

www.ignousite.com

Course Code : BCSL-044

Course Title : Statistical Techniques Lab

Assignment Number : BCA(4)/L-044/Assignment/2022-23

Maximum Marks : 50

Weightage : 25% www.ignousite.com

Last Dates for Submission : 31st October, 2022 (For July Session)

: 15th April, 2023 (For January Session)



- Q1.** Weight of the students of a class of 30 students was recorded. This data is given in the following table. Perform the tasks given in (i) to (iv) on the data given below using a spreadsheet package: (6)

(Weight in Kgs)

45	57	60	59	45	56	67	75	70	39
65	60	47	42	41	53	49	67	42	50
52	44	45	70	58	44	55	41	45	75

- (i) Find the minimum and maximum weight using the spreadsheet formula.
- (ii) Create 5 classes with suitable class intervals and create the frequency distribution. You must use Array formula.
- (iii) Find the percentage of the students whose weight is less than the mean weight. Also find outlier in the data, if any.
- (iv) Draw the histogram for the data given in the table above. Is the data distribution of data normal distribution?

- Q2.** Perform the following tasks using spreadsheet software (you may use spreadsheet function for computing the value of t): (6)

- (i) Find the value of t for the given value of degree of freedom and significance level (alpha):

Degree of freedom	Significance
20	0.05
15	0.01

- (ii) A company manufactures rice bags of 10 kg weight. A sample of 50 such rice bags were taken out of a lot consisting of 2000 rice bags. The mean sample

weight was found to be 9.95 kg having a standard deviation of 0.06 kg. Assuming random sampling and a confidence level of 95%, will you accept the rice bags assuming that the difference in average weight is due to chance sample. Justify your answer. You should perform all computations using a spreadsheet software. Make suitable assumption, if any.

- Q3.** A hardware company produces disk brakes; the diameter of each disk brake is required to be exactly 250 mm. The company has four different machines to produce these disk brakes. Each day five samples of each machine are taken and the diameter of disk brakes is measured. The following tables lists these details:

(10)

The size of disk brakes (in mm)

Sample	Machine Identifier			
	A	B	C	D
1	251	250	249	251
2	252	249	251	252
3	250	249	251	248
4	249	248	251	247
5	249	249	249	251

Perform an ANOVA using any software to test (at 5% level) whether all the four machines are producing the disk brakes with correct length. Justify your findings. Make suitable assumptions, if any.

- Q4.** The rainfall in the first 15 days of July, 2022 is shown in the following table. Use spreadsheet software to find the moving averages for the length of 3 and 5. Also draw suitable graphs of these moving averages.

(6)

Day	Rainfall (mm)
1	50
2	55
3	5
4	12
5	13
6	0
7	0
8	11
9	0
10	0
11	5
12	0
13	110
14	150
15	130

- Q5.** A company packs sugar in a packet of 5 kg. The quality of process of producing the packets of sugar is controlled statistically. To do so, sample of five packets are taken at four different times of the day. Calculate the control limits for mean and range; and plot the control charts using any statistical software. Make suitable assumptions, if any. (6)

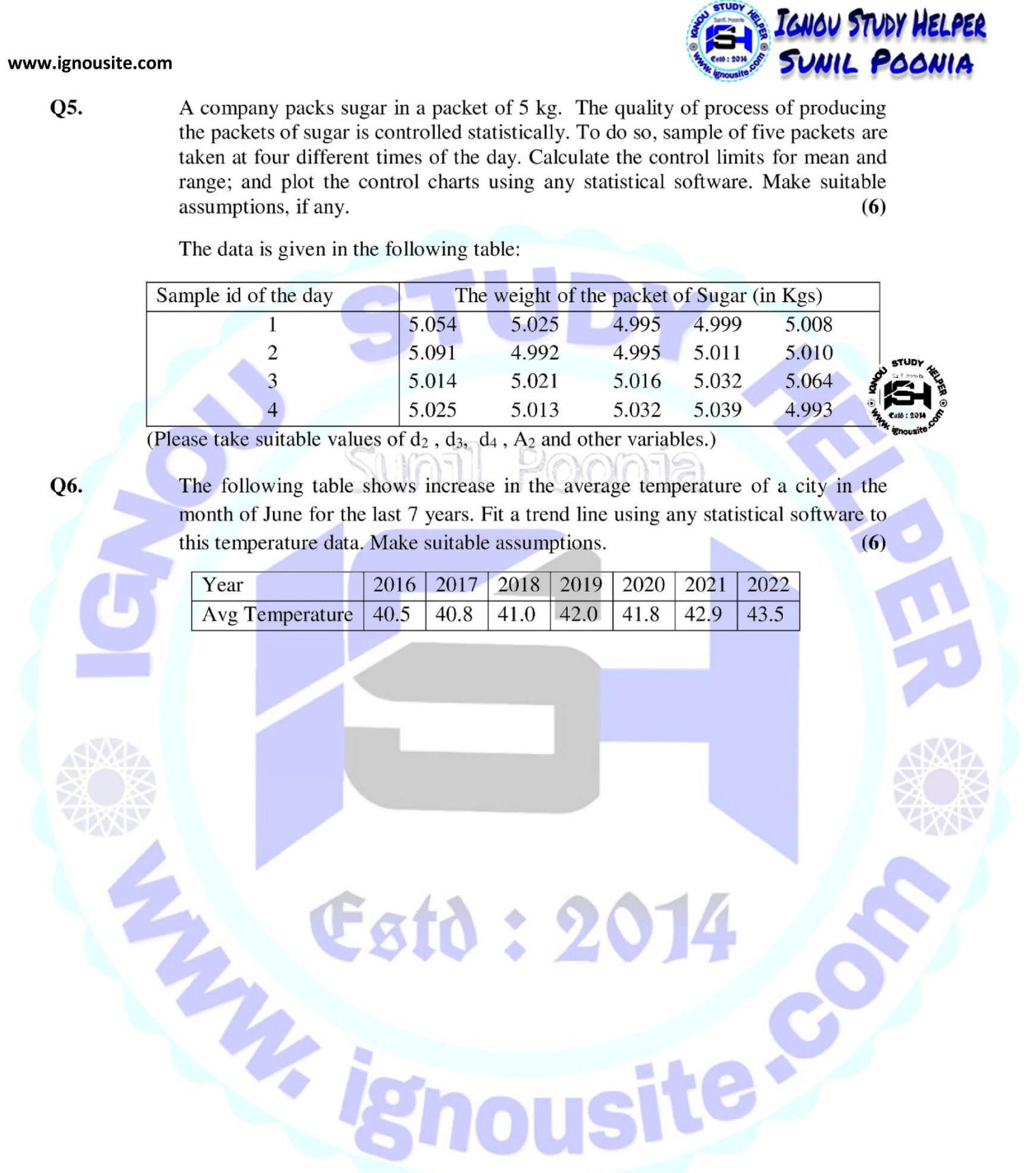
The data is given in the following table:

