Graphical Presentation of Time Series Data

SESSION 5

Structure

5.1 INTRODUCTION

In Lab Session 4, you have learnt how to plot different types of graphs such as frequency bar graph, histogram, frequency polygon and frequency curve for representing frequency distributions. We have also discussed the less than and more than cumulative frequency curves, i.e., ogives for representing less than and more than cumulative frequency distributions, respectively. In Unit 16 of MST-001 (Foundation in Mathematics and Statistics), we have discussed the other graphical methods used for representing the time series data such as line, range and band graphs. You have also learnt how to construct the box plot in Unit 16.

Prerequisite

- Lab Sessions 1 and 3 of MSTL-001 (Basic Statistics Lab).
- Unit 16 of MST-001 (Foundation in Mathematics and Statistics).

In this lab session, we shall explain how to use Excel 2007 to draw three types of graphs, which we use to represent the time series data. These are the line, range and band graphs. We shall also explain how to construct the box plot used for identifying outliers and comparing distributions using Excel 2007.

Objectives

After performing the activities of this session, you should be able to:

- prepare the spreadsheet in MS Excel 2007;
- plot the line, range and band graphs for time series data; and
- construct the box plot.

5.2 PROBLEM DESCRIPTION

1. The data of annual rainfall for the period from 1999 to 2013 are collected to study the rainfall patterns in a state and are given in Table 1.

Table 1: Annual rainfall data

Year	Annual Rainfall (in mm)	Year	Annual Rainfall (in mm)
1999	187.2	2007	173.1
2000	207.9	2008	192.7
2001	171.5	2009	157.3
2002	217.7	2010	244.4
2003	199.5	2011	182.2
2004	185.7	2012	241.5
2005	261.2	2013	232.4
2006	201.5		

Represent the time series data using a suitable graph.

2. The minimum and maximum temperatures during a day are collected for one week. The data are recorded in Table 2.

Table 2: Minimum and maximum temperature data

Days	Maximum Temperature (in ^{O}C)	Minimum Temperature (in ^{O}C)
Monday	40	18
Tuesday	41	20
Wednesday	38	16
Thursday	36	15
Friday	43	21
Saturday	42	20
Sunday	37	17

Represent the time series data using a suitable graph.

3. The production unit of a manufacturing company used five machines to produce glass bottles. The production of each machine for one year is recorded in Table 3.

Table 3: Number of glass bottles produced in one year

Month	Machine I	Machine II	Machine III	Machine IV	Machine V
January	155	135	180	195	165
February	164	144	189	204	174
March	159	139	184	199	169
April	168	148	193	208	178
May	154	134	179	194	164
June	163	143	188	203	173
July	158	138	183	198	168
August	167	147	192	207	177
September	153	133	178	193	163
October	162	142	187	202	172
November	157	137	182	197	167
December	166	146	191	206	176

Plot a suitable graph for this data.

4. A survey was conducted for 30 days to note the number of hours the television was watched in a family. A sample of four families was taken. The number of hours spent in watching TV by the families for 30 days are recorded in Table 4.

Table 4: Number of hours spent in watching TV

Days	Television Viewing Hours			
	Family A	Family B	Family C	Family D
1	5	2	6	3
2	1	2	7	5
3	5	6	6	7
4	2	4	3	3
5	2	3	3	6
6	4	3	2	2
7	2	1	2	5
8	3	2	6	7
9	4	1	5	4
10	5	2	3	6
11	2	3	5	4
12	2	4	4	4
13	5	5	6	7
14	3	6	3	2
15	3	1	4	4
16	3	4	7	6
17	2	3	5	6
18	2	3	2	2
19	5	6	6	8
20	2	3	2	2
21	4	5	5	3
22	4	3	3	5
23	2	4	2	2
24	2	4	5	8
25	2	5	5	3
26	3	4	2	2
27	3	5	5	5
28	5	6	4	7
29	2	5	6	2
30	4	2	2	7

Draw the box plot to compare the hours of TV watched by the families for the data of Table 4.

5.3 LINE GRAPH

In Unit 16 of MST-001, you have learnt how to present time series data using the line graph. We use line graph to represent the values of the variables for different time periods. Here we briefly mention the main steps as follows:

Step 1: We take the time (year, months, weeks, days, etc.) on the horizontal axis and the observed data on the vertical axis.

Step 2: We plot the points against the magnitudes of the observed data corresponding to each successive time period.

Step 3: We join these points by line segments.

Steps in Excel

You have learnt the manual plotting of the line graph in Unit 16 of MST-001. In Problem 1, the annual rainfall data are given for 15 years. We can represent the data of Table 1 by the line graph. Here we describe the procedure of plotting the line graph in MS Excel 2007 as follows:

Step 1: We enter the data given in Table 1 in the Excel 2007 spreadsheet as shown in Fig. 5.1.

	A	B
1	Year	Annual Rainfall (in mm)
2	1999	187.2
3	2000	207.9
4	2001	171.5
5	2002	217.7
6	2003	199.5
7	2004	185.7
8	2005	261.2
9	2006	201.5
10	2007	173.1

Fig. 5.1: Partial screenshot of the spreadsheet for the given data.

Step 2: To obtain the line graph, we refer to Fig. 5.2. We

1. select the data given in Cells B2:B16,
2. click on the ***Insert*** tab,
3. click on the ***Line*** option in the ***Charts*** group, and
4. select a chart subtype ***Line with Markers*** that we wish to use.

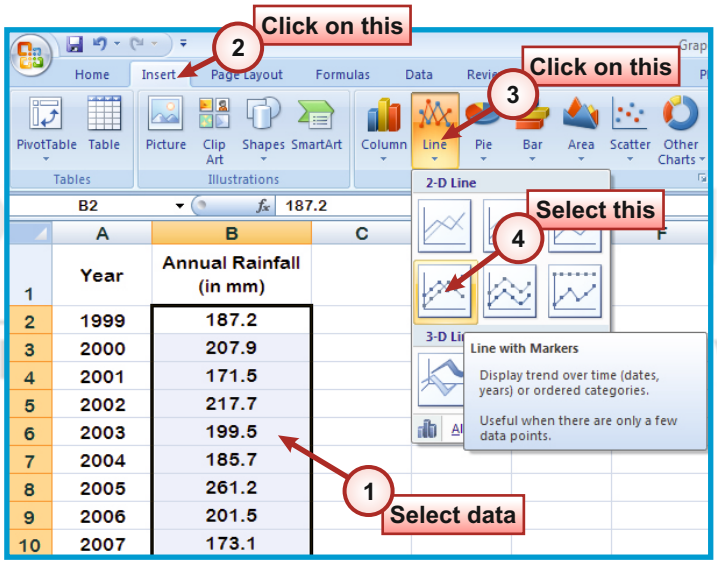


Fig. 5.2

Step 3: When we click on the chart subtype, we get the chart shown in Fig. 5.3. It is called the line graph.

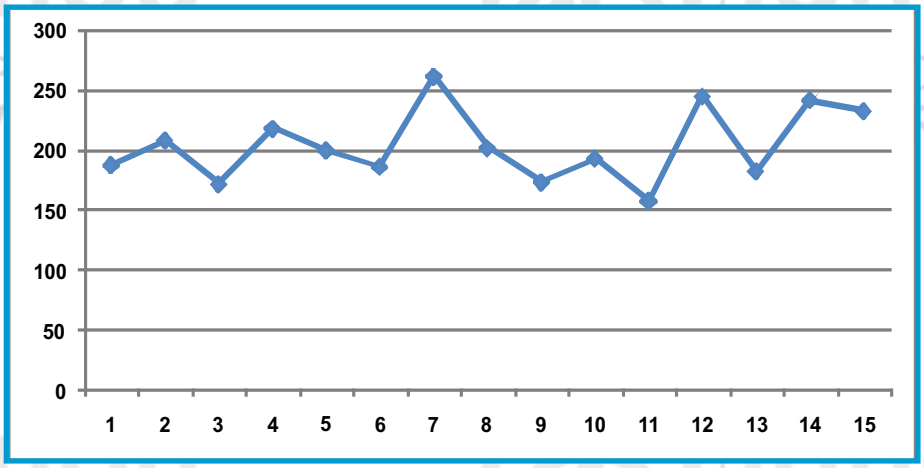


Fig. 5.3

Step 4: We format the chart as explained in Steps 4 to 9 of Sec. 3.3 of Lab Session 3. The resulting line graph is shown in Fig. 5.4.

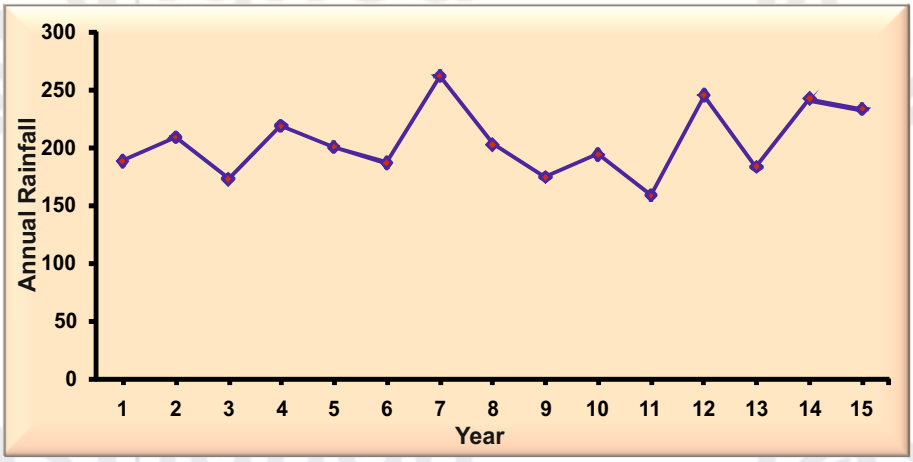


Fig. 5.4

Step 5: We now change the labels of the horizontal axis to years ranging from 1999 to 2013. For this, we follow the steps shown in Fig. 5.5 and

1. select the horizontal axis labels of the chart,
2. click on the ***Design*** tab under ***Chart Tools***, and
3. click on ***Select Data***.

