GOVERNMENT OF KARNATAKA

DEPARTMENT OF COLLEGIATE

AND

TECHNICAL EDUCATION

K.V.T
POLYTECHNC,CHICKBALLAPUR

DETAILED CURRICULUM FOR

COURSE: STATISTICS AND ANALYTICS LAB MANUAL
(FIRST/SECOND SEMESTER)

Course CODE: 20SC02P

(COMMON TO ALL BRANCHES C-20)

Experiment -1

Prepare a questionnaire (closed end) containing 25 questions for a specified problem statement: for example experience of an individual in a restaurant.

Experiment -2

Prepare a Google form for a specified problem statement to collect the dataset. (For example questionnaire to conduct online quiz)

Experiment -3

Send out a survey on your problem statement to number of 50(By Google forms) and collect the data.

Experiment -4

Remove duplicate or irrelevant observations. Remove unwanted observations from the dataset provided, including duplicate observations or irrelevant observations.

Experiment -5

In Microsoft Excel Spread sheet draw the frequency

distribution table for the given data (data set should contain minimum 50 data)

Aim: To prepare Microsoft excel spread sheet and to draw frequency distribution table for the data.

Purpose: A data in excel spread sheet will be helpful for collecting data set and analyses in charts.

Steps involved:

Step1: Enter your data into a worksheet.

Step2: Find the class interval and upper limit of data set.

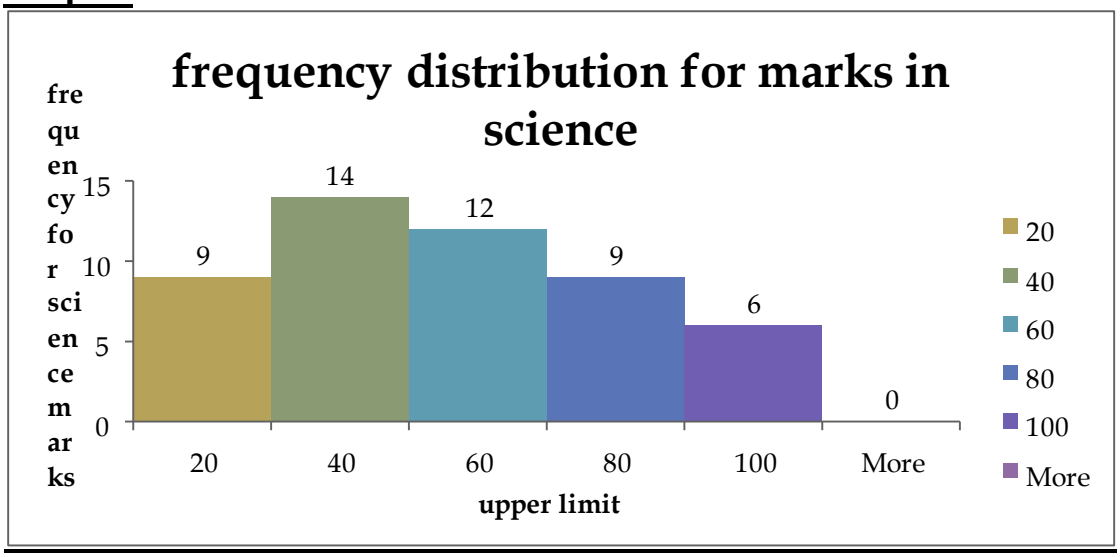
Step3: In the data menu, choose data analysis to open the data analysis dialog box.

Step4: In the data analysis box, select Histogram and click OK button.

Step5: In histogram, enter the data array into the input range box, Bin range (upper limit data), output range, chart output & click OK to display the frequency distribution in charts.

Step6: On Bar charts right click to find format series option then go to series option to make Gap width 0% & vary colors by point using fill button.

Output:



Experiment-6

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

Aim: To prepare Microsoft excel spread sheet to draw the relative frequency distribution table for the data.

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

Steps involved:

Step1: Enter your data into a worksheet.

Step2: Find the class interval and upper limit of data set.

