



Government of Karnataka

Department of Collegiate and Technical Education

Government CPC Polytechnic, Mysuru-570007

DEPART OF SCIENCE

# **STATISTICS AND ANALYTICS LAB MANUAL**

**FOR  
DIPLOMA FIRST/SECOND SEMESTER  
COMMON TO ALL PROGRAMS  
(AS PER C-20 SYLLABUS OF BTE,BENGALURU)**

**PREPARED BY  
RASHMI S  
SENIOR SCALE LECTURER  
DEPARTMENT OF SCIENCE  
GOVT CPC POLYETECHNIC,MYSORE**

**EXPERIMENT-5**

STATISTICS AND ANALYTICS LAB

RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

**AIM: In Microsoft Excel spread sheet draw frequency distribution table for the given data  
(data set should contain minimum 50 data)**

		FREQUENCY DISTRIBUTION TABLE		
Name of Student	Marks of Science	Class interval	Frequency	
Student1	25	0 TO 10	5	=COUNTIFS (B5:B54,">=0", B5:B54,"<10")
Student2	30	10 TO 20	4	=COUNTIFS(B5:B54,">=10",B5:B54,"<20")
Student3	35	20 TO 30	4	=COUNTIFS(B5:B54,">=20",B5:B54,"<30")
Student4	45	30 TO 40	9	=COUNTIFS(B4:B53,">=30",B4:B53,"<40")
Student5	6	40 TO 50	8	=COUNTIFS(B4:B53,">=40",B4:B53,"<50")
Student6	3	50 TO 60	4	=COUNTIFS(B4:B53,">=50",B4:B53,"<60")
Student7	45	60 TO 70	6	=COUNTIFS(B4:B53,">=60",B4:B53,"<70")
Student8	65	70 TO 80	4	=COUNTIFS(B4:B53,">=70",B4:B53,"<80")
Student9	75	80 TO 90	3	=COUNTIFS(B4:B53,">=80",B4:B53,"<90")
Student10	86	90 TO 100	3	=COUNTIFS(B4:B53,">=90",B4:B53,"<=100")
Student11	78	<b>SUM</b>	<b>50</b>	=SUM(E4:E13)
Student12	34			
Student13	0			
Student14	12			
Student15	13			
Student16	45			
Student17	67			
Student18	34			
Student19	67			
Student20	89			
Student21	99			
Student22	88			
Student23	77			
Student24	79			
Student25	69			

29	Student26	60							
30	Student27	40							
31	Student28	30							
32	Student29	100							
33	Student30	100							
34	Student31	23							
35	Student32	45							
36	Student33	45							
37	Student34	30							
38	Student35	56							
39	Student36	55							
40	Student37	44							
41	Student38	56							
42	Student39	67							
43	Student40	56							
44	Student41	34							
45	Student42	32							
46	Student43	44							
47	Student44	33							
48	Student45	22							
49	Student46	21							
50	Student47	11							
51	Student48	12							
52	Student49	1							
53	Student50	2							

**INTERPRETATION:** From the frequency distribution table, we can conclude that

- 1) Out of 50 students, 5 students scored the marks in science in between 0 to 10.
- 2) Out of 50 students, 4 students scored the marks in science in between 10 to 20.
- 3) Out of 50 students, 4 students scored the marks in science in between 20 to 30.
- 4) Out of 50 students, 9 students scored the marks in science in between 30 to 40.
- 5) Out of 50 students, 8 students scored the marks in science in between 40 to 50.
- 6) Out of 50 students, 4 students scored the marks in science in between 50 to 60.
- 7) Out of 50 students, 6 students scored the marks in science in between 60 to 70.
- 8) Out of 50 students, 4 students scored the marks in science in between 70 to 80.
- 9) Out of 50 students, 3 students scored the marks in science in between 80 to 90.
- 10) Out of 50 students, 3 students scored the marks in science in between 90 to 100.
- 11) Out of 50 students, most of the students scored the marks in science in between 30 to 40.
- 12) Out of 50 students, a smaller number of students scored the marks in science in between 80 to 90 and 90 to 100.

**13) PURPOSE:** Tabulation makes the data brief. Therefore, it can be easily presented in the form of graphs.

**EXPERIMENT-6**

STATISTICS AND ANALYTICS LAB

RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

**AIM: In Microsoft Excel spread sheet draw the relative frequency distribution table for the given data(data set should contain minimum 50 data)**

Name of Student	Marks of Science	Class interval	Frequency	Relative frequency
Student1	25	0 TO 10	5	0.1
Student2	30	10 TO 20	4	0.08
Student3	35	20 TO 30	4	0.08
Student4	45	30 TO 40	9	0.18
Student5	6	40 TO 50	8	0.16
Student6	3	50 TO 60	4	0.08
Student7	45	60 TO 70	6	0.12
Student8	65	70 TO 80	4	0.08
Student9	75	80 TO 90	3	0.06
Student10	86	90 TO 100	3	0.06
Student11	78	<b>SUM</b>	<b>50</b>	<b>1</b>
Student12	34			
Student13	0			
Student14	12			
Student15	13			
Student16	45			
Student17	67			
Student18	34			
Student19	67			
Student20	89			
Student21	99			
Student22	88			
Student23	77			
Student24	79			
Student25	69			
Student26	60			
Student27	40			
Student28	30			
Student29	100			

33	Student30	100						
34	Student31	23						
35	Student32	45						
36	Student33	45						
37	Student34	30						
38	Student35	56						
39	Student36	55						
40	Student37	44						
41	Student38	56						
42	Student39	67						
43	Student40	56						
42	Student39	67						
43	Student40	56						
44	Student41	34						
45	Student42	32						
46	Student43	44						
47	Student44	33						
48	Student45	22						
49	Student46	21						
50	Student47	11						
51	Student48	12						
52	Student49	1						
53	Student50	2						

