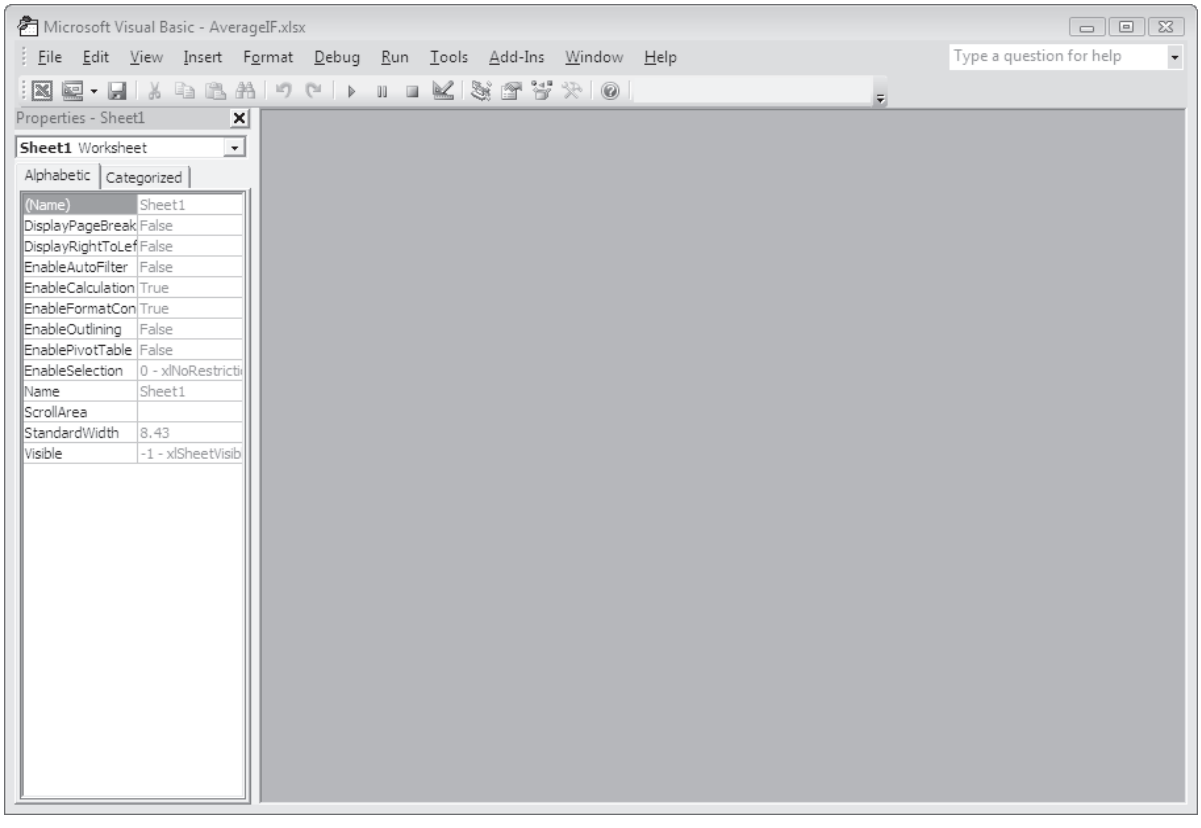


Figure 9.25: VBA Editor

3. Click **Insert** menu, and then select **Module** options. This displays the VBA code module as shown in figure 9.26.

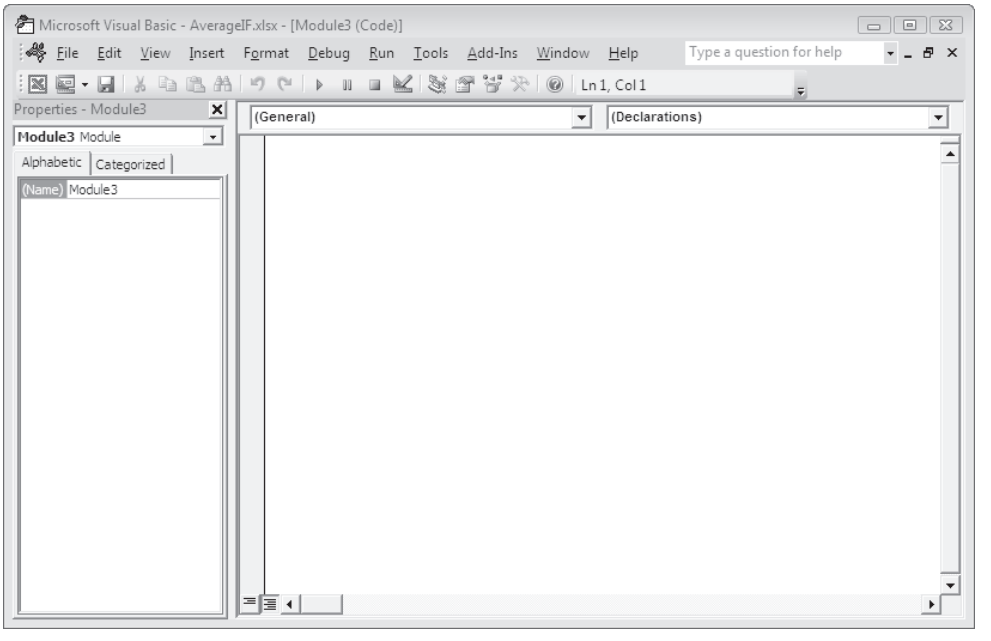


Figure 9.26: VBA Code Module

4. Type the following code in VBA editor.

```
Sub display_footer()
  authorName = InputBox("Enter Author Name", "Add author name to footer")
  With ActiveSheet.PageSetup
     .LeftHeader = "Commission & Sales Data"
     .CenterHeader = ""
     .RightHeader = "Year 2010"
     .LeftFooter = authorName
     .CenterFooter = "@Michgan Limited"
     .RightFooter = "Page 1"
   End With
End Sub
```

The macro will be saved with name "display_footer".

➢ **Running VBA macros**

The following steps demonstrate how to run "display_footer" macro.

