SPSS (Practical) **Semester – III**

18CCP312 **4H – 2C**

Instruction Hours / week: L: 0 T: 0 P : 4 **Marks: Internal: 40 External: 60 Total: 100**

End Semester Exam: 3 Hours

COURSE OBJECTIVES:

To make the students

1. To understand the Importance of SPSS and the features for entering the data according to the variabletype.
2. To understand and apply the descriptive analyticaltools
3. To know the univariate tools and itsapplication
4. To comprehend the application of Bivariateanalysis
5. To understand and compute the multivariate analysis using thepackage.

COURSEOUTCOMES:

Learners should be able to

1. Create datasheet and enter thedata
2. Compute descriptive statistics using the package and graphically represent thedata.
3. Perform univariate and bivariate analysis in the softwarepackage.
4. Perform multivariate analysis in the softwarepackage.
5. Demonstrate capabilities of problem-solving, critical thinking, and communication skills to infer theoutput.

UNIT I Overview and Data Entry

SPSS – Meaning – Scope- Limitation- Data view- Variable view- Data entry procedures- Data editing- Missing

UNIT II Descriptive Statistics

Descriptive statistics – Frequencies Distribution – Diagram –Graphs, Mean, Median, Mode, Skewness – Kurtosis – Standard Deviation.

UNIT III Non parametric and parametric test

Cross tabulation, Chi square, t test, independent sample t test, paired t test.

UNIT IV Analysis of Variance, Bivariate Analysis

ANOVA – One way, Two Way ANOVA, Correlation – Rank correlation – Regression – charts.

UNIT V Multivariate analysis

Factor Analysis, Cluster Analysis and Discriminate analysis.

Use the inbuilt case studies in SPSS for applying the statistical test.

SUGGESTED READINGS:

1. Darren George, Paul Mallery (2016), *IBM SPSS Statistics 23 Step by Step*, Routledge, NewDelhi.
2. Asthana &Braj Bhushan (2017), *Statistics for Social Sciences (With SPSS Applications)*, PHI,New Delhi.
3. Keith Mccormick, Jesus Salcedo, Aaron Poh, *SPSS Statistics for Dummies*, 3rd edition, Wiley, New Delhi.
4. Keith McCormick, Jesus Salcedo, Jon Peck, Andrew Wheeler, Jason Verlen (2017), *SPSS Statistics for Data Analysis and Visualization*, Wiley, NewDelhi.
5. Brian C. Cronk (2016), *How to Use SPSS®: A Step-By-Step Guide to Analysis and Interpretation*, 9thedition, Routledge, NewDelhi.

Ex. No. 1

DATA ENTRY AND MISSING VALUES

Aim

To enter data in to SPSS and to ascertain missing values.

Algorithm

Algorithm

Step 1:	Start the Process
Step 2:	Type the Qualitative Data in variable 1 and do not enter any data in two or three cells
Step 3:	Go to Variable view and in Label mention Gender
Step 4:	Click values button and assign 1 as Male and 2 as female for gender
Step 5:	Select Descriptive Statistics option from Analyze Menu
Step 6:	Select Frequencies option from Descriptive Sub menu
Step 7:	Forward the Variable 1
Step 8:	Click Ok button on Frequency Window
Step 9:	Stop the process

Result

Thus the data have been entered and missing values are identified.

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	6	50.0	60.0	60.0
	Female	4	33.3	40.0	100.0
	Total	10	83.3	100.0	
Missing	System	2	16.7		
Total		12	100.0		

Ex.No.2

DESCRIPTIVE STATISTICS

Aim

To compute Mean, Median, Mode, Standard Deviation, Skewness and Kurtosis

Algorithm

Step 1:	Start the Process
Step 2:	Type the following Quantitative Data 320, 395, 342, 444, 551, 395, 425, 417, 395, 401, 390, 400 in Column Variable 1 of Data view window
Step 3:	Select Descriptive Statistics option from Analyze Menu
Step 4:	Select Frequencies option from Descriptive Sub menu
Step 5:	Forward the Variable 1 data to Variables Window
Step 6:	Select Statistics Command button on Frequencies window
Step 7:	Select Mean, Median and Mode from Central Tendency Option and Standard Deviation from Dispersion option Skewness and Kurtosis from Distribution option and click Continue command button
Step 8:	Click Ok button on Frequency Window
Step 9:	Stop the process

Result

The above statistical analysis has been verified by using SPSS Package.