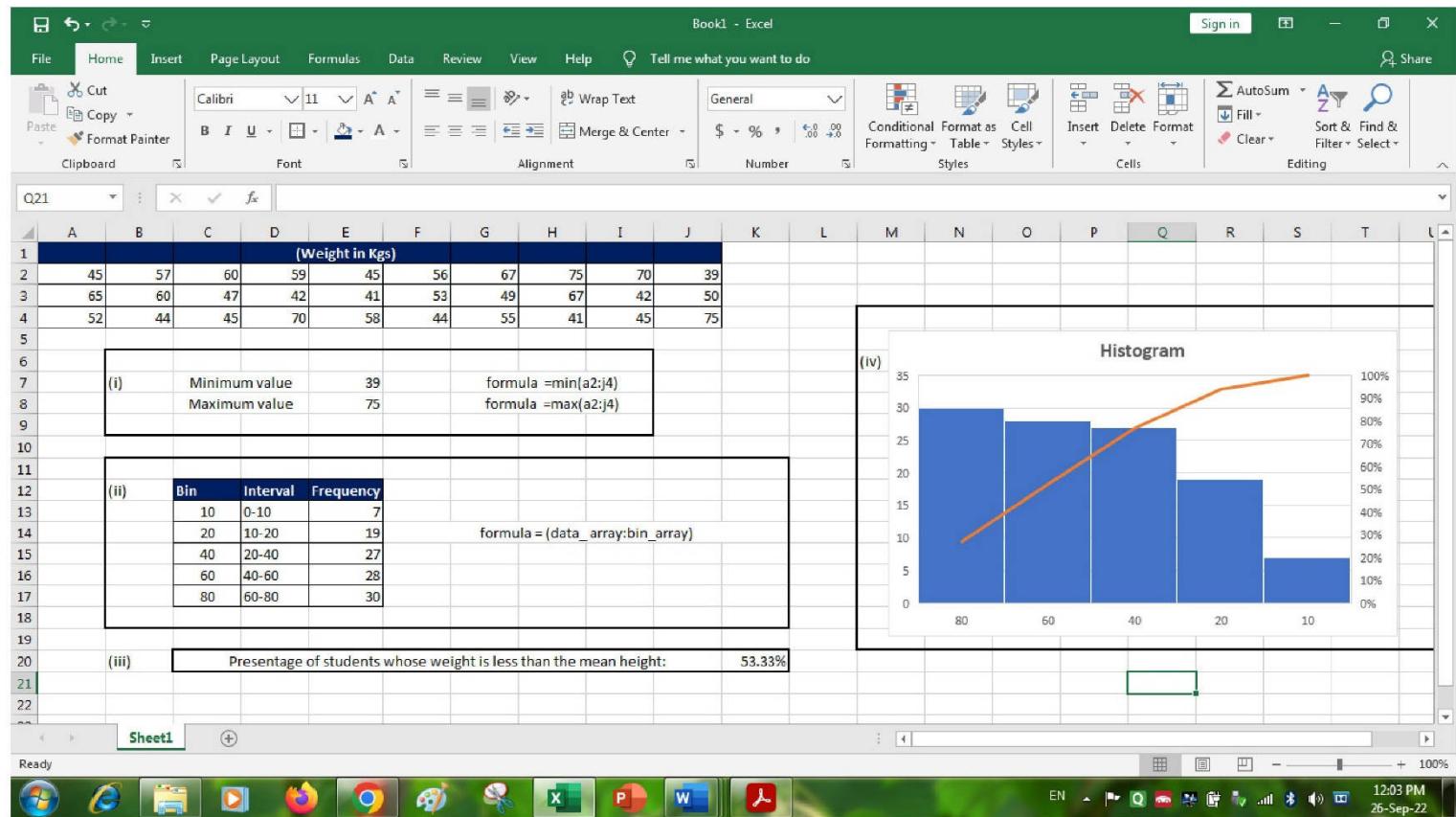
Sample id of the day	The weight of the packet of Sugar (in Kgs)				
1	5.054	5.025	4.995	4.999	5.008
2	5.091	4.992	4.995	5.011	5.010
3	5.014	5.021	5.016	5.032	5.064
4	5.025	5.013	5.032	5.039	4.993

(Please take suitable values of d_2 , d_3 , d_4 , A_2 and other variables.)