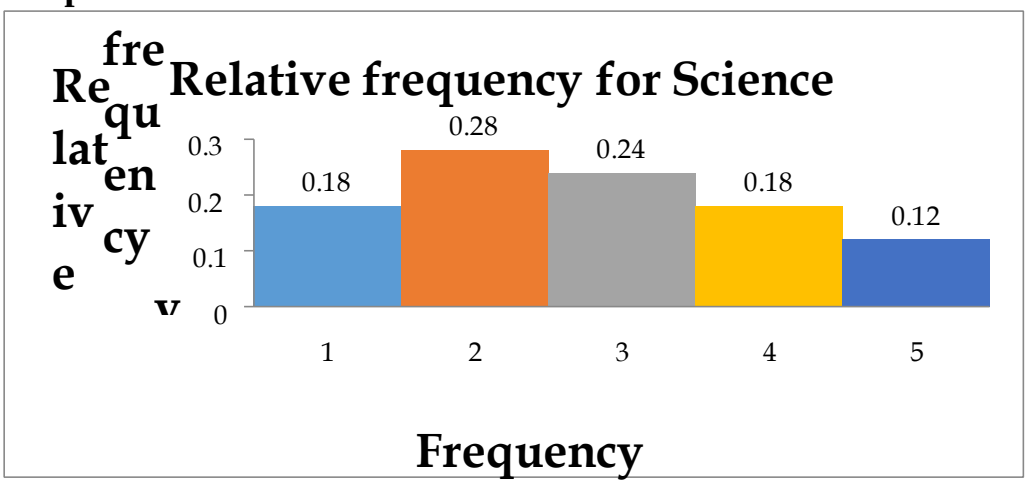
Step3: In the data menu, choose data analysis to open the data analysis dialog box.

Step4: In the data analysis box, select Histogram and click OK button.

Step5: In histogram, enter the data array into the input range box, Bin range (upper limit data), output range, & click OK button.

Step6: To find relative frequency use the relation each frequency value by total frequency, then frequency and relative frequency values to display in charts. (On Bar charts right click to find format series option then go to series option to make Gap width 0% & vary colors by point using fill button).

Output:



Experiment-7

Using Microsoft excel spread sheet plot bar graph for the data collected from 100 people (for example, conduct a survey on the favourite fruit of a person in your locality (restricting to 5 to 6 fruits).Explain the bar graph with minimum 30 words.

Aim: To conduct survey on favorite fruit of 100 persons using excel spread sheet and to plot bar graph for the collected data.

Purpose: The data in excel spread sheet is helpful for collecting data set and analyze in bar graph charts.

Steps involved:

Step1: Enter your data into a worksheet.

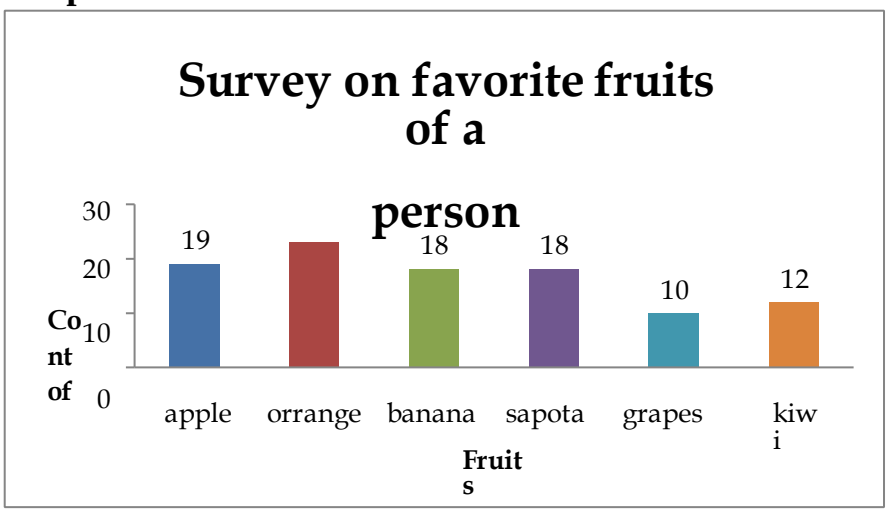
Step2: Count number of fruits with respect to their name to individual person, using Count if relation & find total response.