Output

Statistics

VAR00001

N	Valid	12
	Missing	0
Mean		406.25
Median		397.5
Mode		395.00
Std. Deviation		56.50603
Skewness		1.317
Std. Error of Skewness		.637
Kurtosis		3.906
Std. Error of Kurtosis		1.232

VAR00001

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	320	1	8.3	8.3	8.3
	342	1	8.3	8.3	16.7
	390	1	8.3	8.3	25.0
	395	3	25.0	25.0	50.0
	400	1	8.3	8.3	58.3
	401	1	8.3	8.3	66.7
	417	1	8.3	8.3	75.0
	425	1	8.3	8.3	83.3
	444	1	8.3	8.3	91.7
	551	1	8.3	8.3	100.0
	Total	12	100.0	100.0	

Ex.No.3

Frequency Distribution

Aim

To prepare a frequency distribution table

Algorithm

Step 1:	Start the Process
Step 2:	Type the Qualitative Data in variable 1 and variable 2.
Step 3:	Go to Variable view and in Label mention Gender and Education for Variable 1 and Variable 2 respectively
Step 4:	Click values button and assign 1 as Male and 2 as female for gender. Assign 1 for Under Graduate, 2 for Post Graduate and 3 for Professional for Education.
Step 5:	Select Descriptive Statistics option from Analyze Menu
Step 6:	Select Frequencies option from Descriptive Sub menu
Step 7:	Forward the Variable 1 and Variable 2 data to Variables Window
Step 8:	Click chart command button and select required chart type
Step 9:	Click Ok button on Frequency Window
Step 10:	Stop the process

Result

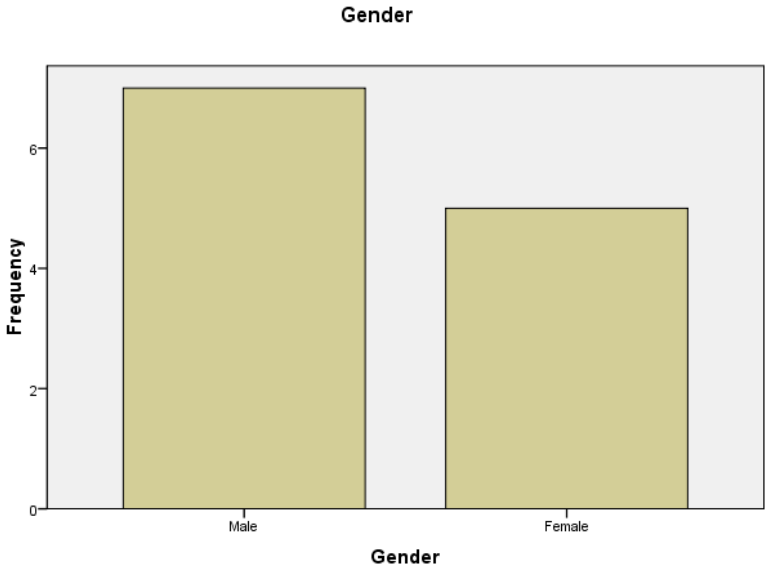
Thus frequency table has been prepared using SPSS Package.

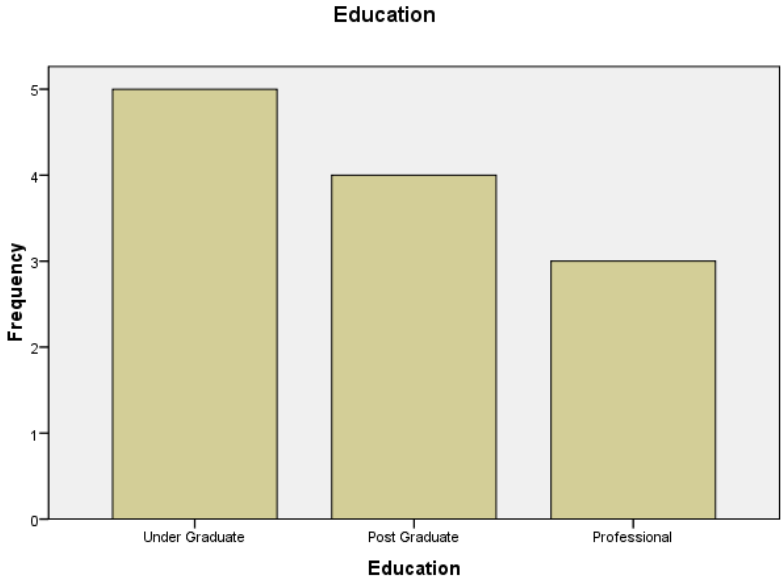
Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	7	58.3	58.3	58.3
	Female	5	41.7	41.7	100.0
	Total	12	100.0	100.0	