- Q6.** The following table shows increase in the average temperature of a city in the month of June for the last 7 years. Fit a trend line using any statistical software to this temperature data. Make suitable assumptions. (6)

Year	2016	2017	2018	2019	2020	2021	2022
Avg Temperature	40.5	40.8	41.0	42.0	41.8	42.9	43.5



Ans.1

Ans.2

(i) Value of t for the given value of degree of freedom and significance level is calculated using the Excel formula:

=TINV(probability/Significance, degree_of_freedom)

In our case,

(a) t-value (for degree of freedom = 20, level of significance = 0.05) = 2.08596344 [=TINV(0.05,20)]

(b) t-value (for degree of freedom = 15, level of significance = 0.1) = 1.75305033 [=TINV(0.1,15)]

File	Home	Insert	Page Layout	Formulas	Data	Review	View	Help	Tell me what you want to do	
Paste	Cut	Copy	Format Painter	Calibri	11	A ⁺ A ⁻	Wrap Text	General		
Clipboard				B I U		Horizontal Alignment	Merge & Center	\$ % , .		
				Font		Vertical Alignment			Number	
K4										
	A	B	C	D	E	F	G	H	I	J
1			Degree of freedom	Significance	t-value					
2			20	0.05	2.08596344					
3			15	0.01	1.75305033					
4										
5										
6										
7										
8										

(ii) The intended weight of the rice bag (μ) = 10 kg.

The sample size (n) = 50

This implies, the value of degree of freedom (Y) = $(50 - 1) = 49$

The sample mean (given) is: $\bar{X} = 9.95$ kg,

And the Standard Deviation: $\sigma = 0.06$

The value of Random Variable: $t = (\bar{X} - \mu) / (\sigma / \sqrt{n}) = -5.8926$

Permissible t-value (for $Y=49$ & $\alpha = 0.05$) = 2.0096 (computed using the formula:

[=TINV(Probability, Degree of Freedom)]

Alignment Number Styles Cells

F	G	H	I	J	K	L	M
(in kg)							
(intended weights)	$\mu =$	10					
(sample size)	$n =$	50		$y =$	49		
(sample mean)	$x =$	9.95					
standard deviation	$\sigma =$	0.06					
value of a random variable	$t =$	-5.8926					
now	degree of freedom	$(n-1) =$	49				
permissible t-value							
for $y = 49$ & $\alpha = 0.05$		2.0096 [(probability degree freedom)]					

Calculated t-value - 5.8926 is less than the t-value (permissible) for the degree of freedom value of 49 & probability of 0.05. So the process is under control i.e. the bearings are acceptable.

Book1 - Excel

Sign in

View Help Tell me what you want to do

Wrap Text Merge & Center Conditional Formatting Format as Table Cell Styles Insert Delete Format Cells AutoSum Fill Sort & Filter Find & Select Editing

F	G	H	I	J	K	L	M	N	O	P
(in kg)										
(intended weights)	$\mu =$	10								
(sample size)	$n =$	50		$y =$	49					
(sample mean)	$x =$	9.95								
standard deviation	$\sigma =$	0.06								
value of a random variable	$t =$	-5.8926								
now	degree of freedom	$(n-1) =$	49							
permissible t-value										
for $y = 49$ & $\alpha = 0.05$		2.0096 [(probability degree freedom)]								

Calculated t-value - 5.8926 is less than the t-value (permissible) for the degree of freedom value of 49 & probability of 0.05.
So the process is under control i.e. the bearings are acceptable.

Ans.3 As per our problem,

Null Hypothesis: There is no significant difference in the production of disk brakes with correct length from all the available machines.

Alternate Hypothesis: There is a significant difference in the production of disk brakes with correct length from all the available machines.

We enter the data in the Excel spread-sheet.

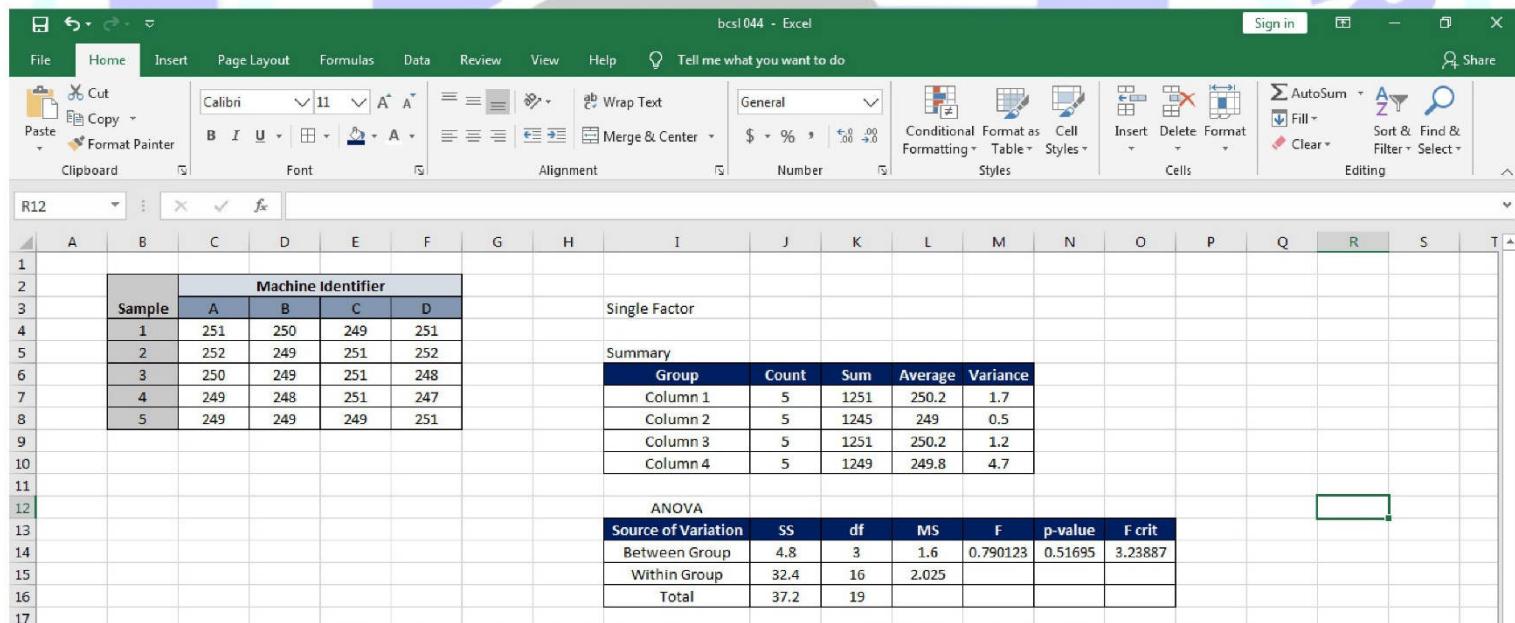
Now we select Data Tab > Data Analysis > select Anova: Single FactorThe Anova: Single Factor dialog box pops up.