1. Click **View Microsoft Excel** toolbar button in VBA editor to display the worksheet as shown in figure 9.27.

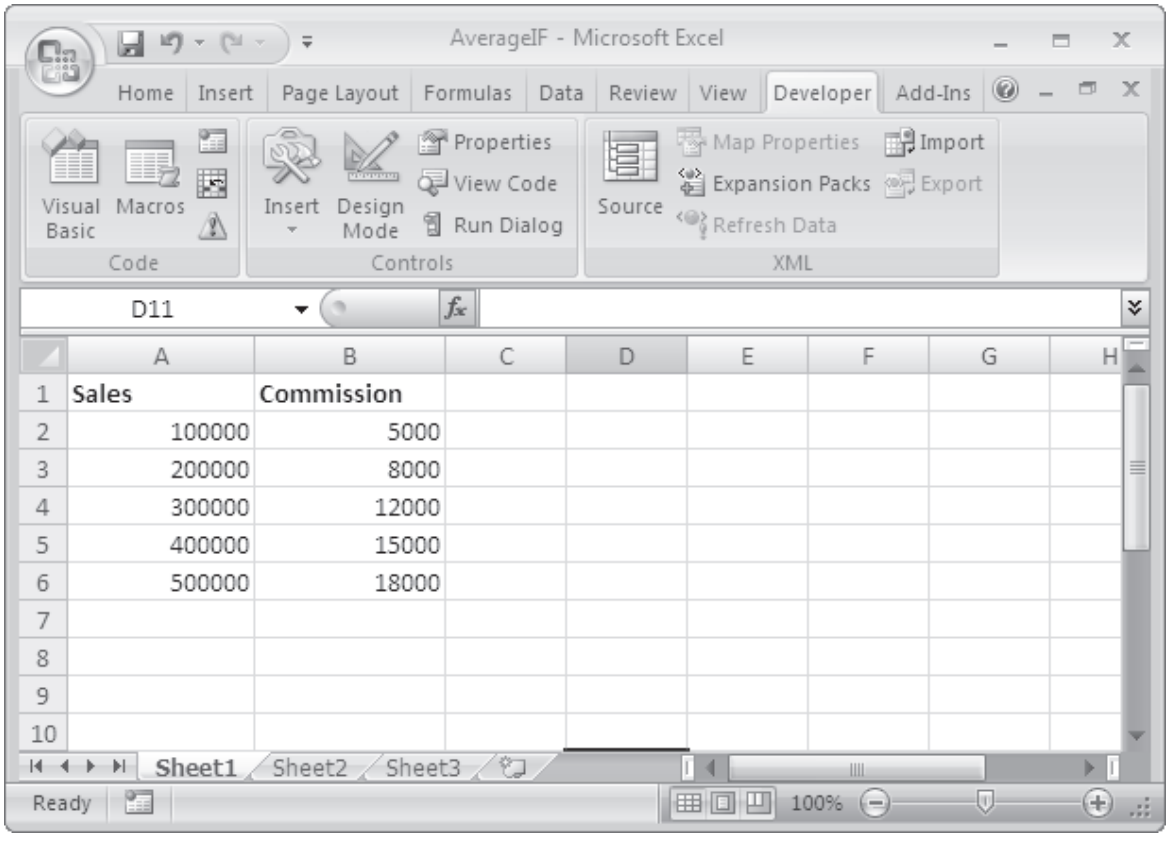


Figure 9.27: Excel Worksheet

2. Click **Macros** option in **Developer** ribbon. This will display **Macro** dialog box for executing a macro as shown in figure 9.28.

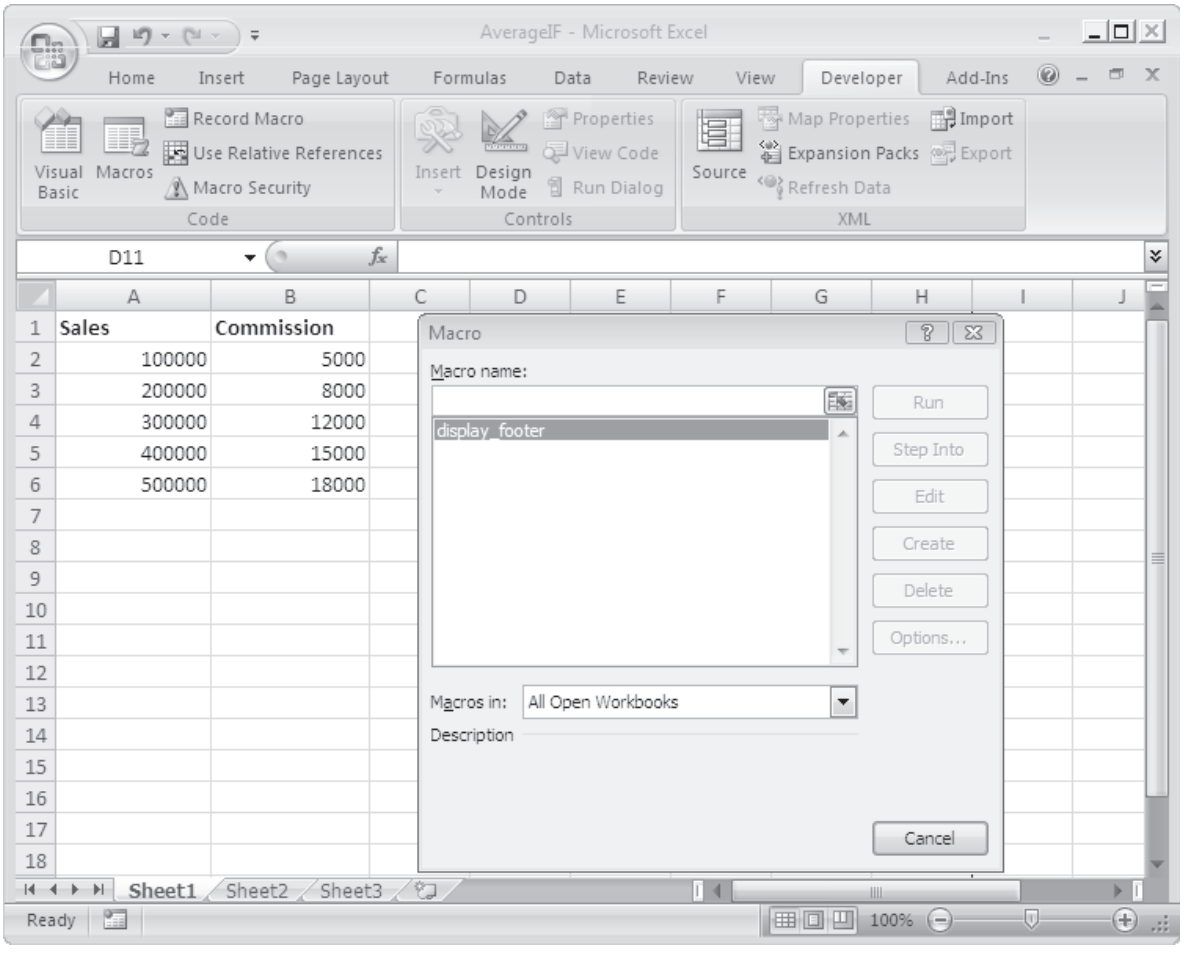


Figure 9.28: Excel Worksheet

3. Select "display_footer" macro, and then click **Run** to execute the macro. This displays the first step in macro, the "Add author name to footer" dialog box as shown in figure 9.29.