Education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Under Graduate	5	41.7	41.7	41.7
	Post Graduate	4	33.3	33.3	75.0
	Professional	3	25.0	25.0	100.0
	Total	12	100.0	100.0	


Gender
Frequency
6
4
2
0
Male
Female
Gender


Education
Frequency
5
4
3
2
1
0
Under Graduate
Post Graduate
Professional
Education

Ex.No.4

Chi-square

Aim

To calculate Chi-square test to find association between two variables

Algorithm

Step 1:	Start the Process
Step 2:	Type two set of Qualitative data on Variable 1 and Variable 2 on Data View Window
Step 3:	Specify names for Variable 1 and Variable 2 on variable view window
Step 4:	Select Descriptive Statistics option from Analyze Menu
Step 5:	Select Crosstabs from Descriptive Sub menu
Step 6:	Forward Variable 1 to Row and Variable 2 to Column options
Step 7:	Click Statistics Command button on Crosstab window, from which select Chi-square option and press continue command button
Step 8:	Click Cell command button on Crosstab window, in which select row on percentage option and press continue command button
Step 9:	Click Ok button on Crosstab window
Step 10:	Stop the process

Result

The Chi-square test result has been verified by using SPSS Package.

Output

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
AOR * SQ	300	100.0%	0	.0%	300	100.0%

AOR * SQ Crosstabulation

			SQ			
			1	2	3	Total
AOR	1	Count	5	75	3	83
		% within AOR	6.0%	90.4%	3.6%	100.0%
	2	Count	29	94	8	131
		% within AOR	22.1%	71.8%	6.1%	100.0%
	3	Count	6	72	8	86
		% within AOR	7.0%	83.7%	9.3%	100.0%
Total		Count	40	241	19	300
		% within AOR	13.3%	80.3%	6.3%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	18.085[a]	4	.001
Likelihood Ratio	18.256	4	.001
Linear-by-Linear Association	.545	1	.460
N of Valid Cases	300		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.26.

Ex.No.5

Independent Sample 't' Test

Aim

To calculate independent 't' test to find whether mean differs between two groups

Algorithm

Step 1:	Start the Process
Step 2:	Type Quantitative data on Variable 1
Step 3:	Group the Quantitative data in to two groups by assigning qualitative values (i.e.) 1,2
Step 4:	Select Compare Means from Analyze Menu
Step 5:	Select Independent Sample 't' Test from Compare Means menu
Step 6:	Forward Quantitative data to Test Variables Option and Qualitative data to Group Variables option
Step 7:	Click Define Groups command button and assign 1 and 2 at Group 1 and Group 2 and click continue command button
Step 8:	Click Ok button on Independent Sample 't' test window
Step 9:	Stop the process

Result

The above statistical analysis has been verified by using SPSS Package.

Output

Group Statistics

	VAR00002	N	Mean	Std. Deviation	Std. Error Mean
VAR00001	1	193	69.6732	13.58475	.97785
	2	107	72.9753	11.50454	1.11219

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
VAR00001	Equal variances assumed	4.822	.029	-2.127	298	.034	-3.30211	1.55281	-6.35798	-.24625
	Equal variances not assumed			-2.230	250.561	.027	-3.30211	1.48093	-6.21877	-.38546

Ex.No.6

Paired 't' Test

Aim

To calculate paired 't' test to find whether student's performance differs before and after training programme

Algorithm

Step 1:	Start the Process
Step 2:	Type Quantitative data on Variable 1 and Variable 2
Step 3:	Select Compare Means from Analyze Menu
Step 4:	Select paired 't' Test from Compare Means menu
Step 5:	Forward Variable 1 and Variable 2 data set to Paired Variables window
Step 6:	Click Ok button on Independent Sample 't' test window
Step 7:	Stop the process

Result

The above statistical analysis has been verified by using SPSS Package.