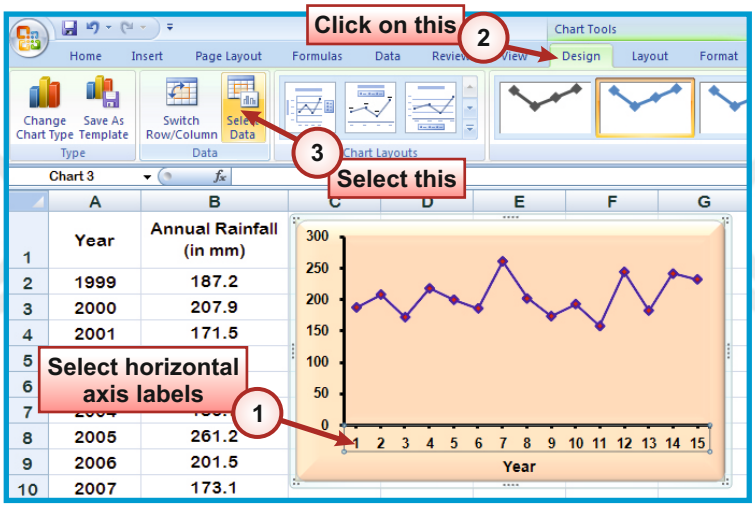


Fig. 5.5

Step 6: A new dialog box as shown in Fig. 5.6 appears. We click on ***Edit*** under ***Horizontal (Category) Axis Labels*** as shown in Fig 5.6.

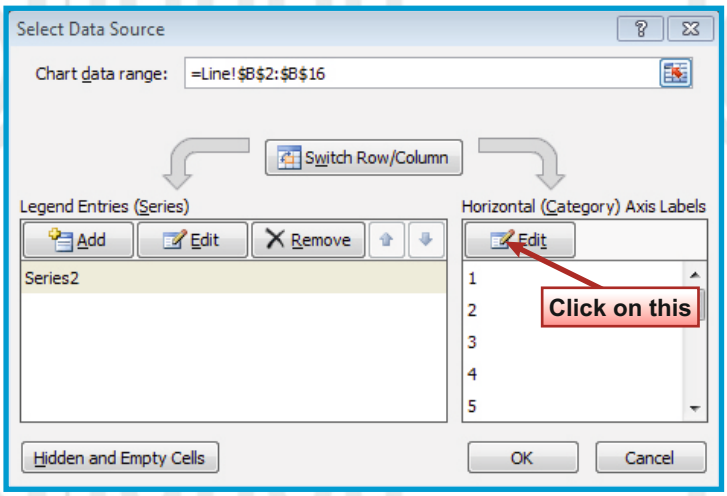
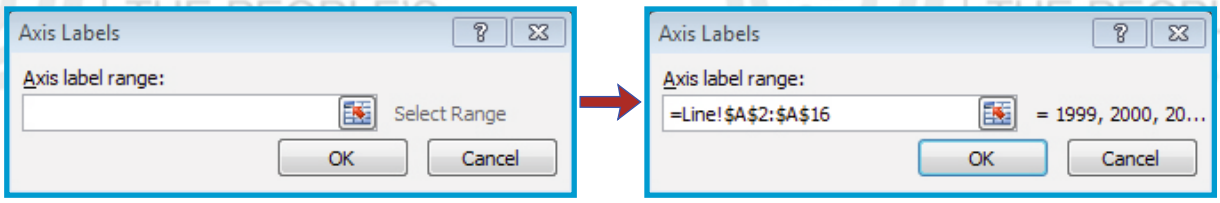


Fig. 5.6

Step 7: Again a new dialog box as shown in Fig. 5.7a appears. We select Cells A2:A16 as ***Axis label range*** as shown in Fig. 5.7b and click on ***OK***.



(a) (b)

Fig. 5.7

Step 8: Now the dialog box shown in Fig. 5.6 changes to the one shown in Fig 5.8 and we click on ***OK***.

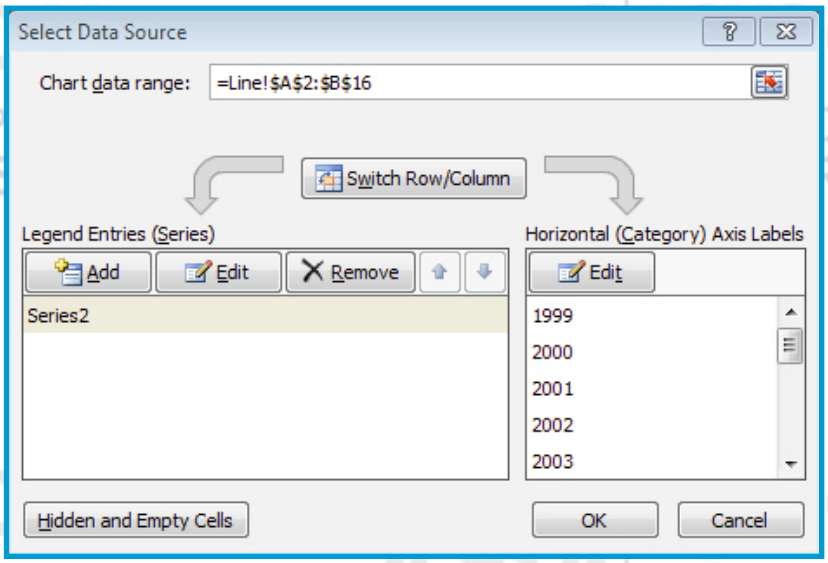


Fig. 5.8

Step 9: The graph with years on the horizontal axis is shown in Fig. 5.9.

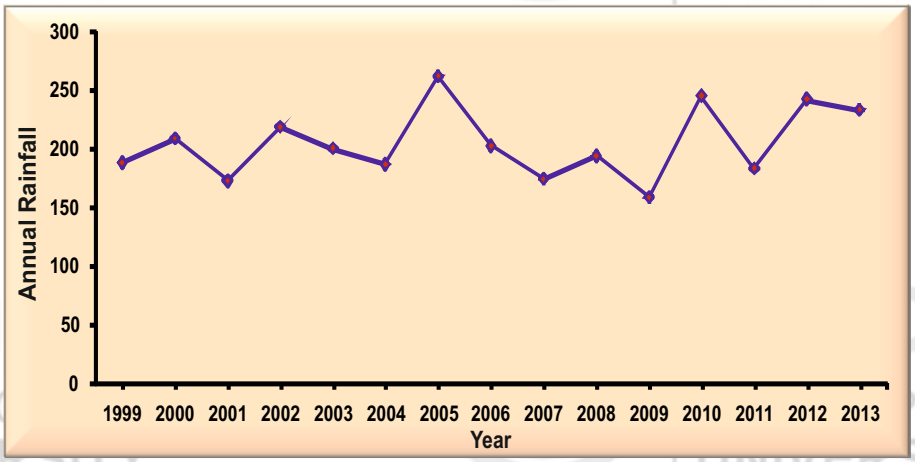


Fig. 5.9

5.4 RANGE GRAPH

In Unit 16 of MST-001, you have learnt how to present data using the range graph. We use the range graph to show the variation and fluctuation between the maximum and minimum values of a variable at the same point of time. For example, the minimum and maximum price of a product, marks of a student, temperature, etc. Here we briefly mention the main steps as follows:

Step 1: We take time on the horizontal axis and corresponding value of the given variable on the vertical axis.

Step 2: We draw two line graphs for the minimum and maximum values of data.

Step 3: The gap between both the curves represents the range of variation. We draw the vertical lines joining two line graphs.

Steps in Excel

You have learnt the manual plotting of range graph in Unit 16 of MST-001. In Problem 2, we can show the variation of the minimum and maximum temperatures using the range graph. Here we describe the procedure of plotting the range graph in MS Excel 2007 as follows:

Step 1: We enter the data given in Table 2 in Excel 2007 spreadsheet as shown in Fig. 5.10.

	A	B	C
1	Days	Minimum Temperature	Maximum Temperature
2	Monday	40	18
3	Tuesday	41	20
4	Wednesday	38	16
5	Thursday	36	15
6	Friday	43	21
7	Saturday	42	20
8	Sunday	37	17
9			

Fig. 5.10: Partial screenshot of the spreadsheet for the given data.

Step 2: To obtain the line graph, we refer to Fig. 5.11. We

1. select the data given in Cells B1:C8,
2. click on the ***Insert*** tab,
3. click on the ***Line*** option in the ***Charts*** group, and
4. select a chart subtype ***Line with Markers*** that we wish to use.

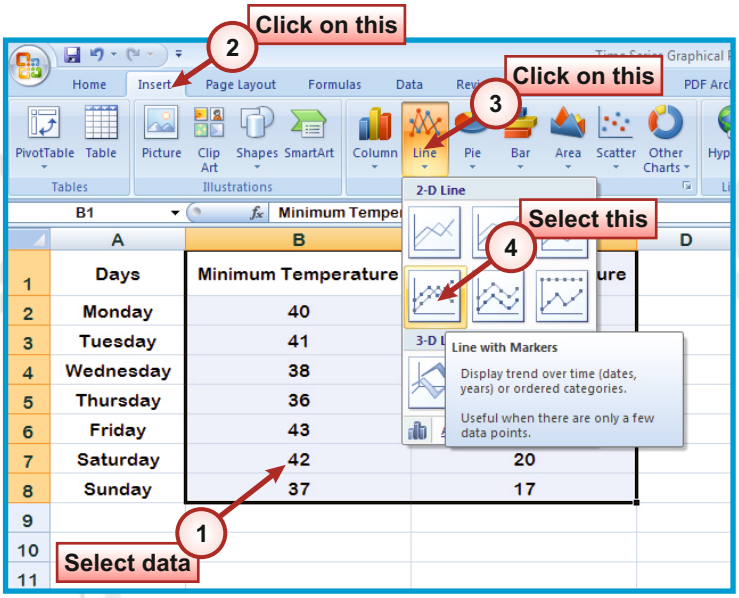


Fig. 5.11

Step 3: The resulting line graph is shown in Fig. 5.12.