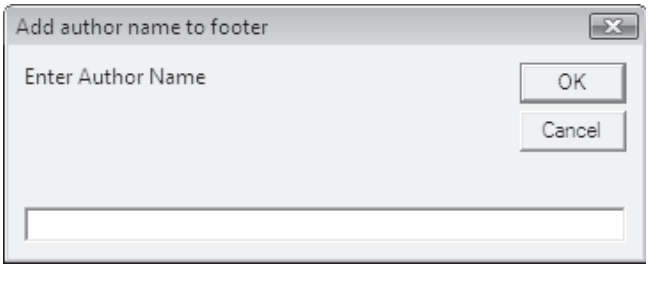


Figure 9.29: Add author name to footer dialog box

4. Type "Patricia Haldane" in the **Enter Author Name** text box, and click **OK**. This will set the other values of header and footer.

5. To view the header and footer of the page, click **Print**, and then select **Print Preview** option. The page is displayed as shown in figure 9.30.

Commission Sales Data Year 2010

Sales	Commission
100000	5000
200000	8000
300000	12000
400000	15000
500000	18000

Patricia Haldane @Michgan Limited Page 1

Figure 9.30: Worksheet with header and footer

9.5.4 Macro Security Issues

Macros offer many benefits. However, an Excel macro can play havoc with your system or your files. Besides, macros can be programmed to run computer viruses. Excel 2007 has introduced a model called **Trust Center**. The Trust Center can be set up to disable macros in any file unless it contains a valid digital signature issued by a reputed Certificate Authority and a trusted source. The center allows you to configure and control all macro settings from a single point.

Figure 9.31 shows the **Trust Center** dialog box that is enabled when you click **Macro Security** button in **Developer** ribbon. To configure default security settings you can select options under **Macro Settings** area. **Trusted Publishers** tab allows you to add a set of publishers with trusted digital signatures. This enables you to run macros from trusted publishers. The **Trusted Locations** tab allows you to specify locations where you can safely stores necessary files that are accessible by macros.

Concepts

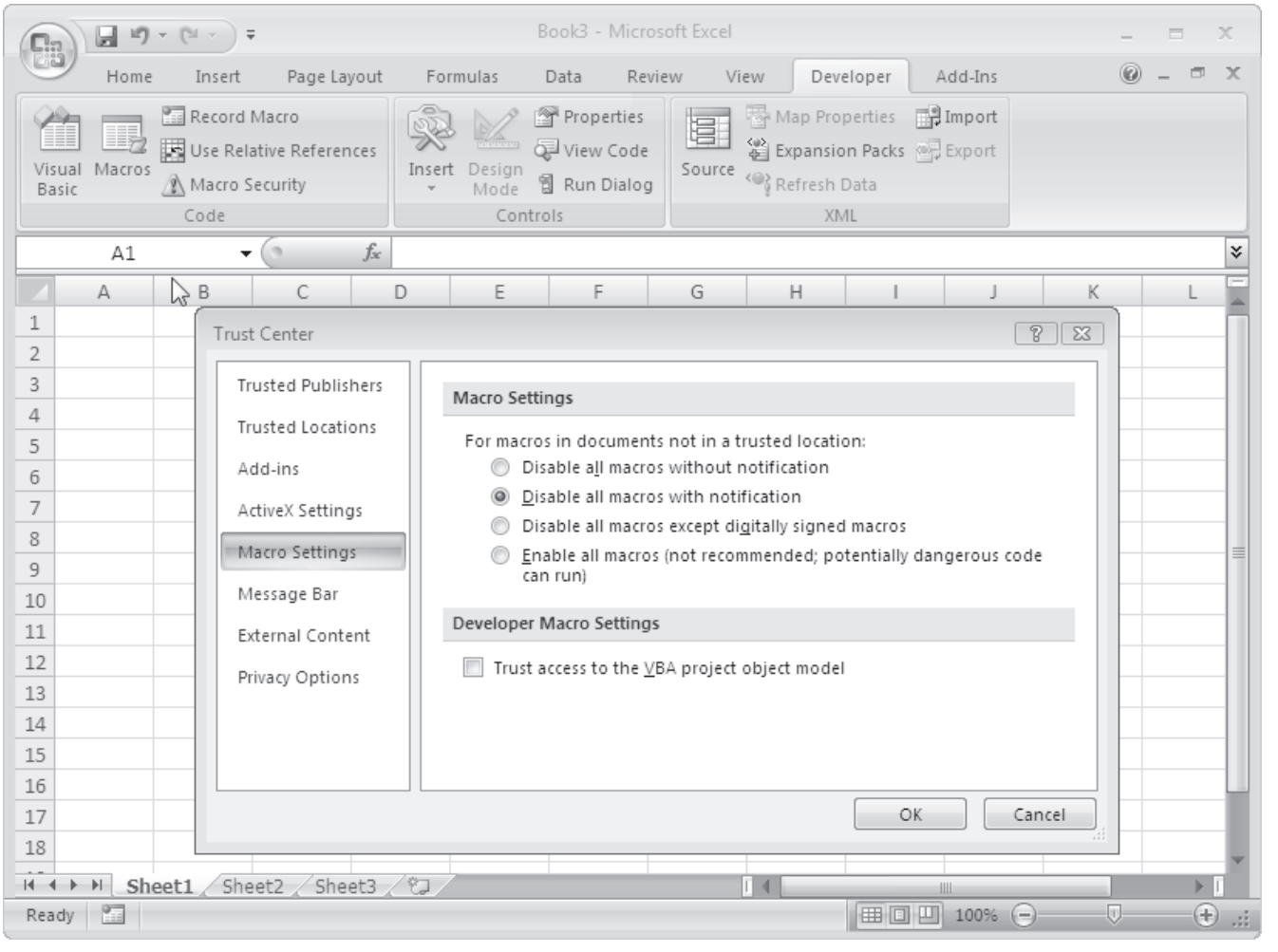


Figure 9.31: Trust Center Dialog Box

V 1.0 © 2008 Aptech Limited

Summary

- Functions in Excel allow you to calculate and perform difficult tasks.
- Functions in Excel can be categorized into lookup, financial, and statistical functions.
- Lookup functions allow you to search for a value in a row or a column within a specified range.
- Financial functions allow you to perform necessary calculations for a loan or an investment.
- MS Excel 2007 provides many powerful financial functions, such as PV(), PMT(), FV() and RATE() that can calculate the present value, the future value, rate of interest, and fixed payment to be made.
- Statistical functions allow you perform complex statistical calculations for a range of data in a spreadsheet.
- Macros allow you to define shortcuts for performing complex and repetitive tasks in a spreadsheet.