Output

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	VAR00001 & VAR00002	9	.784	.012

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	VAR00001 - VAR00002	-1.16667E1	6.72681	2.24227	-16.83735	-6.49598	-5.203	8	.001

Ex.No.7

One Way Analysis of Variance (ANOVA)

Aim

To calculate ANOVA test to find whether mean differs among more than two groups

Algorithm

Step 1:	Start the Process
Step 2:	Type Quantitative data on Variable 1
Step 3:	Group the Quantitative data in to more than two groups by assigning qualitative values (i.e.) 1,2,3
Step 4:	Select Compare Means from Analyze Menu
Step 5:	Select One way ANOVA option from Compare Means menu
Step 6:	Forward Quantitative data to Dependent list and Qualitative data to Factor
Step 7:	Click Option Command button
Step 8:	Click Description option on Statistics and click continue button
Step 9:	Click Ok button on One way ANOVA window
Step 10	Stop the process

Result

The above statistical analysis has been verified by using SPSS Package.

Output

Descriptives

VAR00001

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	83	65.7420	12.40484	1.36161	63.0334	68.4507	35.86	93.10
2	131	73.1511	12.02735	1.05083	71.0722	75.2301	35.86	91.72
3	86	72.2779	13.61429	1.46807	69.3590	75.1968	42.76	93.79
Total	300	70.8510	12.95901	.74819	69.3786	72.3233	35.86	93.79

ANOVA

VAR00001

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3034.594	2	1517.297	9.552	.000
Within Groups	47178.261	297	158.849		
Total	50212.855	299			

Ex.No.8

Two Way Analysis of Variance (ANOVA)

Aim

To calculate Two way ANOVA test to find whether mean differs among more than two groups

Algorithm

Step 1:	Start the Process
Step 2:	Type Quantitative data on Variable 1 (Agriculture Productivity) and Qualitative data in Variable 2 (Soil type) and Variable 3 (Irrigation type)
Step 3:	Group the Quantitative data in to more than two groups by assigning qualitative values (i.e.) 1,2,3
Step 4:	Select General Linear Model from Analyze Menu
Step 5:	Select Univariate option from General Linear Model menu
Step 6:	Forward Quantitative data to Dependent list and Qualitative data to Fixed Factors
Step 7:	Click Model command button and verify Type III is selected in Sum of Squares and click continue button
Step 8:	Click Ok button on univariate window
Step 9:	Stop the process

Result

The above statistical analysis has been verified by using SPSS Package.

Between-Subjects Factors

		N
Soil Type	1	6
	2	6
Irrigation Type	1	5
	2	4
	3	3

Tests of Between-Subjects Effects

Dependent Variable:VAR00001

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	7059.250[a]	5	1411.850	2.938	.111
Intercept	201884.033	1	201884.033	420.057	.000
VAR00002	710.533	1	710.533	1.478	.270
VAR00003	3894.961	2	1947.481	4.052	.077
VAR00002 * VAR00003	3481.008	2	1740.504	3.621	.093
Error	2883.667	6	480.611		
Total	242073.000	12			
Corrected Total	9942.917	11			

a. R Squared = .710 (Adjusted R Squared = .468)

Ex.No.9

Correlation

Aim

To calculate Correlation to find nature of relation between dependent and independent variable

Algorithm

Step 1:	Start the Process
Step 2:	Type Dependent variables on Variable 1 and Independent variables from Variable 2
Step 3:	Select Correlate option from Analyze Menu
Step 4:	Select Bivariate command option from Correlate sub menu
Step 5:	Forward all the variables (Dependent and Independent variables) to variables window
Step 6:	Select Pearson option and Spearman from Correlation Coefficient for running Karl Pearson Correlation and Spearman respectively
Step 7:	Click Ok button on Bivariate Correlation window
Step 8	Stop the process

Result

The above statistical analysis has been verified by using SPSS Package.