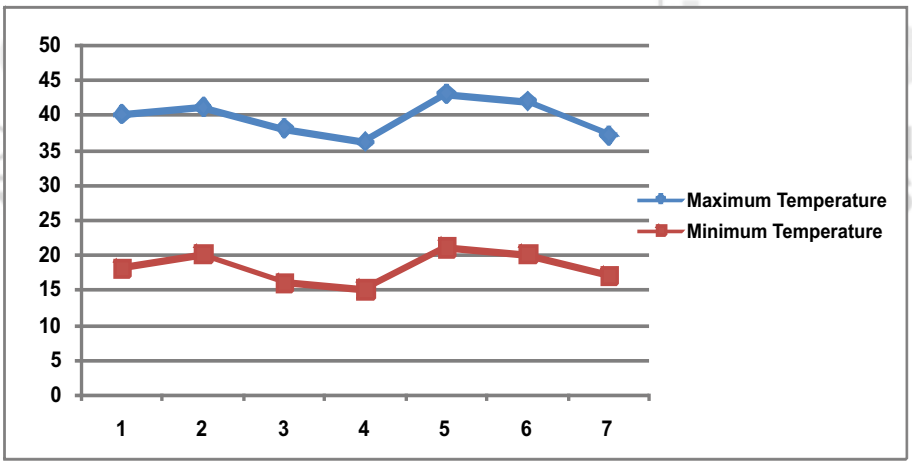


Fig. 5.12

Step 4: We format the graph as explained in Steps 4 to 9 of Sec. 3.3 of Lab Session 3. The resulting graph is shown in Fig. 5.13.

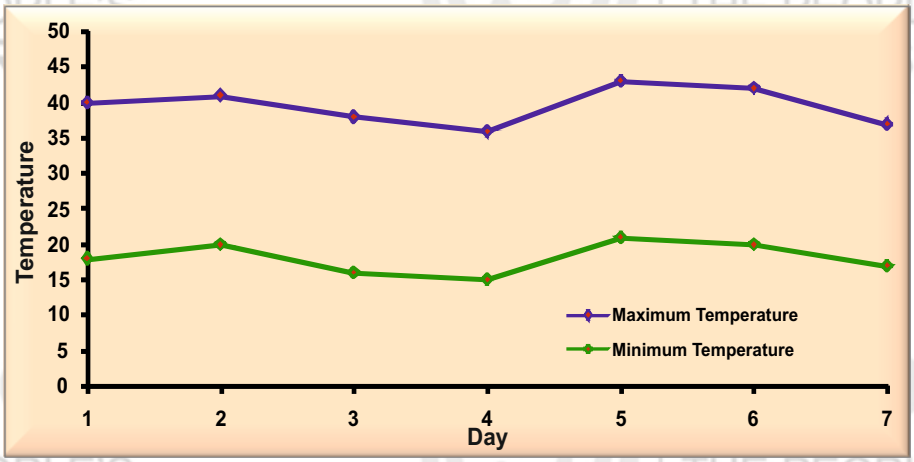


Fig. 5.13

Step 5: We now change the labels of the horizontal axis to days ranging from Monday to Sunday as explained in Steps 5 to 8 of Sec. 5.3. The resulting graph is shown in Fig. 5.14.

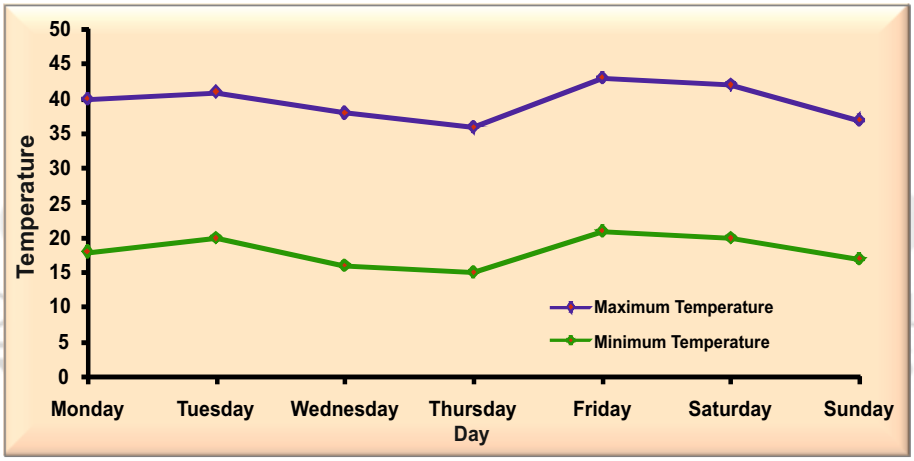


Fig. 5.14

Step 6: To change the look of the line graph shown in Fig. 5.14 as in the range graph, we follow the steps shown in Fig. 5.15 and

1. click on the chart,
2. click on the ***Layout*** tab under ***Chart Tools***, and
3. choose ***High-Low Lines*** under ***Lines*** option.

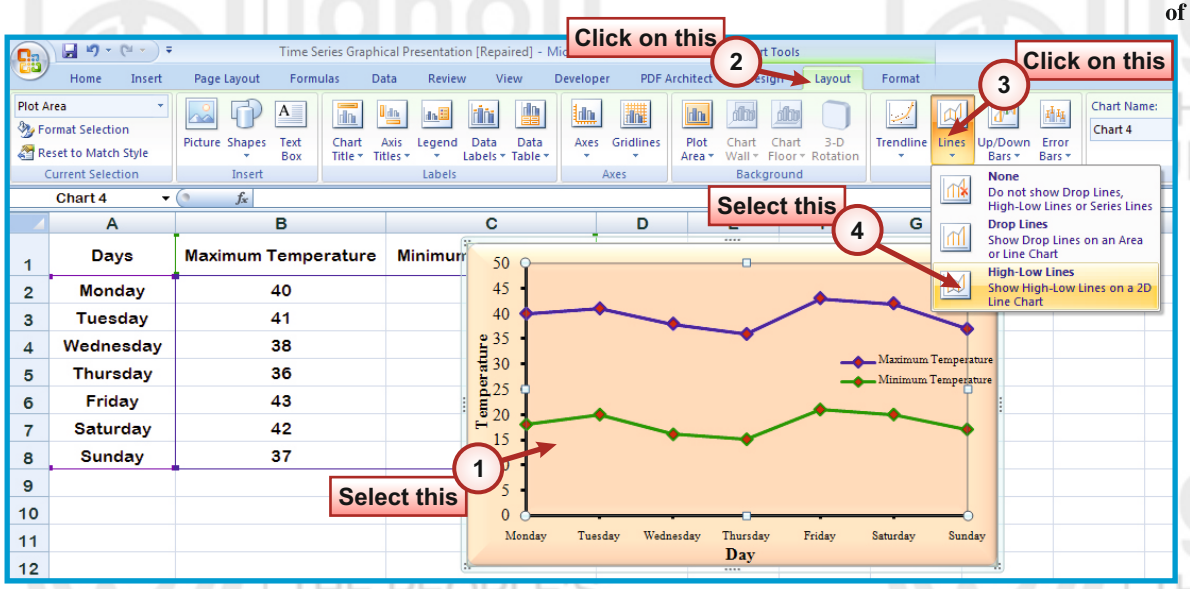


Fig. 5.15

Step 7: The resulting range graph is shown in Fig. 5.16.

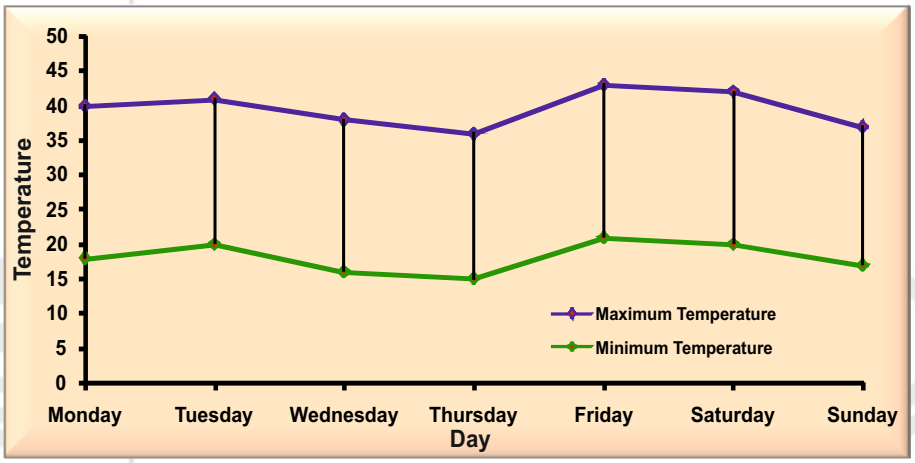


Fig. 5.16

5.5 BAND GRAPH

In Unit 16 of MST-001, you have learnt how to present time series data using the band graph. We use the band graph to show the total components, which may be broken up into sub divisions for each component of the total. We plot the different components of the total as line graphs, one over the other in the form of a series of bands. Here we briefly mention the main steps for drawing the band graph as follows:

Step 1: We take time on the horizontal (X) axis and the variable under consideration on the vertical (Y) axis.

