KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University Established Under Section 3 of UGC Act 1956)
Coimbatore – 641 021.

SEMESTER – II

		L	T	P	C
17PAU211	**CORE – STATISTICAL PACKAGE USING SPSS (PRACTICAL)**	**-**	**-**	**3**	**1**

Practical List

1. Introduction to SPSS Package
2. Working with windows of SPSS
3. Defining variables in variable view window in SPSS
4. Drawing of Simple and multiple bar diagrams in SPSS Package
5. Drawing of Histogram and Pie diagram
6. Calculation of Mean for individual, discrete series using SPSS Package.
7. Mean for continuous series using SPSS Package.
8. Median for individual and discrete series using SPSS Package..
9. Median for continuous series using SPSS Package..
10. Mode for individual and discrete series using SPSS Package..
11. Standard deviation for individual and discrete series using SPSS Package.
12. Coefficient of variation for individual and discrete series using SPSS Package.
13. Karl Pearson's Correlation using SPSS Package.
14. Rank Correlation Coefficient using SPSS Package.

INTRODUCTION TO SPSS PACKAGE

Introduction: What is SPSS?

- Originally it is an acronym of Statistical Package for the Social Science but now it stands for Statistical Product and Service Solutions
- One of the most popular statistical packages which can perform highly complex data manipulation and analysis with simple instructions

The basics of managing data files:

Opening SPSS

Start → All Programs → SPSS Inc → SPSS

The default window will have the data editor.

Saving the data:

To save the data file you created simply click 'file' and click 'save as.' You can save the file in different forms by clicking "Save as type."

WORKING WITH WINDOWS OF SPSS

Windows in SPSS:

The Four Windows:

- Data editor
- Output viewer
- Syntax editor
- Script window

Data Editor

This window is a spreadsheet-like system for defining, entering, editing, and displaying data. Extension of the saved file will be "sav."

Output Viewer

This window displays output and errors. Extension of the saved file will be "spv."

Syntax editor

This window is a text editor for syntax composition. Extension of the saved file will be "sps."

Script Window

This window provides the opportunity to write full-blown programs, in a BASIC-like language. Extension of the saved file will be "sbs."

DEFINING VARIABLES IN VARIABLE VIEW WINDOW IN SPSS

Sheets in Data Editor:

There are two sheets in the window:

1. Data view
2. Variable view

Data View window

This sheet is visible when we first open the Data Editor and this sheet contains the data.

Variable View window

This sheet contains information about the data set that is stored with the dataset.

- **Name**
  - The first character of the variable of the variable name must be alphabetic.
  - Variable names must be unique, and have to be less than 64 characters.
  - Spaces are not allowed.
- **Type**
  - Click on the "type" box. The two basic types of variables that we will use are numeric and string. This column enables us to specify the type of variable.
- **Width**
  - Width allows us to determine the number of characters SPSS will allow to be entered for the variable
- **Decimals**
  - Number of decimals
  - It has to be less than or equal to 16
- **Label**
  - You can specify the details of the variable
  - You can write characters with spaces up to 256 characters

- **Values**
  - This is used and to suggest which numbers represent which categories when the variable represents a category

Defining the value labels

- Click the cell in the values column as
- For the value, and the label, we can put up to 60 characters.
- After defining the values click add and then click OK.

DRAWING OF SIMPLE BAR DIAGRAM

Question:

Guests staying at Marada Inn were asked to rate the quality of their accommodations as being excellent, above average, average, below average or poor. The ratings provided by a sample of 20 guest are,

Below Average	Average	Above Average
Above Average	Above Average	Above Average
Above Average	Below Average	Below Average
Average	Poor	Poor
Above Average	Excellent	Above Average
Average	Above Average	Average
Above Average	Average	

Aim:

To draw the simple bar diagram using SPSS package.