Check Your Progress

1. Which of the following function allows you to search for a value in columns within a given range?

   a. HLOOKUP()

   b. VLOOKUP()

   c. PMT()

   d. PV()

2. Which of the following function can calculate payment of a loan over a fixed period of time?

   a. RATE()

   b. PV()

   c. PMT()

   d. FV()

3. Which of the following function returns the rate of interest of an investment?

   a. RATE()

   b. PV()

   c. PMT()

   d. FV()

V 1.0 © 2008 Aptech Limited

Check Your Progress

4. Which of the following helps you to perform repetitive tasks with a shortcut?

   a. HLOOKUP()

   b. PV()

   c. Macros

   d. PMT()

5. Which of the following function allows you to calculate the standard deviation of a population of data?

   a. STDEV()

   b. STDEVP()

   c. MEDIAN()

   d. MODE()

6. Which of the following function returns middle value in an odd set of numbers?

   a. MEAN()

   b. MODE()

   c. MEDIAN()

   d. AVERAGE()

Check Your Progress

7. You want to find out the value of an investment at the end of a period. Which of the following function allows you to do this?

   a. PMT()

   b. PV()

   c. FV()

   d. RATE()

8. Which of the following function allows you to return the position of an item by matching a user-specified value?

   a. MATCH()

   b. INDEX()

   c. OFFSET()

   d. LOOKUP()

V 1.0 © 2008 Aptech Limited

Session

10 Functions and Macros (Lab)

Objectives

At the end of this session, the student will be able to:

- *Apply functions of MS Excel 2007*
- *Apply Macros in MS Excel 2007*

Part I – 60 Minutes

Exercise 1

A reputed product development company maintains the details of orders in a worksheet as shown in figure 10.1.

	A	B	C	D
1	OrderId	ProductId	Unit Price	Quantity
2	1001	1	23.45	120
3	1002	2	45.89	150
4	1003	3	24.67	180
5	1004	4	54.78	190
6	1005	5	34.78	210
7				

Figure 10.1: Order Details

You have been working for the company as an office assistant for years. Your manager asks you to extract some information based on the given sales data. Hence, you need to perform the following tasks:

1. Find the unit price of an order with ID = 1003.
2. Find the quantity ordered for a product with ID = 2.
3. Find the quantity ordered for an order with ID = 1003 and product ID = 3.
4. Find the price for the order with ID = 1001 and product ID = 1.

Analysis:

The given problems can be solved using the `INDEX()` and `MATCH()` functions. You need to pass the required arguments to get the desired result.

Solution:

1. Find the unit price of an order with ID = 1003.

```
=INDEX(A2:D6,3,3)
```

The `INDEX()` function finds the value at row index 3 and column index 3 within the range A2:D6.

2. Find the quantity ordered for a product with ID = 2.

```
=INDEX(A2:D6,2,4)
```

The `INDEX()` function finds the value at row index 2, and column index 4 within the range A2:D6.

3. Find the quantity ordered for an order with ID = 1003 and product ID = 3.

```
=INDEX(D2:D6,MATCH(1003,IF(COUNTIF(B2:B6,"=3"),A2:A6),0))
```

The calculation starts from the innermost formula. The `COUNTIF()` function finds the occurrence of product ID 3. It returns 1. Since the `COUNTIF()` function returns a value, the `IF()` function assigns the range A2:A6 to the `MATCH()` function. The `MATCH()` function returns the position of 1003 within range A2:A6, that is 3. It is passed to the `INDEX()` function. So, the `INDEX()` function returns the corresponding quantity for third position within the range, D2:D6, that is 180.

4. Find the price for the order with ID = 1001 and product ID = 1.

```
=INDEX(C2:C6,MATCH(1001,IF(COUNTIF(B2:B6,"=1"),A2:A6),0))
```

The solution is almost same as of third problem. Here, the price of the item is returned for order ID = 1001 and product ID = 1, that is, 23.45.

Exercise 2

Radiant Global Enterprises is a reputed company with major interest in manufacturing. It has reputed clients worldwide. Every year it has to submit its audited profit and loss statement to its stakeholders. Figure 10.2 shows the audited profit/ loss statement for last three years. The negative numbers indicate loss.

V 1.0 © 2008 Aptech Limited

	A	B	C	D	E	F	G	H	I	J
1				Yearwise Profit/Loss Account(2007-2009)						
2										
3	Year	Month	Profit/Loss	Year	Month	Profit/Loss	Year	Month	Profit/Loss	
4	2007	January	2100	2008	January	-1234	2009	January	2100	
5	2007	February	2300	2008	February	2345	2009	February	2314	
6	2007	March	1945	2008	March	1200	2009	March	1289	
7	2007	April	-1200	2008	April	2100	2009	April	-567	
8	2007	May	-430	2008	May	-654	2009	May	-1245	
9	2007	June	2400	2008	June	-567	2009	June	-2100	
10	2007	July	-2300	2008	July	-3421	2009	July	3426	
11	2007	August	2456	2008	August	1298	2009	August	895	
12	2007	September	1456	2008	September	2318	2009	September	-5624	
13	2007	October	1345	2008	October	1456	2009	October	2300	
14	2007	November	2340	2008	November	1349	2009	November	3427	
15	2007	December	-3421	2008	December	1289	2009	December	-2189	
16										
17										
18										

Figure 10.2: Audited Profit/ Loss Statement

The accounts manager wants you to find out the following financial data:

- Total profit and loss in 2007
- Average profit in 2008
- Number of months in which profit > $2000 in 2009

Analysis:

You can use the `SUMIF()` function to find out the profit and loss of the company. The `SUMIF()` function will use ">0" as criteria for profit, and "<0" as criteria for loss. You will have to use the `AVERAGEIF()` function with the same criteria for calculating average of profit and loss. The `COUNTIF()` function can be used to find out the number of those months in 2009 with profit > $2000.

