## Correlation and Regression

Variables – Income, Price, and Demand

Analysis 1

a) Predictors: Price
b) Dependent variable: Demand

**Model Summary[b]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .933[a] | .870 | .853 | 7.50000 |

a. Predictors: (Constant), Price

b. Dependent Variable: Demand

**ANOVA[a]**

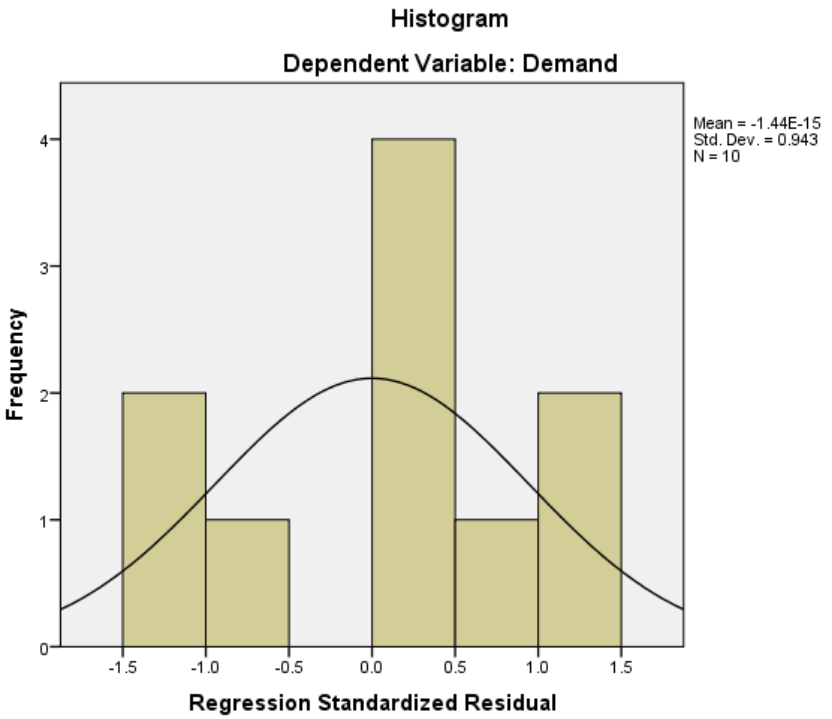
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 3000.000 | 1 | 3000.000 | 53.333 | .000[b] |
| | Residual | 450.000 | 8 | 56.250 | | |
| | Total | 3450.000 | 9 | | | |

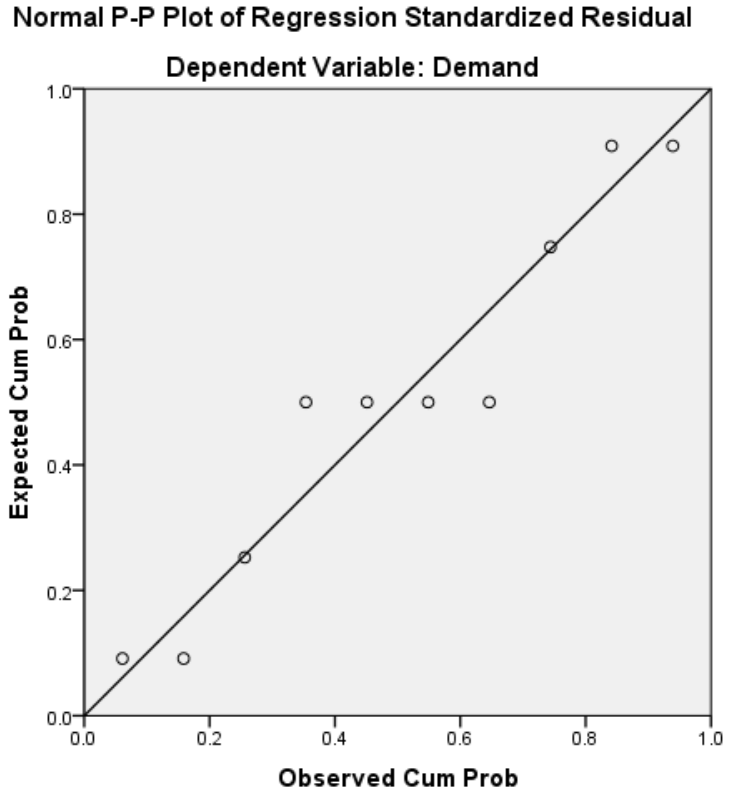
a. Dependent Variable: Demand
b. Predictors: (Constant), Price

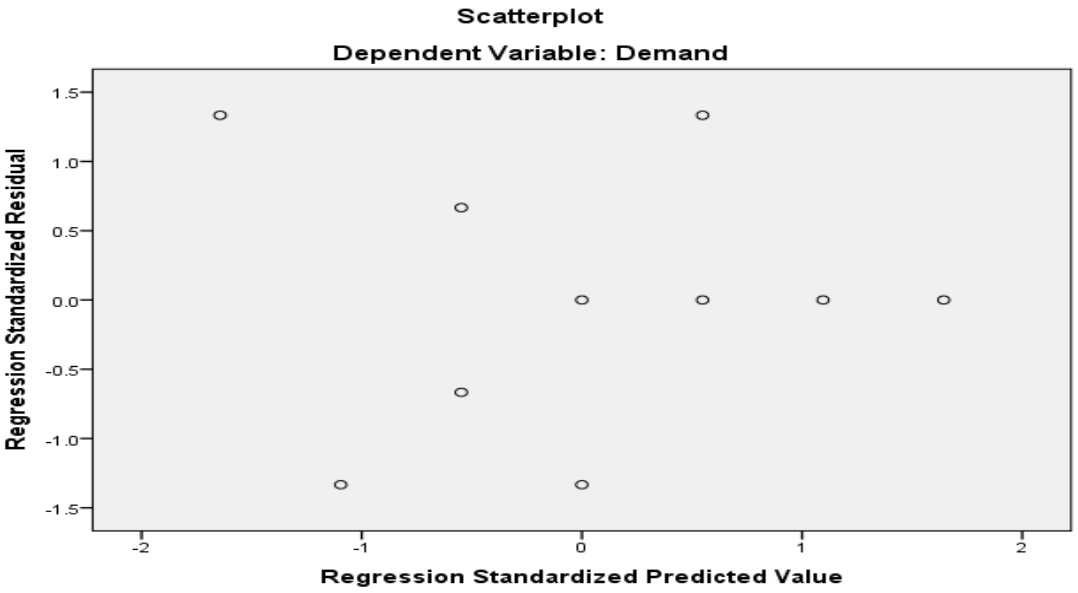
**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 140.000 | 8.551 | | 16.372 | .000 |
| | Price | -10.000 | 1.369 | -.933 | -7.303 | .000 |

a. Dependent Variable: Demand


Histogram
Dependent Variable: Demand
Mean = -1.44E-15
Std. Dev. = 0.943
N = 10
Frequency
Regression Standardized Residual


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: Demand
Expected Cum Prob
Observed Cum Prob




## Inferences:

1. From the **model summary table**, R square value = 0.870. Thus, the explanatory power of the model is 87%.
   Or in other words, Price can explain 87% of the variation in Demand.
2. From the **coefficients table**, the regression equation is:

   Demand = -0.933Price + 140.00 + residual

3. Hypothesis:
   - Ho: There is no significant relationship between the explanatory and dependent variables. (Beta = 0)
   - H1: There is a significant relationship between the explanatory and dependent variables. (Beta is not equal to 0)
4. From the **ANOVA table**, the significance table is less than 0.05. Hence, we reject Ho. This implies that there is a significant relationship between explanatory and dependent variables.
5. From the **Coefficients table**, the significance of the t-test is less than 0.05. Hence the alternative hypothesis is accepted. There is a significant relationship between the explanatory and dependent variables.
6. From the **Histogram chart**, we can see that the histogram is not much skewed. Hence, the residual terms are not exactly normal.
7. From the **Normal Probability Plot**, it is clear that the distribution of residuals is not much far away from the line. It agrees with the Histogram.
8. **Scatter Plot** infers that the variance of the error term is normal. Hence, there is Homoscedasticity.

## Analysis 2

a) Predictors: Income
b) Dependent variable: Demand

**Model Summary[b]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .880[a] | .775 | .747 | 9.84854 |

a. Predictors: (Constant), Income

b. Dependent Variable: Demand

**ANOVA[a]**

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 2674.051 | 1 | 2674.051 | 27.569 | .001[b] |
| | Residual | 775.949 | 8 | 96.994 | | |
| | Total | 3450.000 | 9 | | | |

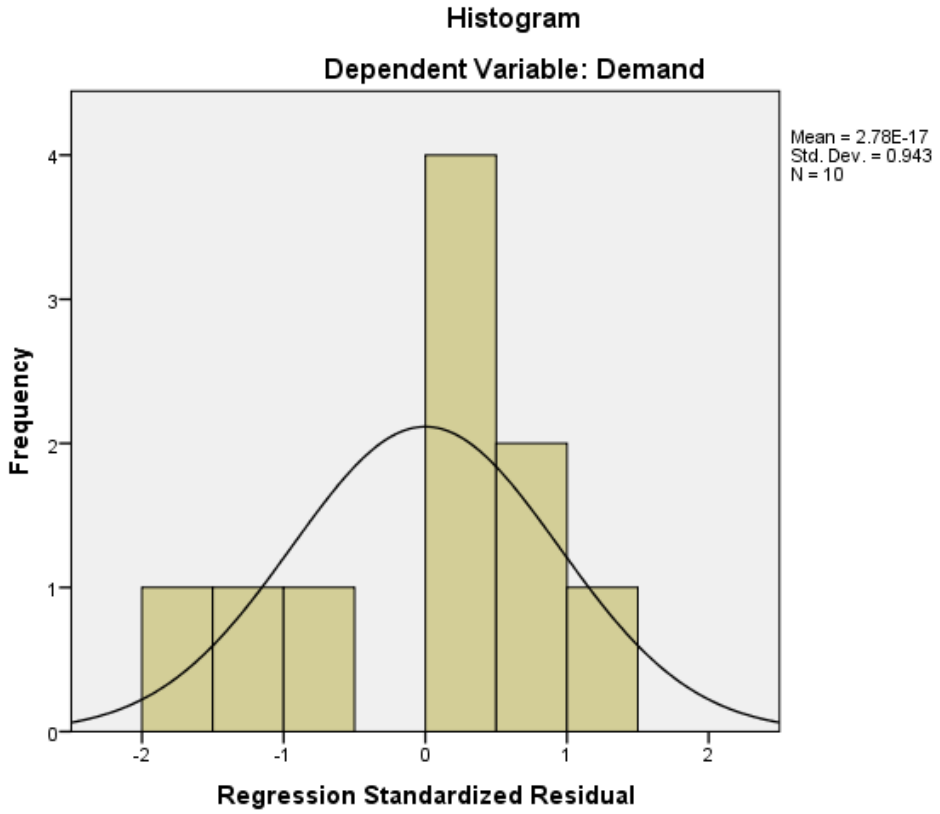
a. Dependent Variable: Demand

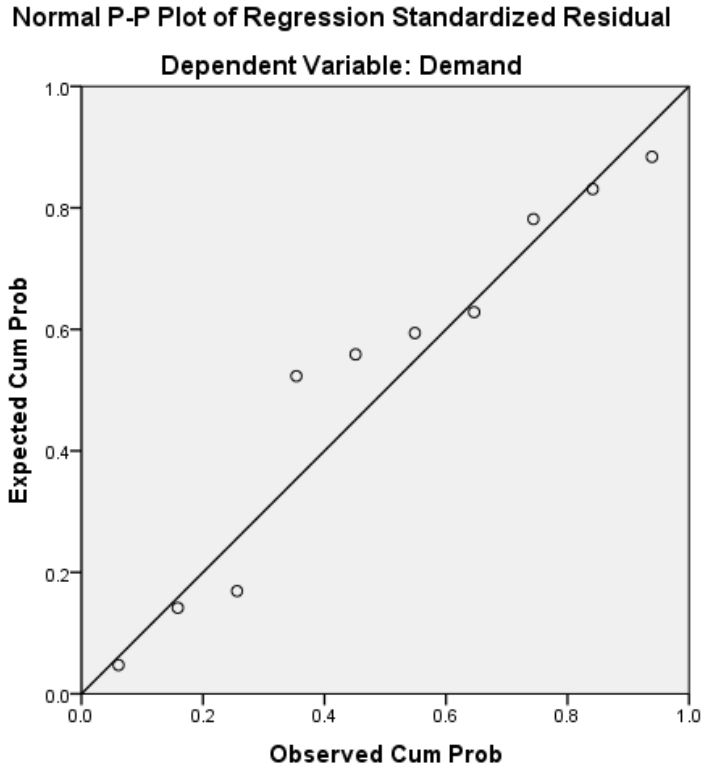
b. Predictors: (Constant), Income

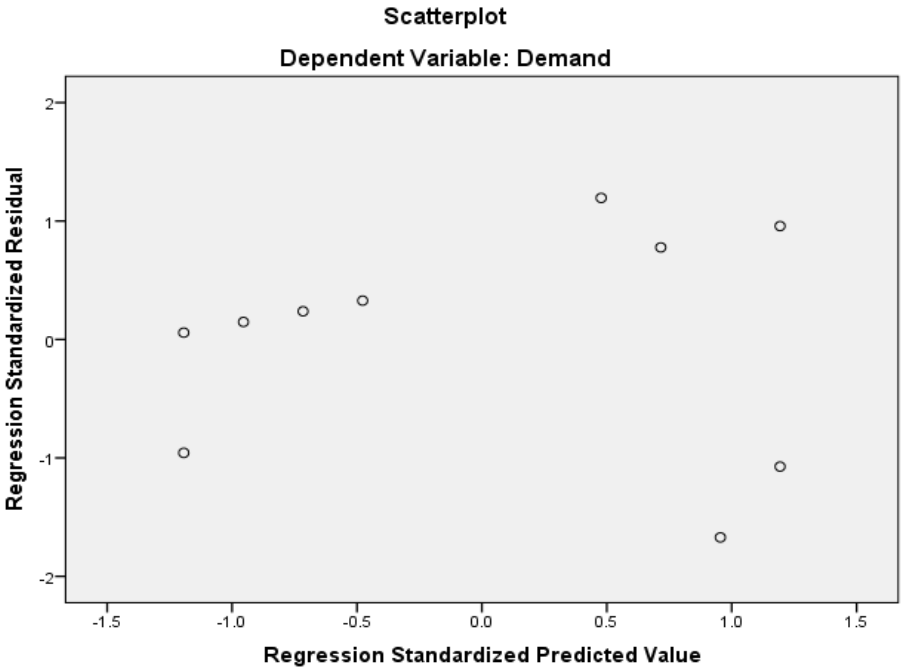
**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 47.089 | 6.999 | | 6.728 | .000 |
| | Income | .041 | .008 | .880 | 5.251 | .001 |

a. Dependent Variable: Demand


Histogram
Dependent Variable: Demand
Mean = 2.78E-17
Std. Dev. = 0.943
N = 10
Frequency
Regression Standardized Residual


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: Demand
Expected Cum Prob
Observed Cum Prob




## Inferences:

1. From the **model summary table**, R square value = 0.775. Thus, the explanatory power of the model is 77.5%.
   Or in other words, Income can explain 77.5% of the variation in Demand.
2. From the **coefficients table**, the regression equation is:

   Demand = 0.88Income + 47.089 + residual

3. Hypothesis:
   - Ho: There is no significant relationship between the explanatory and dependent variables. (Beta = 0)
   - H1: There is a significant relationship between the explanatory and dependent variables. (Beta is not equal to 0)
4. From the **ANOVA table**, the significance table is less than 0.05. Hence, we reject Ho. This implies that there is a significant relationship between explanatory and dependent variables.
5. From the **Coefficients table**, the significance of the t-test is less than 0.05. Hence the alternative hypothesis is accepted. There is a significant relationship between the explanatory and dependent variables.
6. From the **Histogram chart**, we can see that the histogram is not much skewed. Hence, the residual terms are not exactly normal.
7. From the **Normal Probability Plot**, it is clear that the distribution of residuals is not much far away from the line. It agrees with the Histogram.
8. **Scatter Plot** infers that the variance of the error term is normal. Hence, there is Homoscedasticity.

## Analysis 3

## Correlation

Variables: Demand, Price, and Income

**Correlations**

| | | Demand | Price | Income |
|---|---|---|---|---|
| Demand | Pearson Correlation | 1 | -.933** | .880** |
| | Sig. (2-tailed) | | .000 | .001 |
| | N | 10 | 10 | 10 |
| Price | Pearson Correlation | -.933** | 1 | -.857** |
| | Sig. (2-tailed) | .000 | | .002 |
| | N | 10 | 10 | 10 |
| Income | Pearson Correlation | .880** | -.857** | 1 |
| | Sig. (2-tailed) | .001 | .002 | |
| | N | 10 | 10 | 10 |

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

## Inferences:

1. Hypothesis:
    - Ho: There is no association between the variables.
    - H1: There is an association between the variables.
2. From the **Correlations table**, for all the variables, the significance value is less than 0.05. Hence, we accept H1.
3. Between variables:
    - Demand and Price: Pearson Correlation Coefficient is -0.933. Thus, they are negatively correlated meaning as demand increases, price decreases, and vice-versa.
    - Demand and Income: Pearson Correlation Coefficient is 0.880. Thus, they are positively correlated meaning as income increases, demand also increases, and vice-versa.