Step3: Select the data that go into the chart (data cell having name of fruits and count of fruits).

Step4: Select insert menu, recommended bar chart & choose the chart you like.

Step5: Modify the chart by adding title name, axis tittles and format data labels.

Output:



Experiment-8

Using Microsoft excel spread sheet plot pie chart for the data collected from 50 people (for example, conduct a survey on the smokers with respect to their ages in your locality).Explain the pie chart with minimum 30 words.

Aim: To conduct survey on favorite fruit of 100 persons using excel spread sheet and to plot bar graph for the collected data.

Purpose: The data in excel spread sheet is helpful for collecting data set and analyze in bar graph charts.

Steps involved:

Step1: Enter your data into a worksheet.

Step2: Find count of age of smokers with respect to their range.

Step3: Select insert menu to find pivot table & click on pivot table to enter table data & select existing worksheet to display in excel cell.

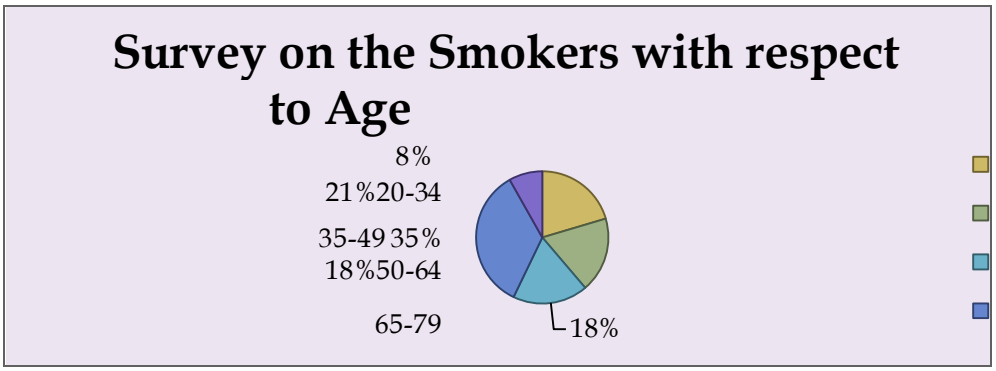
Step4: On right side of screen i.e., in pivot table fields, drag more table values between areas rows and values.

Step5: Right click on any row label values for grouping, i.e., define starting at and ending value with common difference then press ok button.

Step6: Right click on age of smoker values for summarizing data into count, i.e., go to summarize values & click count.

Step7: Select the data that go into the pie chart (data cell having row labels & count of age smokers).

Step8: Select insert menu, recommended pie chart & choose the chart you like.
Step9: Modify the chart by adding title name.
Output:



Experiment-9

Using Microsoft Excel Spread Sheet draw a line graph for the given dataset (The daily Pocket expenses of 206 students in a School).

Aim: To draw line graph for the given dataset.

Purpose: The data in excel spread sheet is helpful for collecting data set and analyze in line graph charts.

Steps involved:

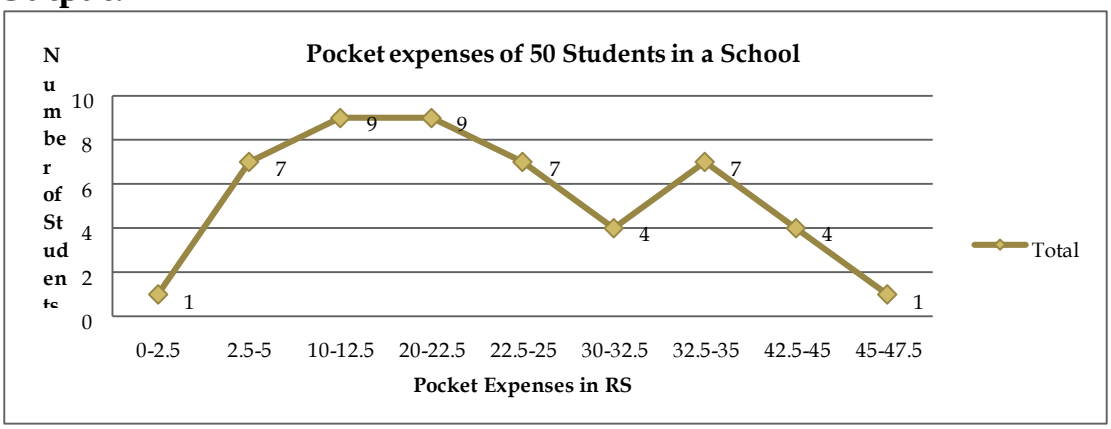
Step1: Enter your data into a worksheet.