Step 2: We plot the points for different time periods corresponding to the first component since here we have more than one component.

Step 3: We join the points by line segments.

Step 4: We now add the values of the first and second components. We plot these values on the graph and join them by line segments.

Step 5: For plotting the third component, we add the first, second and third components and plot them.

Step 6: We repeat this procedure for all components.

Step 7: We differentiate the gaps between the successive lines using different shades, colours, etc., so that it looks like a series of bands.

Steps in Excel

You have learnt the manual plotting of band graph in Unit 16 of MST-001. In Problem 3, we can show the number of glass bottles produced by five machines for different time periods using the band graph. Here we describe the procedure of plotting the band graph in MS Excel 2007 as follows:

Step 1: We enter the data given in Table 3 in Excel 2007 spreadsheet as shown in Fig. 5.17.

	A	B	C	D	E	F
1	Month	Machine I	Machine II	Machine III	Machine IV	Machine V
2	January	155	135	180	195	165
3	February	164	144	189	204	174
4	March	159	139	184	199	169
5	April	168	148	193	208	178
6	May	154	134	179	194	164
7	June	163	143	188	203	173
8	July	158	138	183	198	168
9	August	167	147	192	207	177
10	September	153	133	178	193	163
11	October	162	142	187	202	172
12	November	157	137	182	197	167
13	December	166	146	191	206	176

Fig. 5.17: Partial screenshot of the spreadsheet for the given data.

Step 2: To obtain the band graph, we follow the steps shown in Fig. 5.18. We

1. select the data given in Cells A1:F13,
2. click on the ***Insert*** tab,
3. click on the ***Area*** option in the ***Charts*** group, and
4. select a second chart subtype ***Stacked Area*** that we wish to use.

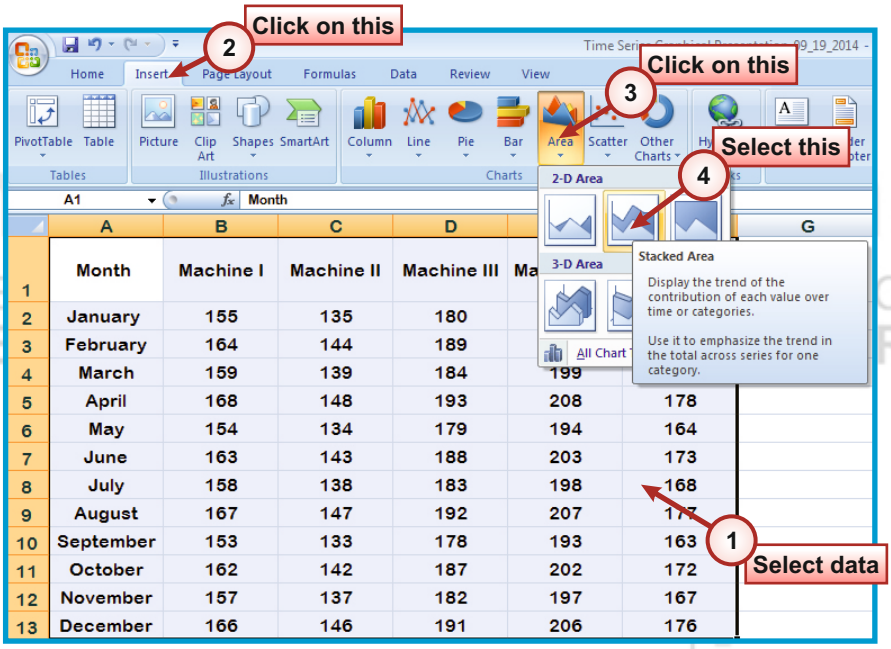


Fig. 5.18

Step 3: The resulting band graph is shown in Fig. 5.19.

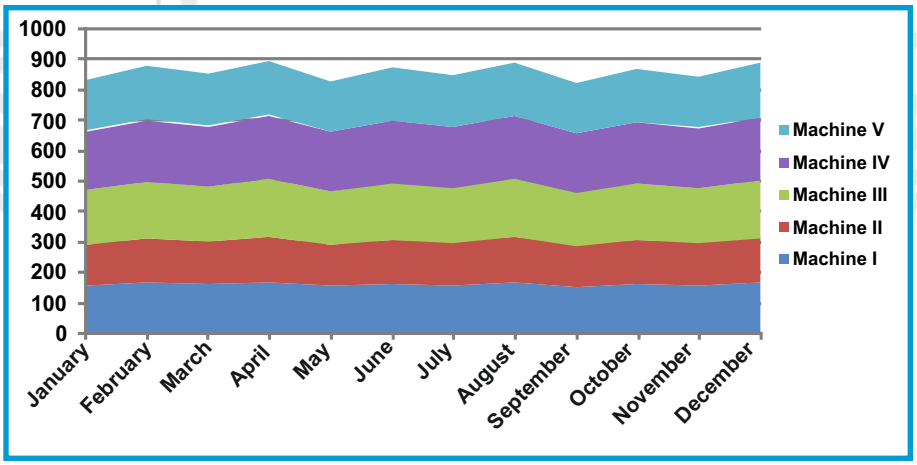


Fig. 5.19

Step 4: We format the graph as explained in Steps 4 to 9 of Sec. 3.3 of Lab Session 3. The resulting graph is shown in Fig. 5.20.

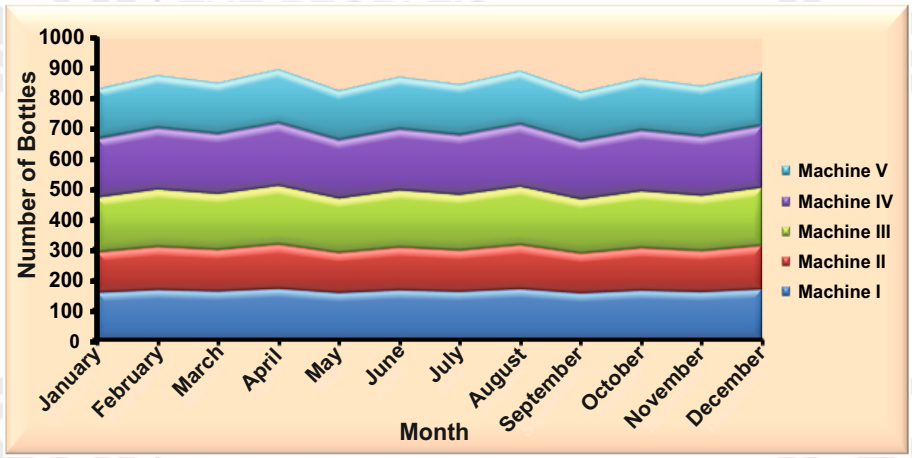


Fig. 5.20

5.6 BOX PLOT

In Unit 16 of MST-001, we have explained the box plot in detail. The box plot is also known as the box-and-whisker plot. We use the box plot to represent the distribution of the data visually. It helps in identifying outliers and comparing distributions. It is more useful when we have a large number of observations and also when we need to compare two or more data sets. The box plot provides a graphical representation of the data through its five number summary, i.e., minimum value, maximum value, median, and the first and third quartiles. Here we briefly mention the main steps as follows:

Step 1: We arrange the data in ascending order.

Step 2: We determine the minimum and maximum value of the data.

Step 3: We also compute the median, the first and third quartiles of the data.

Step 4: We plot the values computed in Steps 2 and 3 on the graph.

Step 5: We draw the box around the median taking the end points at the first and third quartiles.

Step 6: We draw the line from the first quartile to the minimum value and also from third quartile to the maximum value. These lines are called whiskers.

Steps in Excel

You have learnt the manual construction of box plot in Unit 16 of MST-001. Here we describe the procedure of constructing the box plot in MS Excel 2007 as follows:

Step 1: We enter the data given in Table 4 in Excel 2007 spreadsheet as shown in Fig. 5.21.

	A	B	C	D	E
1		Television Viewing Hours			
2	Days	Family A	Family B	Family C	Family D
3	1	5	2	6	3
4	2	1	2	7	5
5	3	5	6	6	7
6	4	2	4	3	3

Fig. 5.21: Partial screenshot of the spreadsheet for the given data.

Step 2: For constructing the box plot, we type "Third Quartile", "Maximum Value", "Minimum Value", "Median" and "First Quartile" in Cells A34:A38 without changing the sequence as shown in Fig. 5.22.

	A
33	
34	Third Quartile
35	Maximum Value
36	Minimum Value
37	Median
38	First Quartile

Fig. 5.22

Step 3: To compute the third quartile for the given data, we

1. select Cell B34,
2. click on ***More Functions*** under the ***Formulas*** tab, and
3. choose the ***Quartile*** function under ***Statistical*** as shown in Fig. 5.23.

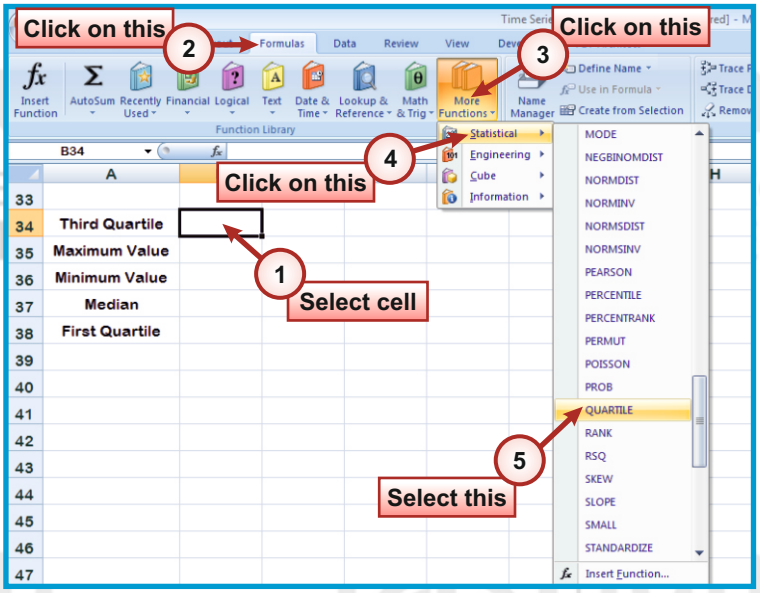


Fig. 5.23

Step 4: When we choose the ***Quartile*** function, the dialog box shown in Fig. 5.24 appears.