Calculation:

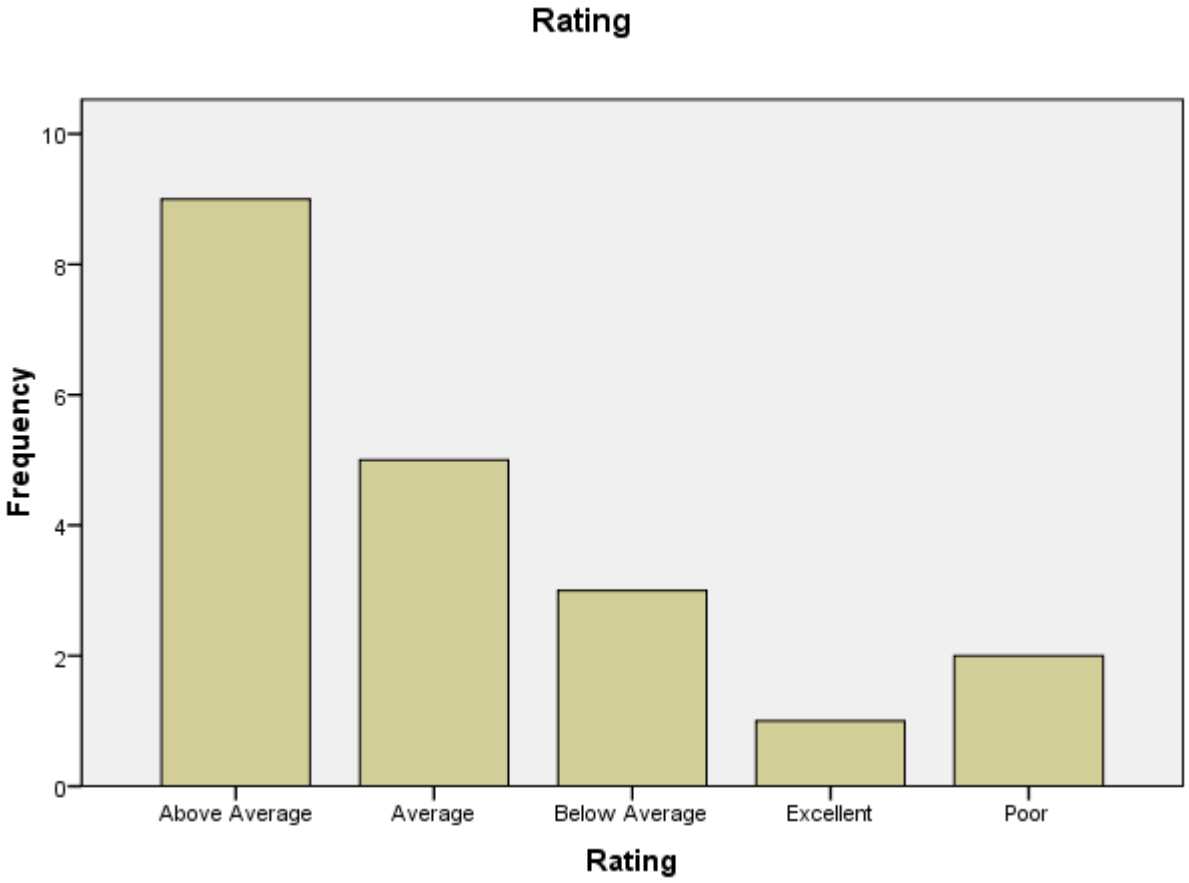
Rating	Frequency
Poor	2
Below Average	3
Average	5
Above Average	9
Excellent	1

SPSS Output:

Statistics

Rating

N	Valid	20
	Missing	0


Rating
Frequency
10
8
6
4
2
0
Above Average
Average
Below Average
Excellent
Poor
Rating

DRAWING OF MULTIPLE BAR DIAGRAM

Question:

The percentage distributions of household income in two regions are:

Income	Region A	Region B
1000	5	4
2000	6	5
3000	4	3
4000	3	2
5000	2	6

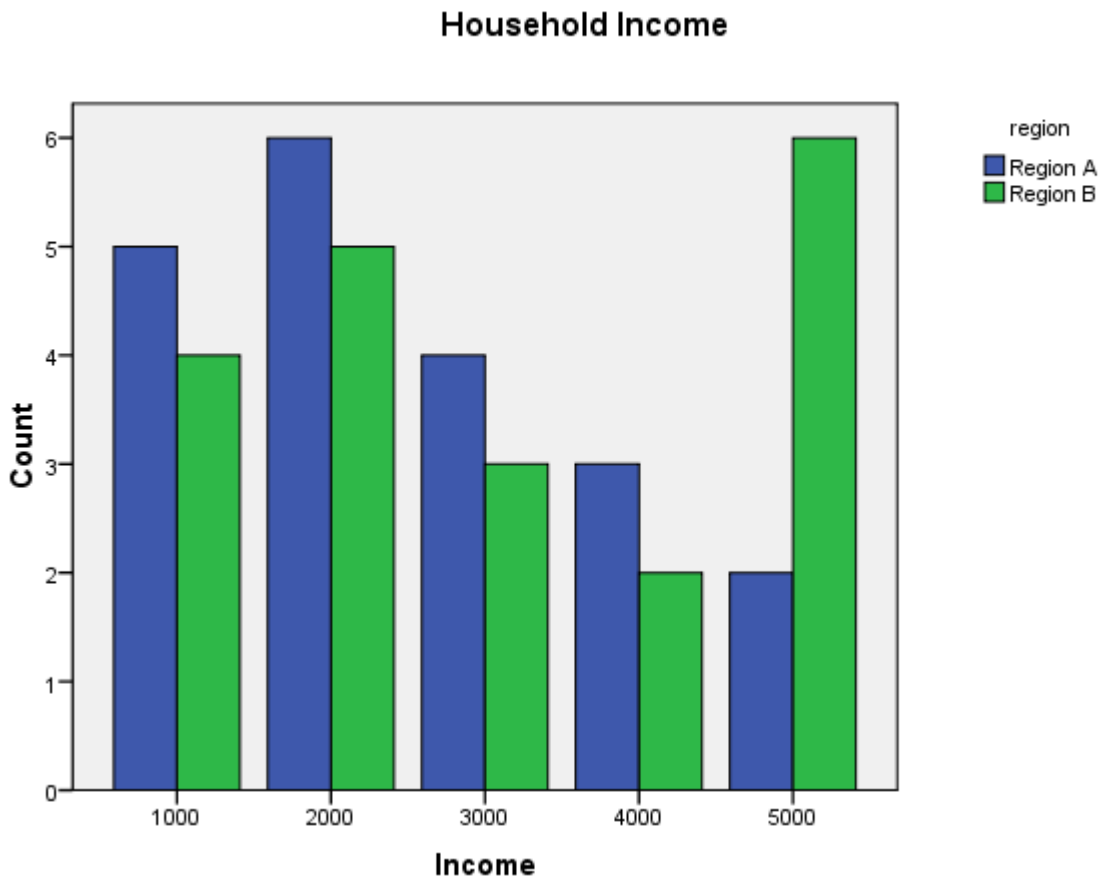
Aim:

To draw multiple bar diagram using SPSS package.

Calculation:

Income	Region A	Region B
1000	5	4
2000	6	5
3000	4	3
4000	3	2
5000	2	6
Total	20	20

SPSS Output:


Household Income
Count
0
1
2
3
4
5
6
1000
2000
3000
4000
5000
Income
region
Region A
Region B

DRAWING OF PIE DIAGRAM

Question:

Grades of 30 students in recent test:

A	B	C	D
5	12	10	3

Aim:

To draw pie chart using SPSS package.

Calculation:

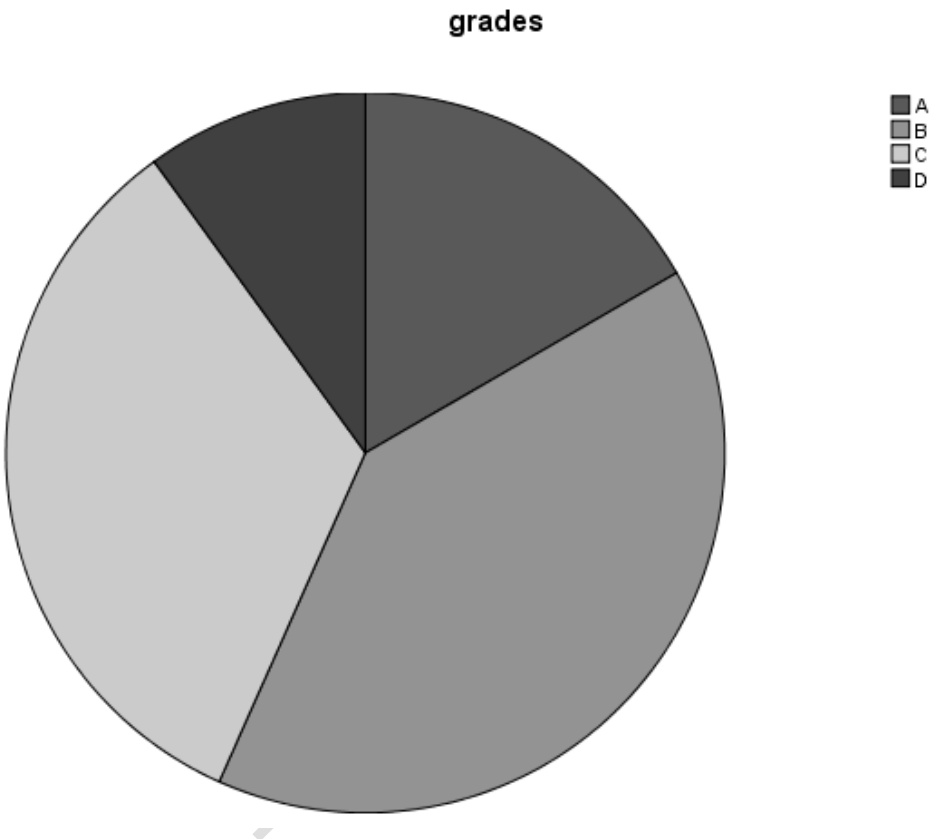
Grades	Frequency	Degrees
A	5	60 degrees
B	12	144 degrees
C	10	120 degrees
D	3	36 degrees
Total	30	360 degrees