We need to enter:

- The Input Range (here \$C\$4:\$F\$8),
- Grouped By Checkbox: (here Columns)
- The Alpha Value (here 0.05 or 5%)
- And the output range (where the Anova chart is going to be drawn):

And then click on OK.

We get the Anova chart (Screenshot attached below)



The screenshot shows an Excel spreadsheet titled "bcs1044 - Excel". The data is organized into three main sections:

- Machine Identifier:** A table with columns A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T. It contains data for five samples (1, 2, 3, 4, 5) across four machine identifiers (A, B, C, D).
- Single Factor:** A summary table with columns Group, Count, Sum, Average, and Variance. The data is as follows:

Group	Count	Sum	Average	Variance
Column 1	5	1251	250.2	1.7
Column 2	5	1245	249	0.5
Column 3	5	1251	250.2	1.2
Column 4	5	1249	249.8	4.7
- ANOVA:** An ANOVA table with columns Source of Variation, SS, df, MS, F, p-value, and F crit. The data is as follows:

Source of Variation	SS	df	MS	F	p-value	F crit
Between Group	4.8	3	1.6	0.790123	0.51695	3.23887
Within Group	32.4	16	2.025			
Total	37.2	19				

Based on p-Value, the Thumb rule states "If p-Value is less than he desired level of significance, reject the null hypothesis, else do not reject the null hypothesis"

In our case the p-value is 0.516952313 (approx) which is much greater than the desired level of significance (0.05). So we must accept the null hypothesis.

That is:

There is no significant difference in the production of disk brakes with correct length from all the available machines.

Ans.4 The daily rainfall data for the first 15 days of the month of July is copied in the tabular format in the Excel Sheet and 2 more columns are added to it :-

(i) 3-Day Moving Average



(ii) 5-Day Moving Average

We can go to Data Analysis > choose Moving Average > select the data and intervals and the output range to get the output.

Now we go to Insert Tab > Charts > Line & choose the necessary Line Chart Style (Line with Marker chosen here) and select the required data.




IGNOU STUDY HELPER
Sunil Poonia
Estd : 2014
www.ignousite.com

bcsl 044 - B

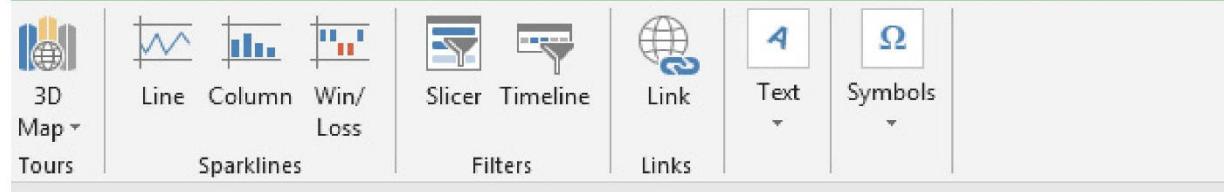
File Home Insert Page Layout Formulas Data Review View Help Tell me what you w

PivotTable Recommended Table Illustrations Get Add-ins My Add-ins Add-ins Recommended Charts PivotChart Charts

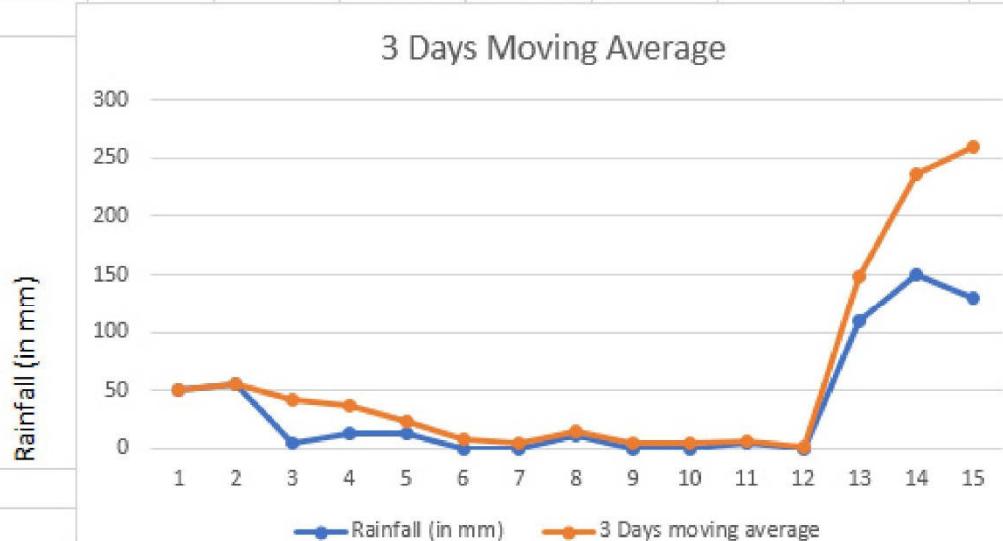
T33

	A	B	C	D	E	F
1	Column 1	Column 2		Column 3	Column 4	
2	Day	Rainfall (in mm)		3 Days moving average	5 Days moving average	
3	1	50				
4	2	55				
5	3	5		36.66666667		
6	4	12		24		
7	5	13		10	27	
8	6	0		8.333333333	17	
9	7	0		4.333333333	6	
10	8	11		3.666666667	7.2	
11	9	0		3.666666667	4.8	
12	10	0		3.666666667	2.2	
13	11	5		1.666666667	3.2	
14	12	0		1.666666667	3.2	
15	13	110		38.33333333	23	
16	14	150		86.66666667	53	
17	15	130		130	79	
18						