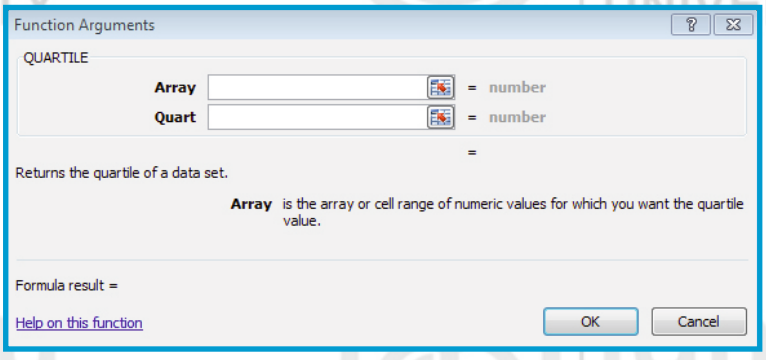


Fig. 5.24

Step 5: We select Cells B3:B32 as ***Array*** and type "3" in ***Quart*** as shown in Fig. 5.25. We click on ***OK*** to obtain the result.

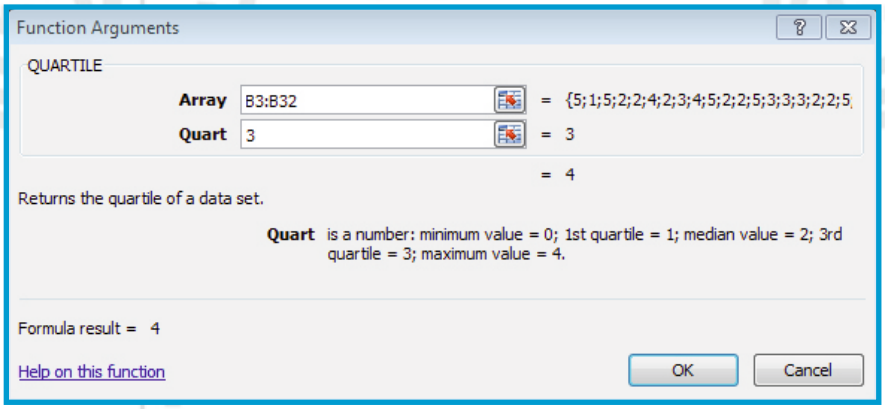


Fig. 5.25

Step 6: We obtain the value of the third quartile, i.e., 4, in Cell B34 as shown in Fig. 5.26.

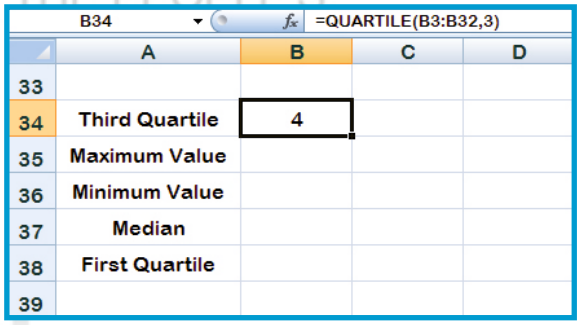


Fig. 5.26

Step 7: After selecting Cell B35, we repeat Step 3 to compute the maximum value. We select Cells B3:B32 as ***Array***, type "4" in ***Quart*** and click on ***OK*** as shown in Fig. 5.27.

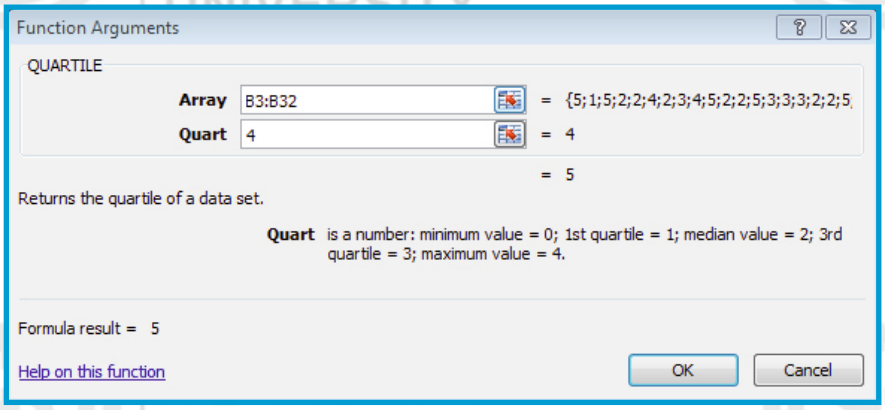


Fig. 5.27

Step 8: The maximum value of the given data is shown in Cell B35 (Fig. 5.28).

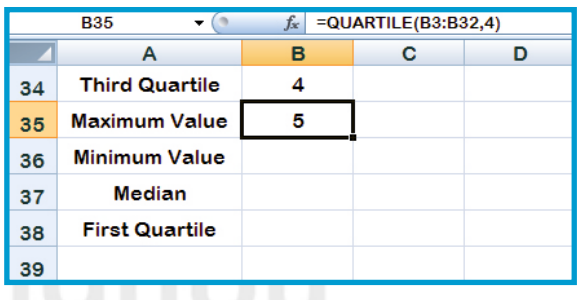


Fig. 5.28

We can also use "=Max(B3:B32)" function instead of "=Quartile (B3:B32,4)" function in Cell B35 to compute the maximum value of the data.

Step 9: After selecting Cell B36, we repeat Step 3 to compute the minimum value. We Select Cells B3:B32 as ***Array***, type "0"in ***Quart*** and click on ***OK*** as shown in Fig. 5.29.

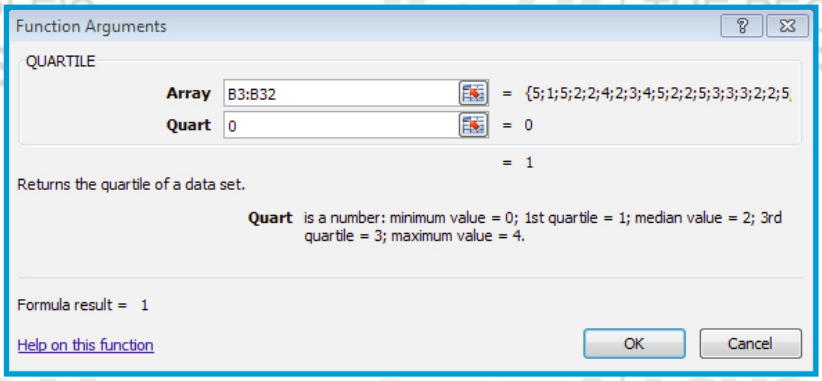


Fig. 5.29

Step 10: The minimum value of the given data is shown in Cell B36 (Fig. 5.30).

We can also use "=Min(B3:B32)" function instead of "=Quartile (B3:B32,0)" function in Cell B36 to compute the minimum value of the data.

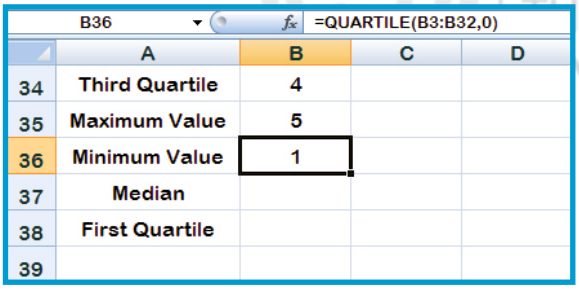


Fig. 5.30

Step 11: After selecting Cell B37, we repeat Step 3 to compute the median of the given data. We select Cells B3:B32 as ***Array***, type "2" in ***Quart*** and click on ***OK*** as shown in Fig. 5.31.

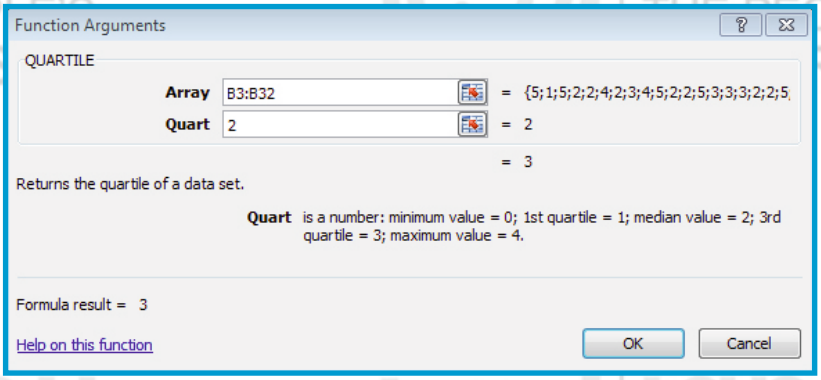


Fig. 5.31

Step 12: The median of the given data is shown in Cell B37 (Fig. 5.32).

We can also use "=Median(B3:B32)" function instead of "=Quartile(B3:B32,2)" function in Cell B7 to compute the median.