Solution:

To calculate the total profit in 2007:

1. Select a cell, C18 in current worksheet, and click Insert Function on the Ribbon as shown in Figure 10.2.

Lab Guide

The **Insert Function** dialog box is displayed as shown in figure 10.3.

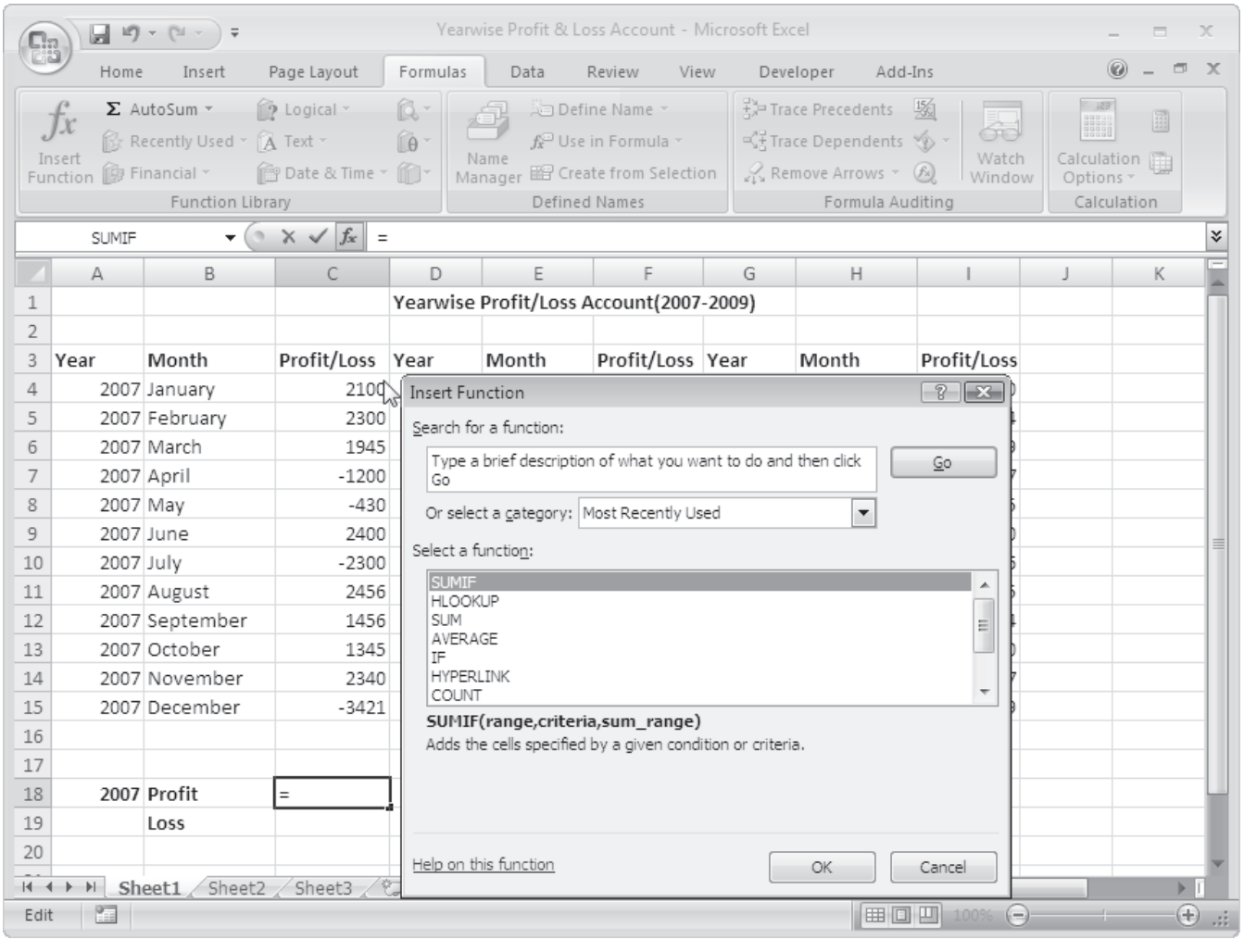
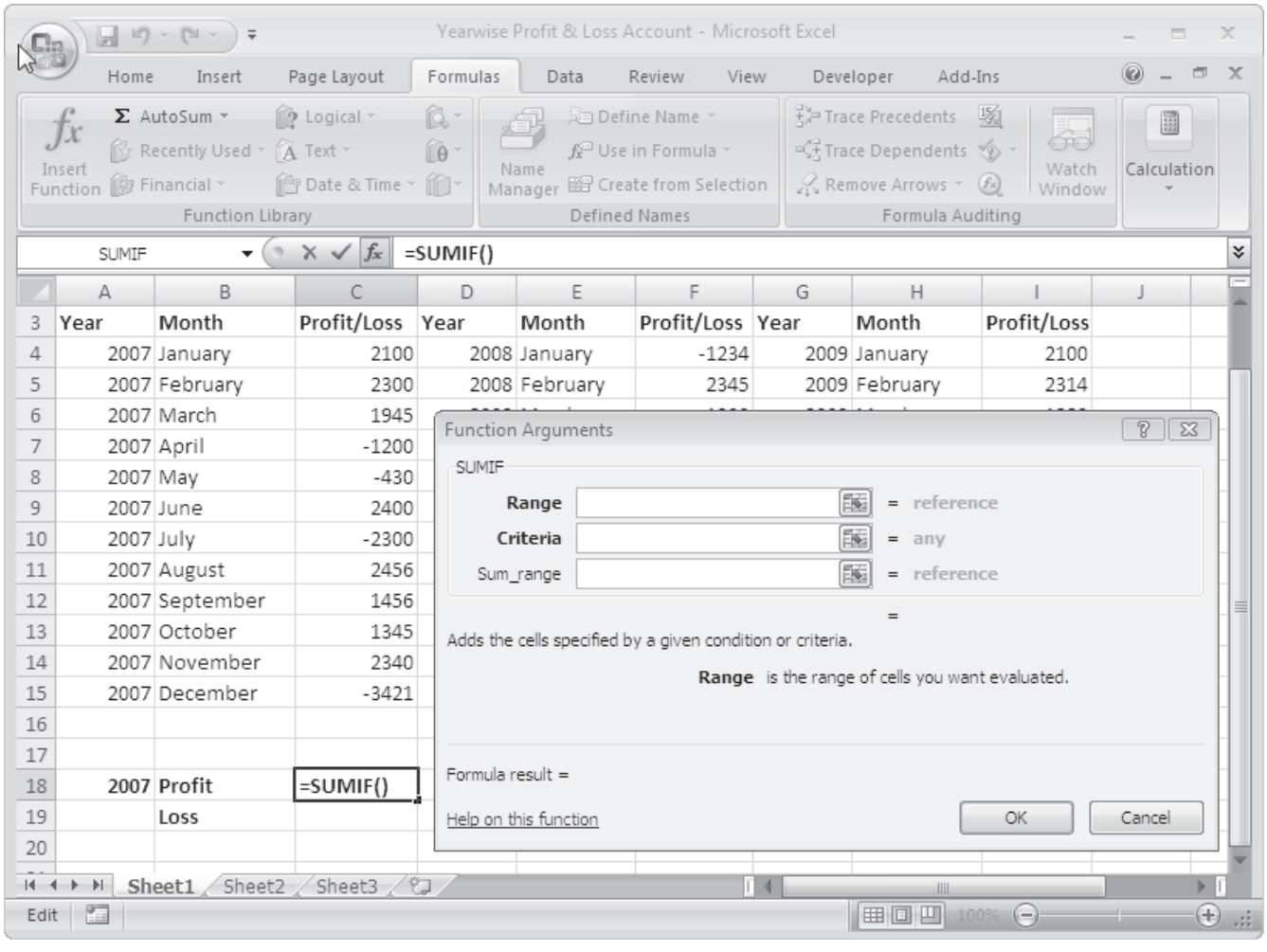