**INTERPRETATION:** From the frequency distribution table, we can conclude that

- 1) Out of 50 students, 5 students scored the marks in science in between 0 to 10 and whose relative frequency is 0.1.
- 2) Out of 50 students, 4 students scored the marks in science in between 10 to 20 and whose relative frequency is 0.08.
- 3) Out of 50 students, 4 students scored the marks in science in between 20 to 30 and whose relative frequency is 0.08.
- 4) Out of 50 students, 9 students scored the marks in science in between 30 to 40 and whose relative frequency is 0.18.
- 5) Out of 50 students, 8 students scored the marks in science in between 40 to 50 and whose relative frequency is 0.16.
- 6) Out of 50 students, 4 students scored the marks in science in between 50 to 60 and whose relative frequency is 0.08.
- 7) Out of 50 students, 6 students scored the marks in science in between 60 to 70 and whose relative frequency is 0.12.
- 8) Out of 50 students, 4 students scored the marks in science in between 70 to 80 and whose relative frequency is 0.08.
- 9) Out of 50 students, 3 students scored the marks in science in between 80 to 90 and whose relative frequency is 0.06.
- 10) Out of 50 students, 3 students scored the marks in science in between 90 to 100 and whose relative frequency is 0.06.

**11)PURPOSE:** The data in excel spread sheet will be helpful for collecting data set and finding relative frequency distribution in charts.

**EXPERIMENT-7**

STATISTICS AND ANALYTICS LAB

RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

**AIM: Using Microsoft Excel Spread Sheet plot Bar Graph for the data collected from 100 people (for example , conduct survey on favorite fruit of a person in your locality)(restricting to 5 to 6 fruits). Explain the graph with minimum 30 words**

Name of Person	Fruit Name	Name of the fruit	Count of Fruits
Person1	apple	apple	19
Person2	orange	orange	23
Person3	banana	banana	18
Person4	sapota	sapota	18
Person5	grapes	grapes	10
Person6	kiwi	kiwi	12
Person7	apple	Total Respose	100
Person8	orange		
Person9	sapota		
Person10	orange		

=COUNTIFS(B4:B103,"apple")

=COUNTIFS(B4:B103,"orange")

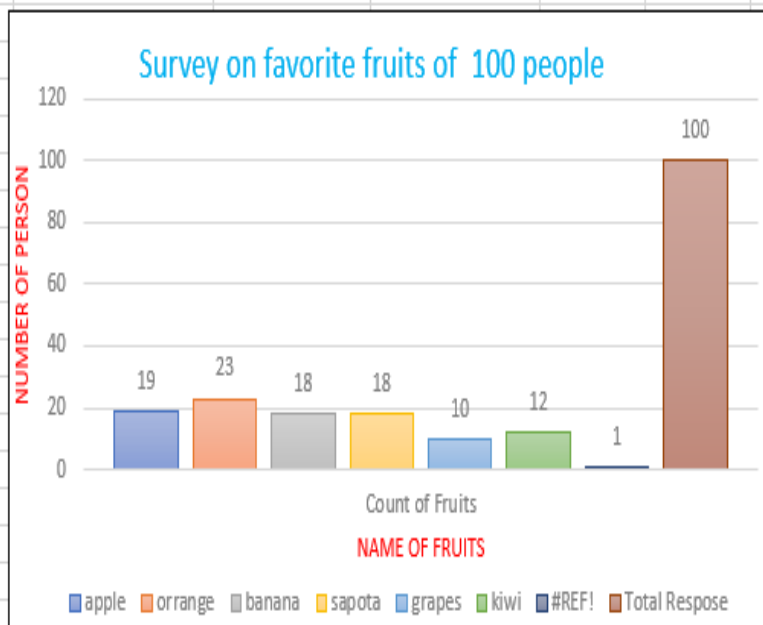
=COUNTIFS(B4:B103,"banana")

=COUNTIFS(B4:B103,"sapota")

=COUNTIFS(B4:B103,"grapes")

=COUNTIFS(B4:B103,"kiwi")

=SUM(E4:E9)



Person11	kiwi
Person12	sapota
Person13	banana
Person14	apple
Person15	sapota
Person16	banana
Person17	orange
Person18	sapota
Person19	apple
Person20	orange
Person21	kiwi
Person22	kiwi
Person23	orange
Person24	sapota
Person25	apple
Person26	sapota
Person27	orange
Person28	kiwi
Person29	orange
Person30	kiwi
Person31	apple

[illegible]

81	Person78	grapes								
82	Person79	banana								
83	Person80	orange								
84	Person81	grapes								
85	Person82	orange								
86	Person83	apple								
87	Person84	sapota								
88	Person85	grapes								
89	Person86	banana								
90	Person87	orange								
91	Person88	apple								
92	Person89	grapes								
93	Person90	banana								
94	Person91	apple								
95	Person92	sapota								
96	Person93	orange								
97	Person94	banana								
98	Person95	apple								
99	Person96	sapota								
100	Person97	orange								
101	Person98	banana								
102	Person99	grapes								
103	Person100	banana								

**INTERPRETATION:** From the bar graph, we can conclude that

- 1) Out of 100 people, 19 people like apple.
- 2) Out of 100 people, 23 people like orange.
- 3) Out of 100 people, 18 people like banana.
- 4) Out of 100 people, 18 people like sapota.
- 5) Out of 100 people, 10 people like grapes.
- 6) Out of 100 people, 12 people like kiwi.
- 7) Out of 100 people, most of the people like apple.
- 8) Out of 100 people, a smaller number of people like grapes.

**9)PURPOSE:** Bar graphs have been in widespread use everywhere from textbooks to newspapers, most audiences understand how to read a bar graph and can grasp the information the graph conveys.