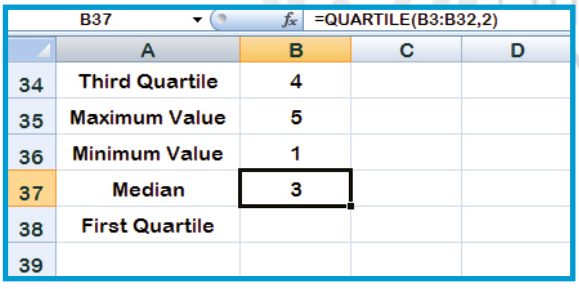


Fig. 5.32

Step 13: After selecting Cell B38, we repeat Step 3 to compute the first quartile. We select Cells B3:B32 as ***Array***, type "1" in ***Quart*** and click on ***OK*** as shown in Fig. 5.33.

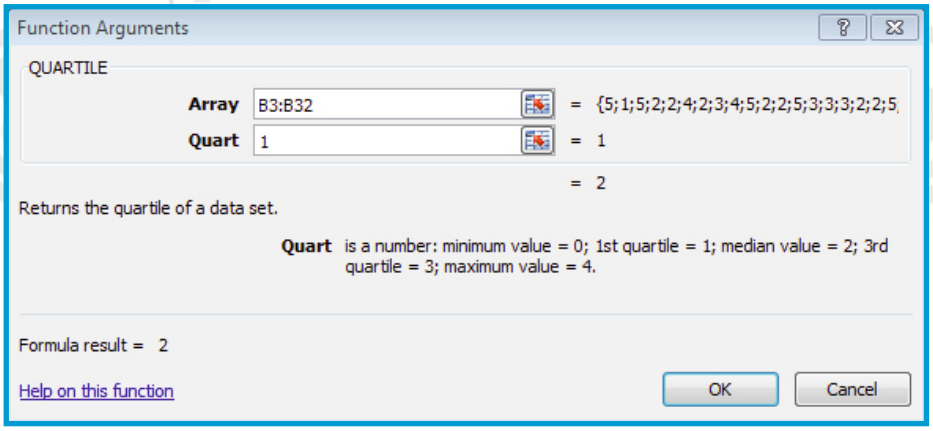


Fig. 5.33

Step 14: The first quartile of the given data is shown in Cell B38 (Fig. 5.34).

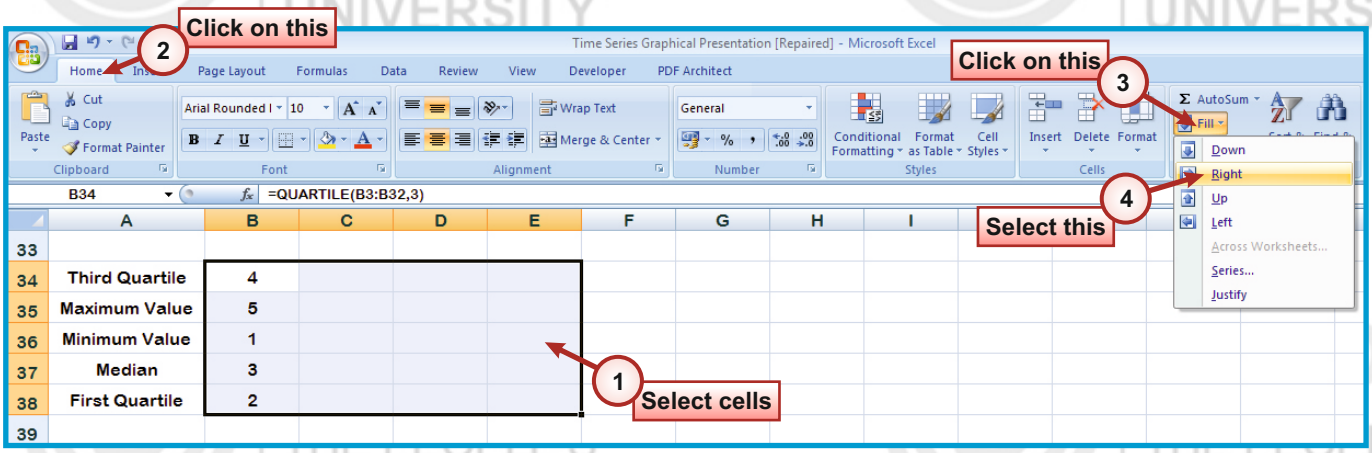
B38 =QUARTILE(B3:B32,1)

	A	B	C	D
34	Third Quartile	4		
35	Maximum Value	5		
36	Minimum Value	1		
37	Median	3		
38	First Quartile	2		
39				

Fig. 5.34

Step 15: To obtain these five values for the remaining families, we follow Fig. 5.35 and

1. select Cells B34:E38,
2. click on the ***Home*** tab, and
3. choose ***Right*** under the ***Fill*** option.



	A	B	C	D	E	F	G	H	I
33									
34	Third Quartile	4							
35	Maximum Value	5							
36	Minimum Value	1							
37	Median	3							
38	First Quartile	2							
39									

	A	B	C	D	E	F
33		Family A	Family B	Family C	Family D	
34	Third Quartile	4	5	5.75	6	
35	Maximum Value	5	6	7	8	
36	Minimum Value	1	1	2	2	
37	Median	3	3.5	4.5	4.5	
38	First Quartile	2	2.25	3	3	
39						

We can also drag Cells B34:B38 right up to column E using mouse.

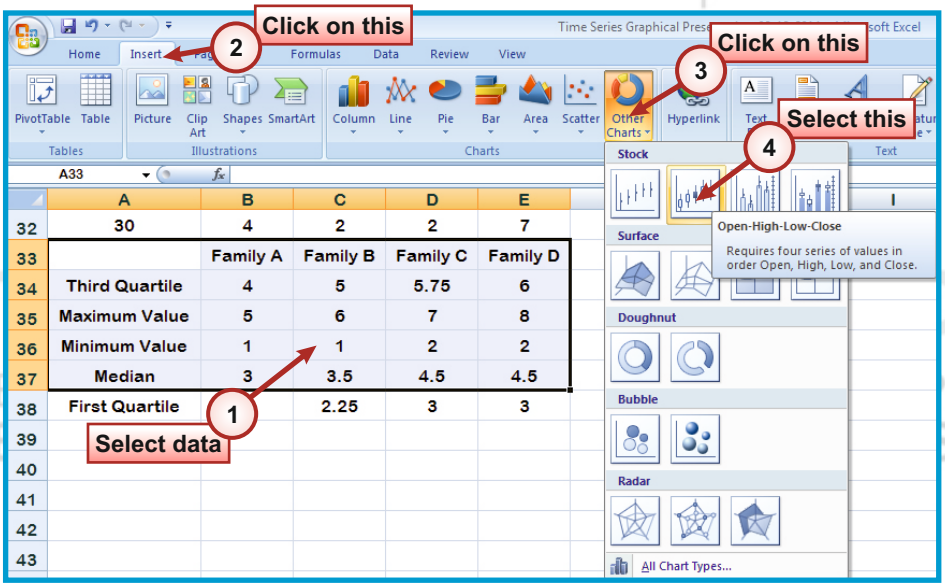
Fig. 5.35

We do not have a built-in chart type in Excel to create box-and-whiskers plots. However, we create it by modifying the *Stock* chart given in Excel. It is a type of *Open-High-Low-Close* chart.

If you have less than 4 columns of data, you will not be able to use this *Stock* chart and you will get an error message. Here is a trick for you. If you have less than four columns (or variables) in the data, you should include additional blank columns in your selection of data while plotting the chart to make it four. The Excel will not plot the blank columns. Once the graph is created, you can use *Select Data* option and reduce the size of your series. It create problem only when we initiate to create the graph.

Step 16: For constructing the box plot, we follow Fig. 5.36 and

1. select Cells A33:E37,
2. click on *Other Charts* under *Insert* tab, and
3. choose second chart subtype *Open-High-Low-Close* under *Stock* option.



	A	B	C	D	E
32	30	4	2	2	7
33		Family A	Family B	Family C	Family D
34	Third Quartile	4	5	5.75	6
35	Maximum Value	5	6	7	8
36	Minimum Value	1	1	2	2
37	Median	3	3.5	4.5	4.5
38	First Quartile		2.25	3	3
39					
40					
41					
42					
43					

Fig. 5.36

Step 17: The resulting graph is shown in Fig. 5.37.

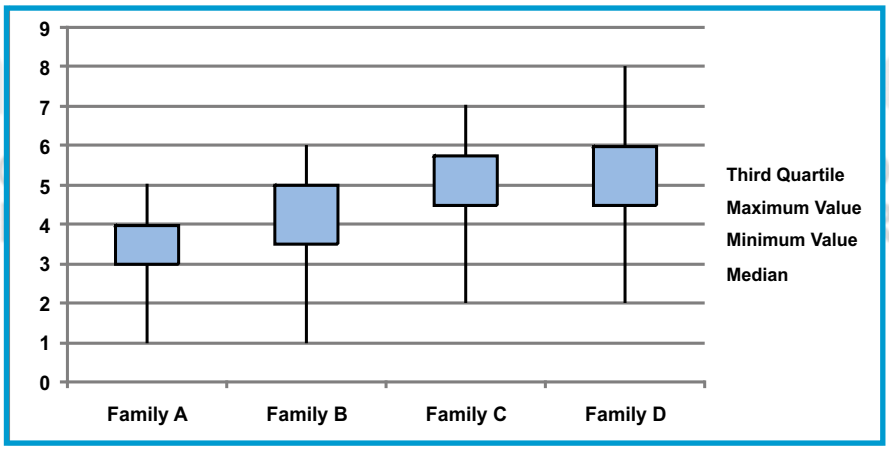


Fig. 5.37

Step 18: We format the graph as explained in Steps 4 to 9 of Sec. 3.3 of Lab Session 3. The resulting graph is shown in Fig. 5.38.