SPSS Output:

Statistics

grades

N	Valid	30
	Missing	0



DRAWING OF HISTOGRAM

Question:

Daily wages of 80 workers in an industry are:

Daily Wages	Number of Workers
0 – 50	8
50 – 100	16
100 – 150	27
150 – 200	19
200 – 250	10

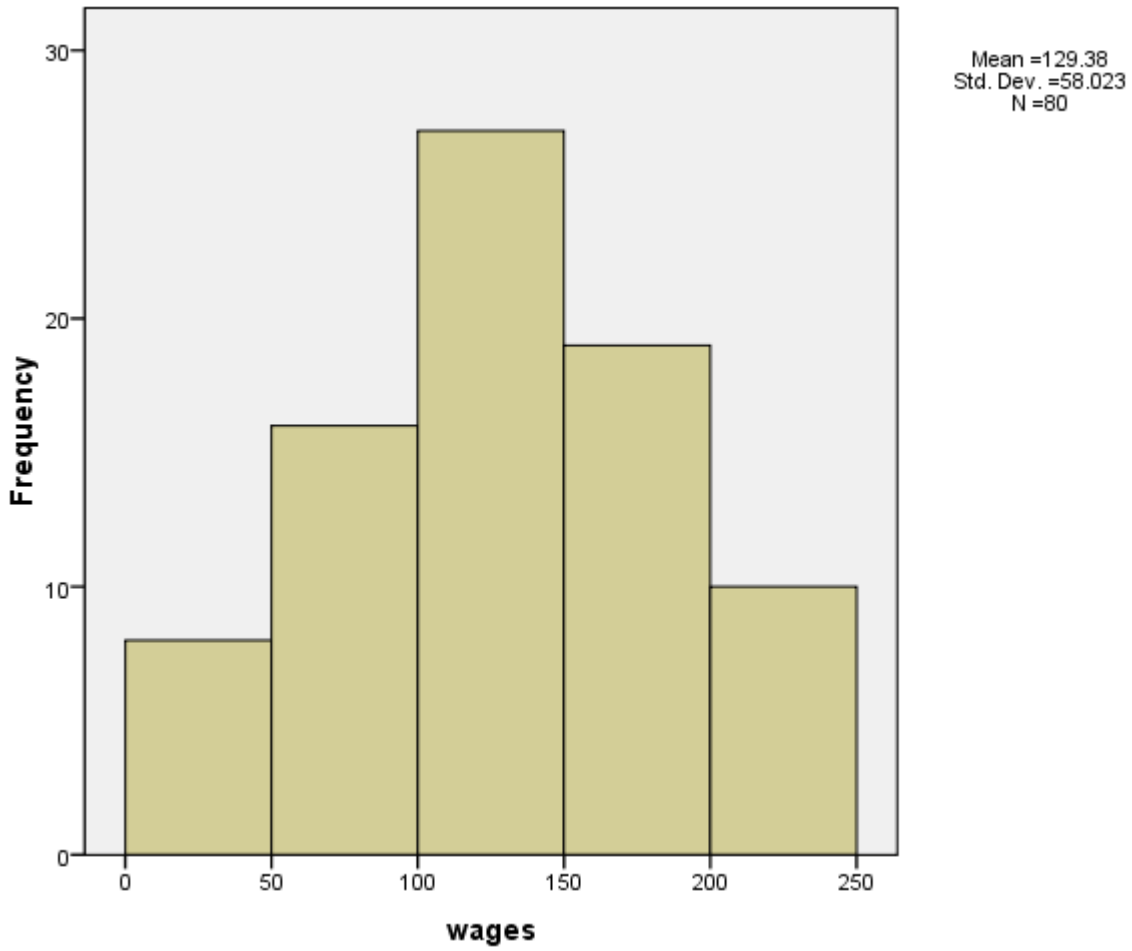
Aim:

To draw histogram using SPSS package.

Calculation:

Daily Wages	m	Number of Workers
0 – 50	25	8
50 – 100	75	16
100 – 150	125	27
150 – 200	175	19
200 – 250	225	10

SPSS Output:


Mean =129.38
Std. Dev. =58.023
N =80
Frequency
30
20
10
0
0
50
100
150
200
250
wages

CALCULATION OF MEAN FOR INDIVIDUAL SERIES

Question:

Calculate the arithmetic mean for the data given below:

X	25	18	27	10	30	42	20	53	20

Aim:

To calculate the arithmetic mean for individual series from the following data using SPSS package.

X	25	18	27	10	30	42	20	53	20

Formula:

Arithmetic Mean $\bar{X} = \Sigma X / N$

where N = number of items.

Calculation:

X
25
18
27
10
30
42
20
53
20

ΣX = 245

Arithmetic Mean = 245/9

=27.22

SPSS Output:

Statistics

X

N	Valid	9
	Missing	0
Mean		27.22

CALCULATION OF MEAN FOR DISCRETE SERIES

Question:

Calculate the arithmetic mean for the following data using SPSS package.

No of members (X)	1	2	3	4	5	6	7	8	9	10	11	12
Frequency (f)	1	3	5	6	10	13	9	5	3	2	2	1

Aim:

To calculate the arithmetic mean for discrete series from the following data using SPSS package.

No of members (X)	1	2	3	4	5	6	7	8	9	10	11	12
Frequency (f)	1	3	5	6	10	13	9	5	3	2	2	1

Formula:

Arithmetic Mean $\bar{X} = \Sigma fX / N$

where N = number of items.

Calculation:

X	f	fX
1	1	1
2	3	6
3	5	15
4	6	24
5	10	50
6	13	78
7	9	63
8	5	40
9	3	27
10	2	20
1	2	22
12	1	12
	$\Sigma f = 60$	$\Sigma fX = 358$

Arithmetic mean $\bar{X} = \Sigma fX/N$

$= 358/60$

$= 5.97$

SPSS Output:

Statistics

X

N	Valid	60
	Missing	0
Mean		5.97

CALCULATION OF MEAN FOR CONTINUOUS SERIES

Question:

Calculate the arithmetic mean for the following data using SPSS package.

Income	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70
Person	6	8	10	12	7	4	3

Aim:

To calculate the arithmetic mean for continuous series from the following data using SPSS package.

Income	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70
Person	6	8	10	12	7	4	3

Formula:

Arithmetic mean $\bar{X} = \Sigma fm/N$

where $N = \Sigma f$

Calculation:

X	m	f	fm
0 – 10	5	6	30
10 – 20	15	8	120
20 – 30	25	10	250
30 – 40	35	12	410
40 – 50	45	7	315
50 – 60	55	4	220
60 – 70	65	3	195
		$\Sigma f = 50$	$\Sigma fm = 1550$

Arithmetic mean $\bar{X} = \Sigma fm/N$

= 1550/50

= 31

SPSS Output:

Statistics

X

N	Valid	50
	Missing	0
Mean		31.00

CALCULATION OF MEDIAN FOR INDIVIDUAL SERIES

Question:

Calculate the median for the data given below:

X	25	18	27	10	30	42	20	53	20

Aim:

To calculate the median for individual series from the following data using SPSS package.