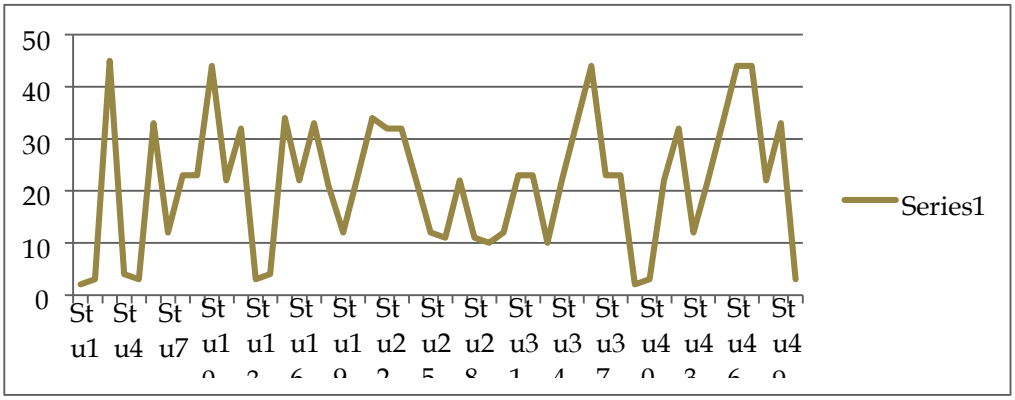
Step2: Select the data that go into the chart (data cell having student name and pocket expenses).

Step3: Select insert menu, recommended line chart & choose the chart you like.

Step5: Modify the chart by adding tittle name, axis tittles and format data labels.

Output:





Experiment -10

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.

Aim: To prepare Microsoft excel spread sheet and to draw frequency polygon and frequency curve for the data.

Purpose: A data in excel spread sheet will be helpful for collecting data set and analyses in charts.

Steps involved:

Step1: Enter your data into a worksheet.

Step2: Find the class interval and upper limit of data set.

Step3: In the data menu, choose data analysis to open the data analysis dialog box.

Step4: In the data analysis box, select Histogram and click OK button.

Step5: In histogram, enter the data array into the input range box, Bin range (upper limit data), output range, cumulative percentage & click OK to display the frequency distribution in charts.(Individual for each subject follow the same steps)

Step6: Add the first row value by `0` in all subjects.

Step7: Select the data that go into the polygon line chart (data cell having Frequency column of all Subjects).

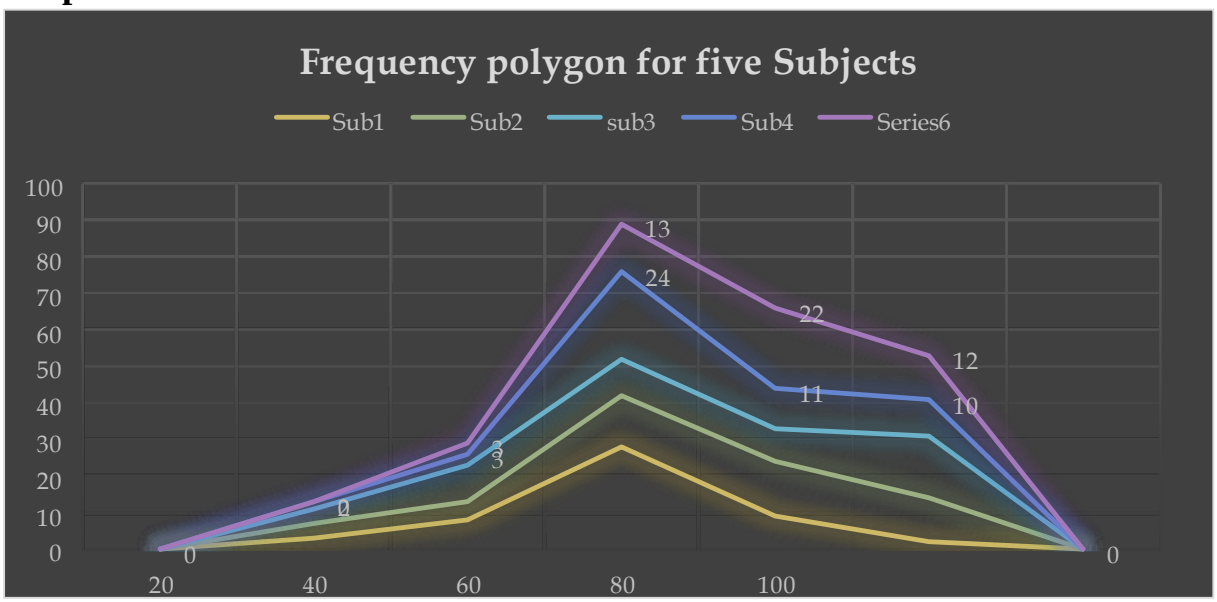
Step8: Select insert menu, go to inserted line chart & choose the chart you like.