For the 3-Day Moving Average, we choose columns 1, 2 & 3 and for 5-Day Moving Average, we choose columns 1, 2 & 4. We also customize the chart by adding necessary chart titles, vertical & horizontal axis titles:

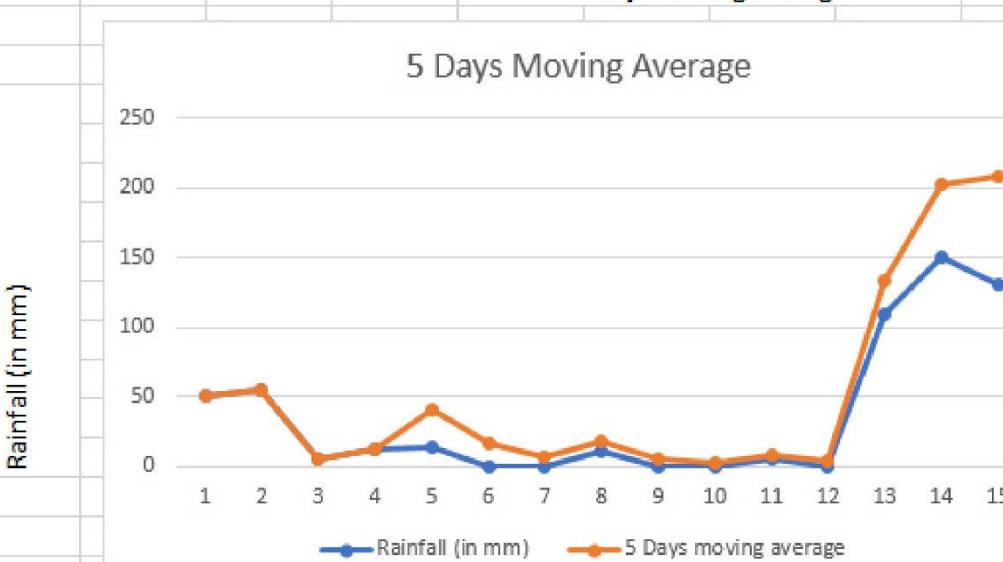


G H I J K L M N O P
3 Days moving average



Days of month

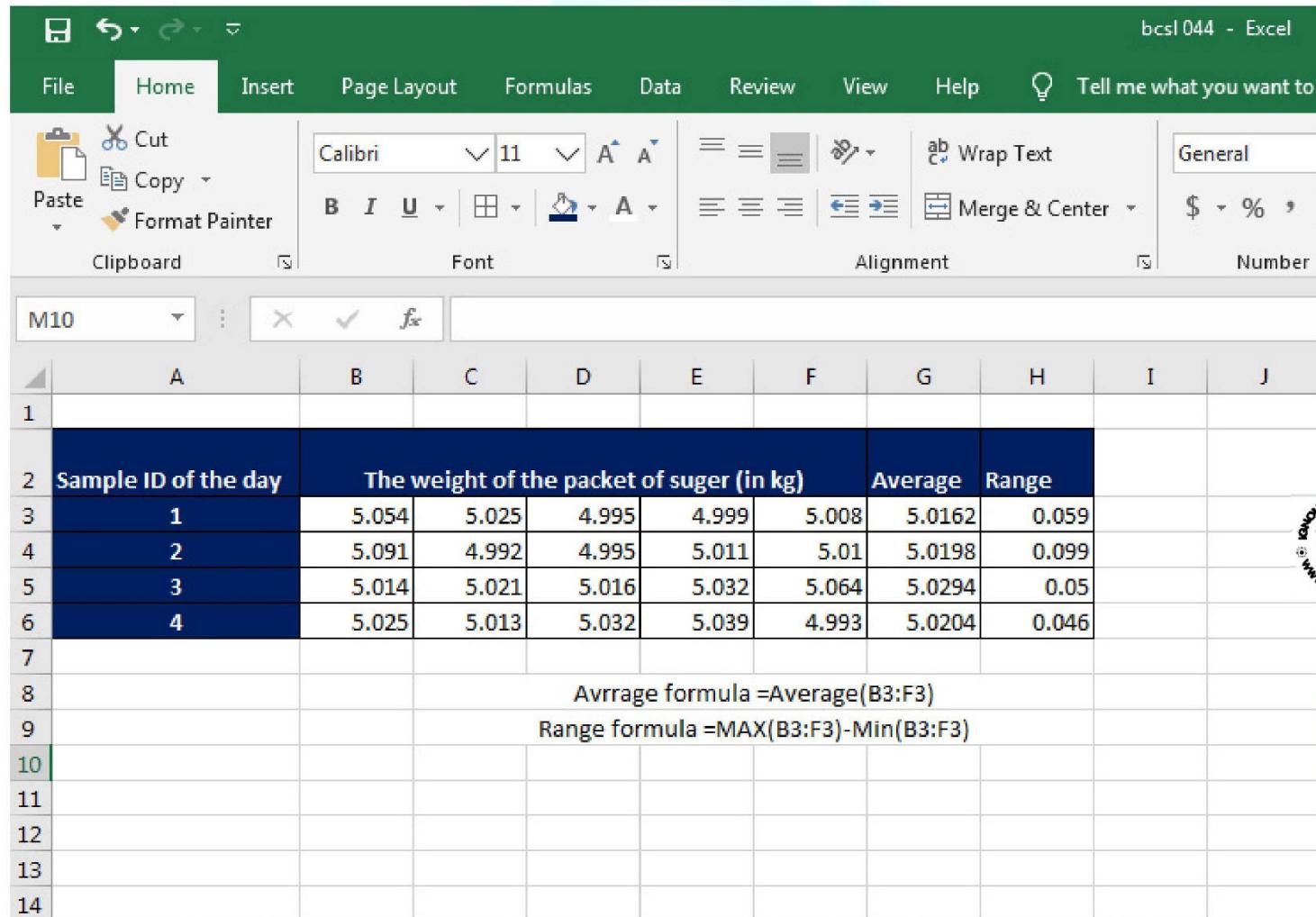
5 Days moving average



Days of month

Ans.5 We enter the data in the Excel spread-sheet and add 2 more columns – Average & Range. The Average (or X- mean) holds the average of the 5 samples for each of the 4 samples in the day and the Range holds the difference between the maximum and the minimum quantity between the 5 samples.

To get the fill the "Average" column, we enter the formulas: =AVERAGE (B3:F3) in the cell G3and so on for the rest of the data in the column G (Average).



	A	B	C	D	E	F	G	H	I	J
1										
2	Sample ID of the day	The weight of the packet of suger (in kg)					Average	Range		
3	1	5.054	5.025	4.995	4.999	5.008	5.0162	0.059		
4	2	5.091	4.992	4.995	5.011	5.01	5.0198	0.099		
5	3	5.014	5.021	5.016	5.032	5.064	5.0294	0.05		
6	4	5.025	5.013	5.032	5.039	4.993	5.0204	0.046		
7										
8	Avrrage formula =Average(B3:F3)									
9	Range formula =MAX(B3:F3)-Min(B3:F3)									
10										
11										
12										
13										
14										

To get the fill the "Range" column, we enter the formulas: =MAX(B3:F3) – MIN(B3:F3) in the cell H3 and so on for the rest of the data in the column H (Range).

On the basis of these mean & range values, \bar{X}' (Mean of Means) and R' (Mean of the ranges) is calculated as below:

=SUM (G3:G6)/COUNT (G3:G6) for \bar{X}

=SUM (H3:H6)/COUNT (H3:H6) for R

Now we compute the Control Limits for Mean and Range separately.

		File		Home		Insert		Page Layout		Formulas		Data		Review		View		Help		Tell me what you want to do			