## Correlation and Regression - 2

## Analysis 1

## Regression

a) Predictors: No. of years of Experience
b) Dependent variable: Starting Salary of a Lecturer

**Model Summary[b]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .877[a] | .769 | .750 | 1.82808 |

a. Predictors: (Constant), No. of Years of Experience

b. Dependent Variable: Starting Salary of a Lecturer (in Rs. thousand per month)

**ANOVA[a]**

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 133.612 | 1 | 133.612 | 39.981 | .000[b] |
| | Residual | 40.103 | 12 | 3.342 | | |
| | Total | 173.714 | 13 | | | |

a. Dependent Variable: Starting Salary of a Lecturer (in Rs. thousand per month)

b. Predictors: (Constant), No. of Years of Experience

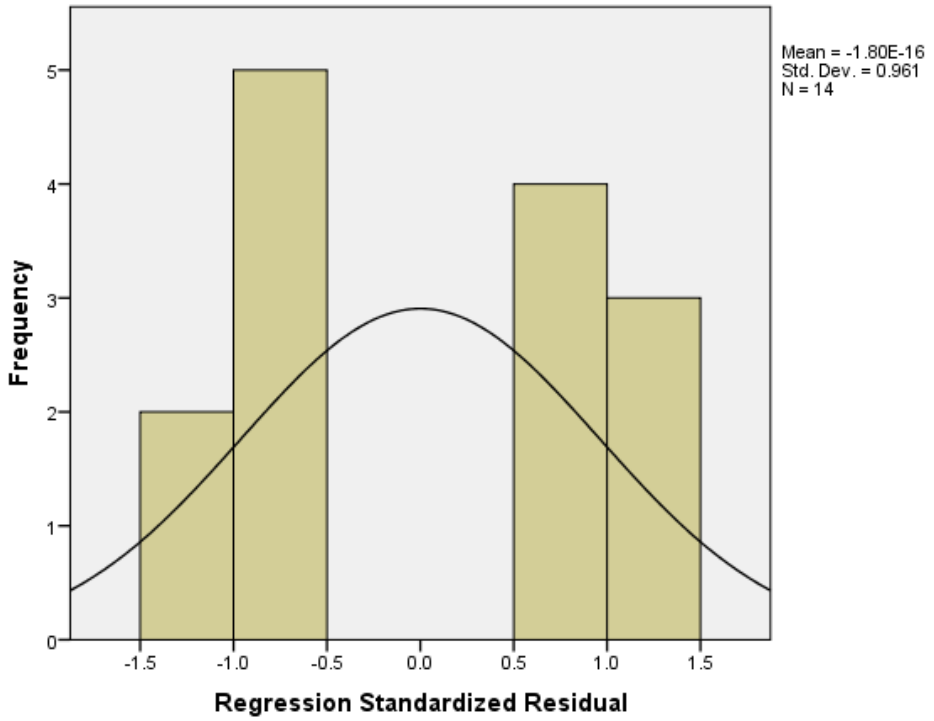
**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 18.964 | 1.092 | | 17.359 | .000 |
| | No. of Years of Experience | 1.545 | .244 | .877 | 6.323 | .000 |

a. Dependent Variable: Starting Salary of a Lecturer (in Rs. thousand per month)

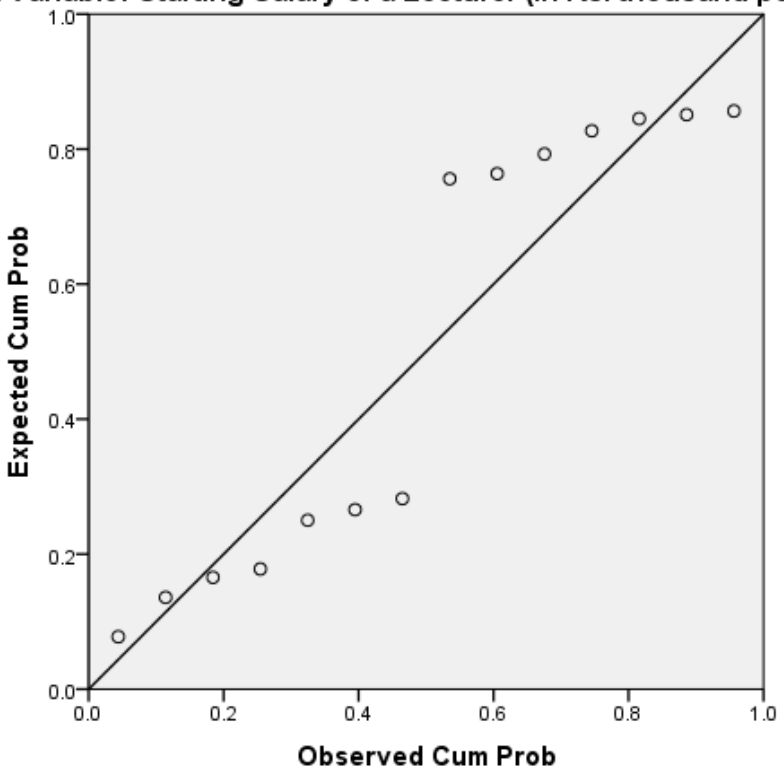
Histogram

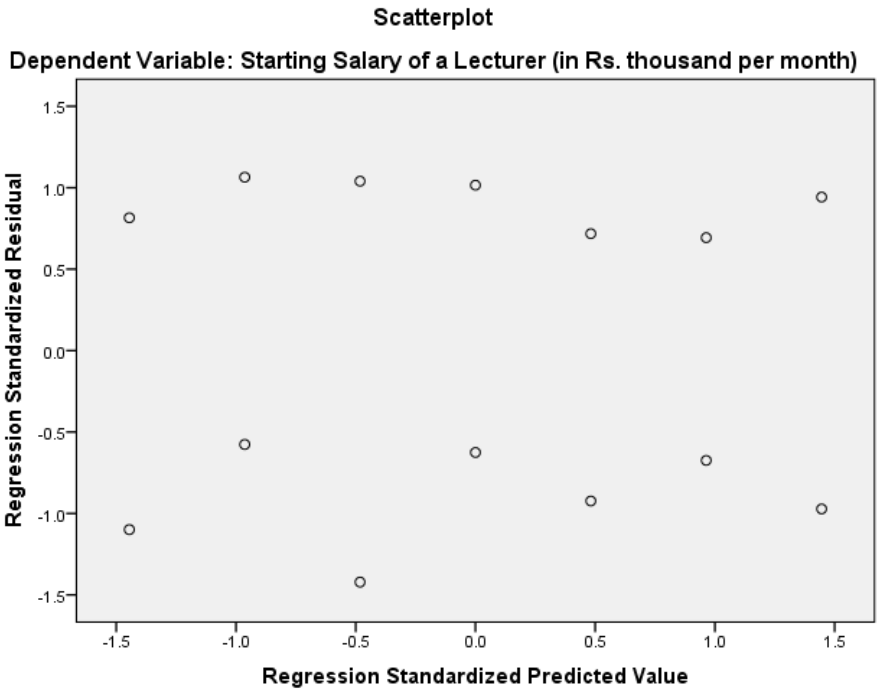
Dependent Variable: Starting Salary of a Lecturer (in Rs. thousand per month)




Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Starting Salary of a Lecturer (in Rs. thousand per month)






## Inferences:

1. From the **model summary table**, R square value = 0.769. Thus, the explanatory power of the model is 76.9%.
   Or in other words, Income can explain 76.9% of the variation in Demand.
2. From the **coefficients table**, the regression equation is:

   Starting Salary of a lecturer = 0.877(no. of years of experience) + 18.964 + residual

3. Hypothesis:
   - Ho: There is no significant relationship between the explanatory and dependent variable. (Beta = 0)
   - H1: There is a significant relationship between the explanatory and dependent variable. (Beta is not equal to 0)
4. From the **ANOVA table**, the significance table is less than 0.05. Hence, we reject Ho. This implies that there is a significant relationship between explanatory and dependent variables.
5. From the **Coefficients table**, the significance of the t-test is less than 0.05. Hence the alternative hypothesis is accepted. There is a significant relationship between the explanatory and dependent variables.
6. From the **Histogram chart**, we can see that the histogram is not much skewed. Hence, the residual terms are not exactly normal.
7. From the **Normal Probability Plot**, it is clear that the distribution of residuals is not much far away from the line. It agrees with the Histogram.
8. **Scatter Plot** infers that the variance of the error term is normal. Hence, there is Homoscedasticity.

## Analysis 2

## Correlation

**Correlations**

| | | Starting Salary of a Lecturer (in Rs. thousand per month) | No. of Years of Experience |
|---|---|---|---|
| Starting Salary of a Lecturer (in Rs. thousand per month) | Pearson Correlation | 1 | .877** |
| | Sig. (2-tailed) | | .000 |
| | N | 14 | 14 |
| No. of Years of Experience | Pearson Correlation | .877** | 1 |
| | Sig. (2-tailed) | .000 | |
| | N | 14 | 14 |

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

## Inferences:

1. Hypothesis:
    - Ho: There is no association between the variables.
    - H1: There is an association between the variables.
2. From the **Correlations table**, for all the variables, the significance value is less than 0.05. Hence, we accept H1.
3. Between variables:
    - Number of years of experience and Starting Salary of a Lecturer: Pearson Correlation Coefficient is 0.877. Thus, they are positively correlated meaning as number of years of experience increases, starting salary of a lecturer also increases, and vice-versa.

## Correlation and Regression - 3

## Analysis 1

## Multiple Regression

a) Predictors: Taste, Nutrition Value, Preference
b) Dependent variable: Preservation Quality

**Model Summary[b]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .899[a] | .807 | .791 | .78233 |

a. Predictors: (Constant), Taste, Nutrition Value, Preference
b. Dependent Variable: Preservation Quality

**ANOVA[a]**

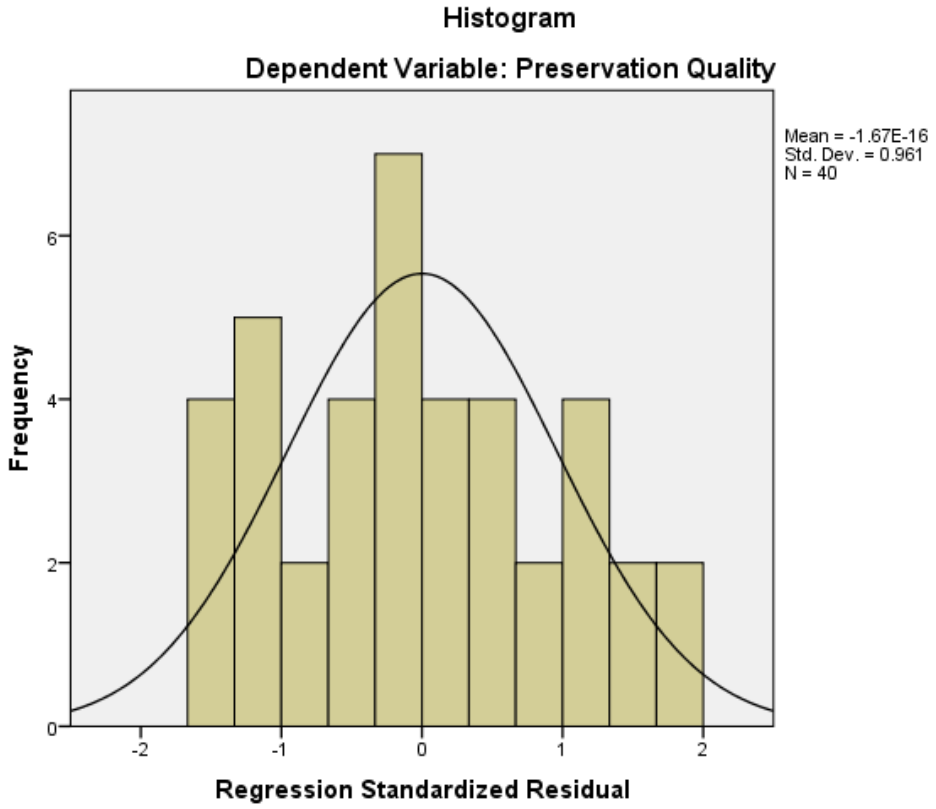
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 92.341 | 3 | 30.780 | 50.291 | .000[b] |
| | Residual | 22.034 | 36 | .612 | | |
| | Total | 114.375 | 39 | | | |

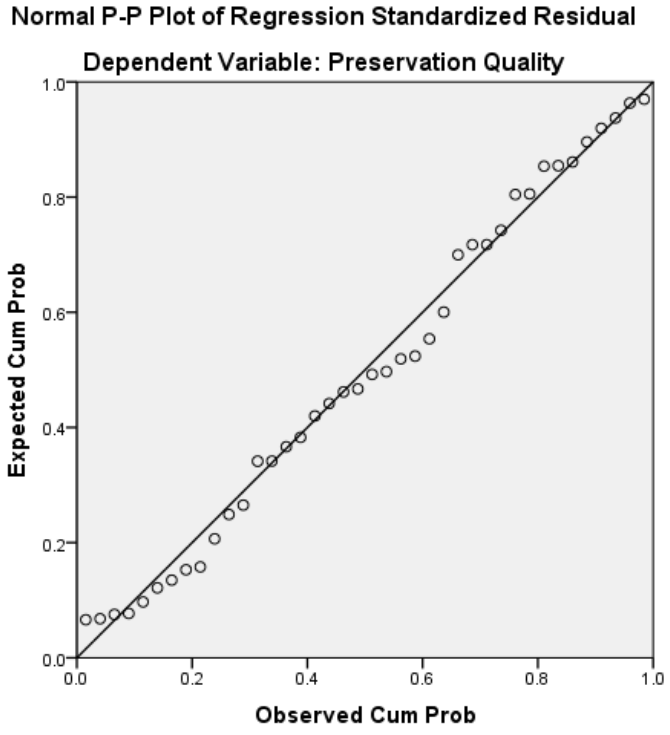
a. Dependent Variable: Preservation Quality
b. Predictors: (Constant), Taste, Nutrition Value, Preference

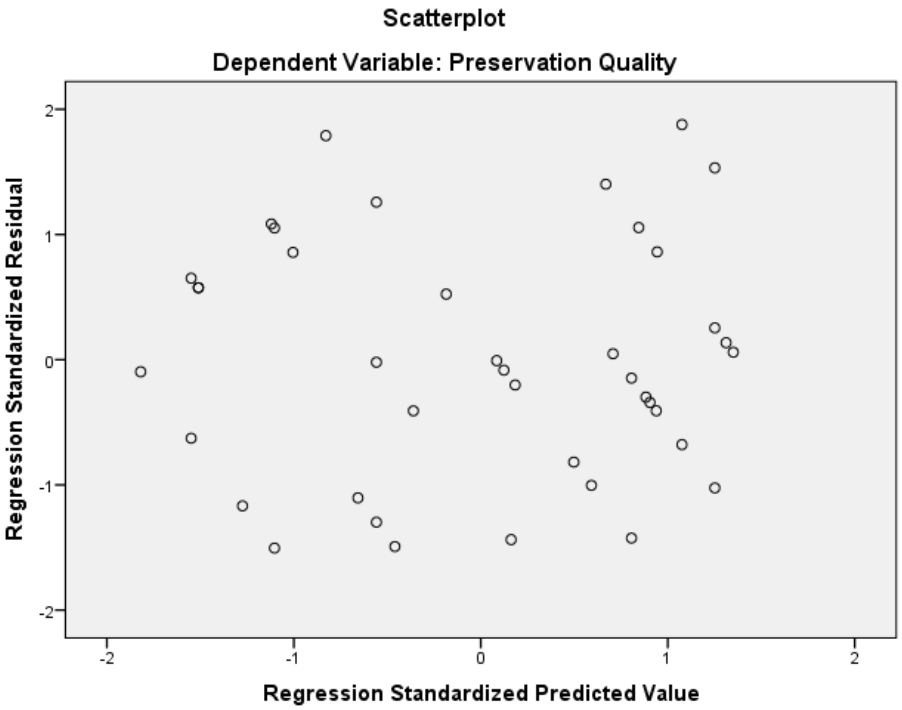
**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | .026 | .363 | | .070 | .944 |
| | Preference | .686 | .147 | .720 | 4.660 | .000 |
| | Nutrition Value | -.059 | .127 | -.060 | -.466 | .644 |
| | Taste | .211 | .114 | .258 | 1.850 | .073 |

a. Dependent Variable: Preservation Quality


Histogram
Dependent Variable: Preservation Quality
Mean = -1.67E-16
Std. Dev. = 0.961
N = 40
Frequency
Regression Standardized Residual


Normal P-P Plot of Regression Standardized Residual
Dependent Variable: Preservation Quality
Expected Cum Prob
Observed Cum Prob




## Inferences:

1. From the **model summary table**, R square value = 0.807. Thus, the explanatory power of the model is 80.7%.
   Or in other words, Income can explain 80.7% of the variation in Demand.
2. From the **coefficients table**, the regression equation is:

Preservation Quality = 0.686(Preference) -0.59(Nutrition Value) +0.211(Taste) + residual

3. Hypothesis:
   - Ho: There is no significant relationship between the explanatory and dependent variable. (Beta = 0)
   - H1: There is a significant relationship between the explanatory and dependent variable. (Beta is not equal to 0)
4. From the **ANOVA table**, the significance table is less than 0.05. Hence, we reject Ho. This implies that there is a significant relationship between explanatory and dependent variables.
5. From the **Coefficients table**, we can infer that each explanatory variable has different significance in the relationship with the dependent variable.
6. From the **Histogram chart**, we can see that the histogram is not much skewed. Hence, the residual terms are almost normal.
7. From the **Normal Probability Plot**, it is clear that the distribution of residuals is not much far away from the line. It agrees with the Histogram.
8. **Scatter Plot** infers that the variance of the error term is normal. Hence, there is Homoscedasticity.

## Analysis 2

## Correlation

**Correlations**

| | | Preference | Nutrition Value | Taste |
|---|---|---|---|---|
| Preference | Pearson Correlation | 1 | .810** | .841** |
| | Sig. (2-tailed) | | .000 | .000 |
| | N | 40 | 40 | 40 |
| Nutrition Value | Pearson Correlation | .810** | 1 | .759** |
| | Sig. (2-tailed) | .000 | | .000 |
| | N | 40 | 40 | 40 |
| Taste | Pearson Correlation | .841** | .759** | 1 |
| | Sig. (2-tailed) | .000 | .000 | |
| | N | 40 | 40 | 40 |

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

## Inferences:

1. Hypothesis:
    - Ho: There is no association between the variables.
    - H1: There is an association between the variables.
2. From the **Correlations table**, for all the variables, the significance value is less than 0.05. Hence, we accept H1.
3. Between variables:
    - Preference and Nutrition Value: Pearson Correlation Coefficient is -0.810. Thus, they are positively correlated meaning as nutrition value increases, preference also increases, and vice-versa.
    - Preference and Taste: Pearson Correlation Coefficient is 0.841. Thus, they are positively correlated meaning as taste increases, preference also increases, and vice-versa.
    - Taste and Nutrition Value: Pearson Correlation Coefficient is 0.759. Thus, they are positively correlated meaning as nutrition value increases, taste also increases, and vice-versa.

# Factor Analysis

**KMO and Bartlett's Test**

| | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .591 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 80.004 |
| | df | 21 |
| | Sig. | .000 |

**Communalities**

| | Initial | Extraction |
|---|---|---|
| Score on Risk Averseness | 1.000 | .598 |
| Score on Returns | 1.000 | .304 |
| Score on Insurance Covers | 1.000 | .612 |
| Score on Tax Rebate | 1.000 | .111 |
| Score on Maturity Time | 1.000 | .624 |
| Score on Credibility of Financial Institution | 1.000 | .725 |
| Score on Easy Accessibility | 1.000 | .631 |

Extraction Method: Principal Component Analysis.

**Total Variance Explained**

| | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| Component | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 2.054 | 29.346 | 29.346 | 2.054 | 29.346 | 29.346 | 2.010 | 28.708 | 28.708 |
| 2 | 1.551 | 22.160 | 51.506 | 1.551 | 22.160 | 51.506 | 1.596 | 22.798 | 51.506 |
| 3 | .970 | 13.857 | 65.363 | | | | | | |
| 4 | .848 | 12.109 | 77.472 | | | | | | |
| 5 | .711 | 10.151 | 87.622 | | | | | | |
| 6 | .490 | 7.003 | 94.626 | | | | | | |
| 7 | .376 | 5.374 | 100.000 | | | | | | |

Extraction Method: Principal Component Analysis.

**Component Matrix[a]**

| | Component | |
|---|---|---|
| | 1 | 2 |
| Score on Risk Averseness | -.176 | .753 |
| Score on Returns | .527 | .160 |
| Score on Insurance Covers | .335 | -.707 |
| Score on Tax Rebate | .309 | .125 |
| Score on Maturity Time | .765 | -.198 |
| Score on Credibility of Financial Institution | .570 | .633 |
| Score on Easy Accessibility | .793 | .047 |

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

**Rotated Component Matrix[a]**

| | Component | |
|---|---|---|
| | 1 | 2 |
| Score on Risk Averseness | .057 | -.771 |
| Score on Returns | .551 | .004 |
| Score on Insurance Covers | .109 | .775 |
| Score on Tax Rebate | .332 | -.027 |
| Score on Maturity Time | .671 | .417 |
| Score on Credibility of Financial Institution | .732 | -.435 |
| Score on Easy Accessibility | .771 | .192 |

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.[a]

a. Rotation converged in 3 iterations.

**Component Transformation Matrix**

| Component | 1 | 2 |
|---|---|---|
| 1 | .955 | .298 |
| 2 | .298 | -.955 |

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

## Inferences:

1. The value of **KMO statistics** is greater than 0.5, indicating that factor analysis could be used for the given set of data.
2. **Bartlett's test of sphericity** testing for the significance of the correlation matrix of the variables indicates that the correlation coefficient matrix is significant as indicated by the p value corresponding to the chi-square statistic. P value is less than 0.05.
3. The correlation coefficient between the factor score and the variables is presented in table called factor matrix or component matrix. The correlation coefficient between the first variable, risk awareness and factor 1 is -0.176.
4. From the **Total Variance Explained** table, there are 2 factors with Eigen value more than 1. As there are 7 variables, the total variance equals seven. The percentage of variance explained by each factor can be computed using eigenvalues.

Percentage of variance explained by factor 1 = (Eigenvalue of factor 1 / sum total of Eigen values) *100 = (2.054 / 7)*100 = 29.346 percent

Percentage of variance explained by factor 2 = (Eigenvalue of factor 2 / sum total of Eigen values) *100 = (1.551 / 7)*100 = 22.16 percent

Total variance explained by both factors = 29.346 + 22.16 = 51.506 percent.

5. The **Communalities Table** indicates how much of each variable is accounted for by the underlying factors taken together. It is a measure of the percentage of variable's variation that is explained by the factors.
   Communality for risk averseness = (0.176)2 + (0.753)2 = 0.598
6. In order to interpret the **Rotated Component Matrix,** a cut-off point is decided. Using 0.7 as cut-off point, the two variables corresponding to factor 1 having a factor loading above 0.7 are credibility of the financial institutions and ease of accessibility.
   The variables corresponding to factor 2 for which the factor loadings are greater than 0.7 are risk averseness and insurance cover.
7. Factor 1 comprising of the credibility of the financial institutions and ease of accessibility could be named as perceived value of service.
   Factor 2 could be named as security factor.

# Factor Analysis - 2

**KMO and Bartlett's Test**

| | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .664 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 342.005 |
| | df | 55 |
| | Sig. | .000 |

**Communalities**

| | Initial | Extraction |
|---|---|---|
| Aerated soft drinks are refreshing | 1.000 | .699 |
| Aerated soft drinks are bad for health | 1.000 | .567 |
| Aerated soft drinks are very convenient to serve | 1.000 | .476 |
| Aerated soft drinks should be avoided with age | 1.000 | .612 |
| Aerated soft drinks are very tasty | 1.000 | .635 |
| Aerated soft drinks are not good for children | 1.000 | .752 |
| Aerated soft drinks should be consumed occasionally | 1.000 | .576 |
| Aerated soft drinks should not be taken in large quantity | 1.000 | .239 |
| Aerated soft drinks are not as good as energy drinks | 1.000 | .158 |
| Aerated soft drinks are better than fruit juices | 1.000 | .886 |
| Recommending aerated drinks to others | 1.000 | .867 |

Extraction Method: Principal Component Analysis.

**Total Variance Explained**

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 3.086 | 28.058 | 28.058 | 3.086 | 28.058 | 28.058 | 2.859 | 25.992 | 25.992 |
| 2 | 2.127 | 19.334 | 47.393 | 2.127 | 19.334 | 47.393 | 1.875 | 17.047 | 43.039 |
| 3 | 1.254 | 11.396 | 58.789 | 1.254 | 11.396 | 58.789 | 1.732 | 15.749 | 58.789 |
| 4 | .991 | 9.008 | 67.796 | | | | | | |
| 5 | .844 | 7.675 | 75.472 | | | | | | |
| 6 | .701 | 6.371 | 81.842 | | | | | | |
| 7 | .555 | 5.050 | 86.892 | | | | | | |
| 8 | .549 | 4.992 | 91.885 | | | | | | |
| 9 | .482 | 4.386 | 96.271 | | | | | | |
| 10 | .271 | 2.462 | 98.732 | | | | | | |
| 11 | .139 | 1.268 | 100.000 | | | | | | |

Extraction Method: Principal Component Analysis.

**Component Matrix[a]**

| | Component | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Aerated soft drinks are refreshing | -.267 | .291 | .737 |
| Aerated soft drinks are bad for health | .702 | .248 | -.115 |
| Aerated soft drinks are very convenient to serve | .124 | .630 | .252 |
| Aerated soft drinks should be avoided with age | .714 | .267 | -.177 |
| Aerated soft drinks are very tasty | -.110 | .722 | .319 |
| Aerated soft drinks are not good for children | .841 | .198 | .079 |
| Aerated soft drinks should be consumed occasionally | .677 | .296 | -.172 |
| Aerated soft drinks should not be taken in large quantity | .484 | .023 | -.066 |
| Aerated soft drinks are not as good as energy drinks | .011 | -.346 | -.196 |
| Aerated soft drinks are better than fruit juices | -.579 | .611 | -.422 |
| Recommending aerated drinks to others | .501 | -.610 | .493 |

Extraction Method: Principal Component Analysis.

a. 3 components extracted.

**Rotated Component Matrix[a]**

| | Component | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Aerated soft drinks are refreshing | -.344 | .224 | .728 |
| Aerated soft drinks are bad for health | .748 | .062 | .060 |
| Aerated soft drinks are very convenient to serve | .249 | -.151 | .625 |
| Aerated soft drinks should be avoided with age | .782 | .012 | .031 |
| Aerated soft drinks are very tasty | .047 | -.257 | .753 |
| Aerated soft drinks are not good for children | .807 | .282 | .146 |
| Aerated soft drinks should be consumed occasionally | .756 | -.017 | .059 |
| Aerated soft drinks should not be taken in large quantity | .465 | .140 | -.059 |
| Aerated soft drinks are not as good as energy drinks | -.050 | .076 | -.387 |
| Aerated soft drinks are better than fruit juices | -.218 | -.894 | .201 |
| Recommending aerated drinks to others | .129 | .910 | -.147 |

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.[a]
a. Rotation converged in 5 iterations.

**Component Transformation Matrix**

| Component | 1 | 2 | 3 |
|---|---|---|---|
| 1 | .909 | .411 | -.063 |
| 2 | .321 | -.597 | .735 |
| 3 | -.264 | .689 | .675 |

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

## Inferences:

1. The value of **KMO statistics** is greater than 0.5, indicating that factor analysis could be used for the given set of data.
2. **Bartlett's test of sphericity** testing for the significance of the correlation matrix of the variables indicates that the correlation coefficient matrix is

significant as indicated by the p value corresponding to the chi-square statistic. P value is less than 0.05.

3. The correlation coefficient between the factor score and the variables is presented in table called factor matrix or **component matrix**. The correlation coefficient between the first variable, Aerated soft drinks are refreshing and factor 1 is -0.267.
4. From the **Total Variance Explained** table, there are 3 factors with Eigen value more than 1. As there are 11 variables, the total variance equals eleven. The percentage of variance explained by each factor can be computed using eigenvalues.

Percentage of variance explained by factor 1 = (Eigenvalue of factor 1 / sum total of Eigen values) *100 = (3.086 / 11)*100 = 28.054 percent

Percentage of variance explained by factor 2 = (Eigenvalue of factor 2 / sum total of Eigen values) *100 = (2.127 / 11)*100 = 19.336 percent

Percentage of variance explained by factor 3 = (Eigenvalue of factor 3 / sum total of Eigen values) *100 = (1.254 / 11)*100 = 11.4 percent

Total variance explained by 3 factors = 28.054 + 19.336 + 11.4 = 58.79 percent.

5. The **Communalities Table** indicates how much of each variable is accounted for by the underlying factors taken together. It is a measure of the percentage of variable's variation that is explained by the factors.
   Communality for Aerated soft drinks are refreshing = $(-0.267)^2 + (0.291)^2 + (0.737)^2 = 0.699$

6. In order to interpret the **Rotated Component Matrix,** a cut-off point is decided. Using 0.7 as cut-off point, the **variables corresponding to factor 1** having a factor loading above 0.7 are Aerated soft drinks are bad for health, Aerated soft drinks should be avoided with age, Aerated soft drinks are not good for children and Aerated soft drinks should be consumed occasionally.
   The **variables corresponding to factor 2** for which the factor loadings are greater than 0.7 are Recommending aerated drinks to others.
   The **variables corresponding to factor 3** for which the factor loadings are greater than 0.7 are Aerated soft drinks are refreshing and Aerated soft drinks are very tasty.

7. Factor 1 could be named as perceived avoiding factors for aerated soft drinks.
   Factor 2 could be named as recommending factor.
   Factor 3 could be named as perceived supporting factors for aerated soft drinks.

# Factor Analysis - 3

**KMO and Bartlett's Test**

| | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .613 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 355.669 |
| | df | 153 |
| | Sig. | .000 |

**Communalities**

| | Initial | Extraction |
|---|---|---|
| Price on Road | 1.000 | .743 |
| Brand Name | 1.000 | .773 |
| Engine Capacity | 1.000 | .650 |
| Looks & Design | 1.000 | .763 |
| Fuel Efficiency | 1.000 | .710 |
| Discount Scheme | 1.000 | .582 |
| Resale Value | 1.000 | .671 |
| After Sale Services | 1.000 | .554 |
| Running and Maintaining Cost | 1.000 | .686 |
| Convenience Features | 1.000 | .493 |
| Purpose of Purchase | 1.000 | .697 |
| Performance Information Available | 1.000 | .587 |
| Driving Pleasure | 1.000 | .635 |
| Car Image & Positioning | 1.000 | .579 |
| Economical | 1.000 | .738 |
| Colours Available | 1.000 | .595 |
| Advertising & Marketing | 1.000 | .463 |
| Safety | 1.000 | .740 |

Extraction Method: Principal Component Analysis.

**Total Variance Explained**

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 3.860 | 21.447 | 21.447 | 3.860 | 21.447 | 21.447 | 2.621 | 14.558 | 14.558 |
| 2 | 2.275 | 12.640 | 34.087 | 2.275 | 12.640 | 34.087 | 2.303 | 12.794 | 27.353 |
| 3 | 1.738 | 9.658 | 43.745 | 1.738 | 9.658 | 43.745 | 1.748 | 9.711 | 37.063 |
| 4 | 1.436 | 7.975 | 51.720 | 1.436 | 7.975 | 51.720 | 1.696 | 9.420 | 46.483 |
| 5 | 1.244 | 6.910 | 58.630 | 1.244 | 6.910 | 58.630 | 1.682 | 9.343 | 55.826 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 6 | 1.104 | 6.131 | 64.761 | 1.104 | 6.131 | 64.761 | 1.608 | 8.936 | 64.761 |
| 7 | .952 | 5.289 | 70.050 | | | | | | |
| 8 | .847 | 4.703 | 74.753 | | | | | | |
| 9 | .777 | 4.316 | 79.069 | | | | | | |
| 10 | .668 | 3.714 | 82.783 | | | | | | |
| 11 | .620 | 3.442 | 86.225 | | | | | | |
| 12 | .532 | 2.953 | 89.178 | | | | | | |
| 13 | .491 | 2.727 | 91.904 | | | | | | |
| 14 | .412 | 2.287 | 94.191 | | | | | | |
| 15 | .312 | 1.735 | 95.926 | | | | | | |
| 16 | .295 | 1.637 | 97.563 | | | | | | |
| 17 | .259 | 1.439 | 99.002 | | | | | | |
| 18 | .180 | .998 | 100.000 | | | | | | |

Extraction Method: Principal Component Analysis.

**Component Matrix[a]**

| | Component | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Price on Road | .110 | .464 | .309 | .349 | .120 | .533 |
| Brand Name | .199 | -.151 | .783 | .133 | -.222 | .175 |
| Engine Capacity | .339 | .267 | .533 | -.332 | .260 | .043 |
| Looks & Design | .161 | -.396 | .567 | .105 | .339 | -.364 |
| Fuel Efficiency | .442 | .549 | .173 | -.369 | -.002 | -.216 |
| Discount Schme | .435 | .268 | -.163 | .531 | -.028 | .107 |
| Resale Value | .442 | .408 | .047 | .462 | .306 | .015 |
| After Sale Services | .506 | -.026 | -.375 | .110 | .326 | -.195 |
| Running and Maintaining Cost | .657 | .409 | -.078 | -.277 | .047 | .043 |
| Convenience Features | .471 | -.460 | -.023 | -.109 | .168 | .136 |
| Purpose of Purchase | .276 | .236 | .287 | .060 | -.429 | -.543 |
| Performance Information Available | .544 | -.029 | -.147 | -.122 | .489 | -.120 |
| Driving Pleasure | .652 | -.428 | -.014 | -.027 | -.153 | -.047 |
| Car Image & Positioning | .551 | -.194 | -.207 | .361 | -.155 | -.201 |
| Economical | .513 | .458 | -.250 | -.197 | -.397 | .080 |
| Colours Available | .496 | -.530 | -.003 | -.034 | -.017 | .258 |
| Advertising & Marketing | .447 | -.202 | .085 | .335 | -.318 | -.036 |
| Safety | .606 | -.284 | -.034 | -.375 | -.193 | .335 |

Extraction Method: Principal Component Analysis.

a. 6 components extracted.

**Rotated Component Matrix[a]**

| | Component | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Price on Road | -.063 | .143 | -.229 | -.149 | .802 | -.026 |
| Brand Name | .278 | .156 | -.587 | .216 | .264 | .459 |
| Engine Capacity | .116 | .668 | -.082 | -.182 | .173 | .346 |
| Looks & Design | .137 | .030 | .060 | .138 | -.059 | .847 |
| Fuel Efficiency | -.081 | .822 | .106 | .109 | .037 | -.049 |
| Discount Schme | .046 | -.001 | .250 | .369 | .588 | -.188 |
| Resale Value | -.084 | .203 | .359 | .191 | .670 | .095 |
| After Sale Services | .201 | .081 | .687 | .157 | .103 | -.018 |
| Running and Maintaining Cost | .230 | .677 | .277 | .074 | .195 | -.232 |
| Convenience Features | .645 | .000 | .221 | -.007 | -.025 | .163 |
| Purpose of Purchase | -.195 | .403 | -.128 | .675 | -.108 | .113 |
| Performance Information Available | .296 | .291 | .614 | -.062 | .082 | .165 |
| Driving Pleasure | .662 | .088 | .161 | .389 | -.072 | .081 |
| Car Image & Positioning | .309 | -.084 | .333 | .591 | .127 | -.033 |
| Economical | .141 | .527 | .054 | .287 | .114 | -.585 |
| Colours Available | .754 | -.083 | .068 | .088 | .026 | .082 |
| Advertising & Marketing | .337 | -.057 | -.041 | .557 | .181 | .038 |
| Safety | .788 | .280 | -.036 | .029 | -.063 | -.186 |

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.[a]
a. Rotation converged in 19 iterations.

**Component Transformation Matrix**

| Component | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | .621 | .485 | .362 | .413 | .279 | -.001 |
| 2 | -.613 | .572 | .012 | -.042 | .413 | -.354 |
| 3 | -.002 | .318 | -.600 | .033 | .186 | .709 |
| 4 | -.206 | -.561 | .091 | .460 | .635 | .140 |
| 5 | -.067 | .009 | .607 | -.590 | .222 | .479 |
| 6 | .439 | -.148 | -.363 | -.517 | .514 | -.350 |

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

## Inferences:

1. The value of **KMO statistics** is greater than 0.5, indicating that factor analysis could be used for the given set of data.
2. **Bartlett's test of sphericity** testing for the significance of the correlation matrix of the variables indicates that the correlation coefficient matrix is significant as indicated by the p value corresponding to the chi-square statistic. P value is less than 0.05.
3. The correlation coefficient between the factor score and the variables is presented in table called factor matrix or **component matrix**. The correlation coefficient between the first variable, Price on Road and factor 1 is 0.110.

4. From the **Total Variance Explained** table, there are 6 factors with Eigen value more than 1. As there are 18 variables, the total variance equals eighteen. The percentage of variance explained by each factor can be computed using eigenvalues.

Percentage of variance explained by factor 1 = (Eigenvalue of factor 1 / sum total of Eigen values) *100 = (3.860 / 18)*100 = 21.444 percent

Percentage of variance explained by factor 2 = (Eigenvalue of factor 2 / sum total of Eigen values) *100 = (2.275 / 18)*100 = 12.638 percent

Percentage of variance explained by factor 3 = (Eigenvalue of factor 3 / sum total of Eigen values) *100 = (1.738 / 18)*100 = 9.656 percent

Percentage of variance explained by factor 4 = (Eigenvalue of factor 4 / sum total of Eigen values) *100 = (1.436 / 18)*100 = 7.978 percent

Percentage of variance explained by factor 5 = (Eigenvalue of factor 5 / sum total of Eigen values) *100 = (1.244 / 18)*100 = 6.911 percent

Percentage of variance explained by factor 6 = (Eigenvalue of factor 6 / sum total of Eigen values) *100 = (1.104 / 18)*100 = 6.133 percent

Total variance explained by both factors = 21.444 + 12.638 + 9.656 + 7.978 + 6.911 + 6.133 = 64.76 percent.

5. The **Communalities Table** indicates how much of each variable is accounted for by the underlying factors taken together. It is a measure of the percentage of variable's variation that is explained by the factors.
   Communality for Price on Road = $(0.110)^2 + (0.464)^2 + (0.309)^2 + (0.349)^2 + (0.120)^2 + (0.533)^2 = 0.743$

6. In order to interpret the **Rotated Component Matrix,** a cut-off point is decided.
   Using 0.7 as cut-off point, the **variables corresponding to factor 1** having a factor loading above 0.7 are Aerated soft drinks are bad for health, Aerated soft drinks should be avoided with age, Aerated soft drinks are not good for children and Aerated soft drinks should be consumed occasionally.
   The **variables corresponding to factor 2** for which the factor loadings are greater than 0.7 are Recommending aerated drinks to others.
   The **variables corresponding to factor 3** for which the factor loadings are greater than 0.7 are Aerated soft drinks are refreshing and Aerated soft drinks are very tasty.

7. Factor 1 could be named as perceived avoiding factors for aerated soft drinks.
   Factor 2 could be named as recommending factor.
   Factor 3 could be named as perceived supporting factors for aerated soft drinks.

# Discriminant Analysis

**Group Statistics**

| Buyer / Non-Buyer | | Mean | Std. Deviation | Valid N (list wise) Unweighted | Weighted |
|---|---|---|---|---|---|
| Non-Buyer | Durability | 4.00 | 2.000 | 9 | 9.000 |
| | Light Weight | 4.33 | 1.803 | 9 | 9.000 |
| | Low Investment | 4.33 | 1.414 | 9 | 9.000 |
| | Rot Resistance | 3.67 | 1.936 | 9 | 9.000 |
| Buyer | Durability | 7.44 | 1.944 | 9 | 9.000 |
| | Light Weight | 5.78 | 1.986 | 9 | 9.000 |
| | Low Investment | 5.22 | 2.108 | 9 | 9.000 |
| | Rot Resistance | 3.44 | 1.424 | 9 | 9.000 |
| Total | Durability | 5.72 | 2.608 | 18 | 18.000 |
| | Light Weight | 5.06 | 1.984 | 18 | 18.000 |
| | Low Investment | 4.78 | 1.801 | 18 | 18.000 |
| | Rot Resistance | 3.56 | 1.653 | 18 | 18.000 |

**Tests of Equality of Group Means**

| | Wilks' Lambda | F | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Durability | .538 | 13.729 | 1 | 16 | .002 |
| Light Weight | .860 | 2.610 | 1 | 16 | .126 |
| Low Investment | .935 | 1.103 | 1 | 16 | .309 |
| Rot Resistance | .995 | .077 | 1 | 16 | .785 |

**Pooled Within-Groups Matrices**

| | | Durability | Light Weight | Low Investment | Rot Resistance |
|---|---|---|---|---|---|
| Correlation | Durability | 1.000 | .633 | .549 | .209 |
| | Light Weight | .633 | 1.000 | .541 | .327 |
| | Low Investment | .549 | .541 | 1.000 | .064 |
| | Rot Resistance | .209 | .327 | .064 | 1.000 |

**Eigenvalues**

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | 1.033[a] | 100.0 | 100.0 | .713 |

a. First 1 canonical discriminant functions were used in the analysis.

**Wilks' Lambda**

| Test of Function(s) | Wilks' Lambda | Chi-square | df | Sig. |
|---|---|---|---|---|
| 1 | .492 | 9.936 | 4 | .042 |

**Standardized Canonical Discriminant Function Coefficients**

| | Function |
|---|---|
| | 1 |
| Durability | 1.219 |
| Light Weight | -.104 |
| Low Investment | -.338 |
| Rot Resistance | -.268 |

**Structure Matrix**

| | Function |
|---|---|
| | 1 |
| Durability | .911 |
| Light Weight | .397 |
| Low Investment | .258 |
| Rot Resistance | -.068 |

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

**Canonical Discriminant Function Coefficients**

| | Function |
|---|---|
| | 1 |
| Durability | .618 |
| Light Weight | -.055 |
| Low Investment | -.188 |
| Rot Resistance | -.157 |
| (Constant) | -1.800 |

Unstandardized coefficients

**Classification Results**[a,c]

| | | Buyer / Non-Buyer | Predicted Group Membership | | Total |
|---|---|---|---|---|---|
| | | | Non-Buyer | Buyer | |
| Original | Count | Non-Buyer | 7 | 2 | 9 |
| | | Buyer | 1 | 8 | 9 |
| | % | Non-Buyer | 77.8 | 22.2 | 100.0 |
| | | Buyer | 11.1 | 88.9 | 100.0 |
| Cross-validated[b] | Count | Non-Buyer | 6 | 3 | 9 |
| | | Buyer | 2 | 7 | 9 |
| | % | Non-Buyer | 66.7 | 33.3 | 100.0 |
| | | Buyer | 22.2 | 77.8 | 100.0 |

a. 83.3% of original grouped cases correctly classified.

b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

c. 72.2% of cross-validated grouped cases correctly classified.

## Inferences:

1. The **Group Statistics** table computes the mean values to get an idea of the differences in the mean score of variables for buyer and non-buyer. The mean score for durability for the buyer group is 7.44, whereas for the non-buyer group it is 4.0. For other characteristics the mean values are closer for buyer and non-buyer. Therefore we can expect that durability could be useful in discriminating between prospective buyers and non-buyers. However, in terms of variability, the standard deviations of variables like low investment and rot resistance seem to vary a lot.

2. From the table of **Tests of Equality of Group Means** table, it is observed that the significant difference in the mean exists for the durability, for which the p value is less than 0.05, the assumed level of significance. There does not seem to be any significant difference in the means of the remaining three characteristics as the p value in each of these cases is greater than 0.05.

3. From the table of **Pooled within-Groups Matrices** table**,** the correlation between any pair of predictor variables does not exceed 0.75. So, there is no problem of multicollinearity.

4. From the table of **Canonical Discriminant Function Coefficients,** the discriminant function is

**Y = -1.800 + 0.618 Durability - 0.055 Light Weight – 0.188 Low Investment -0.157 Rot resistance**

5. From the **Eigenvalues** table, the last column, canonical correlation, is the simple correlation coefficient between the discriminant score and their corresponding group membership (buyer / non buyer).
   The square of the canonical correlation (0.713) is 0.508. This means 50.8 percent of the variance in the discriminating model between a prospective buyer and a non-buyer is due to the changes in the four predictor variables.

6. From the **Wilk's Lambda** table, the value of Wilks' Lambda is 0.492. This parameter takes a value between 0 and 1. Lower the value of Wilks' Lambda, the higher is the significance of the discriminant function. Therefore 0 would be the most preferred one. The statistical significance of Wilks' Lambda is carried out with the chi-square statistic (9.936) and its p value. Since p value is less than 0.05, the discriminant function is significant.

7. From the **Group Centroids** table, we can compute the mean discriminant scores of the buyer and non-buyer groups separately. This is called group centroids. This is -0,958 for a non- buyer and 0.958 for a buyer. This is used for designing a decision rule to classify a customer into a buyer / non-buyer category. Take the average of the two group centroids. It is 0 here.
   Any respondent whose discriminant score is greater than zero would be classified as a prospective buyer, whereas the one with less than zero would be classified as a non-buyer.

8. **The Standardized Canonical Discriminant Function Coefficients** table indicates that durability is the most important characteristic, which discriminates between the buyer and non-buyer group, followed by low investment, rot resistance and light weight.

9. From the **Structure Matrix,** the structural coefficients are obtained by computing the correlation between the discriminant score and each of the independent variables. These are also called discriminant loadings.

10. From the **Classification Results** table, Hit ratio = No. of correct predictions / Total number of cases
    Here, there are 15 correct predictions out of a sample of 18. 7 non buyers predicted as non-buyers, 8 buyers predicted as buyers from a TOTAL OF 15.
    Hit ratio = 15/18 = 83.33 percent.

# Cluster Analysis

**Case Processing Summary[a,b]**

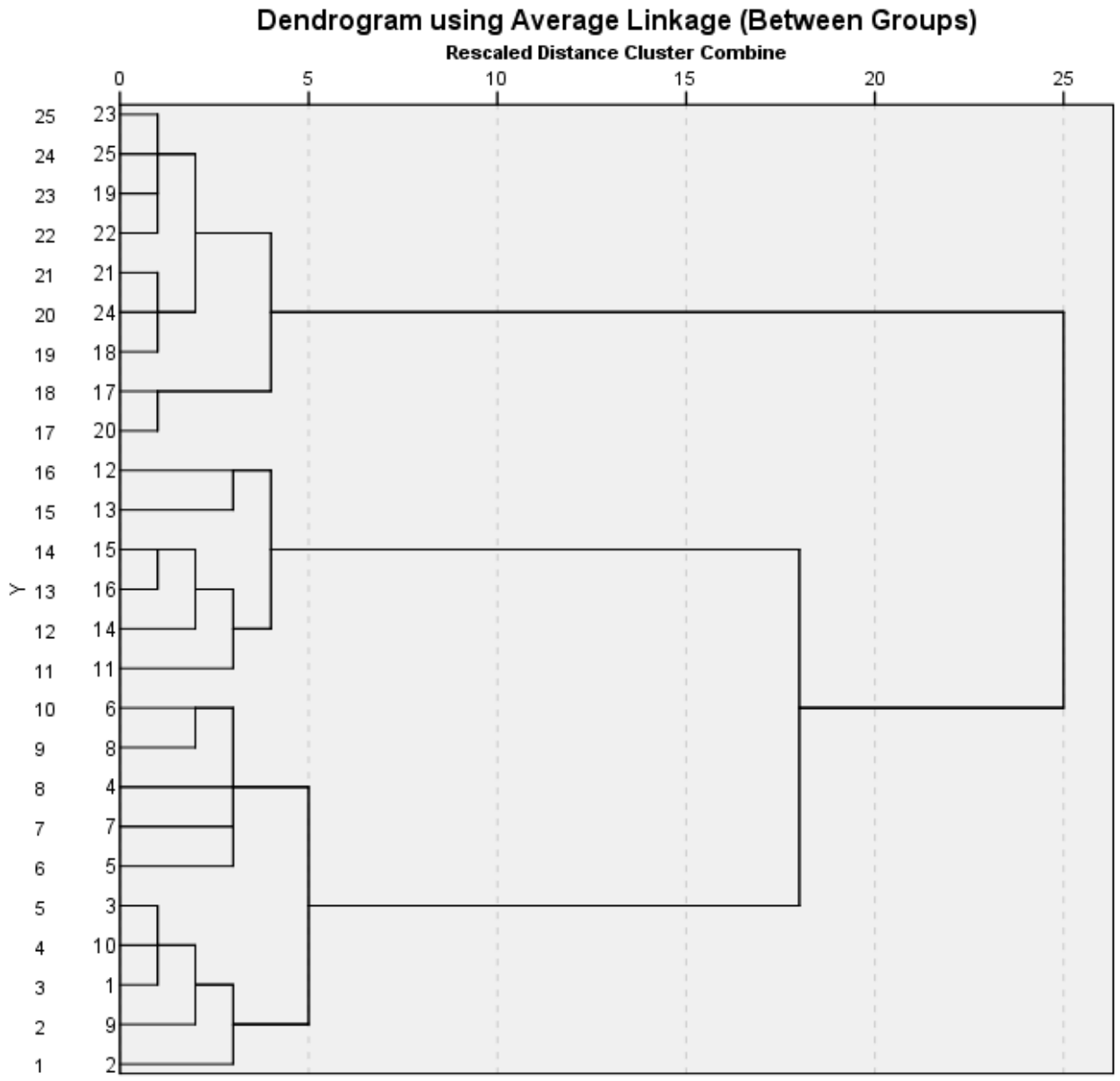
| Cases | | | | | |
|---|---|---|---|---|---|
| Valid | | Missing | | Total | |
| N | Percent | N | Percent | N | Percent |
| 25 | 100.0 | 0 | .0 | 25 | 100.0 |

a. Squared Euclidean Distance used

b. Average Linkage (Between Groups)

**Agglomeration Schedule**

| Stage | Cluster Combined | | Coefficients | Stage Cluster First Appears | | Next Stage |
|---|---|---|---|---|---|---|
| | Cluster 1 | Cluster 2 | | Cluster 1 | Cluster 2 | |
| 1 | 23 | 25 | .000 | 0 | 0 | 8 |
| 2 | 21 | 24 | .000 | 0 | 0 | 4 |
| 3 | 19 | 22 | .000 | 0 | 0 | 8 |
| 4 | 18 | 21 | .000 | 0 | 2 | 10 |
| 5 | 17 | 20 | .000 | 0 | 0 | 20 |
| 6 | 3 | 10 | 1.000 | 0 | 0 | 7 |
| 7 | 1 | 3 | 1.500 | 0 | 6 | 13 |
| 8 | 19 | 23 | 2.000 | 3 | 1 | 10 |
| 9 | 15 | 16 | 2.000 | 0 | 0 | 11 |
| 10 | 18 | 19 | 4.000 | 4 | 8 | 20 |
| 11 | 14 | 15 | 4.000 | 0 | 9 | 18 |
| 12 | 6 | 8 | 4.000 | 0 | 0 | 16 |
| 13 | 1 | 9 | 4.667 | 7 | 0 | 19 |
| 14 | 12 | 13 | 5.000 | 0 | 0 | 21 |
| 15 | 4 | 7 | 5.000 | 0 | 0 | 16 |
| 16 | 4 | 6 | 6.000 | 15 | 12 | 17 |
| 17 | 4 | 5 | 6.250 | 16 | 0 | 22 |
| 18 | 11 | 14 | 6.667 | 0 | 11 | 21 |
| 19 | 1 | 2 | 7.000 | 13 | 0 | 22 |
| 20 | 17 | 18 | 7.857 | 5 | 10 | 24 |
| 21 | 11 | 12 | 8.500 | 18 | 14 | 23 |
| 22 | 1 | 4 | 11.800 | 19 | 17 | 23 |
| 23 | 1 | 11 | 40.667 | 22 | 21 | 24 |
| 24 | 1 | 17 | 59.222 | 23 | 20 | 0 |




## Cluster Membership

| Case Number | Cluster | Distance |
|---|---|---|
| 1 | 3 | 1.706 |
| 2 | 3 | 2.304 |
| 3 | 3 | 1.873 |
| 4 | 3 | 1.646 |
| 5 | 3 | 2.124 |
| 6 | 3 | 2.472 |
| 7 | 3 | 1.453 |
| 8 | 3 | 2.390 |
| 9 | 3 | 2.076 |
| 10 | 3 | 1.646 |
| 11 | 2 | 1.908 |
| 12 | 2 | 1.280 |

| | | |
|---|---|---|
| 13 | 2 | 2.375 |
| 14 | 2 | 1.404 |
| 15 | 2 | 1.624 |
| 16 | 2 | 1.280 |
| 17 | 1 | 2.015 |
| 18 | 1 | .916 |
| 19 | 1 | 1.652 |
| 20 | 1 | 2.015 |
| 21 | 1 | .916 |
| 22 | 1 | 1.652 |
| 23 | 1 | 1.083 |
| 24 | 1 | .916 |
| 25 | 1 | 1.083 |

**Final Cluster Centers**

| | Cluster | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Indian Technology high order | 2 | 2 | 4 |
| Buy Made in India | 2 | 2 | 5 |
| Value for money | 1 | 5 | 3 |
| Convenience over style | 2 | 5 | 2 |
| Don't do waste full expenditure | 1 | 4 | 3 |
| No compromise on safety | 1 | 5 | 3 |
| saver not spender | 1 | 4 | 3 |
| try new things | 5 | 2 | 1 |
| be part of a changing world | 4 | 1 | 1 |

**ANOVA**

| | Cluster | | Error | | | |
|---|---|---|---|---|---|---|
| | Mean Square | df | Mean Square | df | F | Sig. |
| Indian Technology high order | 16.383 | 2 | .420 | 22 | 39.036 | .000 |
| Buy Made in India | 22.426 | 2 | .499 | 22 | 44.896 | .000 |
| Value for money | 18.692 | 2 | .348 | 22 | 53.716 | .000 |
| Convenience over style | 17.106 | 2 | .263 | 22 | 65.008 | .000 |
| Don't do waste full expenditure | 20.653 | 2 | .224 | 22 | 92.103 | .000 |
| No compromise on safety | 21.356 | 2 | .422 | 22 | 50.579 | .000 |
| saver not spender | 18.383 | 2 | .783 | 22 | 23.468 | .000 |
| try new things | 34.752 | 2 | .212 | 22 | 164.223 | .000 |
| be part of a changing world | 25.213 | 2 | .261 | 22 | 96.749 | .000 |

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

**Number of Cases in each Cluster**

| | | |
|---|---|---|
| Cluster | 1 | 9.000 |
| | 2 | 6.000 |
| | 3 | 10.000 |
| Valid | | 25.000 |
| Missing | | .000 |

## Inferences:

1. There are 25 respondents. No missing cases
2. From the **Agglomeration Schedule** table, we can see that 23 and 25 are most similar pair in stage 1. They are joined by 19 in stage 8th stage. 21 and 24 are the next similar pair; they are joined by 18, in the 4th stage. In total 24 clustering stages occur. This agglomeration schedule can be used to see, how many distinct clusters exist.
   We start with the last coefficient when all objects group into a single cluster value (stage 24.) next we subtract the coefficient from the 2 cluster (stage 23) as follows:
   59.222 – 40.667 = 18.55
   Then, we look at the difference between 2 clusters (stage 23) and 3 cluster (stage 22):
   40.667 – 11.800 = 28.867
   The next difference is 11.8 – 8.5 = 3.5
   Thus we can see from the data above that the maximum variation happens when we move from a two cluster solution to a three-cluster solution. So, three clusters are adequate and distinct enough for analysis or 25 respondents who took the survey can be grouped into three distinct clusters.

3. The **Dendogram** clearly gives three clusters that are distinctly different from each other.
4. Examine F values from the ANOVA tables to establish the discriminating power of each clustering variable. As observed from the table, all the variables are significant at the 5 percent level of significance and may be used for the interpretation.

5. From the **Final Cluster Centers** table, we can infer:
   Cluster 1 is high on variables ‘try new things’ and ‘be part of a changing world’.
   Cluster 2 is high on variables ‘Value for money’, ‘Convenience over style’, ‘Don't do waste full expenditure’, ‘No compromise on safety’ and ‘saver not spender’.
   Cluster 3 is high on variables ‘Indian Technology high order’ and ‘Buy Made in India’.

6. Naming clusters:
   Cluster 1 can be innovative consumer
   Cluster 2 can be cautious customer
   Cluster 3 can be patriotic customer

# Cluster Analysis - 2

**Case Processing Summary[b,c]**

| Cases | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Rejected | | | | | |
| Valid | | Missing Value | | Out of Range Binary Value[a] | | Total | |
| N | Percent | N | Percent | N | Percent | N | Percent |
| 40 | 100.0 | 0 | .0 | 0 | .0 | 40 | 100.0 |

a. Value different from both 1 and 0.

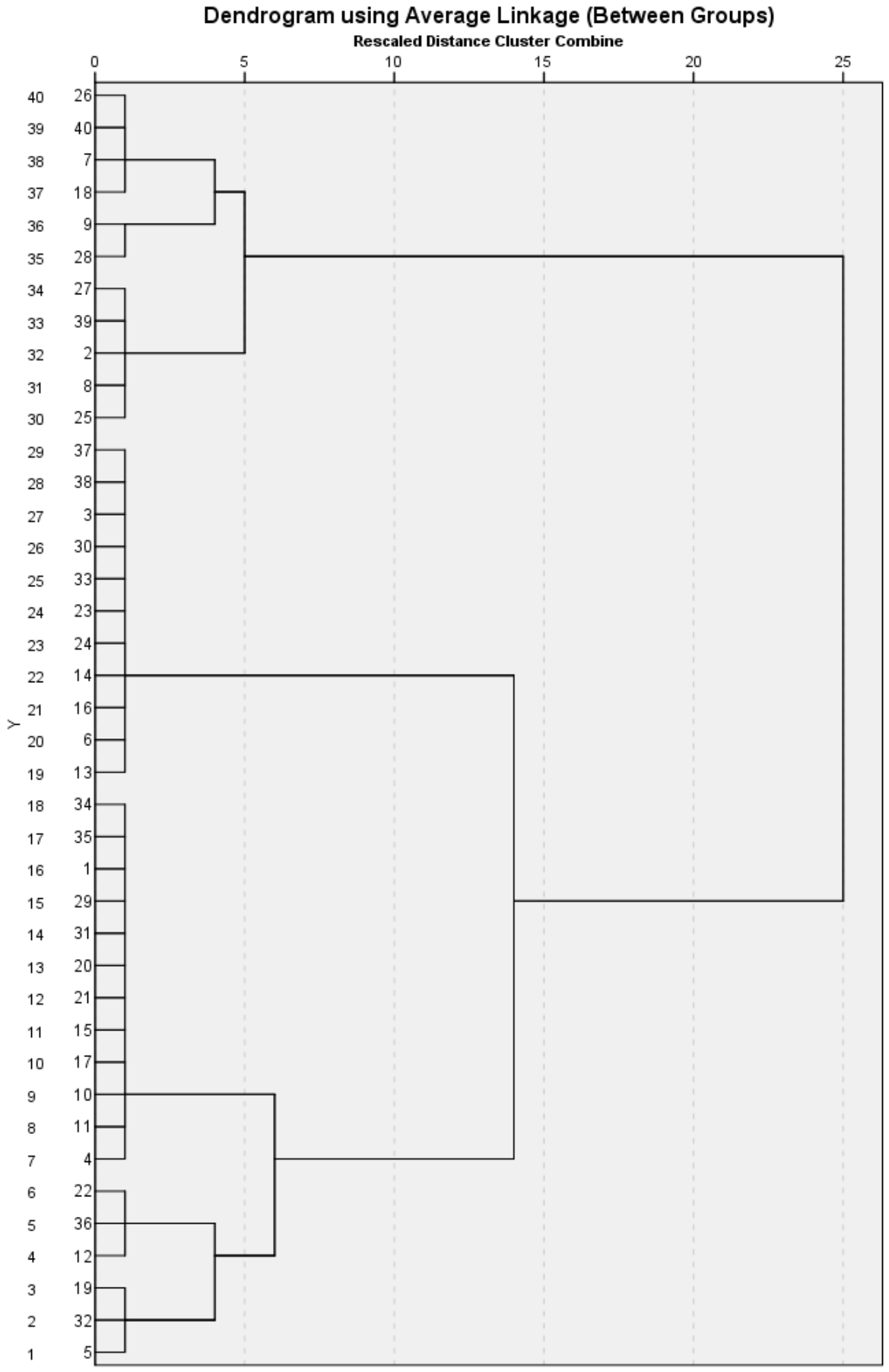
b. Simple matching Measure used

c. Average Linkage (Between Groups)

**Agglomeration Schedule**

| Stage | Cluster Combined | | Coefficients | Stage Cluster First Appears | | Next Stage |
|---|---|---|---|---|---|---|
| | Cluster 1 | Cluster 2 | | Cluster 1 | Cluster 2 | |
| 1 | 26 | 40 | 1.000 | 0 | 0 | 15 |
| 2 | 27 | 39 | 1.000 | 0 | 0 | 14 |
| 3 | 37 | 38 | 1.000 | 0 | 0 | 4 |
| 4 | 3 | 37 | 1.000 | 0 | 3 | 11 |
| 5 | 22 | 36 | 1.000 | 0 | 0 | 19 |
| 6 | 34 | 35 | 1.000 | 0 | 0 | 7 |
| 7 | 1 | 34 | 1.000 | 0 | 6 | 12 |
| 8 | 30 | 33 | 1.000 | 0 | 0 | 11 |
| 9 | 19 | 32 | 1.000 | 0 | 0 | 22 |
| 10 | 29 | 31 | 1.000 | 0 | 0 | 12 |
| 11 | 3 | 30 | 1.000 | 4 | 8 | 18 |
| 12 | 1 | 29 | 1.000 | 7 | 10 | 21 |
| 13 | 9 | 28 | 1.000 | 0 | 0 | 35 |
| 14 | 2 | 27 | 1.000 | 0 | 2 | 31 |
| 15 | 7 | 26 | 1.000 | 0 | 1 | 23 |
| 16 | 8 | 25 | 1.000 | 0 | 0 | 31 |
| 17 | 23 | 24 | 1.000 | 0 | 0 | 18 |
| 18 | 3 | 23 | 1.000 | 11 | 17 | 27 |
| 19 | 12 | 22 | 1.000 | 0 | 5 | 34 |
| 20 | 20 | 21 | 1.000 | 0 | 0 | 21 |
| 21 | 1 | 20 | 1.000 | 12 | 20 | 26 |
| 22 | 5 | 19 | 1.000 | 0 | 9 | 34 |
| 23 | 7 | 18 | 1.000 | 15 | 0 | 35 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 24 | 15 | 17 | 1.000 | 0 | 0 | 26 |
| 25 | 14 | 16 | 1.000 | 0 | 0 | 27 |
| 26 | 1 | 15 | 1.000 | 21 | 24 | 30 |
| 27 | 3 | 14 | 1.000 | 18 | 25 | 32 |
| 28 | 6 | 13 | 1.000 | 0 | 0 | 32 |
| 29 | 10 | 11 | 1.000 | 0 | 0 | 30 |
| 30 | 1 | 10 | 1.000 | 26 | 29 | 33 |
| 31 | 2 | 8 | 1.000 | 14 | 16 | 36 |
| 32 | 3 | 6 | 1.000 | 27 | 28 | 38 |
| 33 | 1 | 4 | 1.000 | 30 | 0 | 37 |
| 34 | 5 | 12 | .889 | 22 | 19 | 37 |
| 35 | 7 | 9 | .889 | 23 | 13 | 36 |
| 36 | 2 | 7 | .852 | 31 | 35 | 39 |
| 37 | 1 | 5 | .833 | 33 | 34 | 38 |
| 38 | 1 | 3 | .611 | 37 | 32 | 39 |
| 39 | 1 | 2 | .258 | 38 | 36 | 0 |


Dendrogram using Average Linkage (Between Groups)
Rescaled Distance Cluster Combine
0
5
10
15
20
25
Y
40 26
39 40
38 7
37 18
36 9
35 28
34 27
33 39
32 2
31 8
30 25
29 37
28 38
27 3
26 30
25 33
24 23
23 24
22 14
21 16
20 6
19 13
18 34
17 35
16 1
15 29
14 31
13 20
12 21
11 15
10 17
9 10
8 11
7 4
6 22
5 36
4 12
3 19
2 32
1 5

**Cluster Membership**

| Case Number | children below 18 | Cluster | Distance |
|---|---|---|---|
| 1 | 1 | 1 | .930 |
| 2 | 0 | 2 | .575 |
| 3 | 1 | 1 | 1.100 |
| 4 | 0 | 1 | .930 |
| 5 | 1 | 1 | .832 |
| 6 | 0 | 1 | 1.100 |
| 7 | 0 | 2 | .490 |
| 8 | 0 | 2 | .575 |
| 9 | 0 | 2 | .936 |
| 10 | 2 | 1 | .930 |
| 11 | 1 | 1 | .930 |
| 12 | 2 | 1 | .852 |
| 13 | 1 | 1 | 1.100 |
| 14 | 0 | 1 | 1.100 |
| 15 | 1 | 1 | .930 |
| 16 | 1 | 1 | 1.100 |
| 17 | 0 | 1 | .930 |
| 18 | 0 | 2 | .490 |
| 19 | 1 | 1 | .832 |
| 20 | 2 | 1 | .930 |
| 21 | 1 | 1 | .930 |
| 22 | 2 | 1 | .852 |
| 23 | 1 | 1 | 1.100 |
| 24 | 0 | 1 | 1.100 |
| 25 | 0 | 2 | .575 |
| 26 | 0 | 2 | .490 |
| 27 | 0 | 2 | .575 |
| 28 | 0 | 2 | .936 |
| 29 | 1 | 1 | .930 |
| 30 | 1 | 1 | 1.100 |
| 31 | 0 | 1 | .930 |
| 32 | 1 | 1 | .832 |
| 33 | 0 | 1 | 1.100 |
| 34 | 2 | 1 | .930 |
| 35 | 1 | 1 | .930 |
| 36 | 2 | 1 | .852 |
| 37 | 1 | 1 | 1.100 |
| 38 | 0 | 1 | 1.100 |
| 39 | 0 | 2 | .575 |
| 40 | 0 | 2 | .490 |

**Final Cluster Centers**

| | Cluster | |
|---|---|---|
| | 1 | 2 |
| bournvita | 1 | 0 |
| milo | 1 | 0 |
| zanduchyawanprash | 0 | 1 |
| dabur red | 0 | 1 |
| dabur blue | 0 | 1 |
| protinex | 0 | 0 |
| horlicks | 1 | 0 |
| baidyanath | 0 | 1 |

| | | |
|---|---|---|
| complan | 1 | 0 |

**ANOVA**

| | Cluster | | Error | | | |
|---|---|---|---|---|---|---|
| | Mean Square | df | Mean Square | df | F | Sig. |
| bournvita | 3.072 | 1 | .180 | 38 | 17.100 | .000 |
| milo | 3.072 | 1 | .180 | 38 | 17.100 | .000 |
| zanduchyawanprash | 5.339 | 1 | .043 | 38 | 123.975 | .000 |
| dabur red | 2.373 | 1 | .072 | 38 | 33.060 | .000 |
| dabur blue | 7.975 | 1 | .000 | 38 | . | . |
| protinex | 1.859 | 1 | .191 | 38 | 9.753 | .003 |
| horlicks | 7.975 | 1 | .000 | 38 | . | . |
| baidyanath | 7.975 | 1 | .000 | 38 | . | . |
| complan | 2.741 | 1 | .185 | 38 | 14.804 | .000 |

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

**Number of Cases in each Cluster**

| | | |
|---|---|---|
| Cluster | 1 | 29.000 |
| | 2 | 11.000 |
| Valid | | 40.000 |
| Missing | | .000 |

## Inferences:

1. There are 40 respondents. No missing cases
2. From the **Agglomeration Schedule** table, we can see that 26 and 40 are most similar pair in stage 1. They are joined with 7 in 15$^{th}$ stage. 27 and 39 are the next similar pair; they are joined with 2, in the 14$^{th}$ stage. In total 39 clustering stages occur.
   We can see from the data above that the maximum variation happens when we move from a one cluster solution to a two-cluster solution. Two clusters are adequate and distinct enough for analysis. Or 40 respondents who took the survey can be grouped into two distinct clusters.

3. The **Dendogram** clearly gives two clusters that are distinctly different from each other.
4. Examine F values from the ANOVA tables to establish the discriminating power of each clustering variable. As observed from the table, all the variables are significant at the 5 percent level of significance and may be used for the interpretation.

5. From the **Final Cluster Centers** table, we can infer:
   Cluster 1 is Bournvita, Milo, Horlicks and Complan.
   Cluster 2 is Zandu Chyawanprash, Dabur red, Dabur Blue and Baidyanath.

6. Naming clusters:
   Cluster 1: Non-Ayurvedic Milk based drinks
   Cluster 2: Ayurvedic health drinks

## Multidimensionality Scaling

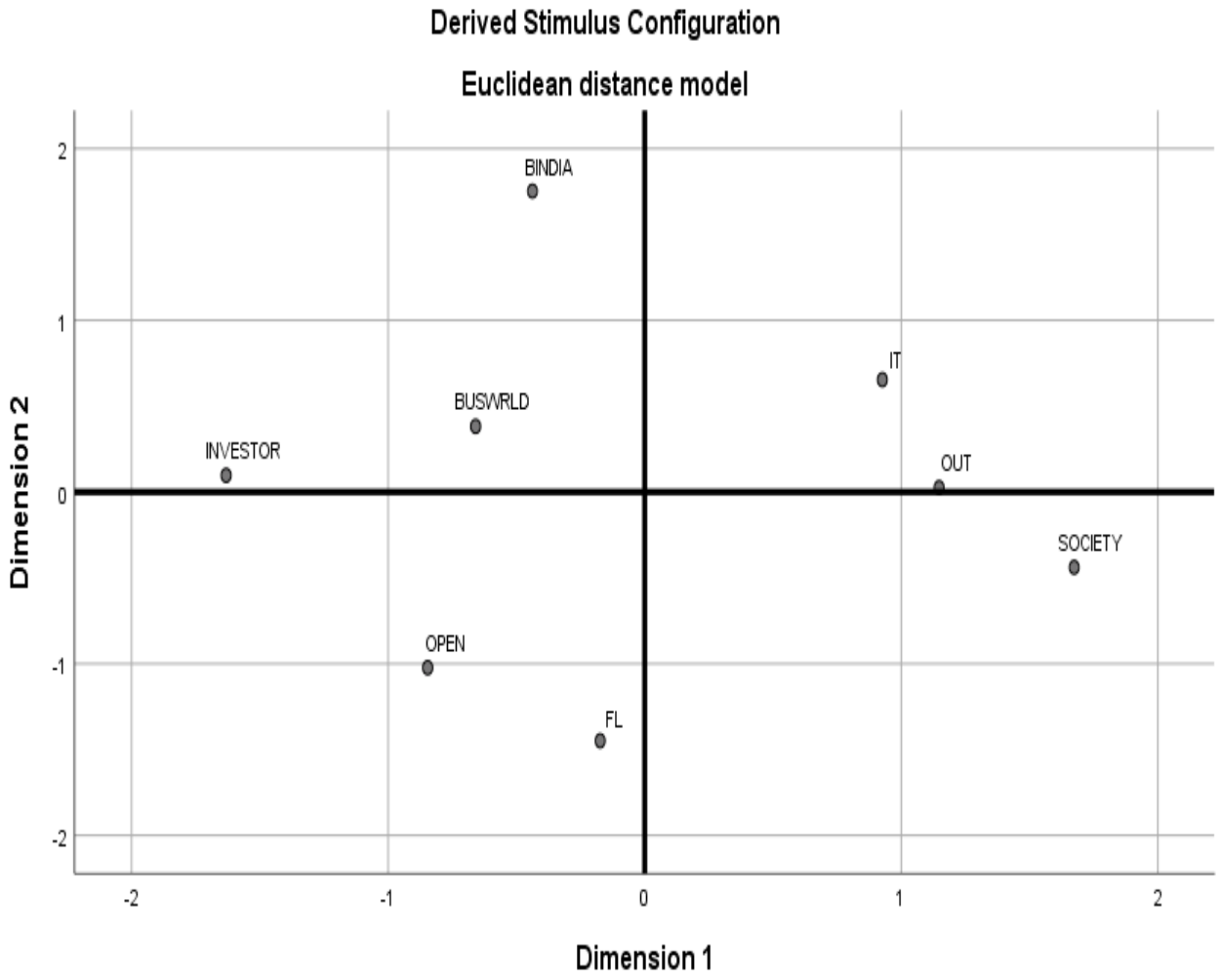
For matrix

Stress = .20997 RSQ = .62649

Configuration derived in 2 dimensions

Stimulus Coordinates

| Stimulus Number | Stimulus Name | Dimension 1 | Dimension 2 |
|---|---|---|---|
| 1 | FL | -.1735 | -1.4487 |
| 2 | SOCIETY | 1.6737 | -.4389 |
| 3 | IT | .9264 | .6538 |
| 4 | OUT | 1.1472 | .0258 |
| 5 | BUSWRLD | -.6590 | .3820 |
| 6 | OPEN | -.8459 | -1.0233 |
| 7 | INVESTOR | -1.6313 | .0970 |
| 8 | BINDIA | -.4375 | 1.7522 |




## Inferences:

1. R Square value

   If the R square value is more than 0.6 for the number of dimensions, the solution is acceptable. In this example, for dimension 2, the RSQ value is 0.62649. Hence we accept a two dimensional solution.

2. The output gives the coordinates of the eight magazines on 2 dimensions. We consider the placement of the magazines and the corresponding coordinates to name the dimensions.

   If we examine the first dimension, we find the society is the highest here, with India today and outlook close together and the last on this dimension is Investor. This seems to be “magazine content” ranging from general interest to specific interest.

   The second dimension has business india at the top and open at the bottom and looking at the placement of the other six magazines, this seems to be “subscription base”, ranging from corporate readership to general readership.

3. Manager's decision

   Thus, based on the similarity analysis, the management can conclude that "society "is a magazine that was of general interest and seemed to be enjoying an uncluttered space. Thus rather than looking at a specialized and a corporate base, the new magazine would be a general interest magazine that will cover on everyday issues. It would not be high on political content like India today or outlook but would focus on lifestyle issues. The name of the monthly magazine would be life & Times.

## Multidimensionality Scaling - 2

```
                For  matrix
    Stress  =   .07677      RSQ =  .95947
```

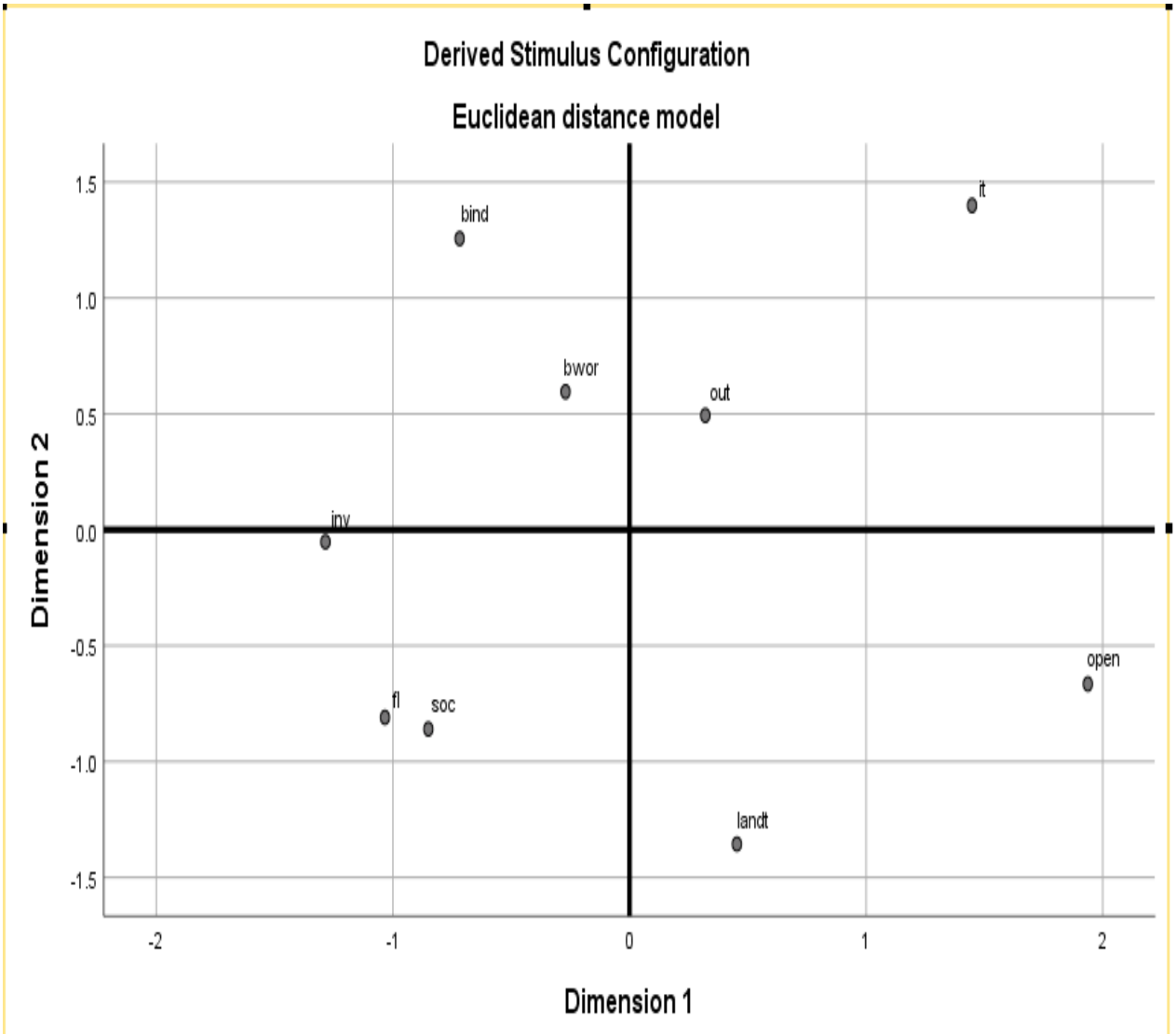
```
           Configuration derived in 2 dimensions
```

```
                    Stimulus Coordinates

                         Dimension

Stimulus   Stimulus      1        2
 Number      Name

    1       it         1.4473   1.3989
    2       out         .3202    .4932
    3       bwor       -.2705    .5955
    4       open       1.9368   -.6654
    5       inv       -1.2852   -.0525
    6       bind       -.7182   1.2562
    7       soc        -.8504   -.8599
    8       fl        -1.0337   -.8097
    9       landt       .4536  -1.3564
```




## Inferences:

1. **R Square value**

   If the R square value is more than 0.6 for the number of dimensions, the solution is acceptable. In this example, for dimension 2, the RSQ value is 0.95947. Hence we accept a two dimensional solution.

2. The output gives the coordinates of the eight magazines on 2 dimensions. We consider the placement of the magazines and the corresponding coordinates to name the dimensions.

   If we examine the first dimension, we find the India today and outlook are found and the last on this dimension is Investor. This seems to be “magazine content” ranging from general interest to specific interest.

   The second dimension has business india at the top and open at the bottom and looking at the placement of the other six magazines, this seems to be “subscription base”, ranging from corporate readership to general readership.

3. Manager's decision

   Thus, based on the similarity analysis, the management can conclude that "society "is a magazine that was of general interest and seemed to be enjoying an uncluttered space. Thus rather than looking at a specialized and a corporate base, the new magazine would be a general interest magazine that will cover on everyday issues. It would not be high on political content like India today or outlook but would focus on lifestyle issues. The name of the monthly magazine would be life & Times.

## Cluster Analysis - 3

**Case Processing Summary[a,b]**

| Cases | | | | | |
|---|---|---|---|---|---|
| Valid | | Missing | | Total | |
| N | Percent | N | Percent | N | Percent |
| 40 | 100.0 | 0 | .0 | 40 | 100.0 |

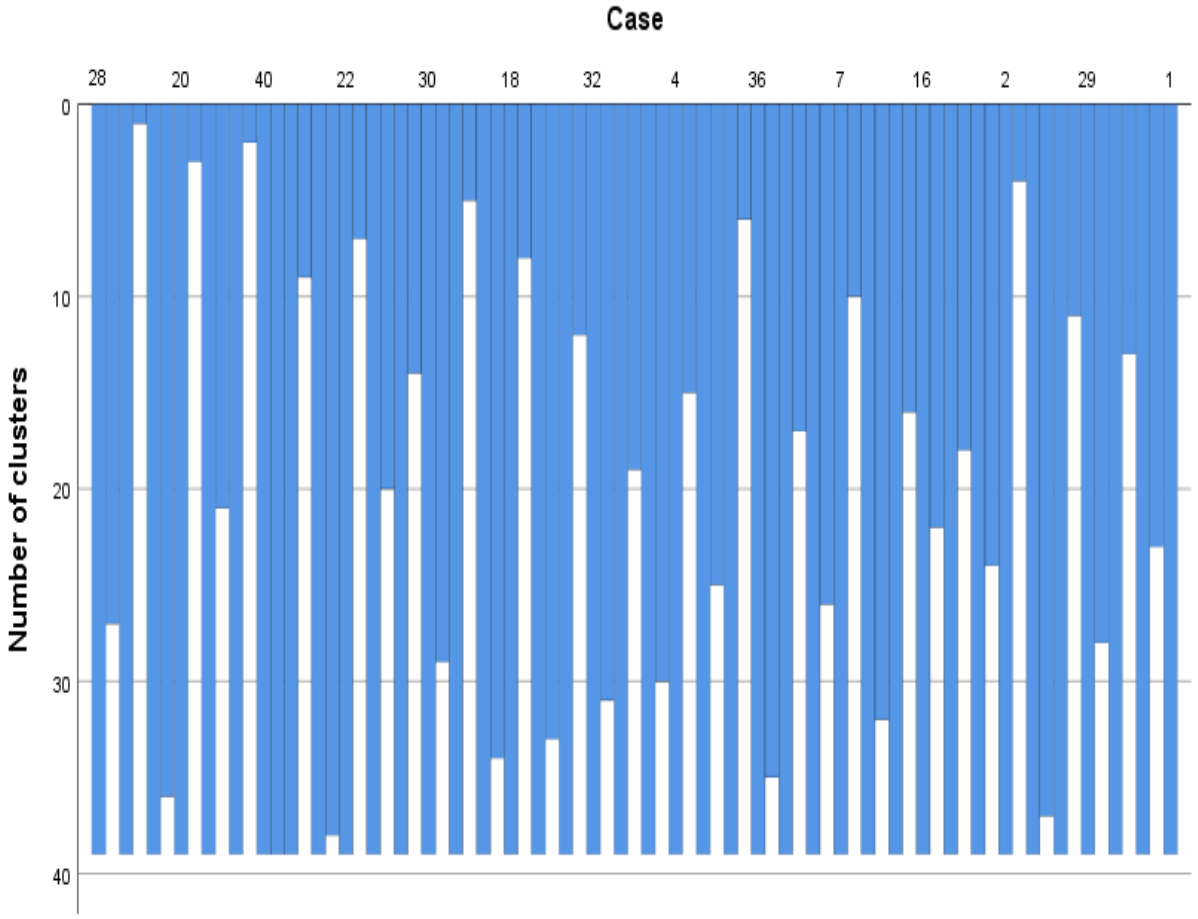
a. Squared Euclidean Distance used

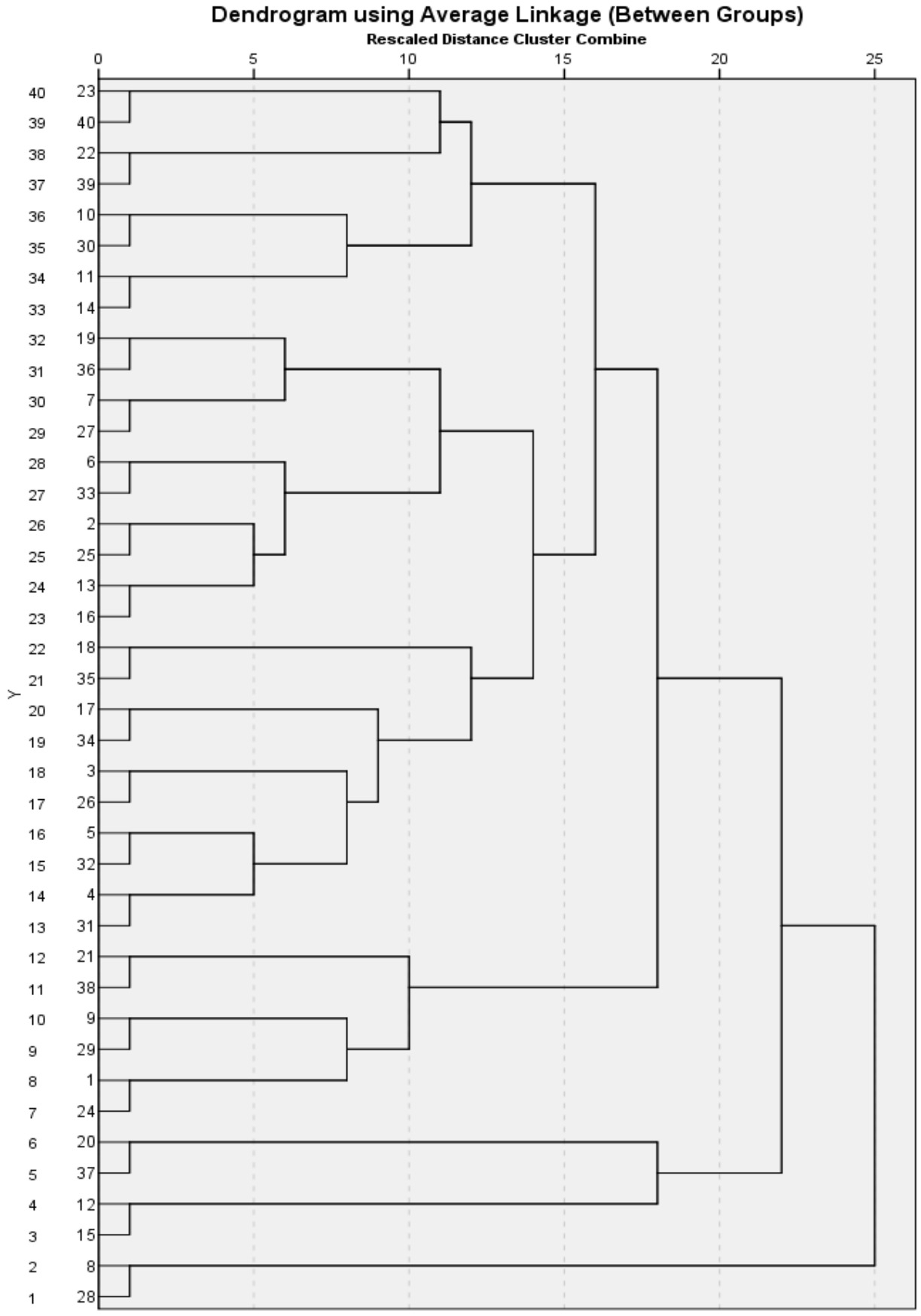
b. Average Linkage (Between Groups)

**Agglomeration Schedule**

| | Cluster Combined | | | Stage Cluster First Appears | | |
|---|---|---|---|---|---|---|
| Stage | Cluster 1 | Cluster 2 | Coefficients | Cluster 1 | Cluster 2 | Next Stage |
| 1 | 23 | 40 | .000 | 0 | 0 | 31 |
| 2 | 22 | 39 | .000 | 0 | 0 | 31 |
| 3 | 21 | 38 | .000 | 0 | 0 | 29 |
| 4 | 20 | 37 | .000 | 0 | 0 | 37 |
| 5 | 19 | 36 | .000 | 0 | 0 | 23 |
| 6 | 18 | 35 | .000 | 0 | 0 | 32 |
| 7 | 17 | 34 | .000 | 0 | 0 | 28 |
| 8 | 6 | 33 | .000 | 0 | 0 | 24 |
| 9 | 5 | 32 | .000 | 0 | 0 | 21 |
| 10 | 4 | 31 | .000 | 0 | 0 | 21 |
| 11 | 10 | 30 | .000 | 0 | 0 | 26 |
| 12 | 9 | 29 | .000 | 0 | 0 | 27 |
| 13 | 8 | 28 | .000 | 0 | 0 | 39 |
| 14 | 7 | 27 | .000 | 0 | 0 | 23 |
| 15 | 3 | 26 | .000 | 0 | 0 | 25 |
| 16 | 2 | 25 | .000 | 0 | 0 | 22 |
| 17 | 1 | 24 | .000 | 0 | 0 | 27 |
| 18 | 13 | 16 | .000 | 0 | 0 | 22 |
| 19 | 12 | 15 | .000 | 0 | 0 | 37 |
| 20 | 11 | 14 | .000 | 0 | 0 | 26 |
| 21 | 4 | 5 | 8.000 | 10 | 9 | 25 |
| 22 | 2 | 13 | 9.000 | 16 | 18 | 24 |
| 23 | 7 | 19 | 11.000 | 14 | 5 | 30 |
| 24 | 2 | 6 | 11.500 | 22 | 8 | 30 |
| 25 | 3 | 4 | 14.000 | 15 | 21 | 28 |
| 26 | 10 | 11 | 15.000 | 11 | 20 | 33 |
| 27 | 1 | 9 | 15.000 | 17 | 12 | 29 |
| 28 | 3 | 17 | 16.667 | 25 | 7 | 32 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 29 | 1 | 21 | 19.500 | 27 | 3 | 36 |
| 30 | 2 | 7 | 20.167 | 24 | 23 | 34 |
| 31 | 22 | 23 | 21.000 | 2 | 1 | 33 |
| 32 | 3 | 18 | 22.000 | 28 | 6 | 34 |
| 33 | 10 | 22 | 23.500 | 26 | 31 | 35 |
| 34 | 2 | 3 | 25.880 | 30 | 32 | 35 |
| 35 | 2 | 10 | 30.450 | 34 | 33 | 36 |
| 36 | 1 | 2 | 33.833 | 29 | 35 | 38 |
| 37 | 12 | 20 | 34.000 | 19 | 4 | 38 |
| 38 | 1 | 12 | 41.294 | 36 | 37 | 39 |
| 39 | 1 | 8 | 49.000 | 38 | 13 | 0 |




Dendrogram using Average Linkage (Between Groups)
Rescaled Distance Cluster Combine
0
5
10
15
20
25
Y
40 23
39 40
38 22
37 39
36 10
35 30
34 11
33 14
32 19
31 36
30 7
29 27
28 6
27 33
26 2
25 25
24 13
23 16
22 18
21 35
20 17
19 34
18 3
17 26
16 5
15 32
14 4
13 31
12 21
11 38
10 9
9 29
8 1
7 24
6 20
5 37
4 12
3 15
2 8
1 28

## Cluster Membership

| Case Number | Cluster | Distance |
|---|---|---|
| 1 | 1 | 3.459 |
| 2 | 2 | 3.241 |
| 3 | 2 | 4.104 |
| 4 | 1 | 3.749 |
| 5 | 1 | 3.550 |
| 6 | 1 | 2.933 |
| 7 | 2 | 3.392 |
| 8 | 2 | 2.839 |
| 9 | 2 | 3.675 |
| 10 | 1 | 3.973 |
| 11 | 1 | 4.271 |
| 12 | 2 | 3.441 |
| 13 | 1 | 2.839 |
| 14 | 1 | 2.587 |
| 15 | 1 | 3.459 |
| 16 | 2 | 4.353 |
| 17 | 2 | 3.392 |
| 18 | 2 | 2.839 |
| 19 | 2 | 3.675 |
| 20 | 1 | 3.973 |
| 21 | 1 | 4.141 |
| 22 | 1 | 4.196 |
| 23 | 1 | 3.406 |
| 24 | 1 | 4.141 |
| 25 | 1 | 4.196 |
| 26 | 1 | 3.406 |
| 27 | 2 | 4.022 |
| 28 | 2 | 4.131 |
| 29 | 2 | 4.022 |
| 30 | 2 | 4.131 |
| 31 | 2 | 4.353 |
| 32 | 1 | 4.271 |
| 33 | 2 | 3.441 |
| 34 | 1 | 2.839 |
| 35 | 1 | 2.587 |
| 36 | 2 | 3.241 |
| 37 | 2 | 4.104 |
| 38 | 1 | 3.749 |
| 39 | 1 | 3.550 |
| 40 | 1 | 2.933 |

## Final Cluster Centers

| | Cluster | |
|---|---|---|
| | 1 | 2 |
| skip breakfast | 2.09 | 2.67 |
| keep ready to eat | 2.18 | 3.67 |
| watch familysoap | 2.55 | 2.78 |

| | | |
|---|---|---|
| breakfastbestmeal | 2.45 | 2.56 |
| skillbasededucation | 2.36 | 3.33 |
| buypackedmilk | 2.00 | 3.78 |
| friendsonfacebook | 2.09 | 2.00 |
| wifeworkstosupport | 2.55 | 3.22 |
| womenempowerment | 2.45 | 3.11 |
| sundayevenings | 2.09 | 2.78 |
| qualityconscious | 2.09 | 4.33 |
| dontmakeweekendmeals | 2.09 | 2.67 |
| mobilephones | 1.64 | 2.22 |
| india is a force | 2.27 | 2.22 |
| quality improvement | 2.00 | 4.00 |

**ANOVA**

| | Cluster | | Error | | | |
|---|---|---|---|---|---|---|
| | Mean Square | df | Mean Square | df | F | Sig. |
| skip breakfast | 3.282 | 1 | .995 | 38 | 3.298 | .077 |
| keep ready to eat | 21.827 | 1 | .612 | 38 | 35.640 | .000 |
| watch familysoap | .534 | 1 | 1.278 | 38 | .418 | .522 |
| breakfastbestmeal | .101 | 1 | 1.313 | 38 | .077 | .783 |
| skillbasededucation | 9.309 | 1 | 1.292 | 38 | 7.206 | .011 |
| buypackedmilk | 31.289 | 1 | .608 | 38 | 51.446 | .000 |
| friendsonfacebook | .082 | 1 | .785 | 38 | .104 | .749 |
| wifeworkstosupport | 4.534 | 1 | 1.383 | 38 | 3.278 | .078 |
| womenempowerment | 4.268 | 1 | .927 | 38 | 4.603 | .038 |
| sundayevenings | 4.671 | 1 | 1.077 | 38 | 4.336 | .044 |
| qualityconscious | 49.782 | 1 | .995 | 38 | 50.021 | .000 |
| dontmakeweekendmeals | 3.282 | 1 | .890 | 38 | 3.688 | .062 |
| mobilephones | 3.398 | 1 | .426 | 38 | 7.970 | .008 |
| india is a force | .025 | 1 | .723 | 38 | .035 | .853 |
| quality improvement | 39.600 | 1 | .737 | 38 | 53.743 | .000 |

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

**Number of Cases in each Cluster**

| | | |
|---|---|---|
| Cluster | 1 | 22.000 |
| | 2 | 18.000 |
| Valid | | 40.000 |
| Missing | | .000 |

## Inferences:

1. There are 40 respondents. No missing cases.

2. From the **Agglomeration Schedule** table, we can see that 23 and 40 are most similar pair in stage 1. They are joined with 22 in 31$^{st}$ stage. 22 and 39 are the next similar pair; they are joined with 23, in the 31$^{st}$ stage. In total 39 clustering stages occur.
   We can see from the data above that the maximum variation happens when we move from a one cluster solution to a two-cluster solution. Two clusters are adequate and distinct enough for analysis. Or 40 respondents who took the survey can be grouped into two distinct clusters.

3. The **Dendogram** clearly gives two clusters that are distinctly different from each other.

4. Examine F values from the **ANOVA table** to establish the discriminating power of each clustering variable. As observed from the table, not all the variables are significant at the 5 percent level of significance and only 4 variables can be used for the interpretation.

5. From the **Final Cluster Centers** table, we cannot infer anything clearly.

6. Even though the **Cluster Membership table,** groups the variables into Cluster 1 & Cluster 2, the Final Cluster Centers table is not so clear on the classifying the variables into clusters.

# Cluster Analysis - 4

## Case Processing Summary[a,b]

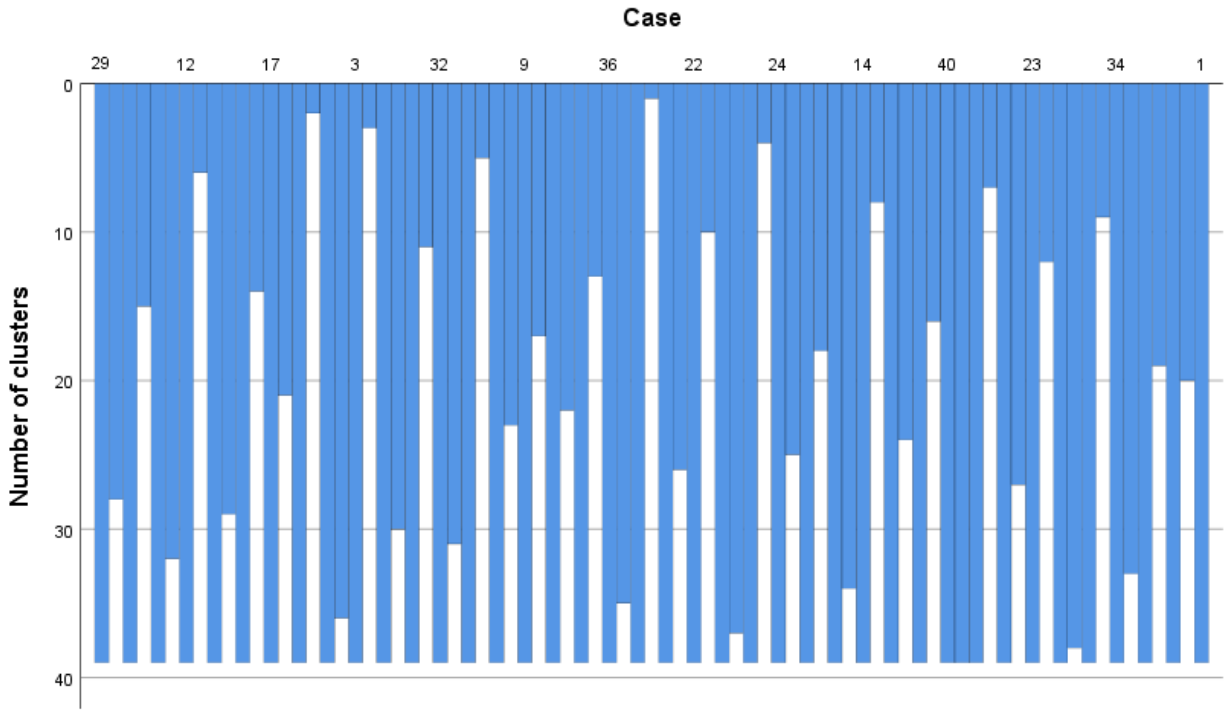
| Cases | | | | | |
|---|---|---|---|---|---|
| Valid | | Missing | | Total | |
| N | Percent | N | Percent | N | Percent |
| 40 | 100.0 | 0 | .0 | 40 | 100.0 |

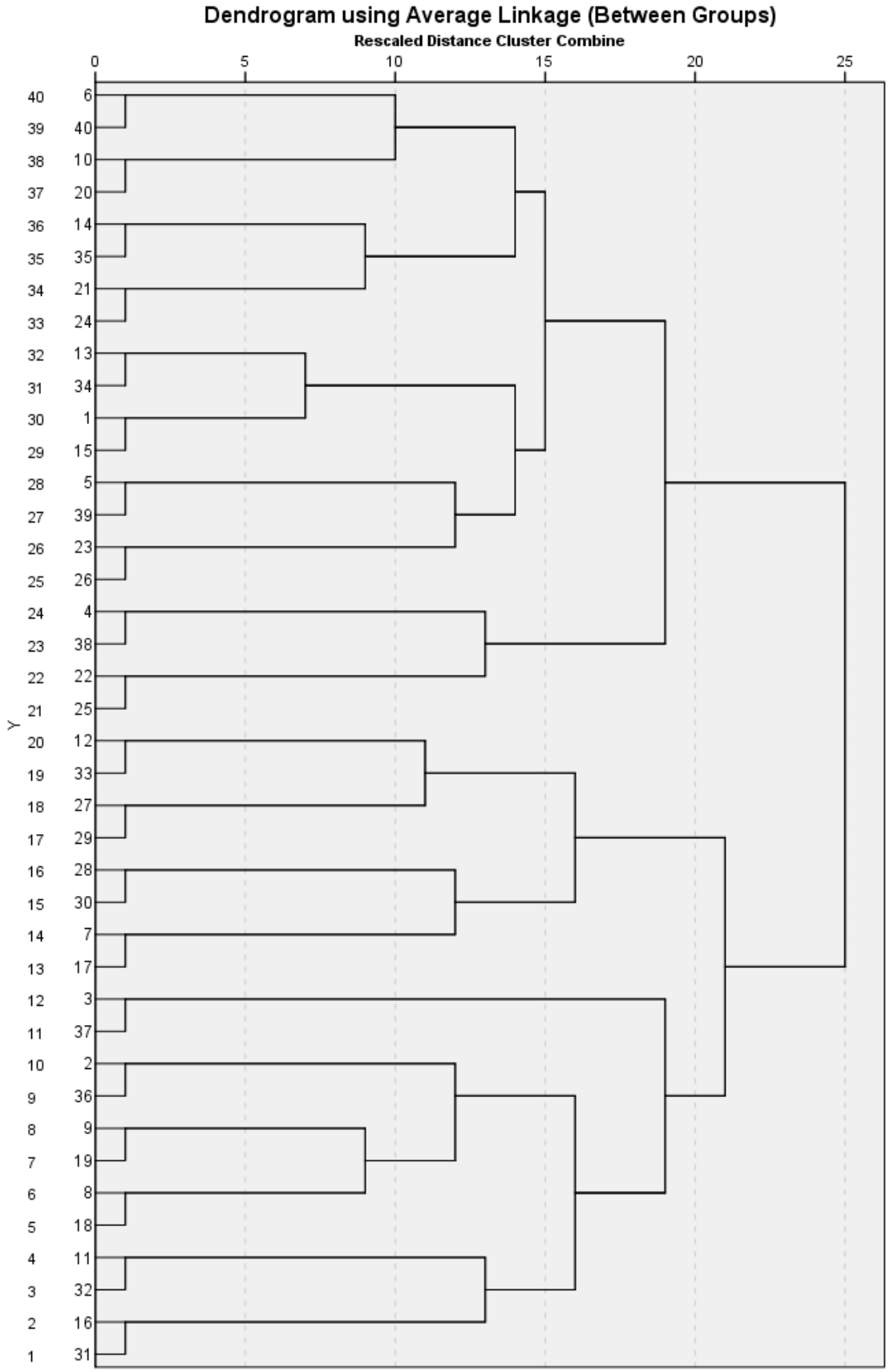
a. Squared Euclidean Distance used

b. Average Linkage (Between Groups)

## Agglomeration Schedule

| | Cluster Combined | | | Stage Cluster First Appears | | |
|---|---|---|---|---|---|---|
| Stage | Cluster 1 | Cluster 2 | Coefficients | Cluster 1 | Cluster 2 | Next Stage |
| 1 | 6 | 40 | .000 | 0 | 0 | 24 |
| 2 | 5 | 39 | .000 | 0 | 0 | 28 |
| 3 | 4 | 38 | .000 | 0 | 0 | 30 |
| 4 | 3 | 37 | .000 | 0 | 0 | 37 |
| 5 | 2 | 36 | .000 | 0 | 0 | 27 |
| 6 | 14 | 35 | .000 | 0 | 0 | 22 |
| 7 | 13 | 34 | .000 | 0 | 0 | 21 |
| 8 | 12 | 33 | .000 | 0 | 0 | 25 |
| 9 | 11 | 32 | .000 | 0 | 0 | 29 |
| 10 | 16 | 31 | .000 | 0 | 0 | 29 |
| 11 | 28 | 30 | .000 | 0 | 0 | 26 |
| 12 | 27 | 29 | .000 | 0 | 0 | 25 |
| 13 | 23 | 26 | .000 | 0 | 0 | 28 |
| 14 | 22 | 25 | .000 | 0 | 0 | 30 |
| 15 | 21 | 24 | .000 | 0 | 0 | 22 |
| 16 | 10 | 20 | .000 | 0 | 0 | 24 |
| 17 | 9 | 19 | .000 | 0 | 0 | 23 |
| 18 | 8 | 18 | .000 | 0 | 0 | 23 |
| 19 | 7 | 17 | .000 | 0 | 0 | 26 |
| 20 | 1 | 15 | .000 | 0 | 0 | 21 |
| 21 | 1 | 13 | 11.000 | 20 | 7 | 31 |
| 22 | 14 | 21 | 15.000 | 6 | 15 | 32 |
| 23 | 8 | 9 | 15.000 | 18 | 17 | 27 |
| 24 | 6 | 10 | 17.000 | 1 | 16 | 32 |
| 25 | 12 | 27 | 19.000 | 8 | 12 | 34 |
| 26 | 7 | 28 | 20.000 | 19 | 11 | 34 |
| 27 | 2 | 8 | 20.500 | 5 | 23 | 35 |
| 28 | 5 | 23 | 21.000 | 2 | 13 | 31 |
| 29 | 11 | 16 | 22.000 | 9 | 10 | 35 |
| 30 | 4 | 22 | 23.000 | 3 | 14 | 36 |
| 31 | 1 | 5 | 23.500 | 21 | 28 | 33 |
| 32 | 6 | 14 | 24.000 | 24 | 22 | 33 |
| 33 | 1 | 6 | 26.375 | 31 | 32 | 36 |
| 34 | 7 | 12 | 28.000 | 26 | 25 | 38 |
| 35 | 2 | 11 | 28.000 | 27 | 29 | 37 |
| 36 | 1 | 4 | 32.750 | 33 | 30 | 39 |
| 37 | 2 | 3 | 32.800 | 35 | 4 | 38 |
| 38 | 2 | 7 | 36.250 | 37 | 34 | 39 |
| 39 | 1 | 2 | 44.420 | 36 | 38 | 0 |


Case
29
12
17
3
32
9
36
22
24
14
40
23
34
1
Number of clusters
0
10
20
30
40


Dendrogram using Average Linkage (Between Groups)
Rescaled Distance Cluster Combine
0
5
10
15
20
25
Y
40 6
39 40
38 10
37 20
36 14
35 35
34 21
33 24
32 13
31 34
30 1
29 15
28 5
27 39
26 23
25 26
24 4
23 38
22 22
21 25
20 12
19 33
18 27
17 29
16 28
15 30
14 7
13 17
12 3
11 37
10 2
9 36
8 9
7 19
6 8
5 18
4 11
3 32
2 16
1 31

## Cluster Membership

| Case Number | Cluster | Distance |
|---|---|---|
| 1 | 1 | 3.230 |
| 2 | 2 | 2.419 |
| 3 | 1 | 2.709 |
| 4 | 1 | 3.011 |
| 5 | 1 | 2.251 |
| 6 | 2 | 2.253 |
| 7 | 2 | 4.264 |
| 8 | 2 | 4.342 |
| 9 | 1 | 4.240 |
| 10 | 1 | 3.500 |
| 11 | 1 | 4.109 |
| 12 | 2 | 4.051 |
| 13 | 2 | 1.934 |
| 14 | 1 | 4.109 |
| 15 | 2 | 4.051 |
| 16 | 2 | 1.934 |
| 17 | 1 | 3.870 |
| 18 | 1 | 3.640 |
| 19 | 2 | 2.762 |
| 20 | 2 | 4.443 |
| 21 | 2 | 4.456 |
| 22 | 1 | 3.665 |
| 23 | 1 | 3.258 |
| 24 | 1 | 3.230 |
| 25 | 2 | 2.419 |
| 26 | 1 | 2.709 |
| 27 | 2 | 4.264 |
| 28 | 2 | 4.342 |
| 29 | 1 | 4.240 |
| 30 | 1 | 3.500 |
| 31 | 1 | 3.011 |
| 32 | 1 | 2.251 |
| 33 | 2 | 2.253 |
| 34 | 1 | 3.870 |
| 35 | 1 | 3.640 |
| 36 | 2 | 2.762 |
| 37 | 2 | 4.443 |
| 38 | 2 | 4.456 |
| 39 | 1 | 3.665 |
| 40 | 1 | 3.258 |

## Final Cluster Centers

| | Cluster | |
|---|---|---|
| | 1 | 2 |
| dont buy product not from an established brand | 3.55 | 2.00 |
| buy only new products tried & tested as safe | 3.09 | 2.33 |
| I know most cosmetic brands | 2.91 | 3.78 |
| one company cannot provide complete personalcare solution | 2.45 | 3.00 |
| plan shopping trips carefully | 2.64 | 3.11 |

| | | |
|---|---|---|
| personal care companies need to do a lot of research | 3.73 | 2.33 |
| important to look good and presentable in todays times | 2.36 | 3.11 |
| like experimenting with new trends and style | 2.73 | 3.78 |
| go by what filmstars endorse | 4.36 | 3.11 |
| what i wear reflects who I am | 2.00 | 3.78 |

**ANOVA**

| | Cluster | | Error | | | |
|---|---|---|---|---|---|---|
| | Mean Square | df | Mean Square | df | F | Sig. |
| dont buy product not from an established brand | 23.645 | 1 | 1.301 | 38 | 18.169 | .000 |
| buy only new products tried & tested as safe | 5.682 | 1 | 1.627 | 38 | 3.493 | .069 |
| I know most cosmetic brands | 7.471 | 1 | 1.288 | 38 | 5.802 | .021 |
| one company cannot provide complete personalcare solution | 2.945 | 1 | 1.512 | 38 | 1.948 | .171 |
| plan shopping trips carefully | 2.231 | 1 | 1.233 | 38 | 1.809 | .187 |
| personal care companies need to do a lot of research | 19.236 | 1 | 1.273 | 38 | 15.114 | .000 |
| important to look good and presentable in todays times | 5.531 | 1 | 1.128 | 38 | 4.903 | .033 |
| like experimenting with new trends and style | 10.925 | 1 | 1.460 | 38 | 7.484 | .009 |
| go by what filmstars endorse | 15.531 | 1 | 1.128 | 38 | 13.767 | .001 |
| what i wear reflects who I am | 31.289 | 1 | 1.029 | 38 | 30.400 | .000 |

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

**Number of Cases in each Cluster**

| | | |
|---|---|---|
| Cluster | 1 | 22.000 |
| | 2 | 18.000 |
| Valid | | 40.000 |
| Missing | | .000 |

## Inferences:

1. There are 40 respondents. No missing cases.

2. From the **Agglomeration Schedule** table, we can see that 6 and 40 are most similar pair in stage 1. They are joined with 10 in 24$^{th}$ stage. 5 and 39 are the next similar pair; they are joined with 23, in the 28$^{th}$ stage. In total 39 clustering stages occur.
   We can see from the data above that the maximum variation happens when we move from a one cluster solution to a two-cluster solution. Two clusters are

adequate and distinct enough for analysis. Or 40 respondents who took the survey can be grouped into two distinct clusters.

3. The **Dendogram** clearly gives two clusters that are distinctly different from each other.

4. Examine F values from the **ANOVA table** to establish the discriminating power of each clustering variable. As observed from the table, not all the variables are significant at the 5 percent level of significance and only 4 variables can be used for the interpretation.

5. From the **Final Cluster Centers** table, we cannot infer anything clearly.

6. Even though the **Cluster Membership table,** groups the variables into Cluster 1 & Cluster 2, the Final Cluster Centers table is not so clear on the classifying the variables into clusters.

## Discriminant Analysis

**Group Statistics**

| Time spent on week days | | Mean | Std. Deviation | Valid N (listwise) Unweighted | Valid N (listwise) Weighted |
|---|---|---|---|---|---|
| Less than 1 hour | Linking with Professional | 2.6875 | 1.46876 | 32 | 32.000 |
| | Messaging | 3.2188 | 1.31332 | 32 | 32.000 |
| | Networking | 3.6875 | 1.22967 | 32 | 32.000 |
| | Make new friends | 3.3438 | 1.35859 | 32 | 32.000 |
| | Promote events | 2.8438 | 1.19432 | 32 | 32.000 |
| | Blogging | 2.6875 | 1.11984 | 32 | 32.000 |
| | News Updates | 2.8438 | 1.11034 | 32 | 32.000 |
| | Games | 3.0313 | 1.33161 | 32 | 32.000 |
| | Education | 3.0000 | 1.21814 | 32 | 32.000 |
| | Photo sharing | 4.0625 | .91361 | 32 | 32.000 |
| | Job seeking | 2.8125 | 1.25563 | 32 | 32.000 |
| | Online dating | 3.0625 | 1.47970 | 32 | 32.000 |
| 1- less than 3 hours | Linking with Professional | 2.6667 | 1.32842 | 18 | 18.000 |
| | Messaging | 3.3333 | 1.08465 | 18 | 18.000 |
| | Networking | 3.7778 | .73208 | 18 | 18.000 |
| | Make new friends | 3.1667 | 1.15045 | 18 | 18.000 |
| | Promote events | 3.3333 | 1.08465 | 18 | 18.000 |
| | Blogging | 3.3889 | 1.03690 | 18 | 18.000 |
| | News Updates | 3.1667 | .92355 | 18 | 18.000 |
| | Games | 3.2778 | 1.27443 | 18 | 18.000 |
| | Education | 3.2778 | 1.12749 | 18 | 18.000 |
| | Photo sharing | 3.8889 | 1.02262 | 18 | 18.000 |
| | Job seeking | 2.6667 | .76696 | 18 | 18.000 |
| | Online dating | 2.6111 | 1.46082 | 18 | 18.000 |
| 3-less than 5 hours | Linking with Professional | 3.6667 | .57735 | 3 | 3.000 |

| | | | | | |
|---|---|---|---|---|---|
| | Messaging | 4.3333 | .57735 | 3 | 3.000 |
| | Networking | 4.6667 | .57735 | 3 | 3.000 |
| | Make new friends | 4.3333 | .57735 | 3 | 3.000 |
| | Promote events | 4.0000 | .00000 | 3 | 3.000 |
| | Blogging | 4.0000 | .00000 | 3 | 3.000 |
| | News Updates | 3.3333 | .57735 | 3 | 3.000 |
| | Games | 3.3333 | .57735 | 3 | 3.000 |
| | Education | 3.0000 | 1.00000 | 3 | 3.000 |
| | Photo sharing | 3.6667 | 1.52753 | 3 | 3.000 |
| | Job seeking | 3.6667 | .57735 | 3 | 3.000 |
| | Online dating | 2.6667 | 1.52753 | 3 | 3.000 |
| More than 5 hours | Linking with Professional | 3.1250 | 1.35620 | 8 | 8.000 |
| | Messaging | 3.5000 | 1.19523 | 8 | 8.000 |
| | Networking | 3.5000 | 1.41421 | 8 | 8.000 |
| | Make new friends | 2.8750 | 1.12599 | 8 | 8.000 |
| | Promote events | 2.8750 | .99103 | 8 | 8.000 |
| | Blogging | 2.7500 | 1.03510 | 8 | 8.000 |
| | News Updates | 2.3750 | 1.06066 | 8 | 8.000 |
| | Games | 3.2500 | 1.28174 | 8 | 8.000 |
| | Education | 2.7500 | 1.28174 | 8 | 8.000 |
| | Photo sharing | 3.8750 | .83452 | 8 | 8.000 |
| | Job seeking | 3.1250 | 1.35620 | 8 | 8.000 |
| | Online dating | 2.8750 | 1.64208 | 8 | 8.000 |
| Total | Linking with Professional | 2.7869 | 1.37979 | 61 | 61.000 |
| | Messaging | 3.3443 | 1.20948 | 61 | 61.000 |
| | Networking | 3.7377 | 1.10908 | 61 | 61.000 |
| | Make new friends | 3.2787 | 1.25341 | 61 | 61.000 |
| | Promote events | 3.0492 | 1.13176 | 61 | 61.000 |
| | Blogging | 2.9672 | 1.11006 | 61 | 61.000 |
| | News Updates | 2.9016 | 1.04411 | 61 | 61.000 |
| | Games | 3.1475 | 1.26275 | 61 | 61.000 |
| | Education | 3.0492 | 1.17511 | 61 | 61.000 |
| | Photo sharing | 3.9672 | .94811 | 61 | 61.000 |
| | Job seeking | 2.8525 | 1.12303 | 61 | 61.000 |
| | Online dating | 2.8852 | 1.47307 | 61 | 61.000 |

**Tests of Equality of Group Means**

| | Wilks' Lambda | F | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Linking with Professional | .967 | .656 | 3 | 57 | .582 |
| Messaging | .959 | .821 | 3 | 57 | .488 |
| Networking | .957 | .847 | 3 | 57 | .474 |
| Make new friends | .947 | 1.065 | 3 | 57 | .371 |
| Promote events | .925 | 1.539 | 3 | 57 | .214 |
| Blogging | .874 | 2.728 | 3 | 57 | .052 |
| News Updates | .937 | 1.287 | 3 | 57 | .288 |
| Games | .990 | .186 | 3 | 57 | .906 |
| Education | .979 | .408 | 3 | 57 | .748 |
| Photo sharing | .986 | .264 | 3 | 57 | .851 |
| Job seeking | .957 | .854 | 3 | 57 | .470 |
| Online dating | .981 | .372 | 3 | 57 | .773 |

## Pooled Within-Groups Matrices

| | | Linking with Professional | Messaging | Networking | Make new friends | Promote events | Blogging | News Updates | Games | Education | Photo sharing | Job seeking | Online dating |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Correlation | Linking with Professional | 1.000 | -.114 | -.128 | -.192 | .277 | .159 | .187 | -.097 | .390 | -.003 | .455 | -.025 |
| | Messaging | -.114 | 1.000 | .815 | .136 | -.109 | .025 | -.080 | -.182 | -.105 | .449 | -.280 | -.167 |
| | Networking | -.128 | .815 | 1.00 | .024 | -.109 | -.053 | .075 | -.076 | -.117 | .413 | -.280 | -.229 |
| | Make new friends | -.192 | .136 | .024 | 1.000 | .108 | .171 | -.078 | .367 | .023 | .056 | .069 | .484 |
| | Promote events | .277 | -.109 | -.109 | .108 | 1.000 | .676 | .418 | .375 | .343 | .177 | .226 | .226 |
| | Blogging | .159 | .025 | -.053 | .171 | .676 | 1.000 | .525 | .420 | .298 | .190 | .186 | .265 |
| | News Updates | .187 | -.080 | .075 | -.078 | .418 | .525 | 1.000 | .476 | .339 | .286 | .271 | .013 |
| | Games | -.097 | -.182 | -.076 | .367 | .375 | .420 | .476 | 1.000 | .298 | .226 | .265 | .441 |
| | Education | .390 | -.105 | -.117 | .023 | .343 | .298 | .339 | .298 | 1.000 | .096 | .299 | .174 |
| | Photo sharing | -.003 | .449 | .413 | .056 | .177 | .190 | .286 | .226 | .096 | 1.000 | .137 | .117 |
| | Job seeking | .455 | -.280 | -.280 | .069 | .226 | .186 | .271 | .265 | .299 | .137 | 1.000 | .279 |
| | Online dating | -.025 | -.167 | -.229 | .484 | .226 | .265 | .013 | .441 | .174 | .117 | .279 | 1.000 |

## Eigenvalues

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | .351[a] | 48.4 | 48.4 | .510 |
| 2 | .254[a] | 35.0 | 83.3 | .450 |
| 3 | .121[a] | 16.7 | 100.0 | .329 |

a. First 3 canonical discriminant functions were used in the analysis.

## Wilks' Lambda

| Test of Function(s) | Wilks' Lambda | Chi-square | df | Sig. |
|---|---|---|---|---|
| 1 through 3 | .526 | 33.367 | 36 | .594 |
| 2 through 3 | .711 | 17.705 | 22 | .723 |
| 3 | .892 | 5.938 | 10 | .820 |

## Standardized Canonical Discriminant Function Coefficients

| | Function 1 | Function 2 | Function 3 |
|---|---|---|---|
| Linking with Professional | .078 | -.363 | .308 |
| Messaging | .450 | -1.147 | -.109 |
| Networking | .326 | .844 | .391 |

| | | | |
|---|---|---|---|
| Make new friends | .154 | .674 | .616 |
| Promote events | .280 | .015 | .105 |
| Blogging | .528 | .231 | -.619 |
| News Updates | -.130 | .605 | .265 |
| Games | .173 | -.923 | -.414 |
| Education | -.173 | .397 | -.352 |
| Photo sharing | -.685 | .172 | .013 |
| Job seeking | .481 | -.263 | .405 |
| Online dating | -.388 | -.046 | .334 |

**Structure Matrix**

| | Function | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Blogging | .559* | .246 | -.391 |
| Promote events | .441* | .181 | -.189 |
| Messaging | .330* | -.085 | .159 |
| Networking | .296* | .183 | .207 |
| Linking with Professional | .238* | -.190 | .211 |
| Photo sharing | -.184* | .052 | .103 |
| News Updates | .207 | .441* | -.167 |
| Make new friends | .205 | .242 | .466* |
| Job seeking | .227 | -.202 | .369* |
| Online dating | -.163 | -.032 | .287* |
| Education | .046 | .226 | -.254* |
| Games | .129 | -.039 | -.170* |

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.
*. Largest absolute correlation between each variable and any discriminant function

**Canonical Discriminant Function Coefficients**

| | Function | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Linking with Professional | .056 | -.261 | .221 |
| Messaging | .371 | -.944 | -.089 |
| Networking | .293 | .758 | .351 |
| Make new friends | .123 | .538 | .492 |
| Promote events | .250 | .014 | .094 |
| Blogging | .496 | .217 | -.582 |
| News Updates | -.126 | .584 | .255 |
| Games | .134 | -.716 | -.321 |
| Education | -.145 | .332 | -.295 |
| Photo sharing | -.709 | .178 | .013 |
| Job seeking | .426 | -.233 | .359 |
| Online dating | -.259 | -.031 | .223 |
| (Constant) | -2.396 | -1.808 | -2.357 |

Unstandardized coefficients

**Classification Results**[a,c]

| | | Time spent on week days | Predicted Group Membership | | | | Total |
|---|---|---|---|---|---|---|---|
| Original | Count | Less than 1 hour | 15 | 7 | 3 | 7 | 32 |
| | | 1- less than 3 hours | 4 | 9 | 2 | 3 | 18 |
| | | 3-less than 5 hours | 0 | 0 | 3 | 0 | 3 |
| | | More than 5 hours | 0 | 0 | 2 | 6 | 8 |
| | % | Less than 1 hour | 46.9 | 21.9 | 9.4 | 21.9 | 100.0 |
| | | 1- less than 3 hours | 22.2 | 50.0 | 11.1 | 16.7 | 100.0 |
| | | 3-less than 5 hours | .0 | .0 | 100.0 | .0 | 100.0 |
| | | More than 5 hours | .0 | .0 | 25.0 | 75.0 | 100.0 |
| Cross-validated[b] | Count | Less than 1 hour | 14 | 7 | 4 | 7 | 32 |
| | | 1- less than 3 hours | 7 | 4 | 2 | 5 | 18 |
| | | 3-less than 5 hours | 0 | 2 | 1 | 0 | 3 |
| | | More than 5 hours | 2 | 2 | 3 | 1 | 8 |
| | % | Less than 1 hour | 43.8 | 21.9 | 12.5 | 21.9 | 100.0 |
| | | 1- less than 3 hours | 38.9 | 22.2 | 11.1 | 27.8 | 100.0 |
| | | 3-less than 5 hours | .0 | 66.7 | 33.3 | .0 | 100.0 |
| | | More than 5 hours | 25.0 | 25.0 | 37.5 | 12.5 | 100.0 |

a. 54.1% of original grouped cases correctly classified.
b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
c. 32.8% of cross-validated grouped cases correctly classified.

## Inferences:

1. The **Group Statistics** table computes the mean values to get an idea of the differences in the mean score of variables for each classifier in Time spent on week days. The mean score for messaging for the less than 1 hour group is 2.6875, whereas for the less than 3 hours group it is 3.333. Therefore we can expect that messaging could be useful in discriminating between prospective buyers and non-buyers.

2. From the table of **Tests of Equality of Group Means** table, there does not seem to be any significant difference in the means of the variables as the p value in each of these cases is greater than 0.05.

3. From the table of **Pooled within-Groups Matrices** table**,** the correlation between any pair of predictor variables does exceed 0.75. So, there is problem of multicollinearity.

4. From the table of **Canonical Discriminant Function Coefficients,** the discriminant function can be determined as Y which will be the function of all the variables.

5. From the **Eigenvalues** table, the last column, canonical correlation, is the simple correlation coefficient between the discriminant score and their corresponding group membership.

6. From the **Wilk's Lambda** table, the value of Wilks' Lambda is 0.594. This parameter takes a value between 0 and 1. Lower the value of Wilks' Lambda, the higher is the significance of the discriminant function. Therefore 0 would be the most preferred one. The statistical significance of Wilks' Lambda is carried out with the chi-square statistic (33.367) and its p value. Since p value is not less than 0.05, the discriminant function is not significant.

7. From the **Group Centroids** table, we can compute the mean discriminant scores of the three groups separately. This is called group centroids.

8. **The Standardized Canonical Discriminant Function Coefficients** table indicates that messaging is the most important characteristic, which discriminates between the groups.

9. From the **Structure Matrix,** the structural coefficients are obtained by computing the correlation between the discriminant score and each of the independent variables. These are also called discriminant loadings.

10. From the **Classification Results** table,
    Hit ratio = No. of correct predictions / Total number of cases
    Hit ratio = 54.1 %

## Discriminant Analysis

**Group Statistics**

| Frequency of eating RTE food | | Mean | Std. Deviation | Valid N (listwise) Unweighted | Valid N (listwise) Weighted |
|---|---|---|---|---|---|
| Rarely | Convenience of use | 4.19 | .624 | 36 | 36.000 |
| | Makes work easy | 4.03 | .609 | 36 | 36.000 |
| | Time saving | 4.11 | .667 | 36 | 36.000 |
| | Ease of availability | 3.47 | 1.183 | 36 | 36.000 |
| | Price reasonability | 2.83 | 1.056 | 36 | 36.000 |
| | Have adequate amount of nutrition/calories | 2.50 | .971 | 36 | 36.000 |
| | Taste as compared to freshly cooked food | 2.03 | .910 | 36 | 36.000 |

| | | | | | |
|---|---|---|---|---|---|
| | Manufacturing according to acceptable quality standards | 3.42 | .692 | 36 | 36.000 |
| | Option while travelling | 4.11 | .667 | 36 | 36.000 |
| | Whether making chapatti takes a significant time while eating RTE curry | 2.83 | .971 | 36 | 36.000 |
| Weekly | Convenience of use | 4.38 | 1.025 | 16 | 16.000 |
| | Makes work easy | 4.19 | .750 | 16 | 16.000 |
| | Time saving | 4.31 | .704 | 16 | 16.000 |
| | Ease of availability | 3.50 | .816 | 16 | 16.000 |
| | Price reasonability | 3.50 | .816 | 16 | 16.000 |
| | Have adequate amount of nutrition/calories | 2.63 | .619 | 16 | 16.000 |
| | Taste as compared to freshly cooked food | 3.06 | 1.124 | 16 | 16.000 |
| | Manufacturing according to acceptable quality standards | 3.19 | .981 | 16 | 16.000 |
| | Option while travelling | 3.69 | .873 | 16 | 16.000 |
| | Whether making chapatti takes a significant time while eating RTE curry | 2.63 | .806 | 16 | 16.000 |
| Regularly | Convenience of use | 3.83 | 1.169 | 6 | 6.000 |
| | Makes work easy | 4.17 | 1.169 | 6 | 6.000 |
| | Time saving | 4.17 | 1.169 | 6 | 6.000 |
| | Ease of availability | 4.00 | 1.095 | 6 | 6.000 |
| | Price reasonability | 3.67 | .516 | 6 | 6.000 |
| | Have adequate amount of nutrition/calories | 2.83 | 1.169 | 6 | 6.000 |
| | Taste as compared to freshly cooked food | 2.67 | 1.033 | 6 | 6.000 |
| | Manufacturing according to acceptable quality standards | 3.17 | .753 | 6 | 6.000 |
| | Option while travelling | 3.50 | 1.378 | 6 | 6.000 |
| | Whether making chapatti takes a significant time while eating RTE curry | 3.00 | .894 | 6 | 6.000 |
| Total | Convenience of use | 4.21 | .811 | 58 | 58.000 |
| | Makes work easy | 4.09 | .708 | 58 | 58.000 |
| | Time saving | 4.17 | .729 | 58 | 58.000 |
| | Ease of availability | 3.53 | 1.080 | 58 | 58.000 |
| | Price reasonability | 3.10 | 1.003 | 58 | 58.000 |
| | Have adequate amount of nutrition/calories | 2.57 | .901 | 58 | 58.000 |
| | Taste as compared to freshly cooked food | 2.38 | 1.073 | 58 | 58.000 |
| | Manufacturing according to acceptable quality standards | 3.33 | .781 | 58 | 58.000 |
| | Option while travelling | 3.93 | .835 | 58 | 58.000 |
| | Whether making chapatti takes a significant time while eating RTE curry | 2.79 | .913 | 58 | 58.000 |

## Tests of Equality of Group Means

| | Wilks' Lambda | F | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Convenience of use | .965 | .983 | 2 | 55 | .381 |
| Makes work easy | .989 | .317 | 2 | 55 | .729 |
| Time saving | .985 | .414 | 2 | 55 | .663 |
| Ease of availability | .978 | .617 | 2 | 55 | .543 |
| Price reasonability | .877 | 3.850 | 2 | 55 | .027 |
| Have adequate amount of nutrition/calories | .986 | .387 | 2 | 55 | .681 |
| Taste as compared to freshly cooked food | .811 | 6.411 | 2 | 55 | .003 |
| Manufacturing according to acceptable quality standards | .978 | .610 | 2 | 55 | .547 |
| Option while travelling | .919 | 2.435 | 2 | 55 | .097 |
| Whether making chapatti takes a significant time while eating RTE curry | .984 | .451 | 2 | 55 | .639 |

## Pooled Within-Groups Matrices

| | | Convenience of use | Makes work easy | Time saving | Ease of availability | Price reasonability | Have adequate amount of nutrition/calories | Taste as compared to freshly cooked food | Manufacturing according to acceptable quality standards | Option while travelling | Whether making chapatti takes a significant time while eating RTE curry |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Correlation | Convenience of use | 1.000 | .589 | .350 | .035 | .066 | -.010 | .162 | -.082 | .071 | .059 |
| | Makes work easy | .589 | 1.000 | .475 | .234 | .159 | .050 | .029 | .189 | .010 | -.185 |
| | Time saving | .350 | .475 | 1.000 | .423 | .245 | .164 | .048 | .164 | .261 | -.039 |
| | Ease of availability | .035 | .234 | .423 | 1.000 | .592 | .101 | -.289 | .115 | .013 | .051 |
| | Price reasonability | .066 | .159 | .245 | .592 | 1.000 | .452 | -.097 | .129 | .051 | -.103 |
| | Have adequate amount of nutrition/calories | -.010 | .050 | .164 | .101 | .452 | 1.000 | .173 | .223 | -.009 | -.287 |
| | Taste as compared to freshly cooked food | .162 | .029 | .048 | -.289 | -.097 | .173 | 1.000 | -.218 | .095 | -.029 |
| | Manufacturing according to acceptabl | -.082 | .189 | .164 | .115 | .129 | .223 | -.218 | 1.000 | .221 | -.335 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | e quality standards | | | | | | | | | |
| | Option while travelling | .071 | .010 | .261 | .013 | .051 | -.009 | .095 | .221 | 1.000 | -.150 |
| | Whether making chapatti takes a significant time while eating RTE curry | .059 | -.185 | -.039 | .051 | -.103 | -.287 | -.029 | -.335 | -.150 | 1.000 |

**Eigenvalues**

| Function | Eigenvalue | % of Variance | Cumulative % | Canonical Correlation |
|---|---|---|---|---|
| 1 | .641[a] | 85.4 | 85.4 | .625 |
| 2 | .110[a] | 14.6 | 100.0 | .315 |

a. First 2 canonical discriminant functions were used in the analysis.

**Wilks' Lambda**

| Test of Function(s) | Wilks' Lambda | Chi-square | df | Sig. |
|---|---|---|---|---|
| 1 through 2 | .549 | 30.274 | 20 | .066 |
| 2 | .901 | 5.263 | 9 | .811 |

**Standardized Canonical Discriminant Function Coefficients**

| | Function | |
|---|---|---|
| | 1 | 2 |
| Convenience of use | -.183 | -.921 |
| Makes work easy | .011 | .741 |
| Time saving | .281 | -.327 |
| Ease of availability | -.259 | .385 |
| Price reasonability | .847 | -.082 |
| Have adequate amount of nutrition/calories | -.471 | .435 |
| Taste as compared to freshly cooked food | .753 | -.072 |
| Manufacturing according to acceptable quality standards | .014 | -.190 |
| Option while travelling | -.556 | .005 |
| Whether making chapatti takes a significant time while eating RTE curry | -.152 | .539 |

**Structure Matrix**

| | Function | |
|---|---|---|
| | 1 | 2 |
| Taste as compared to freshly cooked food | .596* | -.216 |

| | | |
|---|---|---|
| Price reasonability | .456* | .246 |
| Option while travelling | -.354* | -.272 |
| Manufacturing according to acceptable quality standards | -.184* | -.062 |
| Time saving | .142* | -.139 |
| Makes work easy | .134* | .004 |
| Convenience of use | .046 | -.560* |
| Ease of availability | .079 | .409* |
| Whether making chapatti takes a significant time while eating RTE curry | -.084 | .329* |
| Have adequate amount of nutrition/calories | .117 | .220* |

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions

Variables ordered by absolute size of correlation within function.

*. Largest absolute correlation between each variable and any discriminant function

**Canonical Discriminant Function Coefficients**

| | Function | |
|---|---|---|
| | 1 | 2 |
| Convenience of use | -.226 | -1.135 |
| Makes work easy | .016 | 1.034 |
| Time saving | .381 | -.444 |
| Ease of availability | -.238 | .354 |
| Price reasonability | .886 | -.086 |
| Have adequate amount of nutrition/calories | -.517 | .478 |
| Taste as compared to freshly cooked food | .765 | -.073 |
| Manufacturing according to acceptable quality standards | .018 | -.242 |
| Option while travelling | -.683 | .006 |
| Whether making chapatti takes a significant time while eating RTE curry | -.164 | .584 |
| (Constant) | -.020 | -.484 |

Unstandardized coefficients

**Classification Results[a,c]**

| | | Frequency of eating RTE food | Predicted Group Membership | | | |
|---|---|---|---|---|---|---|
| | | | Rarely | Weekly | Regularly | Total |
| Original | Count | Rarely | 26 | 6 | 4 | 36 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | Weekly | 1 | 12 | 3 | 16 |
| | | Regularly | 2 | 1 | 3 | 6 |
| | % | Rarely | 72.2 | 16.7 | 11.1 | 100.0 |
| | | Weekly | 6.3 | 75.0 | 18.8 | 100.0 |
| | | Regularly | 33.3 | 16.7 | 50.0 | 100.0 |
| Cross-validated[b] | Count | Rarely | 23 | 7 | 6 | 36 |
| | | Weekly | 3 | 5 | 8 | 16 |
| | | Regularly | 3 | 2 | 1 | 6 |
| | % | Rarely | 63.9 | 19.4 | 16.7 | 100.0 |
| | | Weekly | 18.8 | 31.3 | 50.0 | 100.0 |
| | | Regularly | 50.0 | 33.3 | 16.7 | 100.0 |

a.70.7% of original grouped cases correctly classified.
b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
c.50.0% of cross-validated grouped cases correctly classified.

## Inferences:

1. The **Group Statistics** table computes the mean values to get an idea of the differences in the mean score of variables for rarely, weekly and regularly. The mean score for convenience of use for the rarely group is 4.19, whereas for the weekly group it is 4.38. For other characteristics the mean values are closer for buyer and non-buyer. Therefore we can expect that convenience of use could be useful in discriminating between rarely, weekly and regularly.

2. From the table of **Tests of Equality of Group Means** table, it is observed that the significant difference in the mean exists for the Taste as compared to freshly cooked food only, for which the p value is less than 0.05, the assumed level of significance. There does not seem to be any significant difference in the means of the remaining characteristics as the p value in each of these cases is greater than 0.05.

3. From the table of **Pooled within-Groups Matrices** table**,** the correlation between any pair of predictor variables does not exceed 0.75. So, there is no problem of multicollinearity.

11. From the table of **Canonical Discriminant Function Coefficients,** the discriminant function can be determined as Y which will be the function of all the variables.

4. From the **Eigenvalues** table, the last column, canonical correlation, is the simple correlation coefficient between the discriminant score and their corresponding group membership

5. From the **Wilk's Lambda** table, the value of Wilks' Lambda is 0.549. This parameter takes a value between 0 and 1. Lower the value of Wilks' Lambda, the higher is the significance of the discriminant function. Therefore 0 would be the most preferred one. The statistical significance of Wilks' Lambda is carried out with the chi-square statistic (30.274) and its p value. Since p value is not less than 0.05, the discriminant function is not significant.

6. From the **Group Centroids** table, we can compute the mean discriminant scores of the rarely, weekly and regularly groups separately. This is called group centroids. This is used for designing a decision rule to classify the variables into groups.

7. **The Standardized Canonical Discriminant Function Coefficients** table indicates that Taste as compared to freshly cooked food is the most important characteristic, which discriminates between rarely, weekly and regularly groups.

8. From the **Structure Matrix,** the structural coefficients are obtained by computing the correlation between the discriminant score and each of the independent variables. These are also called discriminant loadings.

9. From the **Classification Results** table,
   Hit ratio = No. of correct predictions / Total number of cases
   Hit ratio = 70.7 percent.