X	25	18	27	10	30	42	20	53	20

Formula:

Median = Size of [$(N+1)/2$] th item

Where N = number of observation

Calculation:

X
10
18
20
20
25
27
30
42
52

Median = Size of [$(N+1)/2$] th item

= Size of [$(9+1)/2$] th item

= Size of [$10/2$] th item

= Size of 5th item

Median = 25

SPSS Output:

Statistics

X

N	Valid	9
	Missing	0
Median		25.00

CALCULATION OF MEDIAN FOR DISCRETE SERIES

Question:

Calculate the median for the following data using SPSS package.

Wage(Rs)	50	75	100	150	250
No. of labours	8	14	10	5	3

Aim:

To calculate the median for discrete series from the following data using SPSS package.

Wage(Rs)	50	75	100	150	250
No. of labours	8	14	10	5	3

Formula:

Median = Size of [$(N+1)/2$] th item

where N = Σf

Calculation:

x	**F**	**CF**
50	8	8
75	14	22
100	10	32
150	5	37
250	3	40
	ΣF=40	

Median = Size of [$(N+1)/2$] th item

= Size of [$(40+1)/2$] th item

= Size of [$41/2$] th item

= Size of [20.5]th item

= (20^{th} term+21^{st} term)/2

= (75+75)/2

= 150/2

Median = 75

SPSS Output:

Statistics

X

N	Valid	40
	Missing	0
Median		75.00

CALCULATION OF MEDIAN FOR CONTINUOUS SERIES

Question:

Calculate the median for the following data using SPSS package.

Income	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70
Person	6	8	10	12	7	4	3

Aim:

To calculate the median for continuous series from the following data using SPSS package.

Income	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50	50 – 60	60 – 70
Person	6	8	10	12	7	4	3

Formula:

Median = $L+[\frac{N}{2} - cf]/f * i$

Position of median = N/2

where N = Σf

L = lower boundary of the class interval
cf = cumulative frequency of the class preceding the median class interval
f = frequency of the class interval
i = difference of the class interval

Calculation:

X	m	f	cf
0 – 10	5	6	6
10 – 20	15	8	14
20 – 30	25	10	24
30 – 40	35	12	36
40 – 50	45	7	43
50 – 60	55	4	47
60 – 70	65	3	50
		$\Sigma f = 50$	

Median= size of N/2th item

= size of 50/2th item

= size of 25th item

Median= 30-40

Median= L+$[\frac{N}{2} - cf]/f * i$

= 30+((25-24)/12) x10

= 30+ (1/12) x 10

= 30+(5/6)

= 30+0.83

Median=30.83

SPSS Output:

Statistics

X

N	Valid	50
	Missing	0
Median		30.45

CALCULATION OF MODE FOR INDIVIDUAL SERIES

Question:

Calculate the mode for the data given below:

X	25	18	27	10	30	42	20	53	20

Aim:

To calculate the mode for individual data using SPSS package.

X	25	18	27	10	30	42	20	53	20

Calculation:

x	**f**
10	1
18	1
20	2
25	1
27	1
30	1
42	1
53	1

Mode = most frequented value

= 20

SPSS Output:

Statistics

X

N	Valid	9
	Missing	0
Mode		20

CALCULATION OF MODE FOR DISCRETE SERIES

Question:

Calculate the mode for the following data using SPSS package.

Size	4	5	6	7	8	9	10	11	12	13
frequency	2	5	8	9	12	14	14	15	11	13

Aim:

To calculate the mode for discrete series using SPSS package.

Size	4	5	6	7	8	9	10	11	12	13
frequency	2	5	8	9	12	14	14	15	11	13

Calculation:

Grouping Table

size	Frequency I	II	III	IV	V	VI
4	2	7		15		
5	5		13		22	
6	8	17				
7	9		21			29
8	12	26		35		
9	14		**28**		**40**	
10	14	**29**				**43**
11	**15**		26	**40**		
12	11	24			39	
13	13					

Analysis Table:

Columns	size
I	11
II	10,11
III	9,10
IV	10,11,12
V	8,9,10
VI	9,10,11

The number 10 is repeated 5 times.

Therefore The mode is 10

Mode=10

SPSS Output:

Statistics

X

N	Valid	103
	Missing	0
Mode		11

CALCULATION OF STANDARD DEVIATION FOR INDIVIDUAL SERIES

Question:

Calculate the standard deviation for the data given below:

X	25	18	27	10	30	42	20	53	20

Aim:

To calculate the standard deviation for individual series using SPSS package.