Clipboard	Cut	Copy	Paste	Format Painter	Font		Font		Font		Font		Font		Font		Font		Font		Font		
	Wrap Text	Merge & Center	General	Conditional Formatting	Format as Table	Styles																	
O9		X	✓	f _x																			
	A	B	C	D	E	F	G	H	I	J	K	L											
1																							
2	Sample ID of the day	The weight of the packet of suger (in kg)					Average	Range															
3	1	5.054	5.025	4.995	4.999	5.008	5.0162	0.059															
4	2	5.091	4.992	4.995	5.011	5.01	5.0198	0.099															
5	3	5.014	5.021	5.016	5.032	5.064	5.0294	0.05															
6	4	5.025	5.013	5.032	5.039	4.993	5.0204	0.046												Total no. of Sample	4		
7																				X double bar	5.02145		
8		Avrage formula =Average(B3:F3)																		R bar	0.0635		
9		Range formula =MAX(B3:F3)-Min(B3:F3)																					
10																							
11																							

Now we compute the Control Limits for Mean and Range separately.

		File		Home		Insert		Page Layout		Formulas		Data		Review		View		Help		Tell me what you want to do			
Clipboard	Cut	Copy	Paste	Format Painter	Font		Font		Font		Font		Font		Font		Font		Font		Font		
	Wrap Text	Merge & Center	General	Conditional Formatting	Format as Table	Styles																	
U15		X	✓	f _x																			
	A	B	C	D	E	F	G	H	I	J	K												
1																							
2																							
3	Control limit for mean						Talking A2=0.577																
4	UCLx	X double bar+A2 *(R bar)			5.0580895																		
5	LCLx	X double bar-A2 *(R bar)			4.9848105															Talking D4 as 2.11			
6																				Talking D3=0 as sample size is 5 which is <7			
7	Control limit for Range						Here the sample size is 5																
8	UCLR	D4*R bar	0.13399																				
9	LCLR	D3*R bar	0																				
10																							