Step9: Modify the line chart by adding tittle name, axis tittles and format data labels.

Step10: To draw frequency polygon curve, select data of cumulative frequency of all subjects then select insert menu, go to scatter then select scatter with smooth lines.

Step11: Modify the Curve by adding tittle name, axis tittles and format data labels.

Output:



Experiment-11

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). **Aim:** To draw box plot for the given dataset.

Purpose: The data in excel spread sheet is helpful for collecting data set and analyze in line box plot charts.

Steps involved:

Step1: Enter your data into a worksheet.

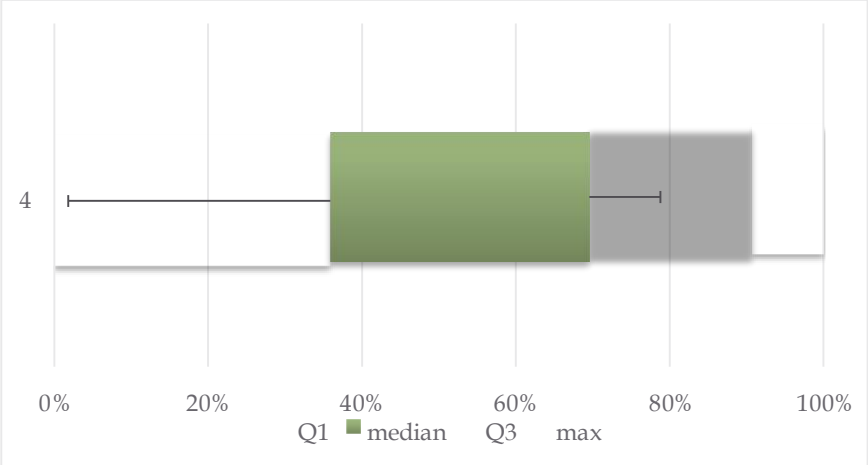
Step2: Select the data which is in passenger's column to find minimum, quartile1, median, quartile3 and maximum values using the following relations

For Minimum value: =MIN (B8:B29)
Quartile value: =QUARTILE (B8:B29,1)
Median value: =MEDIAN (B8:B29)
Quartile value: =QUARTILE (B8:B29,3)
Maximum value: =MAX (B8:B29)

Step3: Find the difference above values by keeping minimum value as constant.

Step4: Select the data that go into the box plot chart (data cell having functions and difference column).

Step5: Select insert menu, recommended box plot chart & choose the chart you like.
Step6: Modify the chart by adding tittle name, axis tittles and format data labels.
Output:



Experiment-12

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

To construct a leaf plot for the given dataset.

Purpose: The data in excel spread sheet is helpful for finding leaf of data values.