X	25	18	27	10	30	42	20	53	20

Formula:

Standard Deviation = $\sqrt{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2}$

Calculation:

X	X^2
25	625
18	324
27	729
10	100
30	900
42	1764
20	400
53	2809
20	400
ΣX=245	ΣX^2=8051

Standard Deviation = $\sqrt{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2}$

$= \sqrt{894.55 - 740.928}$

`

$= \sqrt{153.622}$

= 12.394

SPSS Output:

Statistics

x

N	Valid	9
	Missing	0
Std. Deviation		13.141

CALCULATION OF STANDARD DEVIATION FOR DISCRETE SERIES

Question:

Calculate the standard deviation for the following data.

No of members (X)	1	2	3	4	5	6	7	8	9	10	11	12
Frequency (f)	1	3	5	6	10	13	9	5	3	2	2	1

Aim:

To calculate the standard deviation for discrete series from the following data using SPSS package.

No of members (X)	1	2	3	4	5	6	7	8	9	10	11	12
Frequency (f)	1	3	5	6	10	13	9	5	3	2	2	1

Formula:

Standard Deviation = $\sqrt{\frac{\Sigma fd^2}{\Sigma f} - \left(\frac{\Sigma fd}{\Sigma f}\right)^2}$

Calculation:

x	f	d=x-A (d=x-6)	d^2	fd	$f\,d^2$
1	1	-5	25	-5	25
2	3	-4	16	-12	48
3	5	-3	9	-15	45
4	6	-2	4	-12	24
5	10	-1	1	-10	10
6	13	0	0	0	0
7	9	1	1	9	9
8	5	2	4	10	20
9	3	3	9	9	27
10	2	4	16	8	32
11	2	5	25	10	50
12	1	6	36	6	36
	Σf=60			Σfd=-2	$\Sigma\, fd^2$=326

Standard Deviation = $\sqrt{\frac{\Sigma fd^2}{\Sigma f} - \left(\frac{\Sigma fd}{\Sigma f}\right)^2}$

$$\sigma = \sqrt{\left(\frac{326}{60}\right) - \left(-\frac{2}{60}\right)^2}$$

$\sigma = \sqrt{5.433 - 1.111}$

$\sigma = \sqrt{4.322}$

$\sigma = 2.078$

SPSS Output:

Statistics

X

N	Valid	60
	Missing	0
Std. Deviation		2.350

CALCULATION OF COEFFICIENT OF VARIATION FOR INDIVIDUAL SERIES

Question:

Calculate the coefficient of variation for individual series from the following data:

X	40	22	9	15	20	17	12	11

Aim:

To calculate the coefficient of variation for individual series using SPSS package.

X	40	22	9	15	20	17	12	11

Formula:

Coefficient of variation = $\frac{\sigma}{\bar{X}} \times 100$

where, σ = standard deviation

$\bar{X}$ = mean

Calculation:

x	**X^2**
40	1600
22	484
9	81
15	225
20	400
17	289
12	144
11	121
Σx=146	Σ X^2=3344

N = 8, $\Sigma X = 146$

$\bar{X} = \frac{146}{8} = 18.25$

$\sigma = \sqrt{\frac{\Sigma x^2}{N} - \left(\frac{\Sigma x}{N}\right)^2}$

$= \sqrt{\frac{3344}{8} - \left(\frac{146}{8}\right)^2}$

$= \sqrt{418 - (18.25)^2}$

$= \sqrt{84.9375}$

$\sigma = 9.216$

Coefficient of variation = $\frac{\sigma}{\bar{X}} * 100$

$= \frac{9.216}{18.25} * 100$

c.v= 50.498

SPSS Output:

Statistics

x

N	Valid	8
	Missing	0
Mean		18.25
Std. Deviation		9.852

Coefficient of variation = $\frac{\sigma}{\bar{X}} * 100$

$= \frac{9.85248}{18.2500} * 100$

c.v= 53.986

CALCULATION OF COEFFICIENT OF VARIATION FOR DISCRETE SERIES

Question:

Calculate the coefficient of variation for the discrete series from the following data:

Marks	10	20	30	40	50	60
Number of student	8	12	20	10	7	3

Aim:

To calculate the coefficient of variation for discrete series using SPSS package.