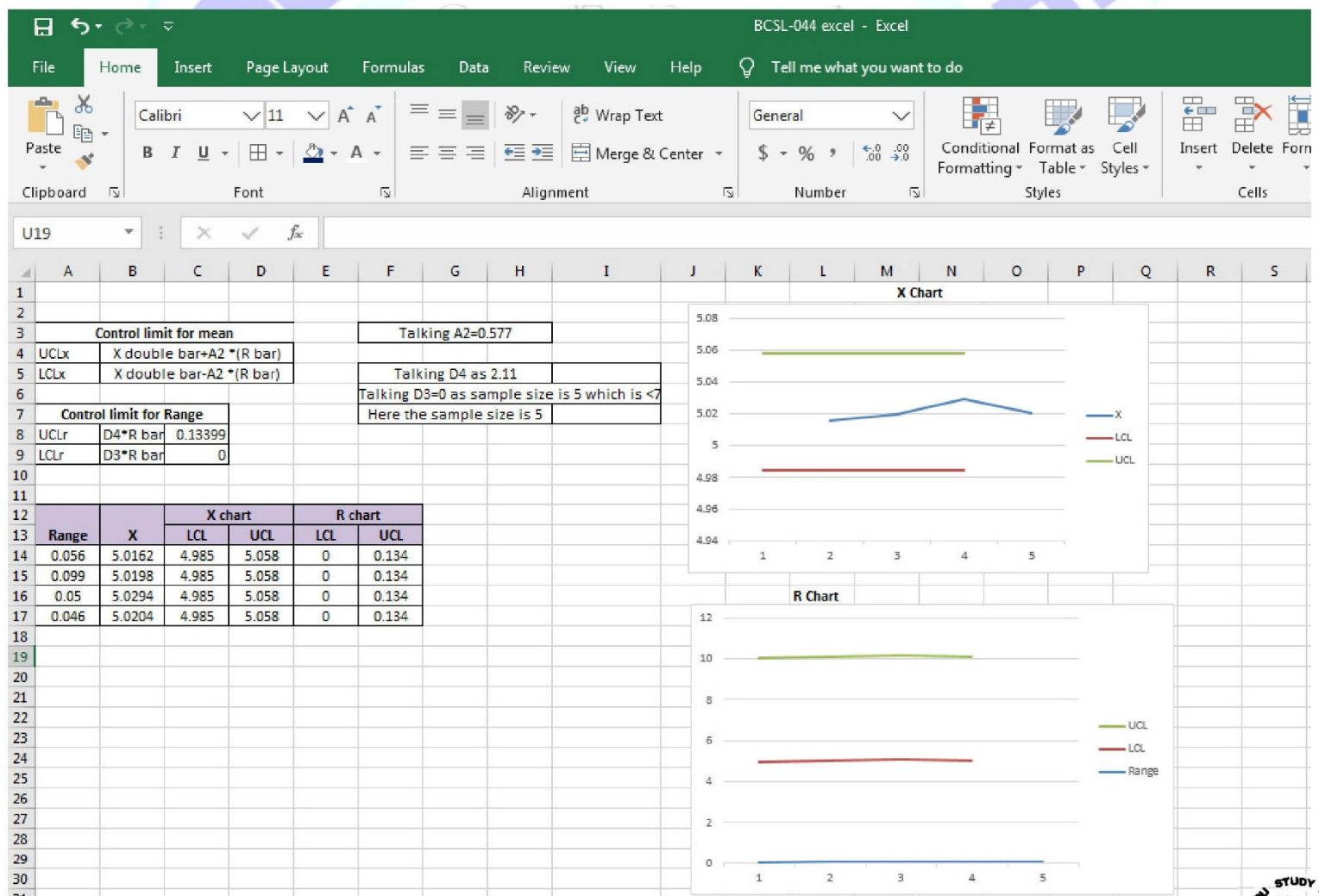
We now Final Table taking the Ranges for each samples, the Means for each samples of the day and the Lower & Upper Control Limits for Both X- Chart & R-Charts as below:



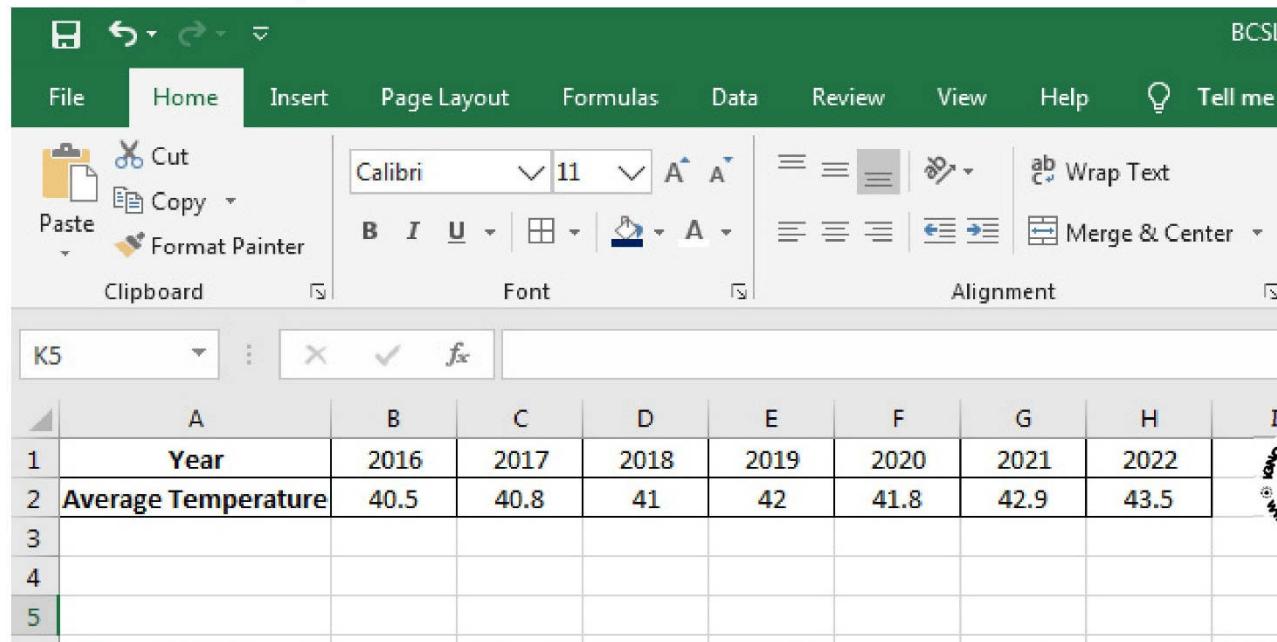
	A	B	C	D	E	F	G
13	Range	X	LCL	UCL	LCL	UCL	
14	0.056	5.0162	4.985	5.058	0	0.134	
15	0.099	5.0198	4.985	5.058	0	0.134	
16	0.05	5.0294	4.985	5.058	0	0.134	
17	0.046	5.0204	4.985	5.058	0	0.134	
18							
19							

Next we draw the two Control Charts which are X-Chart & R-Chart by going to the Insert Tab of Excel > Select Charts > Line Chart and selecting a specific chart.