Karl Pearson Correlation

Correlations

		VAR00001	VAR00002
VAR00001	Pearson Correlation	1	.981**
	Sig. (2-tailed)		.000
	N	12	12
VAR00002	Pearson Correlation	.981**	1
	Sig. (2-tailed)	.000	
	N	12	12

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

Spearman Rank Correlation

Correlations

			VAR00001	VAR00002
Spearman's rho	VAR00001	Correlation Coefficient	1.000	.968**
		Sig. (2-tailed)	.	.000
		N	12	12
	VAR00002	Correlation Coefficient	.968**	1.000
		Sig. (2-tailed)	.000	.
		N	12	12

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

Ex.No.10

Regression

Aim

To ascertain the combined influence of select independent variables over dependent variable, regression test is employed

Algorithm

Step 1:	Start the Process
Step 2:	Type Dependent variables data on variable 1 and independent variables data on variable 2, variable 3 etc.,
Step 3:	Select Regression option from Analyze Menu
Step 4:	Select linear command option from regression sub menu
Step 5:	Forward Dependent variable to Dependent option and Independent variables to Independent window
Step 6:	Click Ok button on Linear Regression window
Step 7:	Stop the process

Result

Thus regression test has been verified by using SPSS Package.

Output

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Sales, Depreciation, PAT, PYD[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Dividend

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.951[a]	.904	.776	.24474

a. Predictors: (Constant), Sales, Depreciation, PAT, PYD

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.695	4	.424	7.074	.070[a]
	Residual	.180	3	.060		
	Total	1.875	7			

a. Predictors: (Constant), Sales, Depreciation, PAT, PYD

b. Dependent Variable: Dividend

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.067	.324		.207	.849
	PYD	.014	.305	.014	.045	.967
	Depreciation	-1.236	.696	-.359	-1.776	.174
	PAT	.883	.217	.838	4.078	.027
	Sales	-.007	.011	-.190	-.647	.564

a. Dependent Variable: Dividend

Ex.No.11

Factor Analysis

Aim

To find out the multi colinearity exists between variables and to find out the important factors (or) variables among the various set of variables introduced

Algorithm

Step 1:	Start the Process
Step 2:	Type qualitative data on data view window
Step 3:	Select Data reduction option from Analyze Menu
Step 4:	Select factor analysis command option from data reduction sub menu
Step 5:	Forward all the variables introduced to variables option
Step 6:	Select Descriptive command window, and select KMO and Bartlett's test of Sphercity option and click continue command button
Step 7:	Select Rotation command button on Factor analysis window, and select varimax option and click continue command button
Step 8:	Click Ok button on Factor Analysis window
Step 9:	Stop the process

Result

Thus factor test has been verified by using SPSS Package.

Output

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.908
Bartlett's Test of Sphericity	Approx. Chi-Square	3.947E3
	df	406
	Sig.	.000

Rotated Component Matrix[a]

	Component						
	1	2	3	4	5	6	7
VAR00001	.585	.192	.452	.085	.235	.045	.307
VAR00002	.325	-.052	.233	.512	.317	-.052	.185
VAR00003	.131	.173	.124	.163	.140	.167	.752
VAR00004	.121	.202	.598	-.160	.102	.038	.066
VAR00005	.258	.233	-.137	.538	.168	.213	.226
VAR00006	.105	.064	.750	.229	.063	.012	-.041
VAR00007	.389	.263	.004	.392	.081	.257	.406
VAR00008	.700	.254	.394	.074	.193	.052	.181
VAR00009	.277	.062	.039	.315	.525	.129	.141
VAR00010	.407	.219	.053	.136	.093	.043	.586
VAR00011	-.040	.169	.569	-.238	-.107	.212	.328
VAR00012	.344	.108	-.059	.690	.033	.253	.065
VAR00013	.138	.328	.651	.136	.076	.072	.039
VAR00014	.796	.106	.058	.259	.133	.213	.102
VAR00015	.016	.134	.043	.267	.641	.012	.361
VAR00016	.140	.642	.348	-.115	-.015	.148	.173
VAR00017	.430	-.012	.102	.417	-.021	.425	.136
VAR00018	.010	.580	.359	.283	.173	.191	.188
VAR00019	.768	.096	-.016	.206	.171	.299	.068
VAR00020	.185	.058	.171	-.106	.737	.084	.082
VAR00021	.121	.655	.114	-.123	.176	-.099	.228
VAR00022	.073	-.156	.204	.382	.196	.603	.234

VAR00023	.213	.658	.191	.344	.036	.151	-.001
VAR00024	.590	.092	.222	.233	.298	.309	.066
VAR00025	.205	.323	-.051	.107	.582	.181	-.126
VAR00026	.037	.477	.231	.124	.248	-.066	.439
VAR00027	.248	.007	.082	.027	.061	.769	.069
VAR00028	.083	.554	.162	.486	.168	-.098	.086
VAR00029	.315	.351	.028	.088	.173	.639	-.017

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 15 iterations.

Ex.No.12

Cluster Analysis

Aim

To group heterogeneous variables into homogenous group

Algorithm

Step 1:	Start the Process
Step 2:	Type quantitative data on data view window for Calorie, Sodium, Salt and Fat
Step 3:	Select Classify from Analyze Menu
Step 4:	Select Hierarchial Cluster analysis form classify menu
Step 5:	Forward all the variables introduced to variables option
Step 6:	Click statistics command button and click Agglomeration Schedule option and click continue
Step 7:	Click plot command button and click Dendrogram option and click continue
Step 8:	Click Method command button and verify Between-groups linkage option from Cluster Method and also verify Squared Euclidean distance from Measure option and click continue
Step 9:	Click Save option and click None option from Cluster Membership and click continue
Step 10:	Click Ok
Step 11:	Stop the process

Result

Thus cluster analysis test has been carried out by using SPSS Package.

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	3	10	124.000	0	0	2
2	3	4	414.000	1	0	6
3	5	9	490.000	0	0	6
4	1	7	540.000	0	0	5
5	1	2	856.000	4	0	9
6	3	5	1471.000	2	3	7
7	3	6	3687.600	6	0	8
8	3	8	5646.333	7	0	9
9	1	3	25908.524	5	8	0

Ex.No.13

Discriminant Analysis

Aim

To find out method of discriminating

Algorithm

Step 1:	Start the Process
Step 2:	Type two set of quantitative data and three set of qualitative data on data view window
Step 3:	Select classify option from Analyse menu
Step 4:	Select Discriminant option from classify option
Step 5:	Forward dichotomous variable to grouping variable
Step 6:	Click Define range and enter 1 and 2
Step 7:	Forward rest of the variables to independent option
Step 8:	Click statistics command button and select Means, Univariate ANOVA and Box's M option from Descriptive option and select Fishers Option from Function Coefficient and click continue button
Step 9:	Click classify command button and select Leave-one-out classification option from Display and click command button
Step 10:	Click Ok button
Step 9:	Stop the process

Result

Thus discriminant analysis test has been verified by using SPSS Package.

Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		8	100.0
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
	Total	0	.0
Total		8	100.0

Group Statistics

VAR00003		Mean	Std. Deviation	Valid N (listwise) Unweighted	Valid N (listwise) Weighted
1	VAR00001	44.6000	10.16366	5	5.000
	VAR00002	13.2000	6.83374	5	5.000
	VAR00004	4.0000	1.41421	5	5.000
	VAR00005	4.2000	.44721	5	5.000
2	VAR00001	35.0000	10.00000	3	3.000
	VAR00002	9.3333	6.02771	3	3.000
	VAR00004	3.6667	.57735	3	3.000
	VAR00005	4.0000	1.73205	3	3.000
Total	VAR00001	41.0000	10.59650	8	8.000
	VAR00002	11.7500	6.40870	8	8.000
	VAR00004	3.8750	1.12599	8	8.000
	VAR00005	4.1250	.99103	8	8.000

Box's Test of Equality of Covariance Matrices

Log Determinants

VAR00003	Rank	Log Determinant
1	4	3.117
2	.[a]	.[b]
Pooled within-groups	4	5.816

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

a. Rank < 3

b. Too few cases to be non-singular

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
VAR00001	.780	1.691	1	6	.241
VAR00002	.902	.648	1	6	.451
VAR00004	.977	.144	1	6	.717
VAR00005	.989	.066	1	6	.806