Marks	10	20	30	40	50	60
Number of student	8	12	20	10	7	3

Formula:

Coefficient of variation = $\frac{\sigma}{\bar{x}} * 100$

Calculation:

x	f	fx	x^2	fx^2
10	8	80	100	800
20	12	240	400	4800
30	20	600	900	18000
40	10	400	1600	16000
50	7	350	2500	17500
60	3	480	3600	28800
	N=65	Σfx=2150		Σ fx^2=85900

$$\overline{X} = \frac{\Sigma X}{N}$$

$$= \frac{2150}{65}$$

$$= 33.07$$

$$\sigma = \sqrt{\frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2}$$

$$= \sqrt{\frac{85900}{65} - \left(\frac{2150}{65}\right)^2}$$

$$= \sqrt{1321.53 - 1094.05}$$

$$= \sqrt{227.48}$$

$$\sigma = 15.082$$

Coefficient of variation = $\frac{\sigma}{\overline{X}} * 100$

$$= \frac{15.082}{33.07} * 100$$

= 45.606

SPSS Output:

Statistics

x

N	Valid	60
	Missing	0
Mean		30.83
Std. Deviation		13.566

Coefficient of variation = $\frac{\sigma}{\bar{X}} * 100$

$= \frac{13.566}{30.83} * 100$

= 44.00

CALCULATION OF KARL PEARSON'S CORRELATION

Question:

Calculate the Karl Pearson coefficient of correlation between two variables X and Y from the following data.

Height of father	64	65	66	67	68	69	70
Height of son	66	67	65	68	70	68	72

Aim:

To calculate the Karl Pearson coefficient of correlation for the following data using SPSS package.

Height of father	64	65	66	67	68	69	70
Height of son	66	67	65	68	70	68	72

Formula:

$$r = \frac{[\Sigma\, dx\, dy]}{\sqrt{[\Sigma(dx)^2 * \Sigma(dy)^2]}}$$

Calculation:

X	dx	dx^2	Y	dy	dy^2	dx dy
64	-3	9	66	-2	4	6
65	-2	4	67	-1	1	2
66	-1	1	65	-3	9	3
67	0	0	68	0	0	0
68	1	1	70	2	4	2
69	2	4	68	0	0	0
70	3	9	72	4	16	12
		$\Sigma(dx)^2=28$			$\Sigma(dy)^2=34$	Σ dx dy=25

$$r = \frac{[\Sigma \text{ dx dy}]}{\sqrt{[\Sigma(dx)^2 * \Sigma(dy)^2]}}$$

$$= \frac{25}{\sqrt{28*34}}$$

$$= 0.81$$

SPSS Output:

Correlations

		x	y
x	Pearson Correlation	1	.810
	Sig. (2-tailed)		.027
	N	7	7
y	Pearson Correlation	.810	1
	Sig. (2-tailed)	.027	
	N	7	7

CALCULATION OF RANK CORRELATION COEFFICIENT

Question:

Calculate rank correlation coefficient for the following data.

Price of Tea	88	90	95	70	60	75	50
Price of Coffee	120	134	150	115	110	140	100

Aim:

To calculate the rank correlation coefficient for the following data using SPSS package.

Price of Tea	88	90	95	70	60	75	50
Price of Coffee	120	134	150	115	110	140	100

Formula:

Rank correlation coefficient = $1-\frac{6\Sigma D^2}{N^3-N}$

Where, D = difference between x and y

N = number of observations

Calculation:

X	Y	RX	RY	D= (RX-RY)	D^2
88	120	3	4	-1	1
90	134	2	3	-1	1
95	150	1	1	0	0
70	115	5	5	0	0
60	110	6	6	0	0
75	140	4	2	2	4
50	100	7	7	0	0
					$\Sigma D^2=6$

Rank correlation coefficient = $1-\frac{6\Sigma D^2}{N^3-N}$

$= 1- [6(6^2)/(7^3-1)]$

$= 1-(36)/336$

$= 0.8928$

SPSS Output:

Correlations

			x	y
Spearman's rho	x	Correlation Coefficient	1.000	.893**
		Sig. (2-tailed)	.	.007
		N	7	7
	y	Correlation Coefficient	.893**	1.000
		Sig. (2-tailed)	.007	.
		N	7	7

**. Correlation is significant at the 0.01 level (2-tailed).