For X-Chart, we take the X column, the LCL & UCL for X-Chart as data & for the R-Chart, we take the Range column, the LCL & UCL for R Chart to draw the specific graph.



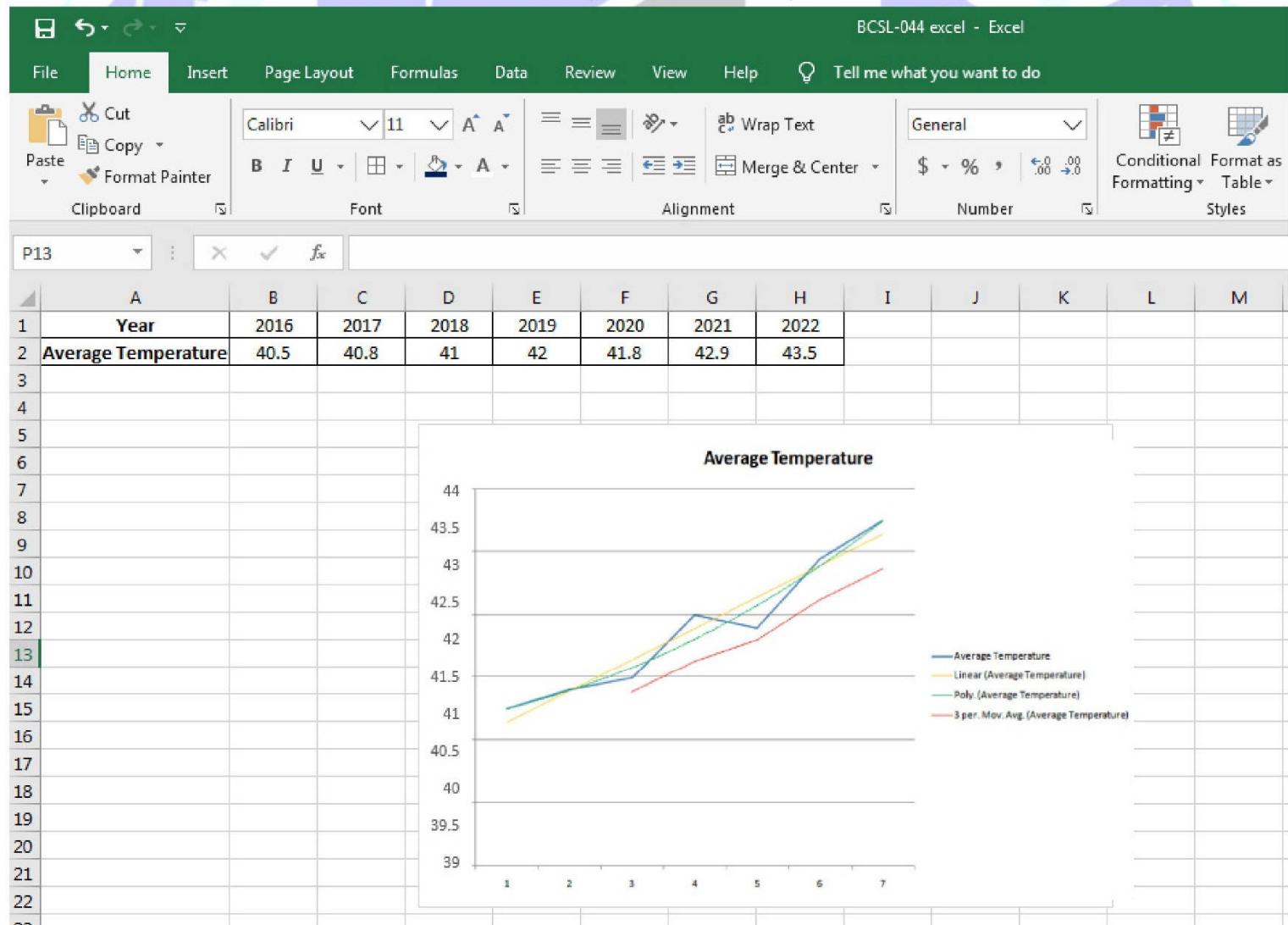
Ans.6 We enter the sample data in the Excel Sheet:



A screenshot of Microsoft Excel showing a table of average temperatures. The table has columns labeled A through H and rows labeled 1 through 5. Row 1 contains "Year" in column A and years from 2016 to 2022 in columns B through H respectively. Row 2 contains "Average Temperature" in column A and values 40.5, 40.8, 41, 42, 41.8, 42.9, and 43.5 in columns B through H respectively. The data is highlighted with a light blue selection.

	A	B	C	D	E	F	G	H
1	Year	2016	2017	2018	2019	2020	2021	2022
2	Average Temperature	40.5	40.8	41	42	41.8	42.9	43.5
3								
4								
5								

Now we choose Insert > Charts > Line Charts and select any one of the 2D Templates.



A screenshot of Microsoft Excel showing a 2D Line Chart titled "Average Temperature". The chart displays four data series over a period of 7 years (2016-2022). The Y-axis ranges from 39 to 44. The X-axis shows the years 1 through 7. The legend identifies the series as follows:

- Average Temperature (Blue line)
- Linear (Ave Temp) (Yellow line)
- Poly (Ave Temp) (Green line)
- 3 per. Mov. Avg. (Ave Temp) (Red line)

The chart shows a general upward trend for all series, with the moving average line being the smoothest and the polynomial line showing the most significant fluctuations.