**EXPERIMENT-8**

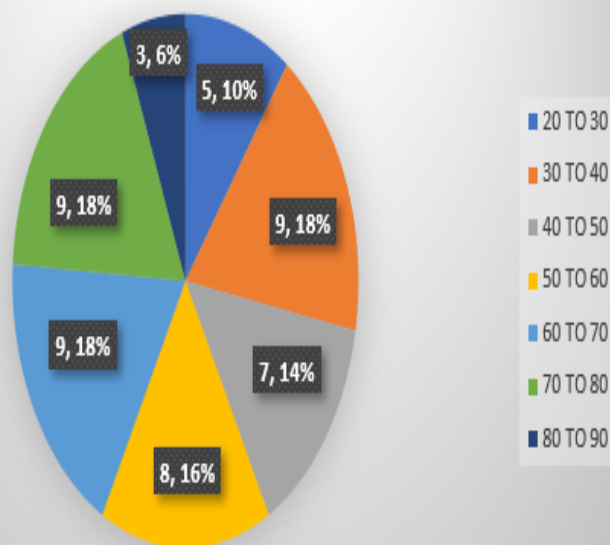
STATISTICS AND ANALYTICS LAB

RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

**AIM:Using Microsoft Excel Spread Sheet plot Pie Chart for data collected from 50 people (For example , Conduct a Survey on smokers with respect to their ages in your locality). Explain the pie chart with minimum 30 words**

Name of smokers	Age of smokers	Class interval	Frequency	
Smoker1	73	20 TO 30	5	=COUNTIFS(B7:B56,">=20",B7:B56,"<30")
Smoker2	25	30 TO 40	9	=COUNTIFS(B7:B56,">=30",B7:B56,"<40")
Smoker3	30	40 TO 50	7	=COUNTIFS(B7:B56,">=40",B7:B56,"<50")
Smoker4	35	50 TO 60	8	=COUNTIFS(B7:B56,">=50",B7:B56,"<60")
Smoker5	26	60 TO 70	9	=COUNTIFS(B7:B56,">=60",B7:B56,"<70")
Smoker6	30	70 TO 80	9	=COUNTIFS(B7:B56,">=70",B7:B56,"<80")
Smoker7	33	80 TO 90	3	=COUNTIFS(B7:B56,">=80",B7:B56,"<=90")
Smoker8	44	SUM	50	=SUM(E7:E13)
Smoker9	22			
Smoker10	65			
Smoker11	34			
Smoker12	56			
Smoker13	46			
Smoker14	57			
Smoker15	58			
Smoker16	68			
Smoker17	68			
Smoker18	76			
Smoker19	20			
Smoker20	57			
Smoker21	75			
Smoker22	45			
Smoker23	54			
Smoker24	34			
Smoker25	32			
Smoker26	23			
Smoker27	56			

**SURVEY ON SMOKERS WITH RESPECT TO THEIR AGE**

34	Smoker28	78									
35	Smoker29	56									
36	Smoker30	78									
37	Smoker31	88									
38	Smoker32	89									
39	Smoker33	67									
40	Smoker34	45									
41	Smoker35	45									
42	Smoker36	67									
43	Smoker37	78									
44	Smoker38	89									
45	Smoker39	67									
46	Smoker40	66									
47	Smoker41	78									
48	Smoker42	66									
49	Smoker43	35									
50	Smoker44	45									
51	Smoker45	33									
52	Smoker46	44									
53	Smoker47	55									
54	Smoker48	66									
55	Smoker49	77									
56	Smoker50	76									

**INTERPRETATION: From the pie chart, we can conclude that**

- 1) Out of 50 smokers, 5 members are in the age in between 20 to 30.
- 2) Out of 50 smokers, 9 members are in the age in between 30 to 40.
- 3) Out of 50 smokers, 7 members are in the age in between 40 to 50.
- 4) Out of 50 smokers, 8 members are in the age in between 50 to 60.
- 5) Out of 50 smokers, 9 members are in the age in between 60 to 70.
- 6) Out of 50 smokers, 9 members are in the age in between 70 to 80.
- 7) Out of 50 smokers, 3 members are in the age in between 80 to 90.
- 8) Out of 50 smokers, most of the smokers are in the age in between 30 to 40, 60 to 70 and 70 to 80.
- 9) Out of 50 smokers, a smaller number of smokers are in the age in between 80 to 90.

**10)PURPOSE:** A pie chart presents data as a simple and easy-to-understand picture. It can be an effective communication tool for even an uninformed audience, because it represents data visually as a fractional part of a whole. Readers or audiences see a data comparison at a glance, enabling them to make an immediate analysis or to understand information quickly.

EXPERIMENT-9

STATISTICS AND ANALYTICS LAB

RASHMI S

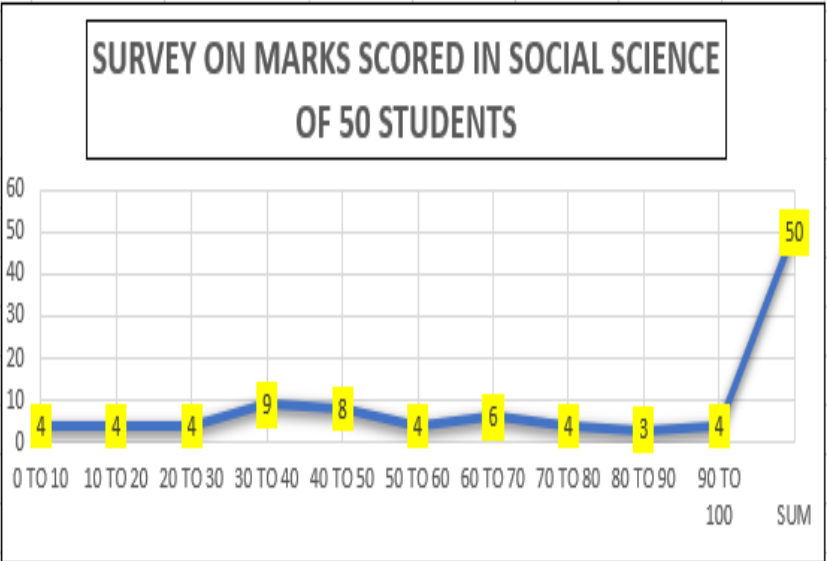
SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC,MYSURU

AIM:Using Microsoft Excel Spread Sheet draw a line graph for the given dataset  
(MARKS SCORED IN SOCIAL SCIENCE SUBJECT)

Name of Student	Marks in Social science
Student1	25
Student2	30
Student3	35
Student4	45
Student5	6
Student6	3
Student7	45
Student8	65
Student9	75
Student10	86
Student11	78
Student12	34
Student13	0
Student14	12
Student15	13
Student16	45
Student17	67
Student18	34
Student19	67
Student20	89
Student21	99
Student22	88
Student23	77
Student24	79