Steps involved:

Step1: Enter your data into a worksheet

Step2: Find maximum and minimum value of data values using the relation
Min: =AVERAGE(Data range)
Max: =MAX(Data range)
Step3: Enter stem value from 1 to 9
Step4: To find leaf of data values use relation,
Leaf: =REPT("0 ",COUNTIF($B$10:$B$34,E16*10+0))&REPT("1

",COUNTIF($B$10:$B$34,E16*10+1))&REPT("2

",COUNTIF($B$10:$B$34,E16*10+2))&REPT("3

",COUNTIF($B$10:$B$34,E16*10+3))&REPT("4

",COUNTIF($B$10:$B$34,E16*10+4))&REPT("5

",COUNTIF($B$10:$B$34,E16*10+5))&REPT("6

",COUNTIF($B$10:$B$34,E16*10+6))&REPT("7

",COUNTIF($B$10:$B$34,E16*10+7))&REPT("8

",COUNTIF($B$10:$B$34,E16*10+8))&REPT("9

",COUNTIF($B$10:$B$34,E16*10+9)) for stem value 1, similarly use the same relation to find leaf to other stem values up to 9th value.

Output:

Stem	leaf
1	0 1 2 3 9
2	5 5
3	2 4 6
4	7 8
5	8 9
6	0 5
7	5 6 7 8
8	9 9
9	2 3 9

13

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

Aim: To find the Mean, Mode and Median for the given data and also represent them in Histogram.

Purpose: The data in excel spread sheet will be helpful for collecting data set and finding mean, mode and median in charts.

Steps involved:

Step1: Enter your data into a worksheet.

Step2: Find mean, mode and median using the default functions.

Mean: =AVERAGE(Data range)

Mode: =MODE(Data range)

Median: =MEDIAN(Data range)

Step3: Select insert menu to find pivot table & click on pivot table to enter table data & select existing worksheet to display in excel cell.

Step4: On right side of screen i.e., in pivot table fields, drag more table values between areas rows and values.

Step5: Right click on any row label values for grouping, i.e., define starting at and ending value with common difference then press ok button.

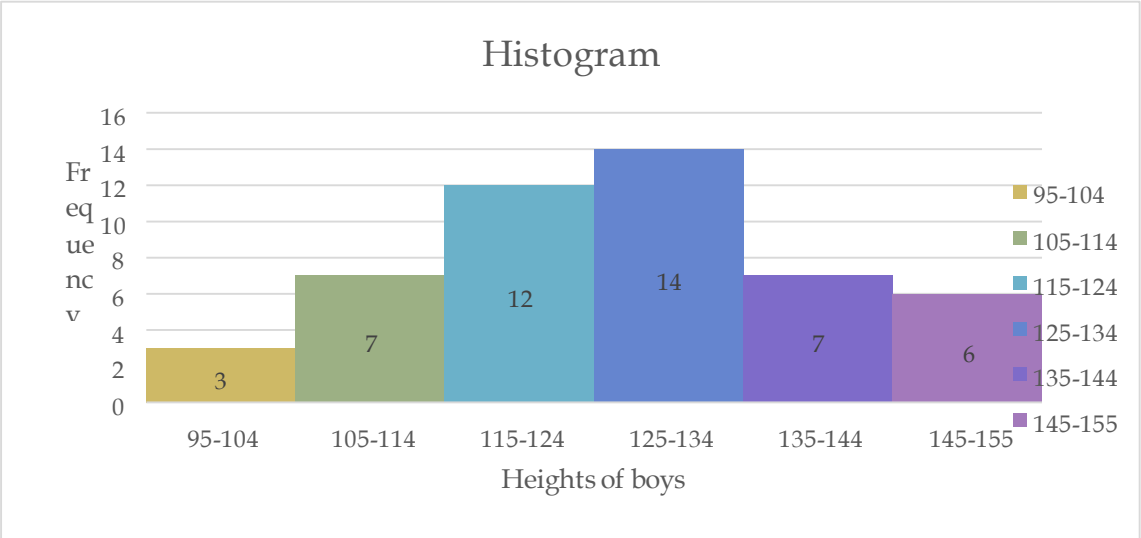
Step6: Right click on age of Height values for summarizing data into count, i.e., go to summarize values & click count.

Step7: Select the data that go into the Histogram chart (data cell having row labels & Height of count).

Step8: Select insert menu, recommended bar plot chart & choose the chart you like.