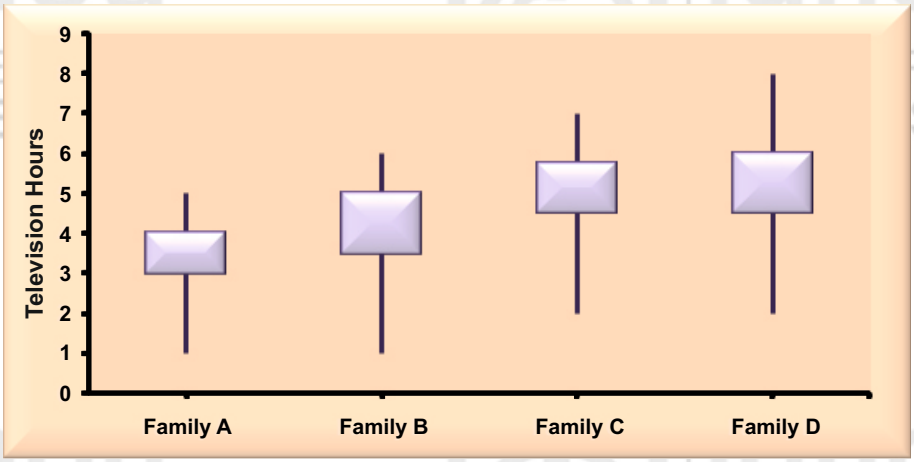


Fig. 5.38

Step 19: We now select the joint points of the lower lines (or whiskers) the boxes as shown in Fig. 5.39 and

1. click on the ***Layout*** tab under ***Chart Tools***, and
2. choose the ***Format Selection*** option.

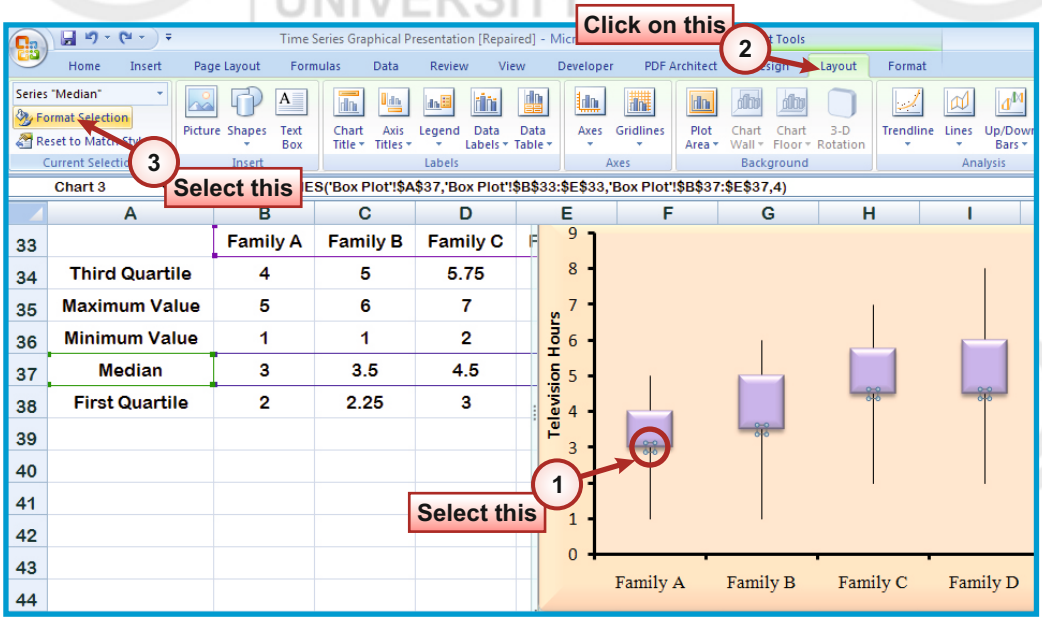


Fig. 5.39

Step 20: The dialog box shown in Fig. 5.40 appears. We now

1. click on ***Marker Options***,
2. tick on the ***Built-in*** option,
3. choose the markers ***Type*** (–) as shown in Fig. 5.40, and
4. increase the ***Size*** as desired.

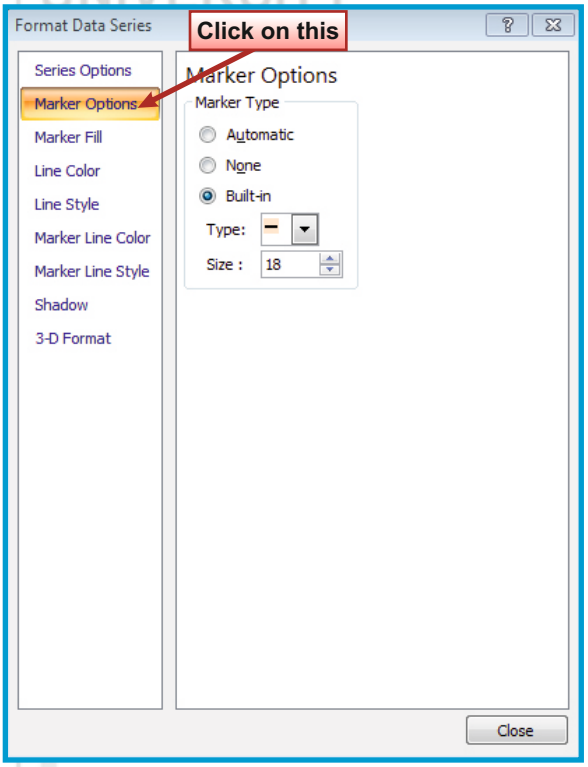


Fig. 5.40

Step 21: The resulting chart with line markers is shown in Fig. 5.41.

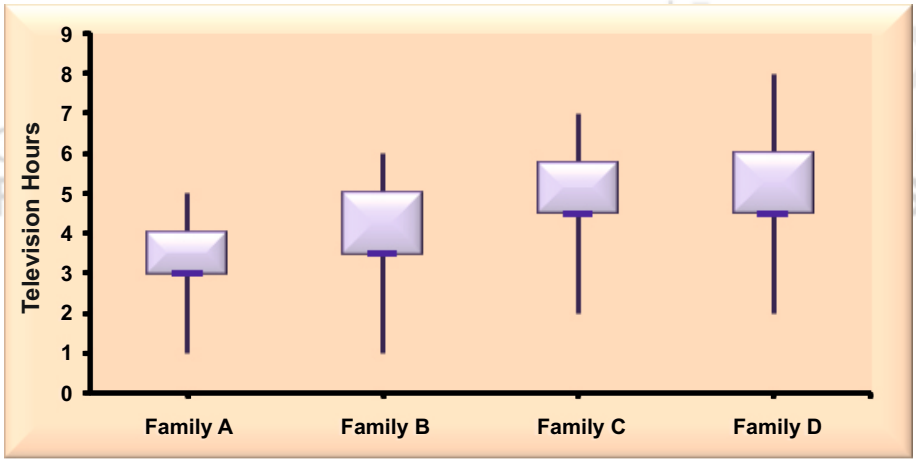


Fig. 5.41

Step 22: The box plot is still not complete. We now follow Fig. 5.42 and

1. click on the chart,
2. click on the ***Design*** tab under ***Chart Tools***, and
3. choose the ***Select Data*** option.

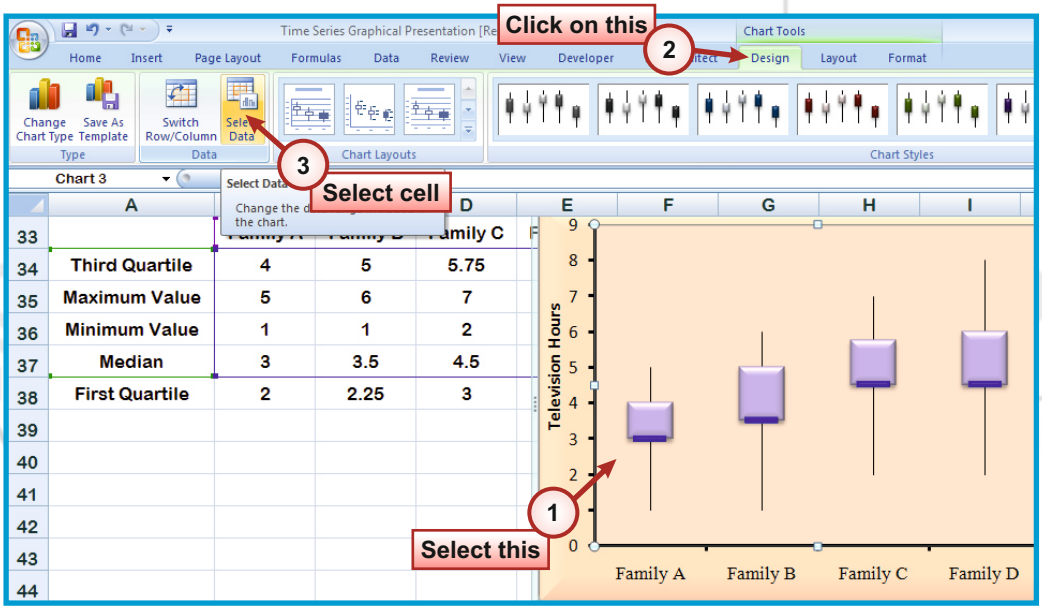


Fig. 5.42

Step 23: The dialog box shown in Fig. 5.43 appears. We click on ***Add***.