FREQUENCY DISTRIBUTION TABLE	
Class interval	Frequency
0 TO 10	4
10 TO 20	4
20 TO 30	4
30 TO 40	9
40 TO 50	8
50 TO 60	4
60 TO 70	6
70 TO 80	4
80 TO 90	3
90 TO 100	4
SUM	50

=COUNTIFS(B7:B56,">=0",B7:B56,"<10")  
=COUNTIFS(B7:B56,">=10",B7:B56,"<20")  
=COUNTIFS(B7:B56,">=20",B7:B56,"<30")  
=COUNTIFS(B7:B56,">=30",B7:B56,"<40")  
=COUNTIFS(B7:B56,">=40",B7:B56,"<50")  
=COUNTIFS(B7:B56,">=50",B7:B56,"<60")  
=COUNTIFS(B7:B56,">=60",B7:B56,"<70")  
=COUNTIFS(B7:B56,">=70",B7:B56,"<80")  
=COUNTIFS(B7:B56,">=80",B7:B56,"<90")  
=COUNTIFS(B7:B56,">=90",B7:B56,"<=100")



[illegible]

**INTERPRETATION:** From the line graph, we can conclude that

- 1) Out of 50 students, 4 students scored the marks in social science in between 0 to 10.
- 2) Out of 50 students, 4 students scored the marks in social science in between 10 to 20.
- 3) Out of 50 students, 4 students scored the marks in social science in between 20 to 30.
- 4) Out of 50 students, 9 students scored the marks in social science in between 30 to 40.
- 5) Out of 50 students, 8 students scored the marks in social science in between 40 to 50.
- 6) Out of 50 students, 4 students scored the marks in social science in between 50 to 60.
- 7) Out of 50 students, 6 students scored the marks in social science in between 60 to 70.
- 8) Out of 50 students, 4 students scored the marks in social science in between 70 to 80.
- 9) Out of 50 students, 3 students scored the marks in social science in between 80 to 90.
- 10) Out of 50 students, 4 students scored the marks in social science in between 90 to 100.
- 11) Out of 50 students, most of the students scored the marks in social science in between 30 to 40.
- 12) Out of 50 students, a smaller number of students scored the marks in social science in between 80 to 90.

**13)PURPOSE:** A line graph, also known as a line chart, is a type of chart used to visualize the value of something over time.

**EXPERIMENT-10**

STATISTICS AND ANALYTICS LAB

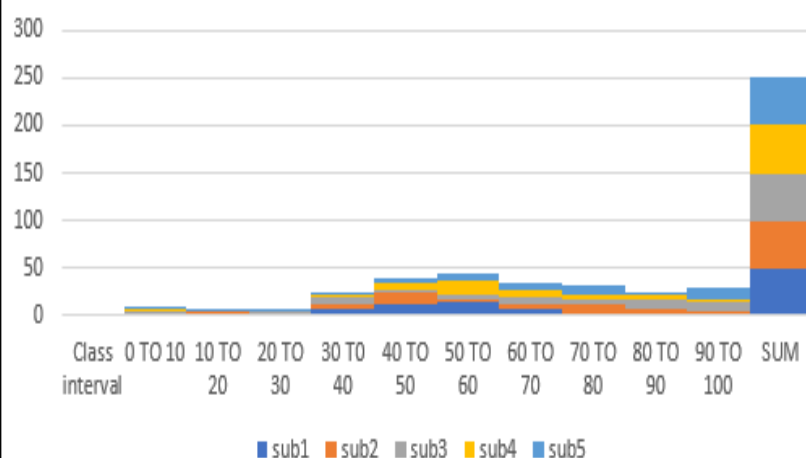
RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

**AIM: Using Microsoft Excel Spread Sheet draw frequency polygon and frequency curve for the data collected from 50 people (For example , Marks obtained by the students in your class in 5 subjects in previous examination). Explain your observations from the graph in minimum 30 words.**

Name of Student	Sub1	Sub2	Sub3	Sub4	Sub5	Class interval	sub1	sub2	sub3	sub4	sub5
Student1	43	35	36	46	56	0 TO 10	0	2	4	1	3
Student2	44	43	88	78	44	10 TO 20	3	2	0	0	1
Student3	34	3	8	90	30	20 TO 30	0	2	2	1	1
Student4	44	4	9	20	40	30 TO 40	8	3	8	2	2
Student5	55	56	7	3	59	40 TO 50	12	12	2	9	5
Student6	67	66	64	90	58	50 TO 60	15	1	7	15	7
Student7	78	67	67	89	56	60 TO 70	7	6	7	7	8
Student8	56	68	56	78	55	70 TO 80	3	10	3	5	10
Student9	45	69	55	78	67	80 TO 90	1	7	9	6	2
Student10	34	87	54	88	67	90 TO 100	1	5	8	4	11
Student11	32	86	54	87	80	SUM	50	50	50	50	50
Student12	12	85	33	86	90						
Student13	13	84	34	85	79						
Student14	14	83	35	95	78						
Student15	56	82	36	77	77						
Student16	45	81	36	44	75						
Student17	76	98	27	66	74						
Student18	88	99	47	33	73						
Student19	99	95	68	55	67						
Student20	56	94	57	44	66						
Student21	45	93	69	67	56						
Student22	34	34	60	56	56						
Student23	56	24	90	57	45						
Student24	76	13	98	58	45						
Student25	46	12	97	59	45						
Student26	45	22	96	60	33						
Student27	39	33	95	50	93						

**SURVEY ON MARKS SCORED IN 5 SUBJECTS OF 50 STUDENTS**



31	Student28	38	44	94	40	90						
32	Student29	37	45	93	30	91						
33	Student30	36	46	82	45	99						
34	Student31	45	47	83	46	92						
35	Student32	47	48	81	47	93						
36	Student33	48	49	84	48	94						
37	Student34	49	48	85	49	95						
38	Student35	50	78	97	51	96						
39	Student36	51	77	77	52	97						
40	Student37	52	76	88	53	88						
41	Student38	53	75	66	55	78						
42	Student39	54	74	8	54	66						
43	Student40	55	73	55	56	78						
44	Student41	56	72	77	57	66						
45	Student42	57	71	33	58	65						
46	Student43	58	70	22	59	64						
47	Student44	59	60	33	66	6						
48	Student45	60	41	44	67	3						
49	Student46	61	42	66	68	2						
50	Student47	62	42	77	69	21						
51	Student48	63	45	88	77	19						
52	Student49	64	66	88	88	78						
53	Student50	65	76	55	99	77						

**INTERPRETATION: From the line graph, we can conclude that**

- 1) Out of 50 students-0,2,4,1 & 3 students scored the marks in sub1, sub2, sub3, sub4 & sub5 respectively in between 0 to 10 marks.
- 2) Out of 50 students-3,2,0,0 & 1 students scored the marks in sub1, sub2, sub3, sub4 & sub5 respectively in between 10 to 20 marks.
- 3) Out of 50 students-0,2,2,1 & 1 students scored the marks in sub1, sub2, sub3, sub4 & sub5 respectively in between 20 to 30 marks.
- 4) Out of 50 students-8,3,8,2 & 2 students scored the marks in sub1, sub2, sub3, sub4 & sub5 respectively in between 30 to 40 marks.
- 5) Out of 50 students-12,12,2,9 & 5 students scored the marks in sub1, sub2, sub3, sub4 & sub5 respectively in between 40 to 50 marks.
- 6) Out of 50 students-15,1,7,15 & 7 students scored the marks in sub1, sub2, sub3, sub4 & sub5 respectively in between 50 to 60 marks.
- 7) Out of 50 students-3,10,3,5 & 10 students scored the marks in sub1, sub2, sub3, sub4 & sub5 respectively in between 60 to 70 marks.
- 8) Out of 50 students-1,7,9,6 & 2 students scored the marks in sub1, sub2, sub3, sub4 & sub5 respectively in between 70 to 80 marks.
- 9) Out of 50 students-1,7,9,6 & 2 students scored the marks in sub1, sub2, sub3, sub4 & sub5 respectively in between 80 to 90 marks.
- 10) Out of 50 students-1,5,8,4 & 11 students scored the marks in sub1, sub2, sub3, sub4 & sub5 respectively in between 90 to 100 marks.

**11)PURPOSE:** Frequency polygons are a graphical device for understanding the shapes of distributions. They serve the same purpose as histograms, but are especially helpful for comparing sets of data.

**EXPERIMENT-11**

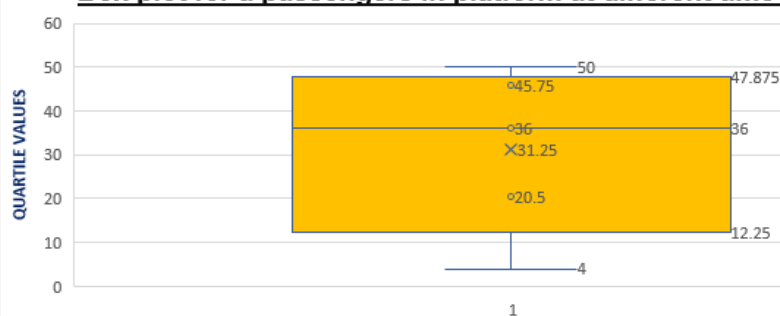
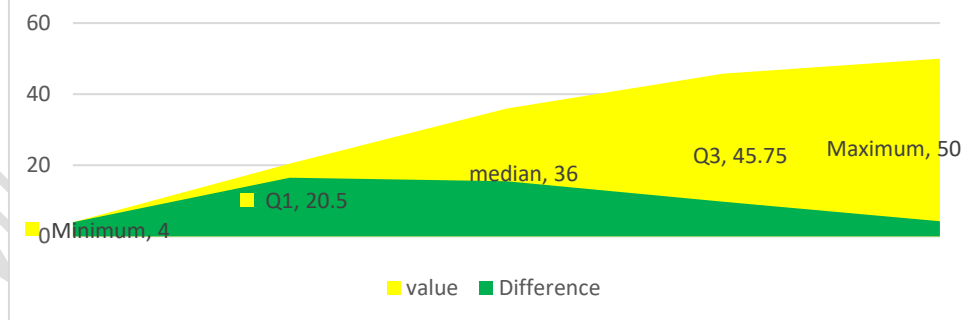
STATISTICS AND ANALYTICS LAB

RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

**AIM: Using Microsoft Excel Spread Sheet construct a box plot for the given dataset. (For example dataset can be the number of passengers in a flat form at different time in a day)**

Timing	passengers		Minimum	value	Difference		Minimum	=MIN(B7:B28)
7:12AM	45		Q1	20.5	16.5	=E8-E7	Quartile1	=QUARTILE(B7:B28,1)
8:20AM	49		median	36	15.5	=E9-E8	Median	=MEDIAN(B7:B28)
9:28AM	50		Q3	45.75	9.75	=E10-E9	Quartile3	=QUARTILE(B7:B28,3)
10:36AM	4		Maximum	50	4.25	=E11-E10	Maximum	=MAX(B7:B28)
11:44AM	50							
12:52PM	34							
2:00PM	25							
3:08PM	34							
4:16PM	12							

**Box plot for a passengers in platform at different time in a day****OR****Area chart plot for a passengers in platform at different time in a day**

**INTERPRETATION:** From the box plot OR area graph, we can conclude that

- 1) Minimum value in data is 4 and its difference is 4
- 2) Quartile 1 is 20.5 and its difference is 16.5
- 3) Median is 36 and its difference is 15.5
- 4) Quartile 3 is 45.75 and its difference is 9.75
- 5) Minimum value in data is 50 and its difference is 4.25

**6) PURPOSE:** Box plot is usually helpful in explanatory data. It indicates the spread out of data based on 5 number summaries namely minimum, Q1 (Quartile 1), Median, Q3 (Quartile 3), and Maximum.



**EXPERIMENT-12**

STATISTICS AND ANALYTICS LAB

RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

**AIM:Using Microsoft Excel Spread Sheet construct a leaf plot for the given dataset.Explain the graph with minimum 30 words**

Name of Student	Marks of maths	stem	leaf						
Student1	10	1	0	1	2	3	9		
Student2	11	2	0	1	4	6			
Student3	12	3	1	3	5	7			
Student4	13	4	0	1	1	3	4		
Student5	19	5	2	4	6				
Student6	20	6							
Student7	21	7							
Student8	24	8	9						
Student9	26	9	2	3	9				
Student10	31								
Student11	33								
Student12	35								
Student13	37								
Student14	40								
Student15	41								
Student16	41								
Student17	43								
Student18	44								
Student19	52								
Student20	54								
Student21	56								
Student22	89								
Student23	92								
Student24	93								
Student25	99								

**INTERPRETATION:** From the stem-leaf plot, we can conclude that

- 1) Out of 25 students, 5 students scored less than 20 marks in mathematics.
- 2) Out of 25 students, 13 students scored the marks in mathematics in between 20 to 40.
- 3) Out of 25 students, 7 students scored the marks in mathematics in between 50 to 90.
- 4) Minimum value is 10.
- 5) Maximum value is 99.

**6)PURPOSE:** A stem-and-leaf display (also known as a stem plot) is a diagram designed to allow you to quickly assess the distribution of a given dataset. It indicates the recurrence of data.



EXPERIMENT-13

STATISTICS AND ANALYTICS LAB

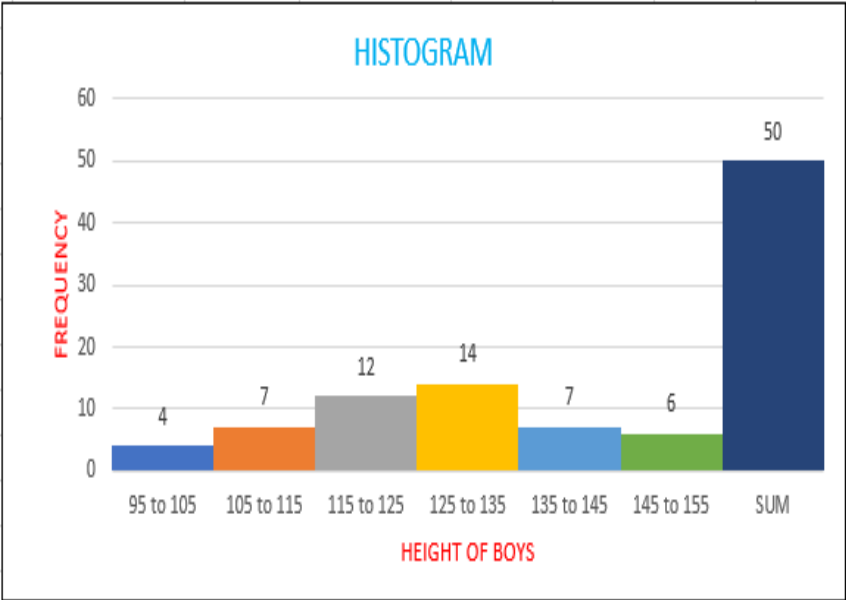
RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC,MYSURU

AIM: Using Microsoft Excel Spread Sheet find the Mean , Mode and Median for the data given and also represent them in Histogram

Name of Boys	Heights in cms	CLASS INTERVAL	Frequency	
Boy1	96	95 to 105	4	=COUNTIFS(B5:B54,">=95",B5:B54,"<105")
Boy2	99	105 to 115	7	=COUNTIFS(B5:B54,">=105",B5:B54,"<115")
Boy3	101	115 to 125	12	=COUNTIFS(B5:B54,">=115",B5:B54,"<125")
Boy4	102	125 to 135	14	=COUNTIFS(B5:B54,">=125",B5:B54,"<135")
Boy5	105	135 to 145	7	=COUNTIFS(B5:B54,">=135",B5:B54,"<145")
Boy6	106	145 to 155	6	=COUNTIFS(B5:B54,">=145",B5:B54,"<=155")
Boy7	108	SUM	50	=SUM(E5:E10)
Boy8	110			

Boy9	110	mean	125.7	=AVERAGE(B5:B54)
Boy10	112	median	126.5	=MEDIAN(B5:B54)
Boy11	113	mode	132	=MODE(B5:B54)
Boy12	115			
Boy13	116			
Boy14	117			
Boy15	118			
Boy16	120			
Boy17	121			
Boy18	122			
Boy19	121			
Boy20	122			
Boy21	123			
Boy22	124			
Boy23	125			
Boy24	120			
Boy25	126			
Boy26	127			
Boy27	128			



32	Boy28	129								
33	Boy29	130								
34	Boy30	131								
35	Boy31	132								
36	Boy32	132								
37	Boy33	133								
38	Boy34	133								
39	Boy35	134								
40	Boy36	135								
41	Boy37	134								
42	Boy38	132								
43	Boy39	136								
44	Boy40	137								
45	Boy41	138								
46	Boy42	139								
47	Boy43	142								
48	Boy44	144								
49	Boy45	145								
50	Boy46	146								
51	Boy47	147								
52	Boy48	148								
53	Boy49	149								
54	Boy50	152								

**INTERPRETATION:** From the histogram, we can conclude that

- 1) Out of 50 boys, 4 boys having a height in between 95 to 105 cm.
- 2) Out of 50 boys, 7 boys having a height in between 105 to 115 cm.
- 3) Out of 50 boys, 12 boys having a height in between 115 to 125 cm.
- 4) Out of 50 boys, 14 boys having a height in between 125 to 135 cm.
- 5) Out of 50 boys, 7 boys having a height in between 135 to 145 cm.
- 6) Out of 50 boys, 6 boys having a height in between 145 to 155 cm.
- 7) Mode is 132
- 8) Median is 126.5

9) **PURPOSE:** The central tendencies Mean, Mode and Median help us understand that has already taken place and predict future values as well.

EXPERIMENT-14

STATISTICS AND ANALYTICS LAB

RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

AIM:Generate a 51 random data sample (26 even and 25 odd number dataset) using Microsoft Excel spread sheet and determine the range and Quartiles.

Even Data set Values		Odd Data set values			
1	10	1	12		
2	13	2	13		
3	22	3	26		
4	24	4	35		
5	23	5	32		
6	24	6	54		
7	33	7	12		
8	33	8	65		
9	34	9	76		
10	34	10	46		
11	34	11	87		
12	34	12	35		
13	45	13	46		
14	67	14	98		
15	78	15	99		
16	88	16	89		
17	24	17	16		
18	56	18	36		
19	98	19	39		
20	35	20	47		
21	82	21	72		
22	49	22	62		
23	18	23	59		
24	23	24	35		
25	90	25	26		
26	37				

Even data set		Odd data set		
MINIMUM	10	MINIMUM	12	=MIN(H5:H29)
QUARTILE1	24	QUARTILE1	32	=QUARTILE(H5:H29,1)
QUARTILE2	34	QUARTILE2	46	=QUARTILE(H5:H29,2)
QUARTILE3	37	QUARTILE3	65	=QUARTILE(H5:H29,3)
MAXIMUM	98	MAXIMUM	99	=MAX(H5:H29)
RANGE	88	RANGE	87	=K12-K8
INTER QUARTILE RANGE(IQR)	13	INTER QUARTILE RANGE(IQR)	33	=K11-K9

RANGE=MAXIMUM-MINIMUM

INTER QUARTILE RANGE(IQR)=Q3-Q1

**INTERPRETATION:** From the above data, we can conclude that

**FOR EVEN DATA SET:**

- 1) Minimum value in data is 10
- 2) Quartile- 1 is 24
- 3) Quartile-2 is 34
- 4) Quartile-3 is 37
- 5) Maximum value in data is 98
- 6) Range is 88
- 7) Inter quartile range is 13

**FOR ODD DATA SET:**

- 1) Minimum value in data is 12
- 2) Quartile- 1 is 32
- 3) Quartile-2 is 46
- 4) Quartile-3 is 65
- 5) Maximum value in data is 99
- 6) Range is 87
- 7) Inter quartile range is 33

**8)PURPOSE:** The quartiles are especially useful when working with data that isn't symmetrically distributed, or a data set that has outliers

**EXPERIMENT-15**

STATISTICS AND ANALYTICS LAB

RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

**AIM: Collect the current yield of a crop from 50 different persons (problem statement can be changed according to priorities of the tutor) in your locality and determine mean deviation and Quartile deviation in Microsoft excel sheet and brief your inference with less than 30 words.**

Person Name	Wheat crop yield(X) in quintal	X-Mean	Absolute	Mean of X	39.1	=AVERAGE(B4:B53)
Person1	15	-24.1	24.1	Q1	16.25	=QUARTILE(B4:B53,1)
Person2	56	16.9	16.9	Q3	56	=QUARTILE(B4:B53,3)
Person3	76	36.9	36.9	Quartile deviation	19.875	=(H5-H4)/2
Person4	45	5.9	5.9	mean deviation (average of absolute)	21.952	=AVERAGE(D4:D53)
Person5	34	-5.1	5.1			
Person6	32	-7.1	7.1	Quartile deviation=(Q3-Q1)/2		
Person7	31	-8.1	8.1	mean deviation= average of absolute		
Person8	12	-27.1	27.1			
Person9	13	-26.1	26.1			
Person10	14	-25.1	25.1			
Person11	15	-24.1	24.1			
Person12	16	-23.1	23.1			
Person13	13	-26.1	26.1			
Person14	4	-35.1	35.1			
Person15	45	5.9	5.9			
Person16	46	6.9	6.9			
Person17	57	17.9	17.9			
Person18	33	-6.1	6.1			
Person19	22	-17.1	17.1			
Person20	33	-6.1	6.1			
Person21	44	4.9	4.9			
Person22	56	16.9	16.9			
Person23	67	27.9	27.9			
Person24	78	38.9	38.9			
Person25	79	39.9	39.9			

29	Person26	70	30.9	30.9								
30	Person27	80	40.9	40.9								
31	Person28	89	49.9	49.9								
32	Person29	98	58.9	58.9								
33	Person30	90	50.9	50.9								
34	Person31	67	27.9	27.9								
35	Person32	56	16.9	16.9								
36	Person33	54	14.9	14.9								
37	Person34	53	13.9	13.9								
38	Person35	52	12.9	12.9								
39	Person36	51	11.9	11.9								
40	Person37	12	-27.1	27.1								
41	Person38	13	-26.1	26.1								
42	Person39	14	-25.1	25.1								
43	Person40	15	-24.1	24.1								
44	Person41	16	-23.1	23.1								
45	Person42	17	-22.1	22.1								
46	Person43	18	-21.1	21.1								
47	Person44	19	-20.1	20.1								
48	Person45	20	-19.1	19.1								
49	Person46	21	-18.1	18.1								
50	Person47	22	-17.1	17.1								
51	Person48	23	-16.1	16.1								
52	Person49	24	-15.1	15.1								
53	Person50	25	-14.1	14.1								

**INTERPRETATION:** From the above data, we can conclude that

- 1) Mean of X is 39.1
- 2) Quartile-1 is 16.25
- 3) Quartile-3 is 56
- 4) Quartile deviation is 19.875
- 5) Mean deviation is 21.952

**6)PURPOSE:** The Mean deviation is an important descriptive statistic that is not frequently encountered in mathematical statistics. This is essentially because while mean deviation has a natural intuitive definition as the "mean deviation from the mean," the introduction of the absolute value makes analytical calculations using this statistic much more complicated than the standard deviation

EXPERIMENT-16

STATISTICS AND ANALYTICS LAB

RASHMI S  
SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

AIM:Collect data of any 2 livestock population from 50 different houses in your locality (problem statement can be changed according to priorities of the tutor) and determine the standard deviationfor both the two separately microsoft excel spread sheet and brief your inference with less than 30 words.

Livestock population		
Houses	Cattle	Buffalos
H1	3	6
H2	4	8
H3	5	9
H4	2	8
H5	4	7
H6	5	7
H7	6	6
H8	7	65
H9	8	67
H10	9	66
H11	89	58
H12	7	58
H13	67	57
H14	66	56
H15	56	55
H16	45	45
H17	34	54
H18	43	44
H19	23	45
H20	34	37
H21	56	36
H22	55	35
H23	23	34
H24	22	33
H25	21	32
H26	12	23

Standard deviation	
Cattle	22.59948 =STDEV.P(B5:B54)
Buffalos	20.58179 =STDEV.P(C5:C54)

31	H27	11	39											
32	H28	10	40											
33	H29	15	41											
34	H30	16	42											
35	H31	17	43											
36	H32	18	44											
37	H33	19	45											
38	H34	20	46											
39	H35	21	47											
40	H36	22	48											
41	H37	23	49											
42	H38	25	50											
43	H39	26	51											
44	H40	4	52											
45	H41	65	53											
46	H42	5	54											
47	H43	6	55											
48	H44	7	76											
49	H45	7	67											
50	H46	6	8											
51	H47	78	3											
52	H48	6	6											
53	H49	58	3											
54	H50	56	2											

**INTERPRETATION:** From the above data, we can conclude that

1) Standard deviation of cattle is 22.59

2) Standard deviation of buffalos is 20.58

**3) PURPOSE:** Standard deviation tells us about the shape of our distribution, how close the individual data values are from the mean value.



**EXPERIMENT-17**

STATISTICS AND ANALYTICS LAB

RASHMI S

SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC, MYSURU

**AIM: Collect the data of two wheeler (with a rider and a pillion) crossing a busy junction in your locality in the peak hours (problem statement can be changed according to priorities of the tutor) and determine the variance of the data in microsoft excel spread sheet and brief your inference with less than 30 words.**

Junction	No. of two wheeler crossing the junction										
J1	56										
J2	23	Variance	10062.55	=VAR.P(B4:B28)							
J3	34										
J4	45										
J5	556										
J6	56										
J7	76										
J8	55										
J9	44										
J10	56										
J11	77										
J12	112										
J13	123										
J14	14										
J15	23										
J16	45										
J17	66										
J18	77										
J19	88										
J20	90										
J21	88										
J22	56										
J23	44										
J24	56										
J25	67										

**INTERPRETATION:** From the above data, we can conclude that

1) Variance is 10062.55

**2) PURPOSE:** Standard deviation and variance tells us about the shape of our distribution, how close the individual data values are from the mean value.

EXPERIMENT-18

	A	B	C	D	E	F	G	H	I	J
	STATISTICS AND ANALYTICS LAB					RASHMI S SENIOR SCALE LECTURER, GOVT CPC POLYTECHNIC,MYSURU				
	AIM:Using Microsoft Excel spread sheet draw a Skewness graph and kurtosis graph for randomly generated dataset.									
	Name of Student		Marks in maths							
	Student1		12	Mean	30.48	=AVERAGE(B4:B28)				
	Student2		8	Median	32	=MEDIAN(B4:B28)				
	Student3		67							
	Student4		45	Kurtosis	-0.94569	=KURT(B4:B28)				
	Student5		6							
	Student6		3							
	Student7		45							
	Student8		13							
	Student9		75							
	Student10		12							
	Student11		32							
	Student12		34							
	Student13		0							
	Student14		1							
	Student15		13							
	Student16		10							
	Student17		67							
	Student18		34							
	Student19		43							
	Student20		41							
	Student21		2							
	Student22		36							
	Student23		77							
	Student24		79							
	Student25		7							

INTERPRETATION: From the above data, we can conclude that

- 1)Mean is 30.48
- 2)Median is 32
- 3)Kurtosis is -0.94569

4)PURPOSE: Skewness and kurtosis is useful in finding the symmetry and peaked Ness of the data distribution.

NOTE: If the mean appears before the median, then declare that the data is negatively skewed and positively skewed otherwise

**PYTHON PROGRAMS****EXPERIMENT-19**

**#WRITE A PYTHON PROGRAM TO CONVERT DECIMAL TO BINARY,OCTAL AND HEXADECIMAL.**

```
dec=int(input("enter number"))
```

```
print(bin(dec))
```

```
print(oct(dec))
```

```
print(hex(dec))
```

**EXPERIMENT-20****#WRITE A PYTHON PROGRAM TO ADD TWO INTEGERS AND TWO STRINGS.**

```
num1=int(input("enter first number"))
```

```
num2=int(input("enter second number"))
```

```
add=num1+num2
```

```
print(add)
```

```
string1="govt cpc polytechnic"
```

```
string2=" mysore"
```

```
add=string1+string2
```

```
print(add)
```

**EXPERIMENT-21****#WRITE A PYTHON PROGRAM TO FIND THE SUM OF FIRST 10 NATURAL NUMBERS****num=10****sum=0****while(num>0):****sum=num+sum****num=num-1****print(sum)**

**EXPERIMENT-22****#WRITE A PYTHON PROGRAM TO FIND WHETHER THE NUMBER IS ODD OR EVEN**

```
num=int(input("enter number"))  
if(num%2)==0:  
    print("the number is even")  
else:  
    print("the number is odd")
```

**EXPERIMENT-23**

**#WRITE A PYTHON PROGRAM TO FIND THE VARIANCE AND STANDARD DEVIATION.**

```
import statistics  
sample=(130,145,128,100,136)  
variance=statistics.pvariance(sample)  
print(variance)  
standarddeviation=statistics.pstdev(sample)  
print(standarddeviation)
```

**EXPERIMENT-24**

**#WRITE A PYTHON PROGRAM TO ENTER THE MARKS OF THE STUDENTS ACROSS THE SUBJECT.**

```
sub1=int(input("kannada"))
sub2=int(input("mathematics"))
sub3=int(input("science"))
sub4=int(input("social"))
sub5=int(input("english"))
add=sub1+sub2+sub3+sub4+sub5
print(add)
percentage=(add*100)/500
print(percentage)
```