Step9: Modify the chart by adding tittle name, axis tittles and format data labels. . (Right click on any one of bar chart then go to data series & make gap width 0%, vary colors by point using fill button).

Output:


Histogram
Frequency
16
14
12
10
8
6
4
2
0
3
7
12
14
7
6
95-104
105-114
115-124
125-134
135-144
145-155
Heights of boys

14

Generate a 50 random data sample (even and odd number dataset) using Microsoft Excel spread sheet and determine the range and Quartiles.

Aim: To generate 50 random data containing even and odd dataset and to determine range and quartile.

Purpose: The data in excel spread sheet will be helpful for collecting data set and finding mean, mode and median in charts.

Steps involved:

Step1: Enter your even and odd data set into a worksheet.

Step2: To find Quartile (Q) i.e., Q1, Q2 and Q3 for even dataset use the following relations.

Q1: QUARTILE(Datarange,1)

Q2: QUARTILE(datarange,2)

Q3: QUARTILE(datarange,3)

Step3: To find Quartile range use the relation IQR: Q3-Q1 (cells) **Step4:**

Similarly we can find for odd data set using the same dataset.

Output:

Even data set	
Q1	23
Q2	34
Q3	37
IQR	14

Odd data set	
Q1	23
Q2	33
Q3	34
IQR	11

Experiment-15

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.

Aim: To find the Mean deviation and quartile deviation using excel sheet.

Purpose: The data in excel spread sheet will be helpful for collecting data set and finding mean deviation and quartile deviation.

Steps involved:

Step1: Enter your data into a worksheet.

Step2: Find mean using the default functions.

Mean: =AVERAGE(Data range)

Step3: Based on mean find mean deviation using the relation

Mean deviation: =B7-$E$7 (B7-crop yield, E7- mean)

Step4: To find absolute value of deviation use

Absolute: =ABS(deviation) and sum the value using default relation.

Step5: To calculate mean deviation use
MD: =H57/50 (ABS value/No. of observations)
Step6: To find Quartile deviation first we need to find Q3 and Q1 using quartile relation. Q1: QUARTILE(Datarange,1) Q3: QUARTILE(datarange,3)
Next we find quartile deviation using relation QD: (Q3-Q1)/2 Where Q3 and Q1 are cell values with respect to excel sheet.**Output:**

Experiment-

Q3	56
Q1	16.25
Quartile deviation	19.875
Mean Deviation	22.025

16

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 deviation of the data in Microsoft excel spread sheet and brief your inference with less than 30 words.

Aim: To collect two livestock population from different houses and to determine the standard deviation of the data in excel.

Purpose: The data in excel spread sheet will be helpful for collecting data set and to find standard deviation of the data in excel.

Steps involved:

Step1: Enter your data into a worksheet having livestock population of cattle and buffalo from different houses.

Step2: To find standard deviation of cattle and buffalo use the following default relation.

Cattle: =STDVES.P(data range) Buffalo:

= STDVES.P(data range) **Output:**

Standard deviation	
Cattle	Buffalos
22.59948	20.58178807

Experiment-

17

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.

Aim: To collect two livestock population from different houses and to determine the variance of the data in excel.

Purpose: The data in excel spread sheet will be helpful for collecting data set and to find variance of the data in excel.

Steps involved:

Step1: Enter your data into a worksheet having number of two wheeler Crossing junction.

Step2: To find variance use the default relation Variance: =VAR.P(data range)

Output:

Variance
10062.5536

Experiment-

18

Using Microsoft Excel spread sheet draw a Skewness graph and kurtosis graph for randomly generated dataset.

Aim: To determine skewness and kurtosis for randomly generated dataset.

Purpose: The data in excel spread sheet will be helpful for collecting data set and to find skewness and kurtosis.

Steps involved:

Step1: Enter your data value into a worksheet

Step2: generate the random values using the default relation =RANDOMBETWEEN(1,100)

Step3: To find skewness and kurtosis go to data then data analysis, select descriptive statistics and click ok button.

Output:

Experiment-

Column1	
Mean	46.1
Standard Error	6.779574587
Median	42.5
Mode	32
Standard Deviation	30.31917927
Sample Variance	919.2526316
Kurtosis	-1.186990977
Skewness	0.272572213
Range	98
Minimum	2
Maximum	100
Sum	922
Count	20

Experiment -19

#Write a python program to convert decimal to binary,octal& hexadecimal.