Figure 10.3: Insert Function Dialog Box

2. Select SUMIF from the Select a function: list, and click OK.

The **Function Arguments** dialog box is displayed as shown in figure 10.4.

V 1.0 © 2008 Aptech Limited



Year	Month	Profit/Loss	Year	Month	Profit/Loss	Year	Month	Profit/Loss
2007	January	2100	2008	January	-1234	2009	January	2100
2007	February	2300	2008	February	2345	2009	February	2314
2007	March	1945						
2007	April	-1200						
2007	May	-430						
2007	June	2400						
2007	July	-2300						
2007	August	2456						
2007	September	1456						
2007	October	1345						
2007	November	2340						
2007	December	-3421						
2007	Profit	=SUMIF()						
	Loss							

Figure 10.4: Function Arguments Dialog Box

3. **Select C4:C15 in Range text box for calculating profit/ loss data in 2007.**

4. **Type “>0” in Criteria text box to include only profit data.**

5. **Click OK.**

The profit in 2007 is calculated to be 16342 as shown in figure 10.5.

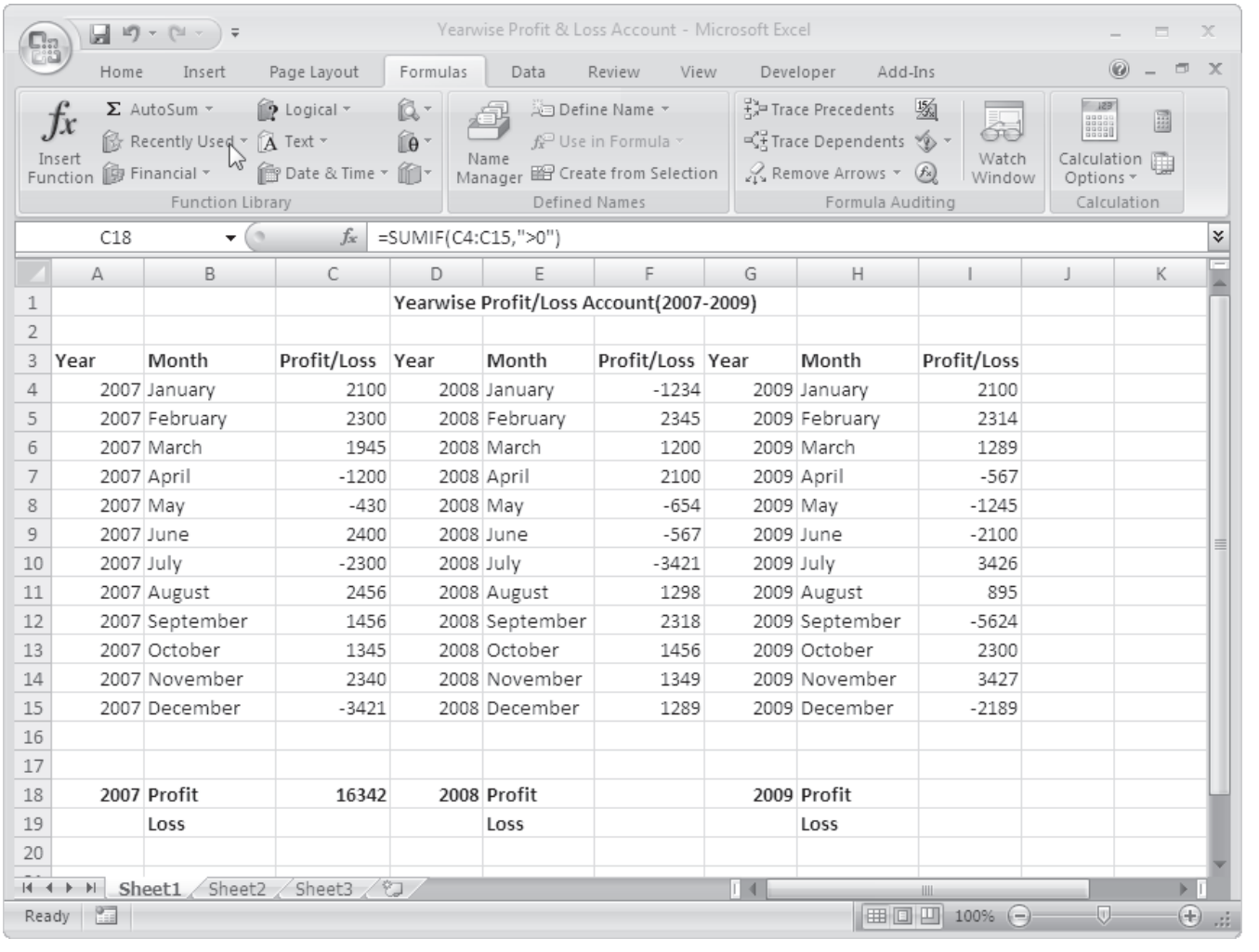


Figure 10.5: Profit in 2007

To calculate the total loss in 2007:

Snippet

The `SUMIF()` function can be used to calculate the total loss for a particular year.

```
=SUMIF(C4:C15, "<0")
```

The total loss in 2007 is calculated to be (-7351).

V 1.0 © 2008 Aptech Limited

To calculate the average profit in 2008:

Snippet

The `AVERAGEIF()` function can be used to calculate the average profit of a company.

```
=AVERAGEIF(F4:F15, ">0")
```

The average profit in 2008 is calculated to 1669.38.

To calculate the number of months with profit > $2000 in 2009:

Snippet

The `COUNTIF()` function is used to calculate the number of months with profit > $2000 in 2009.

```
=COUNTIF(I4:I15, ">2000")
```

The result displayed is 5.

Exercise 3

Problem 1: A person wants to borrow $10,000 from a bank for purchasing a house. He wants to repay the loan in easy installments for 10 months. The bank will charge him at the rate of 8.5% annually. Calculate the amount he will have to pay at the end of every month.

Analysis: The principal amount or the Present Value (PV) is $10,000. The tenure is 10 months. The annual rate of interest is 8.5%. The Future Value (FV) is $0 because the loan amount will be reduced to $0 at the end of the tenure.

Solution:

```
=PMT(8.5%/12, 10, 10000, 0, 0)
```

The monthly payment at the end of every month is calculated to be $1,039.37.

Problem 2: A person is thinking of investing in an insurance policy. The policy needs an initial investment of $60,000. It will return $500 at the beginning of every month for 20 years. The investment will earn 8.5% annually. He wants to find out whether purchasing the policy will be a wise decision.

Analysis: The problem can be solved by calculating the present value of the policy. It determines whether the policy is worth buying. Here, the Future Value is $0.

Solution:

```
=PV(8.5%/12, 240, 500, 0, 1)
```

The Present Value comes out to be $58,023.53. The value is less than the initial investment. So, it is not worth investing in the insurance policy.

Part II 60 Minutes

1. Robert wants to save money for a very special occasion that is going to occur after one year from now. He plans to deposit $200 at the beginning of every month for next 12 months. The annual rate of interest is 8%. How much money will he able to deposit into the account?

   Hint:

   Use `FV()` to solve the problem

2. Mary has set a goal to save $1,00,000 in 20 years by saving a certain amount every month. Assume that the saving will fetch 10% annually. How much will she have to save every month?

   Hint:

   Use `PMT()` to solve the problem. Here, PV = $0 and FV = $1,00,000

3. Consider the following worksheet as shown in figure 10.6.

Book1

	A	B	C	D
1	OrderID	Quantity	Unit Cost	
2	101	105	2.78	
3	102	100	5.78	
4	103	89	6.89	
5	104	78	5.78	
6	105	67	4.89	
7				

Figure 10.6: Order Details

V 1.0 © 2008 Aptech Limited

a. Find the unit cost for an order with OrderId = 103.

Hint:

Use the `VLOOKUP()` function

b. Find the Quantity of an order with OrderId = 105.

Hint:

Use the `VLOOKUP()` function

c. Find the unit cost for an order with OrderId = 104.

Hint:

Use the `OFFSET()` function

d. Find the unit cost for an order with OrderId = 102.

Hint:

Use the `INDEX()` function

e. Find the quantity of an order with OrderId = 101.

Hint:

Use the `LOOKUP()` function

4. Refer to the following spreadsheet as shown in figure 10.7.

	A	B	C	D	E	F
2			Annual Sales Figure			
3	Sales in 2008		Sales in 2009		Sales in 2010	
4	Month	Sales	Month	Sales	Month	Sales
5	January	23000	January	56908	January	57900
6	February	345700	February	54670	February	76300
7	March	56000	March	23190	March	88900
8	April	12678	April	78456	April	75680
9	May	45678	May	81500	May	
10	June	35780	June	87600	June	
11	July	67800	July	34580	July	
12	August	56890	August	96540	August	
13	September	34670	September	76500	September	
14	October	66079	October	56880	October	
15	November	45600	November	74908	November	
16	December	68450	December	86450	December	
17						
18						
19	Average Sales in 2008		71527.08333			
20	Average Sales in 2009		67348.5			
21	Average Sales in 2010		74695			

Figure 10.7: Annual Sales Figure

Create a macro that will format **Annual Sales Figure** in bold, underline with font Times New Roman, 15 points, range A3:F4 with font Courier New, 12 points in bold and underline, range A5:E16 with font verdana, 10 points. Range A19:C21 should be in verdana, 12 points and in bold. The macro also should place rectangular border around range A19:C21.

The macro should open Print Preview dialog box, and from here it should set the Orientation to Landscape. It should then set header and footer with author name, date and time of creation, and organization name.

The macro should then save the spreadsheet with the settings into a print file.

Hint:

Use VBA script for setting header and footer

5. Mr. John wants to make quarterly payments on a personal loan of $10,000 that will be reduced to an amount of $2,000 after 4 years. Rate of Interest is charged annually at a rate of 5% and the fixed payment for the loan must be made at the beginning of each quarter.

   Hint:

   Use **PMT()** to solve the problem

6. The Grow Rich Inc. has issued 15-year no-risk bonds. How much will you need to pay monthly with a maturity value of $10,000,000 with an annual return of 8%?

   Hint:

   Use **PMT()** function

7. How much is $5,000 with a monthly payment $200 worth in 5 years with an annual rate of interest 5.8% that is compounded monthly?

   Hint:

   Use **FV()** function

8. You want to borrow $10,000 for a period of 24 months. At the end of every month, you will have to make a payment of $250. What is the rate of interest?

   Hint:

   Use **RATE()** function

Lab Guide

Do It Yourself

1. Consider monthwise average temperature across places in US as shown in figure 10.8.

	A	B	C	D	E	F
1	Avg Temperature(in F)	Houston	Dallas	New York	Washington	
2	Month					
3	January	45	46.7	36.8	44.5	
4	February	46.5	47.5	38.5	47.3	
5	March	43.7	48.3	39.6	46.4	
6	April	44.7	46.5	43.6	45.6	
7	May	46	44.7	42.5	46.7	
8	June	46	47.4	45.5	44.3	
9	July	43.7	47.3	43.5	46.5	
10	August	47.8	43.8	44.8	45.8	
11	September	46.7	46.5	44.5	46.5	
12	October	45.8	47.3	43.7	47.3	
13	November	38.6	42.4	44.7	45.5	
14	December	32.6	47.5	46.8	43.7	
15						

Figure 10.8: Average Temperature

a. Find the standard deviation of temperature for Houston and New York.

b. Find the standard deviation population of temperature across all cities.

c. Find the average temperature across months in all places.

d. Find cities with temperature more than 46 degrees Fahrenheit for the month of February.

e. Find average temperature for Washington city.

V 1.0 © 2008 Aptech Limited

A

Adjustable Cells

Adjustable cells are cells whose values are changed by the Solver to achieve the desired results.

Analysis ToolPak

The Analysis ToolPak is third party add-in that consists of 19 medium scale statistical analysis tools. These tools can be used to perform statistical functions such as anova, exponential smoothing, t-test, f-test, z-test and so on.

Anova

Analysis of variance is a statistical test that analyses sample data set variances and determines if there is a statistical difference in them.

Answer Report

The Answer report lists the target cell and the changing cells with their original and final values, constraints, and information about the constraints.

Axis Fields

Axis Fields is an area in the PivotTable Field List pane of the PivotChart report. Fields added to this area are displayed as items on the x-axis.

C

Changing Cells

Changing cells display the cell reference for the selected scenario's whose cell values are to be modified.

Column Labels

Column Labels is an area in the PivotTable Field List pane of the PivotTable report. Fields added to this area are displayed as column labels at the top of the report.

Constraints

Constraints refer to limitations imposed on the Adjustable Cells.

Correlation

The Correlation tool computes the correlation coefficients for the given sample data sets.

Correlation coefficients

A correlation coefficient depicts the degree of linear relationship that exists between two variables. The correlation coefficient ranges from +1 to -1. If the variables share a perfect positive linear relationship then there exists a correlation of +1. If the variables share

a perfect negative linear relationship then there exists a correlation of -1. If the variables do not share any linear relationship then there exists a correlation of 0.

Covariance

The Covariance tool computes pair-wise covariance coefficients for two or more sets of data.

Covariance coefficients

A covariance coefficient depicts the degree to which two variables are associated to each other. If an increase or decrease in one variable's value causes a similar increase or decrease in the other variable's value, then it is said that the variables share a positive covariance. On the other hand, if an increase or decrease in one variable's value causes an opposite effect on the other variable, it is said that the variables share a negative covariance.

D

Damping factor

The damping or smoothing factor is associated with the exponential smoothing statistical function. The damping factor represents the error that might have occurred while generating the current set of values. During exponential smoothing, the current values are treated using the damping factor and the resultant values are presented as the predicted or future values.

Data Table

Data Table is a way of viewing different results by altering the value in an input cell that is used in the formula.

Descriptive Statistics

The Descriptive Statistics tool generates a report of univariate statistics for a given set of data.

E

Exponential Smoothing

The Exponential Smoothing tool computes corrected values for a given set of data after smoothing out errors present in the given set of data.

F

Fourier Analysis

The Fourier Analysis tool computes the Discrete Fourier Transform and its inverse for a given set of data.

Frequency domain

A signal or wave is normally represented in the time domain. The signal or wave can also be represented in the frequency domain. In the frequency domain, a signal or wave is plotted on amplitude versus frequency scale.

F-Test

The F-Test tool tests whether the given data set variances differ or not.

G

Goal Seek

The process of determining the value of a cell that results in a specific value returned by the formula.

H

Histogram

The Histogram tool creates a histogram for the given set of data.

Hypothesis

A hypothesis is a tentative assumption made in order to test the validity of an argument. In statistics, the validity of an argument is determined by making two assumptions: first that the argument is wrong and cannot be accepted also referred to as the null hypothesis and second that the argument is correct and can be accepted also referred to the alternate hypothesis. The argument is tested to determine which of the hypothesis can be accepted.

Hypothesized Mean Difference

The hypothesized mean difference factor is associated with the t-Test and z-Test. It is the difference between the mean values of the given data sets; however it reflects an assumption and not the actual value.

L

Labels

Labels are the category fields added to the report.

Legends Filed

Legend Fields is an area in the PivotTable Field List pane of the PivotChart report. Fields added to this area are displayed as data series on the y-axis of the report.

Limits Report

The Limit report lists the target cell and the changing cells with their respective values, lower and upper limits, and target values.

M

Mean

Mean refers to the average of a given set of values.

Moving Average

The Moving Average tool computes the moving averages for a given set of data.

O

Objectives

Objectives refer to the desired result of the problem that the Solver should achieve.

One-input Data Table

A one-input data table is used to see how different values of one variable in one or more formulas will change the results of those formulas in the worksheet.

P

PivotChart Report

PivotChart Report is a graphical representation of a PivotTable report. In other words, it is chart derived from PivotTable report.

PivotTable Report

PivotTable Report is a tool in Excel to summarize, sort, filter, and group data.

R

Random Number Generation

The Random Number Generation tool generates random numbers.

Rank and Percentile

The Rank and Percentile tool computes the rank and percentile for the values of a given set of data.

Regression

The Regression tool performs regression analysis on two sets of data.

Report Filter

The Report Filter is an area in the PivotTable Field List pane of the PivotTable report. The report is filtered using the fields in this area.

Row Labels

Row Labels is an area in the PivotTable Field List pane of the PivotTable report. Fields added to this area are displayed as rows on the left side of the report.

S

Sampling

The Sampling tool generates samples from the given set of data.

Scenario Manager

A scenario is a set of values and formulas that are saved and can be substituted automatically in worksheet cells based on the requirement.

Scenario PivotTable

The Scenario PivotTable report type displays the summary report in the form of a pivot table.

Scenario Summary

The Scenario summary report type displays the summary report in the form of a worksheet outline.

Sensitivity Report

It shows how sensitive the solution is to small changes in the objective function formula or the constraints.

Solver

Solver is a tool that changes the cell value to a specified value, highest value or lowest value for a cell formula.

Statistical significance

In statistics, statistical significance refers to the probability of an argument or event being true or not occurring by chance.

T

Time domain

A signal or wave is normally represented in the time domain. In the time domain, a signal or wave is plotted on amplitude versus time graph.

t-Test

The t-Test tool performs Student's t-Test on two sets of data.

Two-input Data Table

A two-input data table allows you to see the effects of changing the values of two variables (input values) used in a formula.

V

Values

Values is an area in the PivotTable Field List pane of the PivotTable report. Fields added to this area are displayed as a single summarized value in the report.

Variance

Variance represents the average of the squared distances of all the sample values from the mean of that sample.

W

What-If analysis

What-if analysis is the process of changing the values in worksheet cells to see how these changes affects the result of formulas on the formula cells.

Z

z-Test

The z-Test tool performs z-Test on two sets of data with known variances.