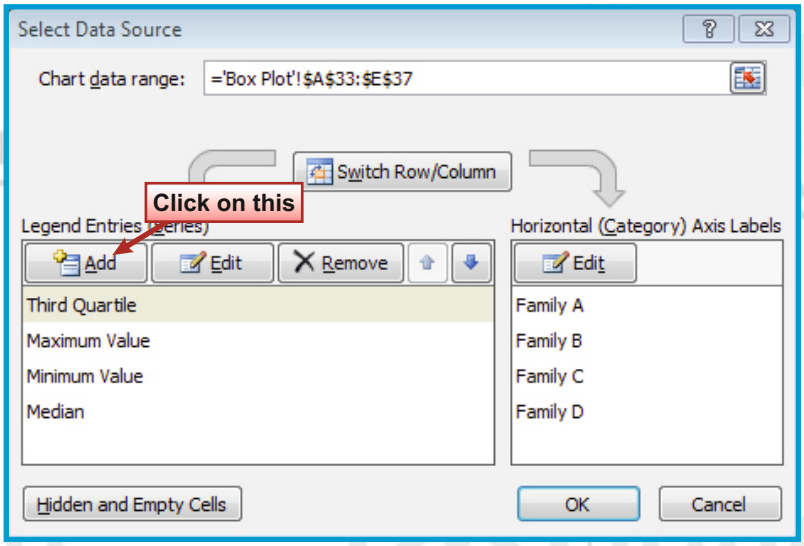
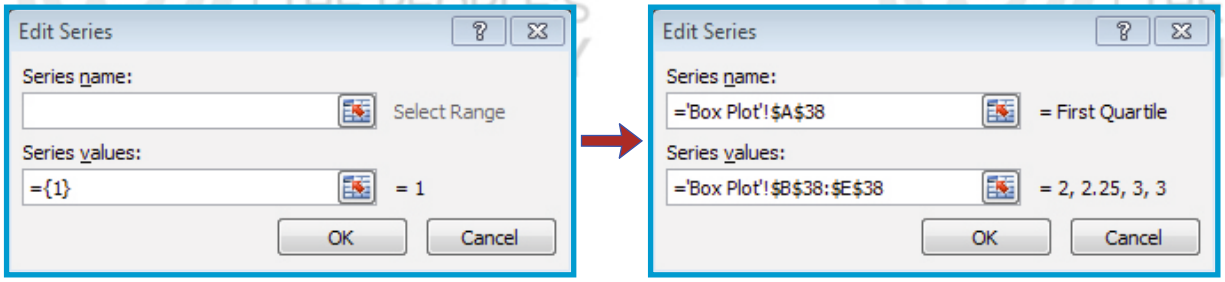


Fig. 5.43

Step 24: One more dialog box shown in Fig 5.44a appears. We select Cell A38 as ***Series name*** and Cells B38:E38 as ***Series values*** and click on ***OK*** as shown in Fig. 5.44b.



(a) (b)

Fig. 5.44

Step 25: The dialog box shown in Fig. 5.43 transforms to the dialog box shown in Fig. 5.45. We now click on ***OK***.

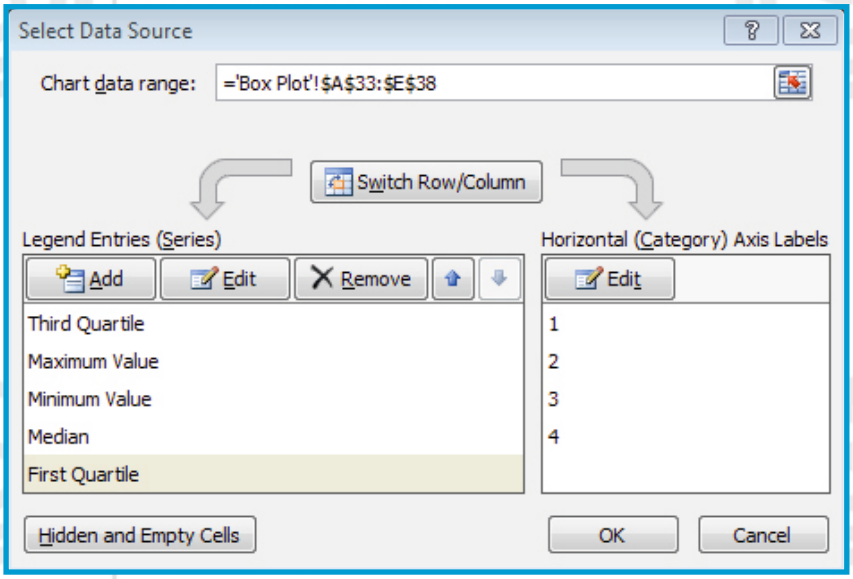


Fig. 5.45

Step 26: The graph shown in Fig. 5.46 is known as the box plot.

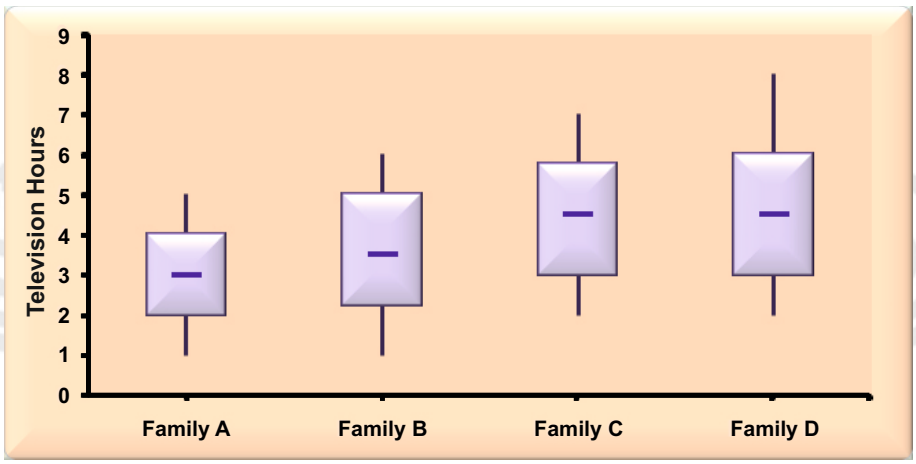


Fig. 5.46

The horizontal line inside the box represents the median. The vertical lines on both sides of the box extending to the minimum and maximum values of the given data are known as whiskers.

You should now apply this method to other problems for practice.

Activity

Construct suitable graphs with the help of MS Excel 2007 for

A1) Example 2 given in Unit 16 of MST-001.

A2) Exercises E1, E2, E6 and E7 given in Unit 16 of MST-001.

Match the results with the manual plotting of the graphs done in Unit 16 of MST-001.

Continuous Assessment 5

1. Suppose the number of cars on a particular road were recorded (in thousands) for 12 months as a part of a hypothetical study of a car manufacturing company in a city. The data are given in Table 5.

Table 5: Number of cars on road

Month	Number of Cars (in thousands)	Month	Number of Cars (in thousands)
1	278	7	418
2	286	8	378
3	398	9	435
4	421	10	498
5	367	11	425
6	387	12	447

Represent the time series data using a suitable graph.

2. The minimum and maximum prices of a particular variety of bag sold by a shopkeeper during a day are collected for fifteen days. The data are recorded in Table 6.

Table 6: Minimum and maximum price data

Days	Minimum Price (in ₹)	Maximum Price (in ₹)
1	800	680
2	820	650
3	775	700
4	720	575
5	860	700
6	840	750
7	740	600
8	780	650
9	800	700
10	750	575
11	780	705
12	800	608
13	775	550
14	750	650
15	840	680

Represent the time series data using a suitable graph.

3. To monitor the pattern of the agricultural production and study the changes in the annual cropping pattern of a field, the data of the yield of four varieties of crops, namely, wheat, rice, sugarcane and corn for the period of 1999 to 2013 were collected and are given in Table 7.

Table 7: Annual yield data for four crops

Year	Yield/Hectare (in tons)			
	Wheat	Rice	Sugarcane	Corn
1999	12.5	10.8	6.9	3.9
2000	14.5	11.5	5.9	3.1
2001	13.6	10.6	5.6	4.9
2002	14.5	11.5	6.5	4.5
2003	15.7	9.2	5.2	4.2
2004	13.5	10.5	5.5	3.9
2005	12.6	12.4	4.6	4.3
2006	13.5	10.5	5.5	3.5
2007	14.2	11.2	6.2	4.2
2008	12.5	9.5	5.8	3.7
2009	11.6	11.7	6.7	4.7
2010	12.5	9.5	4.5	2.9
2011	14.2	7.4	6.2	4.2
2012	13.5	10.5	5.5	3.5
2013	12.6	9.6	4.6	2.6

Plot a suitable graph for this data.

4. A survey was conducted to note the number of packed juices of four different brands sold each day. A sample of twenty-five days was taken. The numbers of juice cartons sold are recorded in Table 8.

Table 8: Data for sales of juice cartons

Days	Number of Sold Juice Cartons			
	Brand A	Brand B	Brand C	Brand D
1	25	10	30	15
2	5	10	35	25
3	25	30	30	35
4	10	20	15	15
5	10	15	15	30
6	20	15	10	10
7	10	5	10	25
8	15	10	30	35
9	20	5	25	20
10	25	10	15	30
11	10	15	25	20
12	10	20	20	20
13	25	25	30	35
14	15	30	15	10
15	15	5	20	20
16	15	20	35	30
17	10	15	25	30
18	10	15	10	10
19	25	30	30	40
20	10	15	10	10
21	20	25	25	15
22	20	15	15	25
23	25	30	20	35
24	10	25	30	10
25	20	10	10	35

Draw the box plot to compare the sales of the given brands of juice.

Home Work: Do It Yourself

1) Follow the steps explained in Secs. 5.3 to 5.6 to represent the data of Tables 1 to 4 graphically. Use a different format for the graphs. Take the screenshots and keep them in your record book.

2) Develop the spreadsheets for the exercises of "Continuous Assessment 5" as explained in this lab session. Take the screenshots of the final spreadsheets and graphs.

3) **Do not forget** to keep the screenshots in your record book as these will contribute to your continuous assessment in the Laboratory.