Input

```
Dec=int(input("Enter the number"))
Print("decimal value of dec","is")
Print(bin(dec),"in binary")
Print(oct(dec),"in octal")
Print(hex(dec),"in hexadecimal")
```

Output:

```
decimal number is :   15
binary value is :   0b1111
octal value is  :   0o17
hexadecimal value is  :   0xf
```

Experiment -20

 # Write a python program to add 2 integers and 2 strings and print the result.

```
Num1=int(input("enter the first number"))
Num2=int(input("enter the second number"))
X=num1+num2
Print("sum",x)
```

Output:

```
Enter first number 3
    Enter second number 18
            Sum 21
```

Experiment -21

 # Write a python program to find the sum of first 10 natural numbers

Input

```
n=int(input("enter the number"))
sum1=0
while(n>0):
    sum1=sum1+n
    n=n-1
print("the sum of first n natural number",sum1)
```

Output:

```
the sum of first 10 natural numbers is :  55
```

Experiment -22

 # Write a python program to find whether the number is odd or even

```
Num=int(input("enter any number"))
Flag=num%2
If flag==0:
 Print(num,"is an even number")
Elif flag==1:
  Print(num,"is an odd number")
Else:
  Print(num,"invalid input")
```

Output:

```
the number is ODD
```

Experiment -23

 # Write a python program to find the variance and standard deviation for the given data.

Input

```
import statistics
data=[526,147,656,123,532]
var=statistics.pvariance(data)
print("the variance is:",var)
sd=statistics.pstdev(data)
print("the standard devition is:",sd)
```

Output:

```
The variance is :  14.8
The standar deviation is :  3.847076812334269
```

Experiment -24

 # Write a python program to display student marks from the record.

```
file = open("student1", "r")
lines = file.readlines()  for
x in lines :  print(x)
file.close()
```

Output:

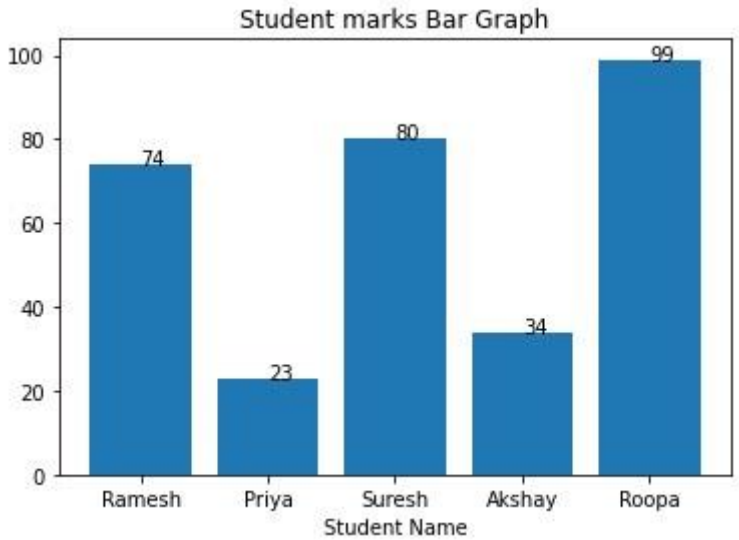
```
Student.txt
Student Name        : Student Marks
Ramesh              : 99
Mruthunjaya         : 90
Manjunath           : 90
Suresh              : 100
Santhosh            : 100
Somanna             : 100
```

Experiment -25

 # write a python program to create a labeled bar graph using matplot.pyplot.

```
Import matplotlib.pyplot as p
X=["m1","m2","m3"]
Y=[100,80,95]
p.bar(x,y,0.4)
p.xlabel("subject")
p.ylabel("marks")
p.show()
```

Output:



Experiment -26

 # Write a python program to create a labeled pie chart using matplotlib.pylot.

```
import matplotlib.pyplot as p
x=[“m1”, “m2”, “m3”]
y=[100,80,95]
p.pie(y)
p.show()
```

Output: